2 The origins of the present investigation: the Sri Lankan O level Examination Project

Background

The analysis which forms the focus of this study was carried out on data collected in 1991, during the final phase of the Sri Lankan O level Evaluation Project. This was a four-year project which set out to evaluate a new national examination of English and to determine whether it was having the effect on classroom teaching that its designers intended.

The present investigation is an extension of the original evaluation project: a re-analysis of data using a conceptual framework and analytical tools that were not available at the time the data was collected. This investigation was motivated by questions emerging from the original project, and it requires knowledge not only of the findings of the project but also of the context in which it took place. This chapter provides an overview of the original project and its findings. The aims are presented, together with the stages which were most relevant to this investigation: identifying the main features of the materials on which the new examination was based, identifying the type of impact the new examination was meant to have, and carrying out a baseline study and a two-year observation programme. The results of the observation programme are presented in some detail, in order to give a clear picture of the type of English teaching that was most prevalent at the time and a wider understanding of the kinds of questions that remained outstanding. The chapter concludes with an explanation of my decision to exploit the data further, utilizing a framework for analysis derived from language testing, assessment in general education, and educational innovation. The theoretical background to this framework is presented in Chapters 3 and 4, my methodology is described in Chapter 5, and the findings of the analysis are presented in Chapters 6–11.

Political and educational setting

Sri Lanka gained its independence in 1948, after 150 years of British rule. English had played an important role during the colonial period, in both the administrative and social structure of the country and its educational system. It was the medium of instruction for the privileged classes from the mid-19th century onwards, and it retained this status even after independence in private denominational schools and in government-run 'central schools' set up for

scholarship students from rural areas (Punchihetti 1994:21). However, political changes in 1956 led to a new language policy and by the end of the decade almost all schools in the country had adopted either Sinhala or Tamil, the two indigenous languages, as their medium of education. English was an ordinary subject on the timetable for the next 20 years, no more important than any other. The teaching was similar to that in many other countries in the 1960s and 70s, with a great deal of emphasis on language form, some attention to reading comprehension, and a neglect of listening, speaking and writing. The Ministry of Education introduced new textbooks in the mid-1970s which reinforced this approach. Every unit followed the same pattern: the presentation of a new language structure through model sentences, two reading passages followed by comprehension questions, and further language practice 'exercises' which were in fact tests of vocabulary and grammar.

There were further political changes in the late 1970s, which led to the opening up of the country to foreign investment and trade (Mosback 1990:53). Tourism became an important industry at home, and at the same time more Sri Lankans were able to travel and work abroad. There were clear incentives for those who were able to communicate in English. The language was also important academically: it had served as the medium of instruction for many university courses (Wijemanne and de Silva 1994:32), but it would now be useful for those who wished to study in other countries. There was a growing need for English to serve as a 'link language' between the majority Sinhala-speaking population and the Tamil-speaking minority.

Policy-makers realized that the country was not equipped to cope with the new demands for the language, and that action was needed to prepare future generations for these opportunities. The Ministry of Education embarked on a large-scale English language teaching reform programme in the early 1980s. The reform was to create change in three areas: materials, teacher training, and testing.

On the materials front, work began on two new textbook series: the *English for Me* series for primary schools (developed from 1980–1985, by a Sri Lankan team with Norwegian financial assistance), and the *English Every Day* series for secondary schools (developed from 1981–1989, by a Sri Lankan team working with British consultants, with financial assistance from the then Overseas Development Administration, now known as the Department for International Development – DFID). Both series stressed the need to communicate in English rather than to just master the grammar. The emphasis was on practical language and there were exercises to develop listening and speaking, two of the skills which had been neglected in the past.

There were attempts to develop both pre-service and in-service training (see Woolger 1994 for details). Much of the in-service work was carried out by members of the textbook design team, with the aim of showing teacher advisers and teachers how to use the new teaching materials (Mosback 1994:55).

The keystone of the reform was to be a new examination. It was recognized from quite early on that it would be difficult to persuade teachers and students to use the new materials and methods if there was not also a change in assessment.

The O level English examination

Students who stayed in school until the end of Year 11 took a series of examinations called 'the O level'. O level results determined whether they would be allowed to continue on to the 'collegiate' level of schooling (the equivalent of the 'Sixth Form' and 'A level' in England and Wales), or whether they had to leave school and find employment. Competition was intense for the few places available at the collegiate level and for the limited number of good jobs that school-leavers could aspire to. Students' O level grades, including English, were amongst the most important in their academic career.

The 'Old Syllabus' English examination, like the teaching materials, focused on language structure and reading comprehension, and it followed the same pattern year after year. Of the 105 'objective' questions, 90 tested some aspect of language form: spelling, vocabulary, prepositions, word order, verb forms, pronouns, etc. There were two reading passages with a total of 15 comprehension questions, many of which tested reading at sentence level only. Candidates could choose to respond to one out of four writing tasks, but it was well known that they could gain good marks without attempting to write at all. Listening and speaking were not tested. (See Appendix 1 for a copy of the 'Old Syllabus' examination, 1987.)

Very few students managed to pass this examination: the pass mark of 35% was usually met by only 20% of the population. There were claims that the examination had 'doubtful validity' (Wijemanne and de Silva 1994:33), but an inspection of both the examination and the textbooks suggests that face and content validity were quite high (see Alderson, Clapham and Wall 1995 for a definition of these terms). However, it was now considered important to produce an examination which matched the new teaching materials. Furthermore, it was argued by several policy-makers, including the Secretary of State for Education, that:

> reformed examinations should and could be used to encourage and engender changes in the attitudes and practices of students and teachers in the English classroom ... a new examination should be developed and used 'as a lever for change', to encourage teachers and students to make a serious and committed response to the evolutionary (and sometimes revolutionary) changes that were already under way through teacher, curriculum, and course development. (Pearson 1994:85)

Development work on the 'New Syllabus' examination began in early 1986, with the first drafting of test specifications (Pearson 1994:91). The

examination was meant to test all four skills, and it would build upon work that had been carried out earlier to find practical ways of testing speaking and listening. (Pearson 1994, and Alderson, Wall and Clapham 1987.) It was necessary to change plans within a year though, following a Ministry announcement that the O level examinations for all subjects should contain an element of continuous (i.e. classroom-based) assessment. The English O level would now consist of two parts: a centrally designed and marked examination paper testing reading and writing, and a series of teacher-designed and assessed classroom activities which would focus on listening and speaking.

Everything about the continuous assessment idea was alien to the Sri Lankan context – from the notion that teachers' judgements about their own pupils could be relied upon (the country had a long tradition of external examining), to the mechanics of conducting in-class assessment and recording the results for the examination authorities. The examination design team now had to devote time to preparing teacher support material for continuous assessment – on top of their work to produce a valid and reliable written paper. Ironically, the continuous assessment programme was withdrawn just after the first administration of the new examination, and none of the marks that teachers had collected were calculated into students' final O level grades. No official reason was given for this decision, but many believed that it had been introduced too hastily, had been too unwieldy, and was too sensitive an issue at a time when there was a great deal of political turmoil in the country. (Some opponents of the government felt that this form of assessment favoured pupils from urban areas and privileged families, who had more exposure to English and more resources and support than pupils from rural areas.) With the withdrawal of continuous assessment the written paper became the sole means of assessing students' English.

The new examination would now only focus on reading and writing. There would be several low-level tasks which any Year 11 student might be expected to complete, several intermediate-level tasks, and several difficult ones. Different text types and tasks would be included each year, and over a number of years the whole syllabus would be covered. This would oblige teachers to cover the whole textbook series rather than focus on content they predicted would appear on the examination. (See Appendix 2 for a copy of the new examination, 1988.)

Teachers received an official orientation to the examination in May 1988, approximately six months before it was due to be administered for the first time. The orientation booklet included the types of texts that students would be expected to cope with, the reading skills they were expected to master, and the types of writing they would be expected to produce. One copy of the booklet was sent to each school, to be shared amongst the members of the teaching staff. Further publications were distributed to schools and to students in 1989 and 1990.

The O level Evaluation Project

Aims of the project

An important feature of the new examination effort was that evaluation was built into the plans from the beginning. The organization which funded much of the development work – the British Overseas Development Administration – also commissioned an evaluation of the examination and its impact on teaching. The project documentation stated that there was a need:

> to establish not only the validity, reliability and practicality of the new examination and associated procedures, but also to assess the extent to which these innovations help to bring about changes in the day-to-day teaching of English, and consequent pupil attainment and ability. (Institute for English Language Education 1988:1)

The project had three aims:

1. to assess the reliability and feasibility of the new O level examination in English, including the arrangements for its administration, the training and monitoring of examiners and the validity of the results,
2. to evaluate and advise on the monitoring, moderation and implementation of continuous assessment in English, and
3. to evaluate the effect on the classroom teaching of English of innovations at O level. (op. cit:2)

The evaluation project was to run from mid-1988 to the end of 1991.

There were two teams involved in the evaluation: a team from Lancaster University, which consisted of a Project Director (Charles Alderson) and a Research Associate (Dianne Wall), and a Sri Lankan team, which consisted of the O level designers, a group of secondary school teachers who would carry out classroom observations, and the British Council Consultant for Testing and Evaluation.

The project was a collaborative effort from the start. The Lancaster team was to design and co-ordinate the evaluation in consultation with the Sri Lankan team, and major decisions were to be taken jointly. A Lancaster team member would visit Sri Lanka twice a year, to investigate the feasibility of various proposals, consult with stakeholders, carry out data collection, and give advice on technical matters. The Sri Lankan team would share their insights into the desirability and practicality of proposed plans, collect data and react to suggestions made after the analysis and interpretation of data. The project had a summative function – to evaluate whether the examination influenced teaching over a given period of time – but it was also meant to provide formative feedback. Findings and recommendations were to be sent to Sri Lanka regularly so that colleagues could make changes in their programme if they agreed that these were necessary.

In the first year the project focused on the first aim – evaluating the examination and associated procedures. I observed the administration of the examination and the training and monitoring of examiners, and carried out a detailed analysis of test results and examiner reliability (see Alderson and Wall 1989 for details). The examination reflected the nature and the spirit of the textbook quite well (content validity) and proved a fairly reliable means of assessing what students were supposed to have learned during their years of English study. There were a number of 'teething troubles', but the design team was able to address problems in content and procedures before the second administration. The team took over some of the evaluation work during the second year and by the third year had sole responsibility for carrying out observations of administration and marking, conducting qualitative and statistical analyses of results, and making recommendations for further improvements of the examination.

The second aim of the project – examining continuous assessment – was dropped when the continuous assessment programme itself was abandoned.

The third aim – measuring the influence of the examination on teaching – was to have become the project's focus during its second year. The design team had carried out a set of observations in mid-1988, before most teachers knew what the new examination would look like or could be affected by its characteristics. Further observations were to be carried out in 1989, to see whether there had been any changes in the way teachers conducted their classes and if this could be traced in any way to the influence of the exam. Unfortunately serious political conflicts forced schools to close for much of that year. It was not until early 1990, when teaching was back to normal and it was safe to travel in most areas of the country, that further observations could be undertaken. This marked the beginning of the 'Impact Study', the two-year observation programme which formed the core of the evaluation project.

Although there was a delay in starting the observation, it had been necessary to begin thinking about how the examination might influence teaching from the earliest days of the project. This meant analysing all of the official statements about the goals of the examination and, because the exam was meant to reinforce the new textbook series, of the textbook as well. These goals are discussed in detail in Chapter 8: Characteristics of the Textbook Series. Only a brief description of the textbook series will be provided at this point to give an idea of what the Impact Study was meant to ascertain.

English Every Day

English Every Day (*EED*) was a five-book series designed by a team of Sri Lankan teachers working with two British Council consultants. The series introduced a number of new ideas into Sri Lankan ELT, both in terms of

content and methodology. It presented texts and topics that were meant to be relevant to Sri Lankan students, tasks that resembled the types of activities they might have to carry out in their second-language environment, and exercises which attempted to develop language skills rather than just grammar. The texts ranged from short messages to informative academic texts in the case of reading, and application forms to quite lengthy reports in the case of writing. The reading skills included skimming and scanning, deducing the meaning of unknown words, picking out the main idea from supporting detail, understanding the communicative function or value of sentences, etc. The writing skills included planning and organizing information, giving information explicitly, transferring information from pictures to reports, and so on. Grammar was also taught but it was meant to play a supporting role rather than lead the students' learning: its importance was acknowledged but it was not presented in a regular or systematic way.

Finally, there were many exercises to develop listening and speaking: role plays, dialogues, picture descriptions, and discussion tasks. It was this attention to oral skills that most distinguished *English Every Day* from earlier materials.

The textbook writers also hoped to encourage innovation in methodology. One of their aims was to persuade teachers to take a less dominant role in the classroom. Another was to give the students opportunities to practise the language in pairs or groups:

> As a general rule, the course writer would make sure that wherever possible, a teaching point would be presented and practised in a pupil-centred communicative activity. (Mosback 1994:57)

The Teacher's Guide which accompanied the first textbook (Year 7) set out the essentials of the new approach and gave suggestions about how to handle the new material. It was hoped that these suggestions, along with the in-service training that was carried out before and at the time the series reached the schools, would change traditional classrooms into a more active places for learning.

As stated earlier, the examination was meant to reinforce the textbooks, both in terms of content and methodology. The evaluation project was to investigate whether this was really happening. Were teachers teaching the way the textbook and examination teams wanted them to? If so, was the examination exerting any influence on their teaching? If not, could the exam be held responsible?

Defining *washback*

The phenomenon which the evaluation project was meant to study was known as *washback*. At the time the project began, very little had been written about

washback in the field of language education. Several testing handbooks offered definitions of washback, but these were very general. A few articles claimed that washback had occurred or would occur, but there was little evidence to support these claims. There was some discussion of the phenomenon in the general education literature, but this literature was not well known to language educators. There is now a greater awareness of the work that has been carried out in other fields and a number of studies have been carried out in language education, but very little of this work was available at the end of the 1980s. (See Chapter 3 for a review of the literature on examination impact.)

The view that I started out with, which was a common view in language education at that time, was that washback was inevitable. It was the 'effect' in a simple cause-and-effect relationship, where an important ('high-stakes') examination was the 'cause'. Washback could either be positive or negative. It was more likely to be positive if the aims and content of the examination matched those of the syllabus (or of the textbook if no syllabus existed). The assumption here was that the syllabus/textbook was always 'good'.

If the washback were positive teachers would be guided by both the textbook and the examination. The following outcomes could be expected:

Table 2.1 Forms that positive washback might take

Type of washback	What would this look like in the classroom?
Content of teaching	Teachers would pay attention to all parts of the textbook, because they would realize that any of the text types or tasks appearing in the textbook might appear on the examination as well.
	They would not give more emphasis to any one skill than the textbook gave, because the weighting in the examination would reflect the weighting in the textbook.
Method of teaching	Teachers would use the general approach and the methods suggested by the Teacher's Guides, as they would recognize these to be effective means of developing the skills that would be assessed on the exam.
Assessing students	Teachers would write tests that would mirror the content of the textbook, because this would also be the content that would appear on the exam.
	They would mark their students' work using the criteria laid down in the textbook guidelines, which would also be the criteria used by examiners when marking the examination.

Washback could also be negative. If the examination and the textbook had different aims or content they might pull in different directions. Teachers would be driven by the examination rather than by the textbook. The outcomes would then be as follows:

Table 2.2. Forms that negative washback might take

Type of washback	What would this look like in the classroom?
Content of teaching	Teachers would not pay attention to all parts of the textbook because they would realise that some skills were not assessed and that it was more useful to concentrate on the ones that were. Even when teaching these skills teachers might neglect some text types or activities, feeling that these never appeared on the examination and were not worth spending time on. Teachers might stop working from the textbook, and begin to use materials which were more related to the examination. These might include teacher-designed materials, past examination papers, or publications designed to help students to prepare for the examination.
Method of teaching	Teachers would use whatever methodology they felt was most expedient to help them to prepare their students for the exam. Some methods might be sacrificed if the teachers felt that they did not achieve results efficiently.
Assessing students	Teachers would write tests which would mirror the content of past examination papers rather than the content of the textbook. They would adapt questions, or would simply 'lift' them, either from past papers or from publications designed to prepare students for the exam. Teachers would adopt the marking criteria used by the exam and would ignore advice in the textbook which went against this way of marking.

It was difficult to predict what the outcomes would be in the case of the new O level examination. The reading and writing tasks resembled the tasks in *English Every Day*, but the lack of oral skills might lead to negative consequences.

The evaluation project used several methods to investigate whether the exam was having the effect that was intended, but the method that distinguished it from the few empirical studies available was classroom observation. Observation took place at two different stages in the project: during the 'Baseline Study', to investigate what teaching looked like before the introduction of the examination, and during the 'Impact Study', to see whether the examination was having any influence on teaching during the first two years of its existence.

The Baseline Study

In order to determine whether the new O level examination was having any impact on teaching, it was first necessary to describe teaching before the examination was introduced. This meant carrying out a *baseline study* – an

investigation which looks at certain conditions before the implementation of an education innovation, to 'help us to monitor any effects that occur during or after "treatment" ' (Weir and Roberts 1994:46). The examination design team managed to visit a number of schools just before the official orientation booklet was distributed. Their intention was to see how teachers were coping with the new textbook and to inform their own process of examination development, but the data they collected also provided a good picture of the type of teaching being carried out before the teachers heard about the new examination. This could serve as a point of comparison for data-gathering exercises in the future.

The design team interviewed teachers at 18 schools and observed classes at 14. This was only a small sample of the schools on the island, but the team was satisfied that it was a representative one. The investigation provided valuable information about how teachers viewed their own teaching and the influences on it. The most important finding was that there was a difference between the claims teachers made about their teaching and the way that they handled their classrooms. They claimed that they had changed their way of teaching once they began using the new textbooks, and that they were now using a 'communicative approach', but their classrooms were very formal and students spent much of their time listening to the teacher or practising language form rather than developing the skills that the textbook promoted. The team felt that many of the teachers had not understood what they were supposed to be teaching, and that more than half had not achieved the objectives of the material they were working with. The teachers' methods were not noticeably different from the teacher-centred methods which the textbook designers had originally tried to discourage.

In addition to investigating what typical lessons looked like, the team tried to find out how much teachers knew about the examination. Half the teachers felt that the official circulars from the Ministry were confusing. A quarter of them had just received the official introduction to the examination but they did not have a clear idea of what the new exam would look like. Only one teacher out of 18 understood correctly what would be tested and how, but she knew this because she had previously worked with the design team. About three quarters of the teachers expected to receive more guidance in future, and half expected to receive 'model papers' that would give them a clear idea of the text types and item types that would be tested. (This finding was significant because the design team had already decided that they would not provide a 'model paper', since this might give teachers the idea that specific text types and item types would appear on the examination and might itself create exam-oriented teaching.)

The observations showed that teachers had begun to teach the content of the new textbooks but that the new ideas about methodology had not yet taken hold. The teachers knew almost nothing about the new examination, which

would be given for the first time in six months' time. The observations provided a reasonable baseline from which to measure whether the textbook or, more importantly for the evaluation project, the examination, would have any influence on teaching in the years to come. (Details of this study can be found in Alderson and Wall 1989.)

The Impact Study: 1990–1991

The aim of the Impact Study was to find out whether the examination was having any impact on teaching during the first two years of its existence.

At the core of this study were seven Sri Lankan teachers who had agreed to be part of the research team. Six of them had attended a training course in Lancaster, where they had learned about classroom observation, and several had contributed to the design of the observation/interview schedule which would be used in the early part of the study. The seventh member of the team was a field supervisor for a teacher training college who was familiar with both the new teaching materials and the examination. The teachers were based in five different parts of the country: Colombo and the surrounding area, Wennapuwa, Ratnapura, Tangalle and Bandarawela. I visited each of them in their home areas in early 1990, to discuss their role in the Impact Study and to explain the importance of their participation to the principals of the schools they taught in. Getting the principals' support was an important step because the teachers would need to leave their own classes in order to visit other, sometimes quite distant, schools. They would need to be absent from their own classes for up to 25 days a year.

Each of the observers agreed to monitor seven schools. This gave a total of 49 schools across the country. The observers were in charge of selecting their own sample, but it was understood that this had to be representative of the schools in their area. Each sample was to include urban and rural schools, Sinhala- and Tamil-medium, and schools with different levels of achievement. The observers were to observe each school six times, once every term for two academic years. They were to observe one Year 11 class (the final year before the O level exam) during every visit, and were to continue to observe the same teacher, if possible, throughout the length of the study.

The observers were to interview the target teachers before they taught their lessons, in order to establish what materials they were intending to use and how they intended to use them. They would then observe the lessons and record what they saw and their impressions of whether the teaching was influenced by the new materials or by the examination. They would interview the teachers again after the lessons, to ask for their impressions of the lessons and to find out why they had decided to teach in the way they had. They would also ask teachers for their opinions of the textbook and the examination, and

try to find out whether they felt that the exam was influencing particular aspects of their teaching: the way they chose the content of their lessons, the methodology they adopted, the way they designed classroom tests and the way they marked their students.

Note: Throughout this study the term *content* will refer to the type of knowledge that teachers were trying to transmit to their students (e.g. the form of a specific grammar structure, or facts relating to a particular topic), or to the general skill they were focusing on (e.g. reading, listening). *Methodology* will refer to either the general approach the teachers adopted (e.g. grammar-translation or 'communicative') or the specific techniques they used. Did they, for example, ask students to translate reading passages from English into their mother tongue? Did they ask them to read them aloud? Or did they encourage them to look for certain types of information in certain places, to cut down the time they needed to read a text for a particular purpose?

The observers would record all of this information on an observation/interview schedule which had been specially designed for the study. They would send their completed schedules to Lancaster, where I would analyse their responses and attempt to draw conclusions about the relationship between the teaching that was taking place and the examination. I would record my impressions of the data-gathering process and send my reactions back to Sri Lanka, along with instructions for the next round of observations.

The first round of observations took place in May/June 1990. This was practically equivalent to another baseline measure since very little normal teaching had taken place since the first administration of the exam 14 months before. The comments made by the observers and teachers gave useful insights into the kinds of teaching and learning that were taking place and into the teachers' views of materials and the exam. They also provided a clear idea of the conditions that many teachers were working in which might prevent washback from occurring in the near future. Amongst these conditions were the shortage of Teacher's Guides for the new textbook series, inadequate understanding of what the series was trying to achieve, a lack of familiarity with the examination due to problems in the distribution of exam support materials, crowded teaching halls often shared by as many as seven or eight different classes, and lack of teaching resources such as books, duplicating facilities etc.

Feedback from the observers led to several modifications in the observation/interview schedule: more questions were added about the school, teaching conditions, and the impact of the textbook, and more attention was given to the teacher's point of view rather than to just the observer's. A revised version of the schedule was used in Rounds 2 and 3 (July and November 1990). In December my UK colleague and I visited all the observers, to ask for their impressions of the revised schedule and their opinions of how to

improve the data-gathering process for the remainder of the study. The product of these consultations was a third version of the schedule, which asked for the same sorts of data but made the collection procedure more systematic and the analysis more straightforward than in earlier rounds. The major changes were a reduction in the number of open-ended questions and the addition of checklists which listed the kinds of text types and activities that would appear in the classroom if the examination were having an effect on the teaching. (A copy of the revised observation/interview schedule can be found in Appendix 3.)

Rounds 4, 5 and 6 took place in 1991, the first 'normal' academic year that much of Sri Lanka had experienced since before the introduction of the exam. (The school closures in 1988 and 1989 had had a 'knock-on' effect into 1990, but many areas of the country were able to enjoy a full three terms of teaching in 1991.) The size of the observation sample fluctuated during this year, due to the resignation of one of the original observation team members, the arrival of several new members, and the difficulties that all observers had getting leave to carry out the research. At its largest the sample contained 64 schools; at its smallest it contained 36.

A full analysis of the observations is available elsewhere (Alderson and Wall 1992) and a table comparing the rounds is presented in Appendix 4. What follows is a brief description of two of the rounds – a round of 'ordinary teaching' and a round of 'examination preparation' – to provide readers with the background necessary to follow the discussion in Chapters 6–11.

Round 5 – 'Ordinary teaching'

Round 5 took place in June and July 1991, five months before the examination. The observers visited Year 11 classrooms in 64 schools. They observed classes which were using the textbook and classes which were using other material.

Lessons where teachers were using the textbook: content

Approximately 75% of the teachers were using the textbook on the day of the observation. About 70% of these were working on material which resembled the exam, but almost half of them were taking their content (texts and exercises) straight from the textbook. They added nothing of their own to the lessons. Almost all of the teachers stated that they had chosen their content because it was 'next in line' in the textbook.

The other half were using texts from the textbook but were adding their own questions or activities. The questions were the sort that might appear on the exam, but this did not mean that the exam was influencing them. The teachers' desire to check their students' comprehension was understandable,

and the fact that they used certain question types more than others (mainly short-answer questions) might have been because they were not familiar with any others. Only a quarter of these teachers reported that they were preparing their students for the exam on the day of the observation; almost all the others reported that they were using the material that was 'next in line'.

It was important to consider how much attention the teachers paid to each skill area. Table 2.3 shows that the percentage of teachers focusing on reading was much greater than the percentage for any other skill, especially listening. This could have been the influence of the examination since roughly half of the exam paper was devoted to reading; however, this would not explain why writing, which also accounted for half the paper, was the focus of so few classes. An analysis of the textbook units that the teachers were covering provided a possible explanation for the discrepancy. The third column of Table 2.3 presents the percentage of unit exercises devoted to each of the skills:

Table 2.3 Round 5 – Comparison of teachers and textbook exercises by skill

Skill	% of teachers focusing on this skill	% of unit exercises devoted to this skill
Reading	52	40
Writing	27	20
Listening	5	10
Speaking	17	25
Language form	10	5

(Note: The figures in the second column add up to more than 100% because some teachers covered more than one skill in a lesson.)

The percentage of teachers focusing on each skill and the percentage of unit exercises devoted to each skill was roughly the same. Only the figure for reading seemed out of line, but this could be accounted for by the few teachers who said they were preparing their students for the exam rather than teaching what was next in line. There was little evidence from these observations that the teachers were modifying the content of the textbook because of the exam or that they were emphasizing any skills more than the textbook would have suggested, with the possible exception of reading.

Lessons where teachers were using the textbook: methodology

The Impact Study also looked into whether the exam was having any effect on the teachers' methodology. There was no evidence in Round 5 (or any of the earlier rounds) that the exam was affecting the approach to the teaching of reading of the teachers who were using the textbook. Their methods were similar to the methods that were observed in the Baseline Study classrooms, which resembled the kind of teaching that the textbook designers had tried to discourage. The examination design team had hoped that by specifying skills like 'Skimming to obtain the gist' and 'Finding specific information' in the official exam support documents and by including fairly long passages in the exam itself they could convince teachers to give their students training in reading quickly. This could take many forms – e.g. giving students hints on how to locate important points or timing them to encourage them to read more quickly. There were no occurrences of these techniques, however, nor of others that emphasized speed. Teachers were teaching their students to understand all the words and structures in every passage.

The Teacher's Guides to the textbook series may have been responsible for this state of affairs as they did not provide clear alternatives to a word-by-word approach to reading. The Guide for Years 10/11 provided 'pre-reading' and 'scanning' questions, but it did not explain how to handle the questions in the classroom. It referred to a procedure called 'Finding Out', but this procedure, which had been explained in the Guide to Year 7, encouraged teachers to develop reading by giving background information to the topic, clarifying difficult structures or vocabulary, reading the passage aloud, having the students read it silently for five minutes, reading it aloud again, and getting the students to read it aloud – all before asking the students to answer any content questions! This method might have been suitable for students at the start of their secondary school studies, but it did not seem appropriate for O level students, who had to read longer, more complex passages and needed to do everything more quickly.

Many teachers provided even more support than the Teacher's Guide to Year 7 recommended, up to the point of 'spoon-feeding'. Teachers were observed explaining *all* of the difficult words of the passage (see Teacher's Guide: 'The pupil's own skill of guessing intelligently from context and relating mutually explanatory parts of the passage MUST be given scope'), dissecting passages sentence by sentence ('They must learn to focus their attention on the GENERAL message in the first instance and not on minor details'), and doing considerable amounts of explaining in the first language ('Don't kill their interest by giving them everything "on a plate" in advance'). Students had to read at the pace of the group rather than at their own pace.

Nor was there evidence of washback on the methodology used for teaching writing. It had been hoped that the exam might reinforce suggestions given in the Teacher's Guides to get students to think about what should go into a piece of writing (e.g. Teacher's Guide to Years 10/11, pages 46, 52 and 59). In general, though, the teachers did not give students the chance to work things out for themselves. As with reading they tended to give more guidance than the Teacher's Guides recommended, even to the point of interference.

To summarize, there was no evidence of washback on the methodology in classes where the teachers were using the textbook. In fact, the way teachers presented content and got students to practise might have run contrary to some of the principles of the textbook series. There are several possible explanations for why this was happening, but one which emerged as early as the Baseline Study was that many teachers might not have understood the methodology of the textbook in the first place. Why else would they have claimed that they were using a 'communicative' approach when the observations were showing something quite different? (A more detailed analysis of the textbook series and Teacher's Guides will be presented in Chapter 8, and of the characteristics of the teachers in Chapter 11.)

Lessons where teachers were using other materials: content

There were 17 classes (25%) where the teachers were using other materials and there seemed to be washback on the content of all of them. The first indicator of this was the teachers' own admission that they had chosen content in order to prepare their students for examinations or tests: two-thirds referred to the O level examination, and one-third referred to teacher-made tests. It is known from other studies (Wall 1994) that teacher-made tests often resembled the O level examination.

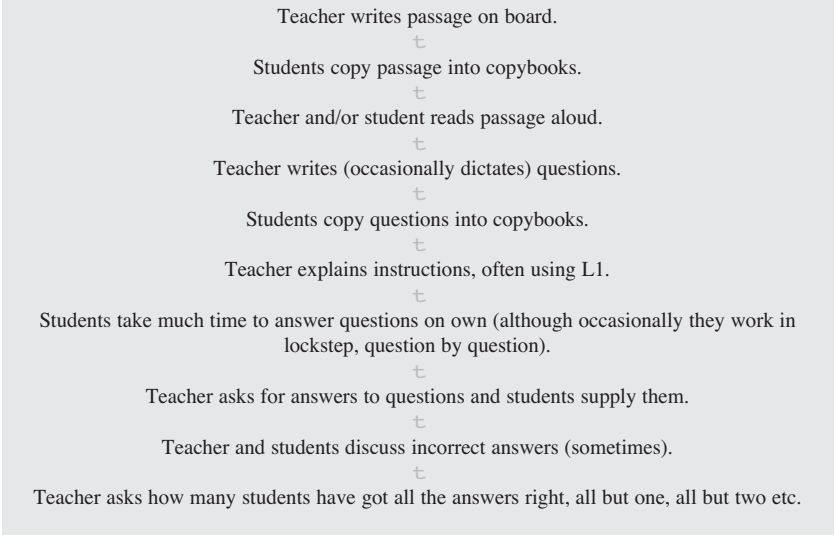
The second indicator was the emphasis that the teachers placed on written skills: 40% focused on reading, 40% on writing, and 20% on language form. There was almost no listening or speaking to be seen in these classes apart from students listening to the teacher and answering the teacher's questions.

The third indicator was the type and source of materials that were being used in the classrooms. Nearly all the reading passages and writing tasks were taken from past examination papers, official exam support material, or commercial examination preparation books. In classes concentrating on language form, the tasks were designed by teachers but the task types had all appeared on past exams.

Lessons where teachers were using other materials: methodology

The methodology of these lessons was similar to the methodology followed by teachers working with the textbook. The main difference was that these teachers had to spend a lot of time copying reading passages and comprehension questions on to the board because students rarely had their own copies of test preparation material. The typical pattern for teaching reading can be seen in Figure 2.1.

Figure 2.1 Typical pattern for teaching reading



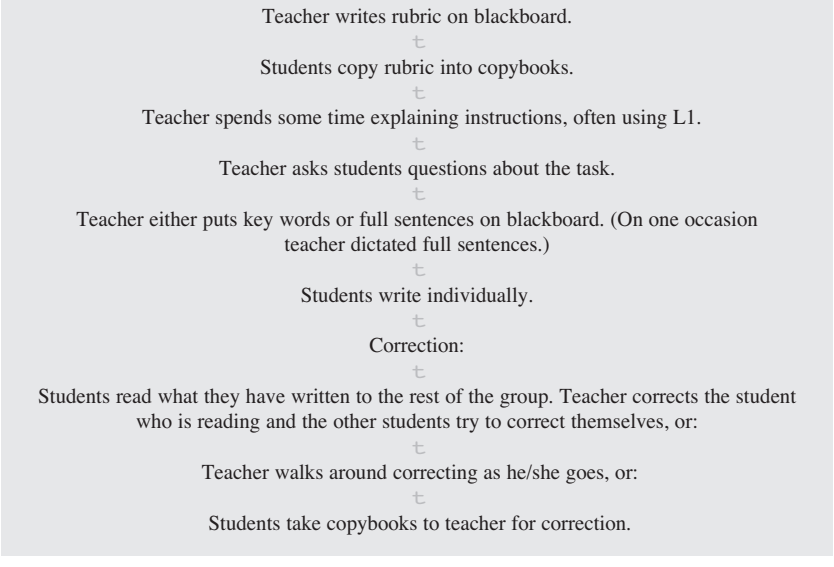
The teachers occasionally dissected texts before they asked their students to answer the questions – explaining or translating the difficult words, paraphrasing or translating difficult sentences. However, this process was often missing when it would have been most useful – when students gave incorrect answers and needed help to find out how they had gone wrong. Students' answers were discussed in only 60% of the lessons. In the remainder they were merely accepted or rejected before the teacher went on to the next question.

This way of teaching obviously precluded skimming and scanning. Students usually read through the passage several times (as the teacher was writing it on the board, as they were copying it into their copybooks, and as

they heard it being read aloud or read it aloud themselves) before they read and copied the questions. There was no opportunity for students to practise selective reading. When they were allowed to work on their own there were no strict time limits, so they did not get used to the idea of having to read quickly. When they worked in lockstep, some did not have a chance to read at all as others who worked more quickly were often asked by the teacher to supply answers to the whole group.

The pattern for writing lessons was less rigid but still visible, as can be seen in Figure 2.2.

Figure 2.2 Typical pattern for teaching writing



- In only one class were students allowed to brainstorm in small groups.
- There was no correction in several classes because the writing itself took up the whole period.
- Teachers did not discuss marking criteria in any of the classes.

Although the content of the writing lessons matched that of past examination papers, the methodology bore little resemblance to what students would have to do when sitting the exam. The biggest difference was in the amount of support that the teachers gave the students, which included writing out sentences for them to copy. The students did not get many opportunities to think independently (either alone or with their classmates) and there was no

evidence that they could name, much less explain, the criteria that would be used to judge them on the O level examination.

To summarize, there was no evidence of impact on the methodology used by teachers when they used materials from sources other than the textbook. The routines they followed when getting students to practise reading and writing did not include the types of activities which the textbook series wished to encourage and which the examination was meant to be reinforcing.

Round 6 – 'Examination preparation'

Round 6 took place in October and November 1991, approximately one month before the examination. The observers visited Year 11 classes in 41 schools. Only 12 (29%) of these teachers were using the textbook on the day of the observation; the other 29 (71%) were using different material. As was the case in Round 5, it was easier to identify examination impact in classes where the teachers were using other sorts of material.

Lessons where teachers were using the textbook: content

The content of these lessons resembled the content of the examination in 11 out of the 12 classrooms. Most of the teachers made no changes to the content of the textbook: the two who did added comprehension questions, as had occurred in Round 5. These teachers were not necessarily under the influence of the examination however: if the textbooks did not contain enough exercises then it would be natural for teachers to want to add more.

Most of the teachers were working on one of the last three units of Book 11. Table 2.4 shows the percentage of teachers paying attention to each of the four skills, as well as the percentage of exercises in the last three units which were intended to develop these skills.

Table 2.4 Round 6 – Comparison of teachers and textbook exercises by skill

Skill	% of teachers focusing on this skill	% of unit exercises devoted to this skill
Reading	58	32
Writing	33	27
Listening	0	7
Speaking	8	27

The teachers were clearly paying less attention to listening and speaking than the textbook intended, and much more attention to reading. The emphasis on reading might have indicated impact from the examination, but it was difficult to prove this given that most of the teachers reported that they were teaching the lesson which came next in line rather than doing exam preparation.

As with Round 5, there was little evidence of washback on the content of lessons where the textbook was being used.

Lessons where teachers were using the textbook: methodology

The methodology used in Round 6 was similar to that in Round 5, all of the earlier rounds and the baseline observations. The comments made in the previous section about the inappropriateness of the methodology when compared to the goals of the exam and the textbook also apply here.

Lessons where teachers were using other materials: content

About 70% of the teachers were using other materials on the day of the observations. All of them stated that they were doing 'exam preparation', and nearly all referred to the O level exam rather than to internal year-end tests.

A quarter of these teachers were using teacher-made materials and about half were using commercial publications designed to help teachers and students to prepare for the examination. The fact that so many teachers used commercial examination preparation materials is an obvious impact of the examination. A few teachers were using past examination papers, and only one was using official exam support material. It is important to note how little use was made of the official support materials, despite the fact that they were informative and well written. This may have been due to problems in distribution (only half the teachers owned or had access to either of the two official booklets), or it may have been that the commercial materials held some attraction we could not see during these observations.

The skills that these teachers were concentrating on during Round 6 were reading (52% of the sample), writing (31%) and language form (17%). No-one was working on listening and speaking. The students listened only to their teachers reading aloud or explaining the lessons. The most common form of 'speaking' was answering questions asked by the teacher, although occasionally some students were asked to read aloud.

The text types and exercise or task types being used were clearly related to the exam. The two most commonly used text types had appeared frequently on the exam, and all of the task types had appeared on past papers. In the writing classes students were practising the filling in of application forms in all but

two of the classes. This type of writing had appeared on the new exam every year. In most of the language form classes students were doing transformation exercises. Again, this type of task had appeared on the new exam every year.

It was clear that the exam was having an impact on the content of classes where the textbook was not being used, with the most obvious effect being the lack of listening and speaking, and the attention given to text and task types which had appeared frequently on past exams.

Lessons where teachers were using other materials: methodology

The methodology that was being used by the teachers who were teaching reading was similar to that used in Round 5. In approximately half the reading classes the students did not have copies of the passages they were supposed to study. Large amounts of class time were spent on writing, with the teacher transferring a text and questions from a past paper or a commercial publication on to the blackboard and the students copying from the board into their copybooks. Sometimes the students spent so much time copying that there was not enough time left for answering the questions or for checking whether the answers were correct.

In the other half of the reading classes the teachers had either borrowed sets of books which came earlier in the textbook series (students had to return their books to the school at the end of each year), or had collected money from the students to pay for the duplication of past papers, or they had asked the students to buy copies of the commercial publications. Duplicating was less expensive than getting students to buy books; however, both options were beyond the means of many families. Only one teacher had found a way around the problem of providing supplementary texts in poor areas: she brought in authentic texts from newspapers and distributed different texts to different students, allowing them to read and answer questions at their own pace and then providing answers for each student individually. This is the kind of activity that the Teacher's Guides to the textbooks and the examination support materials could have been recommending, but were not. This may be because the textbook team and the examination team did not see this type of teacher support as part of their duties. It is unlikely that they were unaware of the shortage of resources, but they may have assumed that teachers were more able to come up with solutions than they really were.

The methodology for writing classes and for language form classes was again similar to that in Round 5. Some groups managed to get through several tasks in one period if the tasks required only short responses. It was difficult, however, to deal with longer tasks like those found towards the middle and the end of the exam because there would be too much input material to copy. This meant that students from poorer families and in schools with fewer resources

were not always able to do certain types of exam practice, and were therefore less likely than their counterparts in better-off schools to do well on certain parts of the examination.

Results of the observations

Rounds 5 and 6, in combination with the earlier rounds of observations, seemed to suggest the following:

Content of teaching:

- Most teachers followed the textbook during the first two terms of the year. They worked their way through the materials, unit by unit. They may have believed that a thorough approach would help their students to do well on the exam, but there was no evidence from the observations to prove this.
- The teachers paid more attention to reading and writing than to listening and speaking, even when they were covering the textbook carefully. This might have been because the textbook itself paid less attention to these skills, or they may not have wished to cover certain skills if they did not appear on the examination.
- There was little visible exam impact on the content of reading, writing or grammar lessons if teachers were using the textbook. Teachers occasionally added comprehension questions to the day's lesson, but this might have been to compensate for a lack of suitable exercises in the textbooks rather than because of the exam. It is important to note, however, that the changes that they did make were always in the direction of the exam.
- There was obvious impact from the exam when the teachers were using materials other than the textbook. They themselves said that they were doing examination practice and the materials they used were designed for this purpose.
- The third term was openly examination-oriented. Teachers finished or abandoned their textbooks and began intensive work with past papers and commercial publications to prepare their students for the exam. No attention was given to listening and speaking. Much attention was given to topics and text types that teachers believed might appear on the examination.

Methodology:

- There was no relationship between the methodology that teachers used, whatever the time of year, and the methodology which was recommended in the textbook or which might have seemed suitable for developing the

skills that were needed in the examination. It is worth noting here that when the observers were asked to judge the effectiveness of the teachers they had visited, they judged fewer than half to be effective. They were not convinced that the teachers understood the basic principles of the textbook or that they were in command of communicative teaching techniques.

The observations gave a clear picture of what was happening in the classroom but they could not cast much light on the teachers' reasons for doing what they were doing. It was therefore necessary to gather other types of data in order to understand what was happening at a deeper level.

The need for complementary data

Although the observations revealed a great deal about the relationship between teaching and the examination, they also gave rise to questions like the following:

- Many teachers reported that they were teaching the lesson which was next in line in the textbook. Why were they doing this?
- Some teachers reported that they concentrated on certain parts of the textbook in order to prepare their students for the exam. Which parts did they think were most important?
- Were there any topics or activities that teachers consistently omitted?
- If teachers brought in supplementary material, what were they trying to achieve?
- What was the attraction of the commercial publications that so many teachers were using in Rounds 5 and 6?
- How much did the teachers really understand of the aims of the textbook series?
- How much did they really know about the exam?
- How much influence did they feel the exam had had on the way they chose their content and methodology, and the way they designed and marked their tests?
- Did they believe the exam influenced their teaching in Years 9 and 10, when the exam was still a long way off?
- What other factors influenced the way that teachers conducted their lessons?

Various methods for collecting data were used in earlier stages of the evaluation project: *document analysis*, in order to find out about the aims

of the ELT reform programme and the examination development activities; *materials analysis*, in order to determine whether the exam had had any influence on teacher-made tests; and *questionnaires*, in order to find out how teachers, teacher advisers and examiners were reacting to the examination. The first two methods were not suitable for delving into the 'why' of teachers' actions. Questionnaires were a possibility but they were problematic for at least two reasons. The first was that respondents sometimes did not understand the questions and/or gave answers that were not relevant. The second was that their responses were sometimes difficult to interpret and it was not possible to probe to get a better idea of what they were saying. Using *interviews* would provide a way of overcoming these problems. (Details of the research procedure are presented in Chapter 5: Methodology.)

I conducted ten group interviews in November 1991, with 64 teachers in seven different regions of the country. The results of the first analysis of the data were incorporated into the final report for the project (May 1992). Although I had confidence in the first analysis, I remained curious about whether I had taken from the data as much as the teachers had to offer. I later decided to carry out a new analysis of the same data, using ideas about examination impact that I was becoming aware of in general education, as well as ideas on educational innovation that I had learned about during my involvement in a number of ELT projects in the mid-1990s.

In Chapter 3, I will discuss the educational literature on examination impact, and in Chapter 4, I will discuss a number of important issues in the area of educational innovation. These chapters, along with the description of my methodology in Chapter 5, will provide the foundations for the new analysis of 'old data' which makes up Chapters 6–11.