

The cases: Monitoring and evaluation

Case 15: Measuring the impact of the PROSPER Project in Romania: A learning experience

Mirela Bardi

The project

PROSPER (Project for Special Purpose English in Romania) was set up in 1991 with the broad aim of improving the level of English proficiency of future personnel in key sectors of the Romanian economy and public life such as engineering, economics, public administration, and medicine. Upgrading and diversifying the teaching/learning of English for Specific Purposes (ESP) in major tertiary education institutions represented a major step in achieving the project aim. Through its diverse professional activities, the project has had effects and implications which have extended well beyond the language classroom.

The project brought together teachers from various institutions and enabled them to work together to accomplish tasks that were relevant to everybody working in ESP, regardless of the institution of origin. Participation in project events has helped teachers to develop skills in areas such as syllabus and course design, textbook writing, classroom research, teacher training, and educational management. Working together to achieve common goals created a professional culture among ESP teachers, who shared the same values and confronted the same set of problems in their local environments. The project proposed its own structures, activities, channels of communication, roles, and responsibilities. It did so at a time when Romanian institutions were beginning to strive to generate new modes of working which were more flexible and open to novelty. Above all, the project created opportunities for learning and development, and aimed to make the need for continuous improvement irreversible. The project members had the freedom and responsibility to make the project work for themselves.

Stage – monitoring and evaluation

This case study focuses on the monitoring and evaluation of the project. This was the most complex and challenging project activity, with the study of impact being carried out over a period of three-and-a-half years (1996–99). The study aimed to capture evidence of expected and unexpected project impact on the teaching/learning

process and on the major stakeholders involved: teachers, students, foreign languages departments, the British Council as the managing institution, and members of the ESP teaching profession in Romania.

This case study sets out to discuss the main aspects of the impact study process and to highlight the lessons learnt by the project team from the actual experience of conducting the study. I will argue that the lessons derived from this experience were both personal and institutional, in the sense that aspects of process management may inform the practice of managing innovation projects.

Story

We aimed to do a study of impact, rather than a standard end-of-project evaluation, because we felt that this latter kind of evaluation was, in a sense, too narrow and maybe too much linked to the original objectives of the project. For such a vast project as PROSPER (spanning over nine years and involving 16 universities), a study aiming to document impact on the numerous categories of stakeholders seemed to be more appropriate and more likely to yield valuable information. Although doing an evaluation was included among the project outputs, the decision to do an insider-led evaluation and the shape of the study depended on the way the project had evolved since its initial design. The collaborative project culture of consulting and involving project members in decision making about project events, and the need to continue and diversify the professional development of the project team were the main reasons for the choice of participatory evaluation, which was regarded as a new learning opportunity.

Rationale and design

The design of the study took account of two main features that differentiate impact studies from other forms of evaluation: the need to gather comparable data and to discuss unexpected ripple effects. In our case, the comparable data came from two sources. One was the baseline study done at the beginning of the project, which describes the situation as it was before the project influence began; the other was the data we collected during the impact study from institutions which were not involved in PROSPER. Impact was, therefore, measured both in terms of how the PROSPER group of institutions had changed when compared with the pre-PROSPER profile and in terms of how the impact of PROSPER had made that group different from groups of teachers who were not affected by the project.

The impact study became a project in its own right. It was an attempt to conduct participatory evaluation, with the project members designing the study framework and carrying out all the stages in the research process: drafting, trialling and revising the instruments, collecting the data, and analysing and writing up the results. However, the complexity of the study represented a considerable challenge for the team, which was entirely unprepared to do evaluation research on a large scale. As the impact study consultant observed at the time of the study publication: 'There is virtually no literature in language education on the design of an actual impact study' (Alderson 1999). The situation may have changed in the meantime, but the lessons we learnt while doing our impact study will hopefully be useful to others who may want to replicate the process or aspects of it.

The impact study also involved outsiders in various roles. The key presence was the impact study consultant, who provided professional and emotional support to a team lacking research training. Critical friends were also called on along the way, and their opinions and advice helped the team distance themselves from their own work and consider it through the perspective of a potential readership.

Methodology

The project team identified the areas of project impact through a brainstorming exercise. In terms of the teaching/learning experience, PROSPER was expected to have had an impact on teachers, students, classroom processes, materials and resources generated by the project members, and testing procedures. Impact was expected to have crossed the border through contacts with ESP project teams in other countries.

Data collection instruments were designed and administered by the impact study team, with one department being responsible for a particular instrument. We started with the collection of quantitative data through a set of five questionnaires (teacher, student, ex-student, employer, and management questionnaires) and a classroom observation schedule, and continued with qualitative data (case studies of graduates and group discussions with members of the core PROSPER departments). Research was carried out in PROSPER and non-PROSPER institutions, and the response rate was very high, as a result of direct administration during institutional visits by members of the impact study team. The following table provides a snapshot of the research instruments and the number of respondents for each of those:

Research instrument	No. of respondents	Response rate
Teacher questionnaire	104 PROSPER and 51 non-PROSPER teachers	92.4%
Classroom observation schedule	59 PROSPER and 25 non-PROSPER classes	100%
Student questionnaire	1,039 PROSPER and 449 non-PROSPER students	100%
Ex-student questionnaire	101 PROSPER and 51 non-PROSPER ex-students	92.7%
Employer questionnaire	46 respondents	76.6%
Case studies of graduates	4 PROSPER and 4 non-PROSPER graduates	
Departmental group discussions	8 institutions (98 PROSPER teachers participated)	92.8%
Comparative analysis of teaching materials	2 PROSPER, 2 pre-PROSPER, and 1 non-PROSPER locally developed textbooks	
Comparative analysis of tests	58 PROSPER, 15 pre-PROSPER, and 17 non-PROSPER tests	
Statements of ripple effects	6 statements from language centres directors, department managers, and managers of ESP projects in the region	

Table 1: Research instruments and respondents

Analysis of the data involved discussing the meaning of the results in terms of the presence or absence of expected impact, and thus trying to detect trends in the innovation process. The team returned to the assumptions underpinning the instruments and reconstructed the rationale of designing the instruments in that particular way. Writing up the study required refocusing and refining the interpretation as well as re-evaluating our project experience. In this recursive process, the innovation was reflected upon and redefined in terms of its meaning for the project team, while efforts were also made to communicate it to the wider professional community.

Lessons

As I have mentioned already, the process of carrying out the PROSPER impact study was essentially a learning experience for the project members – both in terms of their professional practice and of project management. Through its learning outcomes, it consolidated the expertise of the project team, as well as the process of implementing change.

Naturally, the process was not smooth. It involved co-ordination of several institutional teams (each in charge of developing one research instrument) and merging presentation and analysis of all results into a coherent study. Work in small institutional teams alternated with whole-team meetings where everybody had a chance to understand how their work fitted into the wider picture and how the study was coming together. All concerns about the research process - design and administration of instruments, and analysis and interpretation of results – were voiced and addressed during the whole-team meetings. The presence of the impact study consultant at these large meetings facilitated the search for solutions. In retrospect, one can confidently say that seeking solutions together was a form of experiential learning. Team members developed their research skills by engaging with every aspect of the research process in a continuous problem-solving process. The value of experiential learning will be discussed further in this section.

Responsibility and project ownership

The team became involved in the study out of a sense of responsibility towards the project and its successful completion, as well as out of a sense of responsibility towards their own professional development. In our experience, responsibility and ownership are closely connected, and they highlight the complex ties that can exist between projects and project teams, ties that cannot be ‘managed’ through a technical process. Managing projects is very much about building commitment and strong relationships among people and, from a manager’s point of view, it is worthwhile investing in generating and maintaining commitment. Ownership takes time to build, and it was the whole participatory project culture, reinforced over years, that had created that strong sense of ownership.

Interest in the task

The team commitment was fuelled by the interest in the task itself. The task was perceived as interesting and worthwhile but also as contributing to project cohesion. It was an appropriate end-of-project task and an opportunity for project members to find out whether their beliefs about the success of PROSPER were

actually confirmed by data. A useful lesson for project management is that projects need to give the opportunity to the project teams to investigate and reflect on what they have achieved by taking part in the project. Project evaluation can be carried out as a developmental rather than an accountability exercise. In our case, working together to carry out a task that was perceived as valuable by all team members and for whose implementation everyone had clear responsibilities strengthened the feeling of team belonging and identity.

Learning outcomes and learning sources

The team motivation to carry out the study was reinforced by a sense of progress and accomplishment both in relation to the study itself and to team members' own development. The team members developed new skills and new ways of thinking about teaching and research, and increased their awareness of teaching circumstances in many other institutions. While we cannot predict the range of skills that every team member developed, my insider knowledge enables me to make a fair assessment of the impact study experience learning outcomes. The following table indicates the range of skills that participants may have developed, together with the possible sources of learning.

By engaging in the research process, sharing ideas, and discussing work with the whole team, and by talking to other teachers and observing their classes, participants were able to extend and enhance the following skills:

- Professional skills
 - teaching skills
 - research skills
 - critical thinking
 - project management skills

- Social skills
 - communication skills
 - interpersonal skills
 - teamwork skills
 - strategic competence
 - coping with challenging/unexpected situations

- Awareness of
 - the PROSPER team
 - other teachers and teaching circumstances
 - one's own abilities.

Participants learnt for themselves, but also for the group, in the sense of developing abilities for collaborative work. I believe I can justifiably claim that one major aspect of learning was reinforcing the sense of group identity and of confidence in what people can achieve by working as a group.

The role of the team

Learning from experience is another major theme of the impact study process. The team learnt by conducting the research process, but also by making decisions about the process. Uncertainty was present along the road, but the moments when solutions were being sought turned out to be valuable learning experiences.

The team itself was a major source of learning: learning from peer opinions, from team discussions and decision-making, as well as from the good work or mistakes of colleagues. The presence of the team was often a source of energy, because all the members were motivated by a common purpose. We need to be aware, though, that sometimes, feelings of group belonging may actually limit the freedom of individuals, precisely because of the desire to belong and to be accepted. In project management terms, our experience indicated that teams can be both a source of strength or confidence and a source of pressure for some of their members.

Process management

Accomplishment of such a complex task required relentless co-ordination. Apart from the challenge of conducting valuable research, there was also the challenge of managing the process in a way that ensured the quality and coherence of the work. One key aspect of the management process was the provision of support whenever the team or individual members needed it. Support often came through a joint quest for answers (what some of the team members called 'indirect training'), rather than through offering solutions, but what mattered was addressing the issues as they came up. It needs to be said that the project managers were members of the research team and fulfilled both a professional and an administrative role. Involvement of project managers in the research process increased the team's confidence in their ability to supervise the process competently, while making the managers themselves more knowledgeable about the demands of the research process and the needs of the team. In terms of project management lessons, one could say that, in the case of very demanding tasks, it is important for project managers to have an overall view of progress and to pull the strings together when needed, but they can only do so if they fully participate in task completion. Accomplishment of their managing role (planning and organisation of meetings, setting deadlines, and communicating with team members) was, in our case, facilitated by full professional involvement in the research process.

Managing change and managing learning

Change and learning are two key themes that describe the process and outcomes of conducting the study by the project team. They often appear in tandem in the literature of change management, where change seems to be identified with learning experiences:

Change is fundamentally and profoundly about learning – the two are linked in a way that makes them symbiotic. (O'Sullivan and West-Burnham 1998: 45).

In our experience, an essential aspect of change management is to create opportunities for those involved to develop, to gain new insights into themselves and their circumstances, and to empower them to improve those circumstances, if they feel the need to do so. Change becomes sustainable when participants

internalise the experience and translate it into the way they work, into patterns of behaviour and belief, and into perceptions of their role. This process of assimilating or rejecting elements of innovation is essentially a process of learning, and carrying out an impact study was a worthwhile and rewarding learning experience for the PROSPER team.

The following are some of the key management lessons derived from our experience:

1. Project ownership and responsibility are constantly created through the whole project culture of participation, sharing, and transparency. They are the result and also evidence of a participatory culture which needs to be built throughout the project lifetime.
2. Project managers have a crucial role in facilitating communication among team members and in ensuring that there is complete transparency, for all team members, about the implementation process. Communication and transparency generate better understanding and facilitate the process of task implementation.
3. It is very difficult to plan for learning, as learning outcomes of complex project activities are wider than initially envisaged. The role of management is to suggest tasks that are perceived as meaningful by the project teams and to facilitate participation and involvement.
4. Projects that respond to a perceived priority need are more likely to mobilise team effort and commitment.
5. Project teams do not emerge by mandate or team-building exercises, but are consolidated in the process of working together to accomplish a task perceived as meaningful/rewarding.
6. Delegation of responsibility to project teams has an essential role in their learning, as more responsibility generates more commitment and eventually more learning.

Innovations and change can be planned in a more or less rational, linear way, with identification of problems and needs, the planning of action, implementation of the plans, and evaluation of results (Fullan and Stiegelbauer 1991; Burnes 2000). Although these stages are useful in guiding our thinking, it is simplistic to assume that implementing an innovation is a matter of following the predesigned steps in a process. Essential variables such as the values and attitudes of project members can determine unexpected courses of actions and outcomes.

Change can, therefore, be regarded less as a rational, technicist process and more as contextually sensitive and involving human variables. The management of change will be specific to each situation, will require recognition of the fact that change involves human variables such as trust, confidence, and feelings of insecurity, and will move towards a more people-oriented, participatory approach to project management. On the basis of the PROSPER experience, I can, therefore, suggest a paradigm of project management that regards the process of change as a process of learning, not just a method of changing organisational structures and practices.

Case 16: The challenge of monitoring and evaluation in Sri Lanka

Lesley Dick

The project

English was rejected in post-colonial Sri Lanka in favour of a Sinhala Only policy. Currently, though, as Lo Bianco argues in his Sri Lankan language policy history, ‘... the English comeback is remarkable. The language that once provoked fragmentation is now a key to reconciliation’ (Lo Bianco 2011: 52).

Further, as Kennett confirms, ‘...recent amendments to the constitution have brought English back as a link language and English continues to be used as the language of management in public service and of international development’ (Kennett 2010: 319). As Popovici discusses elsewhere in this volume, English literacy creates access to education, information, and technology, creates social equality and cohesion, and promotes local and international communication. In Sri Lanka, inequality of English teaching in urban and rural areas is evident: rural English teachers lack English language proficiency, up-to-date methodology, and often suitable materials and facilities.

The Council for Business with Britain (CBB) English Language Teaching (ELT) Project is an in-service programme for English language teachers designed by the British Council and carried out in association with the Sri Lankan Ministry of Education. A non-profit, non-political organisation focused on expanding bilateral trade links between Sri Lanka and the United Kingdom, CBB has HSBC as principal sponsor, and the project demonstrates its corporate social responsibility. Its goal is to upgrade the English skills of pupils; this is to be achieved through equipping teachers with the skills to deliver high-quality English teaching in their real-life low-resource classrooms. Additionally, the project seeks to redress the imbalance in quality between rural and urban English education with its design as a distance-learning training project that can benefit teachers in rural and urban areas equally. Importantly, from the Ministry of Education’s perspective, the project works within existing school structures and curricula, focusing on long-term, sustainable development, rather than short-term, expensive, resource-based inputs.

Stage – monitoring and evaluation

This case study focuses on the monitoring and evaluation process in the project. The design of the project, which started in 2005, has, after an initial change from short-term weekend workshops to distance teacher training course, changed little from the pilot. The most challenging aspect has been the successful monitoring and evaluation of the impact on classroom teaching skills and on pupil learning. Lessons learnt here point to the importance of Kirkpatrick's four levels of evaluation (Kirkpatrick 1998).

Story

The project narrative is one of a continuing valiant attempt at monitoring and evaluating. The challenges of the four phases summarised in Table 1 will be presented in turn.

	Phase 1	Phase 2	Phase 3	Phase 4
Impact	Aug-Oct 2005	2006-2007	2007-2010	2010-2012
RESCs ¹ involved	6	8	26	28
Mentors trained	0	16	40	1
Teachers trained	150	48	378	323
Pupils reached	12,000	3,840	31,200	42,043

Table 1: Project phases

Phase 1

CBB first approached the British Council in April 2005, with a desire to improve school leavers' English language skills. The project was short-term and low-cost. The 2005 Hornby School *English in the Workplace* was cascaded in six rural areas through weekend workshops targeting secondary English language teachers. The project aimed to train two teachers to cascade the workshops further. It was partly successful: high turnout in some areas; in others, low. Monitoring and evaluation were, in Kirkpatrick terms, level 1 ('reaction'): response questionnaires and attendance, plus informal evaluation by British Council and local trainers. While questionnaires were overwhelmingly positive, low turnout was disappointing. Informal response revealed the root cause: the timing (weekend) clashed with private teaching and led to loss of earnings.

Phase 2

For Phase 2 the British Council proposed a pilot which adapted their existing Teaching Knowledge Test (TKT) Course run by the Teaching Centre as Sunday teacher training workshops for local teachers, into a distance teacher training course. The RESCs, with their national geographic coverage, became the obvious choice as training centres. With CBB consultation, eight districts were selected on the basis of geographic spread. Two RESC trainers from each were chosen to complete

¹ Regional English Support Centres. These are funded by the Ministry of Education and their role is to support primary and secondary teachers in their district in training and resources.

the six-month face-to-face course in Colombo and Kandy, receive mentor training, and then mentor six teachers on the distance course. On the basis of monitoring and evaluation, CBB would decide on whether or not to expand the project.

The primary focus was on measuring the impact on the participants, in Kirkpatrick terms, level 2 ('knowledge'). Monitoring was based on attendance and quality of support for the distance course. Geographic participant spread made it challenging to evaluate. The future mentors attended a TKT face-to-face workshop every Sunday. The British Council trainers could monitor attendance and workshop performance, and report any problems quickly.

Monitoring the distance course was more problematic. It was imperative that British Council trainers could evaluate mentoring skills and study session supervision. British Council trainers visited RESCs once per module to ensure effective mentoring by the RESC trainers and provide them with opinion to ensure their development. Invaluable opinion was collected on materials, tasks, and study sessions. As a direct result of the RESC visits, two extra face-to-face workshops were delivered, portfolio tasks were reworked, two distance course units were revised, and a longer 10-hour mentoring course was developed.

Questionnaires were also used for monitoring both courses. While informative, they tended to be skewed by the inherent Sri Lankan reluctance to give negative opinion. This, plus the distance course participants' stated isolation, prompted the project manager to explore ways of both collecting more accurate opinion and creating a more collective spirit. A focus group at the annual ELT conference became the forum for participants to share views and to collect opinion. TKT as an international guideline and a motivator for participant success was revealed here. To counteract mentor isolation, weekly telephone contact was maintained, ensuring issues were swiftly responded to. This frequent contact built strong relationships and led to more effective monitoring.

Teacher and mentor evaluation used TKT exam performance as the measurement of success. The aim was 80 per cent with band 3 or above and this was easily achieved on the pilot. While this measured 'knowledge', it did not measure 'behaviour' (Kirkpatrick level 3). To demonstrate this, a portfolio was added: for each unit, teachers completed a portfolio task designed to apply the learning to the classroom and evaluate it. Portfolios were monitored by mentors, submitted to the British Council after the course, evaluated, and awarded a pass, weak pass, or strong pass depending on task completion and quality. While the portfolios were evidence of classroom impact, there was no direct evaluation of classroom or pupil performance. Nevertheless, exam results were excellent; portfolios demonstrated evidence of classroom impact; opinion questionnaires were positive; and CBB agreed to extend funding for three years.

Phase 3

Phase 3 saw the project expand from eight to 26 RESCs and from 16 to 40 mentors, who supervised the training of 378 teachers. With this expansion came monitoring and evaluation challenges.

Project coverage increased geographically. Monitoring every RESC module by Colombo-based British Council trainers was financially unrealistic. To reduce costs and to provide evidence of classroom impact, two changes were made. Peer moderation visits replaced British Council trainer visits in 2008. Mentors submitted visit reports for project manager moderation. This proved successful (and much cheaper) and encouraged RESC co-operation. Classroom observation (Kirkpatrick level 3) was also added. Mentors, after becoming a standard part of mentor training, observed three teachers each, before and after the course. Initially seen as developmental and general evidence of progress, these lesson observation reports were later compared to reveal behaviour change, for example, increases in learner-centredness. These were included in the portfolio for end-of-course evaluation.

The funders wanted RESCs in the north and east of Sri Lanka included; the security situation (the civil war having ended in May 2009), therefore, impacted on monitoring and evaluation. In the east, future mentors were traveling to Kandy from Batticaloa and to Colombo from Trincomalee every weekend; ensuring their safety was imperative. In the north, travel to and from Jaffna was severely restricted for local residents. In 2008, the Jaffna RESC co-ordinators commenced the distance course; experienced trainers, they completed the course with the teachers they were mentoring. The delivery of the materials before the course and the administration of the exam after the course by British Council trainers were the only times the RESC was monitored.

Alongside expansion, costs were increasing. The funders wanted better impact measurement and more return on investment. However, they argued against TKT as an evaluation tool because of the exam fees (the most expensive part of the project, despite discounts from Cambridge and the British Council). The case for keeping TKT as an international guideline of participant success was argued strongly on the grounds that the previous questionnaire and focus groups' results revealed the high motivation provided by the Cambridge exam.

While the primary evaluation tool focused on the impact on teachers themselves, the funders wanted evidence of direct impact on the pupils' English language skills. Three evaluation methods were proposed. Testing pupils of project teachers before and after the course using practice KET/PET (Key English Test/Preliminary English Test) tests was rejected as too expensive, because of the oral testing fee of British Council teachers. The O and A level results of schools involved in the project were rejected, as the low numbers of project teachers in any school meant their direct impact on results was impossible to determine. The third method involved collecting school data on the project teachers' classes and comparing this with parallel grades: any noticeable improvement in average marks could be attributed, at least in part, to training. This was accepted.

No primary data had been collected on project teachers' schools or grades. It was now added to application forms and a more detailed database kept. In 2009, the British Council approached the Ministry of Education to assist in obtaining student results for the 2008–09 year end school term English exam, for a sample of project teachers. A data collection form was produced for grade 9 and 10 project teachers

and those in parallel classes. Two sets of results were compared and analysed: one for December 2008 exams before training; one for December 2009 exams after training. Unfortunately, the project design did not readily lend itself to direct classroom impact evaluation of pupil performance. Aside from teaching methodology, other factors could have influenced pupils' performance: lack of an examined speaking component (since the project emphasised speaking skills); general disruption to teaching schedules during the year and teacher absence; lack of exam credibility; poor exam design; and internal issues leading to statistical 'manipulation'. Nevertheless, the conclusion was that the CBB ELT programme led to a positive effect on examination performance by pupils in the two grades. The data supplied, though, led to such a small sample size that conclusions could only be tentative.

Phase 4

In this phase, we saw the inclusion of the RESC in Vavuniya, an area devastated by the civil war. A total of 155 teachers sat TKT in January 2011. Between July 2011 and March 2012, 28 RESCs will run the distance course for 168 teachers.

The monitoring and evaluation of Phase 4 reflects finally the four levels of Kirkpatrick. Level 1 ('reaction') is evaluated through questionnaires. Teachers' beliefs that the distance course is relevant, appropriate, and of good quality, that self-access materials are easy to access and use, that they will apply course methods and have an action plan showing how and when, and that they will recommend the course to their peers are measured. Level 2 ('knowledge') is evaluated through self-assessment and test. In addition to the aim of 80 per cent achieving TKT band 3, their confidence in using communicative methodology and in mentor support is measured. Level 3 ('behaviour') is measured through questionnaire and observation. Questionnaires provide evidence of teachers adding two new activities to classroom practice and an increase in the percentage of classroom English used by teachers and pupils. End-of-course observations hope to demonstrate more use of communicative methodology and tasks than earlier observations. An extra observation three months after the course to evaluate evidence of settled behaviour is also included. Finally, level 4 ('results') is evaluated. At the recruitment stage, a sample of teachers in schools with parallel classes to theirs have their pupils sit the Oxford Placement Test with an additional speaking component before and after the course. By using control groups, as Phillips argues (2003: 117), the impact of the training can be isolated. These results will be compared to give a more accurate measurement of pupil impact. Although there are problems involved in using experimental and control groups for the evaluation of project outcomes in social change projects, as reported by Pawson and Tilley (1997), there is a political need to provide hard numbers.

Lessons

This project story reveals the importance of having a detailed monitoring and evaluation plan, complete with needs and objectives matched to data collection methods, and a schedule in place from project inception. Zikri, in her discussion of the Egyptian curriculum development project in this volume, concurs. A very careful definition of project scope (and, therefore, of evaluation scope) at the

outset is essential. This is also confirmed by Mathew in this volume. Similarly, the need for baseline evaluation in the opening stage of the project is crucial. On the one hand, if your goal is to improve the English skills of pupils, then evaluate that. Evaluating anything else will not reveal true impact. To do this, data collection must be planned and organised from the outset. Here primary data was not initially collected on schools and grades, and thus pupil evaluation was slowed down. On the other hand, although improving student performance might be a long term goal, the immediate project outcome is only going to be improved teacher performance. While this latter is observable and open to evaluation within a reasonable time frame, evaluating student outcomes is (a) very long-term and (b) involves so many variables that are outside the project's control that it is very difficult to evaluate it meaningfully.

Data collected and human resources used should be exploited fully. Here, mentors observed teachers and wrote lesson reports before and after the course, but this data was initially only used as a progress record. Only much later were closer comparisons of the observation reports done on individual teachers to reveal behaviour change. The difference between behaviour change and settled behaviour change was not considered until Phase 4. Had it been built in earlier, stronger impact data would have resulted.

Data collected must be reliable, valid, and accurate. Considerable time and effort were devoted to collecting data on pupil performance from the Ministry of Education without a convincing enough result. Despite the use of a control group, too many other factors could have impacted on pupil evaluation. Although attempts are being made to use control groups more effectively, their use could be seen as a fundamental flaw.

Regular and good methods of communication are essential: while response questionnaires can provide valuable information, face-to-face communication in the form of focus groups can deliver more useful and unpredictable information. This was evident in the informal response in Phase 1, which allowed the project manager to assess accurately the root cause of low turnout, and in the information gleaned from the focus groups in Phase 2. Only the information on the importance the teachers attached to the University of Cambridge exam prevented a potentially disastrous situation from developing when CBB wanted to jettison the TKT exam.

An equally valuable lesson is the unexpected benefits of regular monitoring. While the RESC visits were developed as a mentor performance check, they became a valuable mentor support vehicle, a strong relationship builder of trust between RESCs, Zonal Districts of Education, and the British Council, an information and resource basket to dip into, and a contributor to the strengthening of the respect local teachers had for RESC staff.

This project discussion hopefully points the way for more effective future project monitoring and evaluation.

Case 17: Understanding washback: A case study of a new exam in India

Rama Mathew

The project

The Central Board of Secondary Education (CBSE) undertook a major curriculum renewal project in 1988, namely the CBSE-ELT Project, for English Course A (meant for English-medium students) at the secondary level. The main objective of the project was to improve teaching and learning of English in Classes IX and X with a focus on the development of language skills in communicative situations. The special feature of the project was the intensive involvement of teachers from CBSE schools in all aspects of curriculum development, i.e. designing the new syllabus, the preparation of teaching materials, the new testing scheme and sample papers, and the training manual for orienting teachers to the new curriculum. The new *Interact in English* package¹ was introduced in Class IX in 1993. Before the launch of the new course, about 2,500 teachers (out of about 8,000 teachers who taught at this level in the country) were trained through a 10-day INSET programme by teachers trained for the purpose. The new Class X Board exam was held in March 1995.

Stage – monitoring and evaluation

The Central Institute of English and Foreign Languages (CIEFL), Hyderabad, in keeping with its mandate to improve the standards of teaching English and foreign languages and to conduct advanced training and research in language and literature in India, undertook a national-level study from 1993–97 to monitor, evaluate, and support the CBSE's curriculum renewal project. The national outreach afforded by the CBSE network and its engagement with curriculum renewal were of interest to the Institute. This phase of the project, known as the Curriculum Implementation Study² (CIS), was located in the Evaluation Department at CIEFL, which provided professional expertise with assistance from other colleagues in the Institute. The case study presented here focuses on the monitoring phase of the project.

¹ This consisted of three books: the Main Course Book (MCB), the Work Book (WB), and the Literature Reader (LR), along with the Teachers' Book (TB) and listening materials on cassettes with the tape scripts in the TB.

² The project was funded by the UK Department for International Development (DfID), through the British Council India, while CIS was funded by the CBSE and DfID.

Story

The Curriculum Implementation Study

The main objective of CIS was to study how the different aspects of the curriculum, the new textbooks, the testing scheme, and teacher education unfolded in different types of classrooms, including the washback effect of the Board exam, and to provide continuing support for a more effective implementation of the curriculum. It also enabled teachers to become researchers in their own classrooms with a view to helping them develop professionally.

Key features of CIS

CIS was essentially insider/teacher-oriented and involved teachers in the studying of the different aspects of curriculum-as-reality (how the curriculum unfolded in everyday classroom activities), as opposed to curriculum-as-intention (the syllabus/textbooks as documents) or curriculum-as-product (the language proficiency of students revealed through end-of-course exams). Teachers gathered ongoing reports from classrooms, teachers, students, principals, and parents about the different aspects of the curriculum through a variety of procedures, i.e. classroom observations, interviews, questionnaires, field notes, informal and formal discussions, and examination of student scripts and documents. The study involved all the stakeholders in order to help them to understand the curriculum holistically from different perspectives and in the process to raise their awareness about and illuminate curricular processes. It was also believed that this kind of involvement would give them a sense of ownership and ensure sustainability beyond the life of the project.

The study was developmental with built-in flexibility, which allowed changes and extensions, as and when a need was perceived. Each stage was built on the previous one, on the basis of the evidence collected and lessons learnt. The short needs-based programmes that were organised responded immediately to the needs, felt and expressed, in a particular region, for a particular group, for a particular purpose. Further, the research design made it possible for teachers (and other stakeholders) to take part in the study at any stage, in various capacities, including carrying out action research in their own classrooms. It disseminated the interim results of the study at regular intervals to various stakeholders in ways that were intelligible, meaningful, and usable. The study also made an outsider perspective possible through the involvement of faculty members and research scholars from CIEFL, ELT institutes, and university departments, who looked critically at the different components of the curriculum.

The Final Report gives details of the entire CIS (CBSE-ELT Curriculum Implementation Study 1997). The case study presented here focuses on one aspect, i.e. ways in which the new assessment scheme was received by schools and teachers, and the nature and extent of the washback on different aspects of the curriculum and on stakeholders.

The new assessment scheme

The Board exam at the end of Class X was a paper-pencil test for 100 marks. The breakdown was as follows:

Section	Types of texts/tasks	Book(s) that help	Marks
Reading	3 unseen texts with questions (factual, literary, discursive)	MCB	30
Writing	4 tasks (2 short and 2 long)	MCB	30
Grammar	4 tasks	WB	15
Literature	5 tasks	LR	25

Table 1: Breakdown of Board exam

In Class IX, Continuous Assessment (CA) was assigned 60 per cent, which included informal and formal assessment of conversation (listening and speaking) skills, assignments, unit tests, and half-yearly exams, while the final exam was assigned 40 per cent. This was recorded in students' report cards and communicated to them; for Class X, however, although the same kind of oral assessment was recommended, the final Board exam was the sole indicator of students' success in the exam.

The story unfolds

The baseline study of Class X (1993–1994) showed that teachers and students were quite unhappy with the 'old' curriculum. The teaching/learning in the classroom and out-of-class activities revealed a strong negative washback of the 'old' heavily content-oriented/memory-based exam. Teachers and students were engaged in the monotonous task of revising, learning by heart teacher-interpreted content, and rehearsing mock exams. Obedience and discipline were expected of a Board class, just as it was the responsibility of the Board-class teacher to complete the syllabus as soon as possible so that revision could begin. Since the exam pressure left them with no alternative but to co-operate with each other to fulfil their self-assigned duties, *there was no time for achieving (language) objectives*³. However, everyone strongly felt that a learner-centred, skills-based curriculum, with tests that tested language skills would provide the much needed challenge, 'self-learning' techniques, and variety, even for the weak learners. Both teachers and students were clearly unhappy that they had missed the chance of being part of the 'new' curriculum. This, then, was the CBSE climate when the new curriculum was introduced.

The study of Class IX showed a consistent and definite shift towards a communicatively oriented classroom from year 1 (1993–94) to year 3 (1995–96) in different types of schools in different regions of the country. While teachers developed a slightly better understanding of the rationale of the communicative approach *vis-à-vis* the role of textbooks in year 3 compared with year 1, the idea of MCB and WB as resource books that aided skills development and were, therefore, useful for doing

³ Text in italics is a direct quotation from the data.

Sections A and B in the exam paper was not apparent in teachers' reported data or actual work in the classroom. MCB was seen to have too many activities that were lengthy and repetitive and, therefore, arbitrarily dropped.

Alongside this, efforts on the part of teachers, rather than students, at meeting the exam demands a whole year later were clearly visible. Writing was more product-oriented and received less time and attention than other areas. Literature and grammar classes were content/form-focused, largely teacher-led (with teacher explanations and marginal student participation), which students, on the whole, found dull and boring. The time spent on literature, which increased to about 40 per cent in year 3, was disproportionate to what the exam mandated (25 per cent), with a lot of explicit exam preparation. Completing the syllabus, i.e. the prescribed textbooks, was seen by teachers and principals as a main measure of teacher competence.

While everyone appreciated the new assessment scheme for its inherent worth, classroom-based assessment by and large did not concretise the concept of formative evaluation. Teachers' poor understanding of the area and lack of necessary skills combined with inadequate administrative support seemed to affect the implementation of the scheme to quite an extent. That said, all the stakeholders seemed quite positive in their attitude to change when they made suggestions for better implementation: more teacher orientation, better infrastructure, change to be introduced from primary classes, oral exams in Class X Board exam, among others.

The first year of Class X (before the new exam) revealed quite a communicatively oriented class, although all the stakeholders expressed a good deal of anxiety about the exam and spent a substantial amount of their time and effort in revising the lessons and rehearsing mock exams. The study of Class X after the exam (1995–96) revealed exam preparation as the most crucial activity which was separate from teaching; since oral assessment was not part of the final assessment, it was not done or done half-heartedly, especially in privileged schools where the students' language proficiency was high.

Problems in the exam and in its different uses surfaced as the study progressed:

1. The exam scheme had about 45 per cent devoted to rehearsed tasks.
2. The different versions of question papers meant for Delhi/other regions and for outside India were not equivalent.
3. All those who were trained in marking did not mark exam scripts and vice versa.
4. Teacher-made tests used for classroom-based assessment were poorly constructed, although they mirrored the exam pattern in every respect.
5. The Board's policy of providing only total scores, as opposed to section totals, did not give the necessary response to the stakeholders about what kind of teaching/learning in class resulted in what kind of scores.

The needs-based workshops, such as analysis of teacher-made tests, item writing/banking, assessment of oral skills, and marking writing, succeeded in

problematising the issue and developed prototype materials and tasks. Apart, however, from identifying key people who could take the ideology of CIS forward, reaching out to thousands of teachers spread across the country was beyond its brief. Therefore, the washback of a less skills-based exam but a largely skills-based course was not entirely positive, as had been expected, although all the stakeholders were quite happy with the new curriculum, which, in turn, reflected the extent to which they accepted the change.

This was only a partial picture that emerged in a majority of urban/well-resourced⁴ schools; it was clear that the new curriculum/exam had not addressed the problems of weak students, especially in remote areas and under-resourced schools. In regions such as Arunachal Pradesh, where only the CBSE's Course A is available (which most people concede is not relevant to their needs), the situation was one of a distorted version of the new curriculum and lower scores in the first Board exam. If less proficient learners had to pass and a measure of teacher accountability was the pass percentage in his or her class, recourse to a results-oriented methodology was the only proven alternative. Therefore the MCB was not used until late, i.e. February, or not touched at all, since it was outside the syllabus. Teachers felt quite confident, from past experience, in handling the texts from the LR for exam purposes. The request for more exam-type formats, i.e. integrated grammar tasks, was therefore not without reason. This was corroborated by the analysis of answer scripts of the first Board exam: it showed that students with low scores from regions such as Guwahati and Allahabad had scored almost all of the 25 marks in the literature section and the remainder in B3 and B4 writing tasks which demanded prior knowledge of texts from MCB, but had left out Reading and Grammar sections.

A positive washback of the Class X curriculum was, to some extent, evident in Class XI. Teachers and students wanted their classes to continue to be communicative. They did not like to go back to the old style of teaching, where the meaning of the text, fixed by the teacher, was explained until all the students understood it. Teachers found themselves unable to use old books in new ways, because of the inherent limitations of the texts for adapting for communicative tasks/exercises. The washback of the slightly revised Class XII exam, however, was not apparent. Therefore, what everyone desperately wanted, before all the good things done in Classes IX and X were undone, was a change to books that promoted the development of language skills and an exam scheme that incorporated an oral component. The urgency of this demand also reaffirmed the degree of acceptance the stakeholders had shown of the new curriculum in Classes IX and X.

In sum, then, it appeared that the new curriculum, given the constraints of the new exam, was not entirely satisfactory, especially for weak learners, as opposed to what was envisaged through a communicative curriculum. The conflict that emerged because of dichotomies in the seen-unseen components of the question

⁴ Well-resourced schools and under-resourced schools are not necessarily dichotomous, but represent a continuum in terms of infrastructure (for example, tape recorders in working condition, cyclostyling facilities), adequate number of teachers with high motivation and language competence, students who are not first generation learners, and school management that believes in quality education and provides support to teachers.

paper, language-literature issues, lengthy question-cum-answer papers only for English, among others, had neutralised, to quite an extent, the work done during the project and CIS. A recommendation made to the CBSE by the study was that concerted efforts should be made to support the enormous amount of progress that motivated and committed teachers were making, before the negative impact of the changed curriculum became firmly established.

Lessons

Given the magnitude and complexity of the curriculum change, which marked a major departure from secondary Boards' textbook revision exercises, the project demonstrated a model of curriculum reform at the secondary level. The teacher-as-researcher approach to monitoring and supporting the curriculum change in different types of schools for different types of students, with all stakeholders participating in it willingly, was a unique experiment that other curriculum projects could adopt with suitable modifications.

There seemed to be at least two related issues that CBSE should have addressed: first, the assessment scheme did not capture what transpired in class, thus making the test a de facto curriculum (Shohamy 1997), different in nature and scope from the official curriculum. This negated the impact considerably, hampering, in the process, the confidence the stakeholders had gradually built up. Even if a sudden switch to a completely skills-based exam was not practical, a change in stages should have been conceived of and carried out, but this was not on CBSE's agenda. Further, since the needs of the weak learner had not been adequately met, the question of how the Board, with its reputation of being elitist, should/can be equally successful with all types of schools and regions had to be examined. Alternatively, was this a realistic goal and, if not, should CBSE's outreach be redefined? The CIS illuminated the macro and micro aspects of curriculum implementation, especially from a bottom-up perspective, and, therefore, the stakeholders, it seemed, were equipped to take on such a challenge.

The second factor which hampered effective implementation was the stage at which the innovation was introduced. Given the high-stakes nature of a single, summative, paper-pencil exam at the end of secondary school, Class IX, everyone felt, was too late for any radical change to be initiated. They saw the change as piecemeal and felt constrained to resort to time-tested, results-oriented methodologies. An area where the impact of the Board exam was felt most was the way the scheme percolated to lower classes, especially in private schools, since it was beyond the purview of CBSE. The adoption of the three textbooks (MCB, WB, and LR) for the same exam scheme from Class I onwards, and the use by scores of private publishers of the project team members as materials writers and trainers, to ensure the curriculum worked in practice, give some indication of how the innovation spread. Private schools with more educated and aware parents were, thus, more influential in turning the innovation to their advantage. The weak learner and teachers in under-resourced schools in this enterprise were just passive recipients of the innovation.

A clear message from this experience for the future is the following: it is essential that a Board clearly define, at the beginning of the project, the scope of such a project and ensure that the agreed brief is satisfactorily completed; in addition, it will have to respond to the expressed and felt needs of its stakeholders during the project and incorporate changes along the way, by going beyond a *logical framework* (see Tribble 2000 for details), especially when the project is located within a stakeholder approach to curriculum renewal. Here, it was definitely possible to introduce an evaluation scheme that included a component of oral skills in its high-stakes test, in keeping with the communicative curriculum which emphasised oral skills in the classroom, by which teachers and other stakeholders were more than persuaded. Similarly extending the innovation to lower classes in a systematic, project mode was a necessity without which everyone felt let down. This, then, was a case of lost opportunity, which one needs to be guarded against in future endeavours.

One aspect of the philosophy of CIS which, in my view, was significant was the high level of engagement of the main project 'users'. The project provided the space for teachers and other stakeholders (including me, as Project Director) to interpret the curriculum changes according to their own local realities, to feed those interpretations directly into needs-based workshops, and to provide their personal insights during events such as review workshops, and national and international seminars. The international seminar held in 1995 stood out for its exuberance, professional commitment, and 'passion', as a participant put it, and was an example of how the insider-teachers wrested the podium away from the outside-experts and shared their myriad (research) experiences in an academic forum. For a project of this magnitude to become an evolving curriculum (as opposed to a static one), teachers' (and other stakeholders') voices need to be heard and acted on. While the notion of collaboration within and outside schools was kept alive during CIS, what kind of communication network would need to be set up after the life of the project? It seemed that time-bound projects had to finish their work and move on. This was more true of CBSE and the British Council than CIEFL, an organisation with continuing research interests. That CBSE had expended a lot of its time and attention on English and now had to think of other subjects was not in keeping with the principles of a grass-roots-level project. There were the issues of the roles and responsibilities required of the three organisations in order to ensure sustainability that should have been negotiated early on the project.

The lessons learnt from the exercise, however, can feed into a more sensitive and viable curriculum than has been possible thus far. Indeed, the formative nature of the study has activated the notion of dynamism in the curriculum and has accommodated the changing context, rather than treating it as a top-down prescription to be executed in its entirety. The individual strengths of teachers will need to be co-ordinated and supported at school, and the school in turn will need to be supported by outside agencies, if the notion of 'evaluation as a development tool' is to be concretised in meaningful contexts. It seems that the gap between curriculum-as-intention and curriculum-as-reality can be closed only through a research and development approach to one's own teaching and that long-term improvement of education hinges on the effective utilisation of curriculum research.

Postscript

The CBSE changed its exam scheme in 2005 with the following breakdown: Reading (20), Writing (30), Grammar (20), and Literature (30). It was more memory-based than in 1995, a clear regression in 10 years. A tracer study done in 2001, three years after this study officially ended, revealed that, although there was no longer any dialogue between different stakeholders, some teachers continued to create learner-centred classrooms and adopt a research approach to their teaching. They were system-free, silent innovators working in isolation (Mathew 2006).

Case 18: Redirecting a curriculum development project in Egypt

Mona Zikri

The project

Egypt's longest-ever ELT project ran from 1975 to 1991. It was sponsored by the British Council, UCLA (University of California, Los Angeles), the Binational Fulbright Commission, and the Faculty of Education (FoE) of Ain Shams University. It focused on an ad hoc committee proposal for 'developing ELT in Egypt' through the 'revision of curricula and textbooks in conformity with the language needs of Egypt and individual Egyptians' ('Ad Hoc Proposal 1975: 4). A key new body, the Centre for Developing English Language Teaching (CDELTE), was established and affiliated to the English Department of the FoE.

Three curriculum development projects can be identified. They were classified as CDP I: materials development, CDP II: research, and CDP III: a staff development scheme.

However, the design of the project lacked a needs assessment, well-defined objectives, and adequate planning. Also, the project's change theory had three flaws:

- the adoption of a prescriptive approach by expatriate experts
- the rationale for reform's focus on 'product', not 'process'
- the failure to recognise the 'social context' of the reform.

The CDPs succeeded in developing curricula, but failed on two fundamentals of reform: building ownership and improving capacities. The new curricula were not adopted. Recognition of the importance of the social context of reform led to a redirection of the project in its final years. Limited success was achieved with the improved performance of professionals, but there was no impact on systemic change. CDELTE did not achieve ELT reform in Egypt, although it served temporarily as a professional development hub.

¹ Documents no longer exist but are quoted from Zikri 1991. They will be cited as (Doc: date) without bibliographical citation.

Stage – monitoring and evaluation

This case study focuses on the monitoring and evaluation (M & E) system for CDP. An adequate M & E system is crucial to a project's success and sustainability, especially one with challenges of design and implementation. In fact, a truly iterative project need not be ideal at the design stage, but should have enough flexibility to evolve according to needs, guided by assessment findings. This case study examines the validity of formative and summative evaluations throughout the 16 years of the project, and their success and failure in solving problems and redirecting project goals.

Story

The low impact of continuing evaluation on project outcomes

Although the project had models of formative and summative evaluations, as well as output and implementation assessment, systematic M & E was not yet a common practice when the project began. The validity of the assessment practices, their impact on progress, and the project's sustainability and results will be examined in this paper.

Evaluation is just one of a set of problem-solving interventions, but it is a crucial tool for adjusting plans and making progress towards positive results. During its first decade, CDELTA evaluations had little impact on progress, lacking tactics and strategies, and, in the absence of clear objectives, there were no standards to measure progress against. In some cases, problems were identified, but no action was taken to address them. Eventually, an ethnographic approach to evaluation was adopted, providing insight into and solutions to issues, and supporting the project to enable it to achieve limited success. The features of the evaluations were:

- **Inadequate tools/lack of instruments:** In 1978, the testing expert specified revision issues, which in fact denoted the absence of formative evaluations throughout CDP I, since there existed 'no clear defined measurable objectives, no criterion of mastery established, no concurrent programmes were available, no proficiency measures...' (Henning: Doc: 1978). In 1979, the first evaluation report stated that shortcomings in staff assessment in relation to subject matter expertise were being 'informally made'. No mechanisms were set for formative evaluation through the 'continual approval' of stakeholders (Bowers et al.: Doc: 1979). In general, evaluation results were based on inconclusive evidence. For example, there was a conclusion that students' attitudes towards English had changed after the trial because 'they often used English in the corridors' (Hudson and Melia 1981: 144), without any comparison with the state before the trial.
- **Disregard for Egyptian involvement:** Project continuity depended on the English department's acceptance of the courses. However, two reviews dated 28 May 1981 show a dispute between the Centre and the department chair, the 'Principal Investigator' of the project, who rejected the materials, stating that senior staff could not support materials they had not been allowed to examine, let alone develop. The British team leader disregarded the allegations,

admitted the atmosphere was 'worrying', but decided to proceed with publishing irrespective of the department's decisions. Senior staff members from Ain Shams, Assiut, and Tanta universities confirmed the department's complaints that expatriates did not acknowledge context, regulations, or the local work culture (Hudson and Melia op.cit.). This terminated collaboration between CDELT and the department.

- **Evaluations not supporting the adjustments:** Evaluations reflected issues that resulted from a lack of proper project design, but a lack of planning persisted throughout the three stages, with plans and objectives sometimes appearing after materials were actually produced. Needs and context analyses were absent. Implementation was built on assumptions that were tested by trial and error. Thus we find:
 1. **CDP I:** The British team leader reported that the project and scope of teaching materials were poorly defined. Initial bad planning had led to constant confusion about aims and objectives (Melia: Doc: 1982). D. Loos, a team member, stated: 'In piloting our materials this year, we had to face the fact that our materials were going to be used under conditions which we considered to be less than ideal... Our initial response was (1) the materials were not designed for such use...' (Loos et al 1983: 34). However, this negative evaluation was not acted on, and the project's chance for enhanced impact was lost.
 2. **CDP II:** No materials revision or summative reviews had been conducted for first and second year materials. These would have supported the design of a 'research paper' project to develop the second phase of the CDP. After the completion of 24 research papers, an evaluator reported that the research was 'too theoretical and impractical to support developing materials' (Nelson: Doc: 1984).
 3. **CDP III:** Poor design, constant lack of planning, and unsupportive evaluation negatively affected project results right up to the end. However, a summative assessment found materials to be of high quality and Egyptian professionals to be well versed in the materials and methodology. It recommended that research attachments should be replicated, as local expertise was the only way to disseminate the use of the materials (Robinson: Doc: 1990). In fact, this echoed the views the Department Chair expressed in 1981. The British team leader, in his concluding report, confirmed these findings, but raised the problem that such work could not be sustained, since planning for building local expertise should have been developed earlier (Smith: Doc: 1990).
- **Limitations of external evaluations:** J. Morley and B. Coffey's report praised the materials but, in assessing their applicability, their views were divorced from the realities of the education context (Doc: 1982). For example, the materials were found to be incompatible with 'non-ideal' situations, while in fact the norm for Egyptian classes was 'non-ideal'. This undermined the validity of the materials. The sole Egyptian evaluation by G. Saad Eldin, Y. El Ezabi, and A. Farghali was generally positive towards the materials, with some criticisms and reservations. However, issues were defined and solutions proposed, but not implemented (Doc: 1982).

Evaluations did not capture implementation issues

Evaluations focused on the curriculum products but did not capture the flaws in the implementation process. In fact, CDELTA had potential success factors that were not utilised, including the potential for sustainability and institutionalisation as an Egyptian entity, an ideal duration of 16 years to effect change, distinguished ELT professionals supporting the projects, and the establishment of partnerships. Implementation issues included:

- The work management style was prescriptive. The foreign ‘experts’, researchers from outside the system and culture, saw themselves as innovators with sole control of design and implementation, in isolation from stakeholders and cultural constraints.
- Plans for sustainability were not inbuilt. Handover was on the agenda, but the key roles were unquestionably played by expatriates, with no plans to develop Egyptians to take over in these roles. Finally, it became evident that the transfer of posts to Egyptians was the way to Egyptianise CDELTA (Holliday: Doc: 90).
- A high turnover of expatriates supported a rigid adherence to short term plans. British Council experts held two- or three-year contracts. Newcomers picked up from where their predecessors had left off, guided by and following up on short-term plans with insufficient time or institutional memory to interact with the context. Consequently, issues persisted throughout. It was not until the final phase, when the expatriate expert had an extended contract of five-years, that issues were identified and addressed.

A shift in assessment approach redirected the project

As British involvement was ending, the project still faced the dilemma of how to disseminate and institutionalise the materials. Sustainability, capacity development, and human attributes were emerging issues. As the social constraints became evident, the initial assumption that curriculum development (CD) was the route to reform became invalid. A. Holliday, language co-ordinator (1985-90), set new objectives, changing the focus from curriculum reform to developing Egyptian ‘change agents’. Changed work practices ensured that Egyptian ‘colleagues’ and expatriate ‘consultants’ collaborated, and the value of capacity development, partnership, and ownership became evident. Understandings emerged about the power of social context, cultural biases, and socio-physical conditions.

Holliday introduced new monitoring practices, collecting information by ‘ethnographic classroom observations, discussions, consensus building, and stakeholders’ opinions, and means and needs analysis of the social context’ (Holliday: Doc: 1985, 1987, 1988, 1989, 1990), with results that supported the redirection from CD to staff development.

Practices included:

- First-hand on-site investigations. This included observing as well as teaching classes. New vital observable phenomena emerged concerning students’ behaviour and local staff status.

- Feedback. Egyptian staff provided opinion as experts in their fields.
- Assessment of materials. Visiting specialists and local senior academic staff were involved in discussions and agreement on materials and approaches.

Lessons

Throughout 1983–93, research by different CDELТ team members reflected lessons learnt, many of which formed new principles for international ELТ support.

CDP I: Bowers (1983) recognises that successful design and implementation must take into consideration the following lessons learnt:

- The uniqueness of project context. The measures for innovation which were acceptable in one context were not appropriate for another. Thus, every CDP represents a unique attempt to initiate change. It is necessary to consider the unique physical, cultural, linguistic, and psychological variables in project design and implementation.
- The relationship between design, planning, and monitoring. In order to compensate for the short planning phase of projects and cope with problems arising during implementation, Bowers advocates that monitoring and evaluation should drive a flexible process of adjustment.
- The negative impact of conflict. Bowers describes the CDELТ crisis as ‘disputes over project management’ that negatively impacted on results. He recognises that CDPs that fall into the trap of excluding stakeholders will fail on this essential human factor, more than on any other.
- Cultural interaction between expatriates and host cultures. It is necessary, in a multilingual world, where ‘our (the West’s) contemporary ethic values and practices [have a dominant status], to mediate between these and alternative ethics and ways of acting, both generally and in educational terms... [and] to make common our knowledge and share our uncertainties’ (Bowers 1985: 255-57).
- Clear objectives and stakeholder collaboration. Projects must have a managed set of priorities, objectives, timescales, and balances between external expertise and local inputs.

CDP III: CDELТ offered multiple lessons, as reflected in Holliday’s research (Doc: 1987, 1988, etc. and Holliday 1991, 1997, and 2005). The impact of social constraints in project development was crucial. He argued for ‘appropriate methodologies for different social contexts of education’; advocated shifting the ‘destructive notion of cultural imperialism’ to a more collaborative analogy of a ‘marketplace’ where ‘all partners are equal’ (Holliday 1997: 7). Lessons learnt included the impossibility of conducting CDPs without regard to the host’s social context, and ownership. Appropriate project design necessitates information collection and continuing ethnographic means analysis. Sustainability requires good planning, built into projects, for a local counterpart to take over.

Conclusion

Egyptian academics who were involved with the Centre during the last project phase have managed to evolve some small-scale professional networking and adopt improved ELT practices, but there was no reform infrastructure to support long-lasting systemic changes in ELT.

Broad lessons learnt can be summed up as follows:

1. Project design must be flexible and based on needs.
2. Reform is a process with 'results', not a product with 'outputs'.
3. Crucial to success are issues of ownership, capacity development, and incorporation of the social context through 'appropriate methodologies'.
4. Valid monitoring and evaluation systems should inform project designers, implementers, and policy makers on how best to achieve results.