CSE ENGLISH EXAMINATIONS:
AN EVALUATION OF THE PROCEDURES EMPLOYED
BY THE EAST MIDLAND REGIONAL EXAMINATIONS
BOARD TO ASSESS ORAL AND WRITTEN ENGLISH

A thesis submitted by VALERIE BROOKS for the degree of Doctor of
Philosophy in the University of Leicester.

April, 1984.
CONTENTS

Acknowledgments

Introduction

1. Examining English: the Nature of the Problem
2. The Present Enquiry: its Nature and Purpose

Chapter 1: The Early Years (1850-1947)

1. The Earliest Public Examinations
2. The Situation by 1868: the Report of the Schools Inquiry Commission
3. The "Persistently Perennial" Nature of these Problems
4. The Recommendations of the Schools Inquiry Commission
5. The Positive Effects of External Examinations
6. The Report of the British Association for the Advancement of Science (1903)
7. The Creation of the Board of Education (1899)
8. The Report of the Consultative Committee on Examinations in Secondary Schools (1911)
9. The Introduction of the School Certificate Examination (1917)
10. An Assessment of the School Certificate Examination (1917-1950)
11. The 1944 Education Act
12. Mounting Hostility Towards External Examinations (1929-1943)
13. The Norwood Report (1943)
15. Concluding Remarks

Chapter 2: The Immediate Context (1947-1965)

1. The Provisions made by the SSEC (1947)
2. Evidence of the Failure of the 1947 Provisions
4. The Modern School: the Original Concept
5. The Unrealised Ideal of Parity of Esteem
6. "The Days of Wandering"
7. "'Contracting in' to the 'Competition-Examination-Success' System"
8. School-Based Assessments
9. The Beloe Report and the Modern Schools
10. The Beloe Report's Fears Justified
11. The Beloe Report on Internal Examinations
12. The Beloe Report on GCE and Non-GCE External Examinations
13. CSE: a New Concept of Examining
Chapter 3: Continuous Assessment

A. Introduction

1. The Rationale for this Investigation
2. The Divisiveness of External Examinations
3. Examinations as Sampling Procedures

B. The Findings

1. Introduction
2. Mark Weightings
3. Question 15(i): Motivation
4. Question 15(vi): the Dual Role of Teacher-Assessor
5. Increased Professional Dialogue
6. Obstacles to Full Involvement
7. Question 15(ii): Internal Validity
8. Question 15(iii): a Fairer Method of Assessment?
9. A Threat to the Validity of Continuous Assessment
10. Question 15(iv): Pupils' Workload
11. The Size of Course Work Requirements
12. Question 15(v): Teachers' Workload
13. The Importance of Extraneous Factors
14. Question 15(x): Comparability
15. Moderation
16. The Intrusion of Personal Factors
17. Potential for Abuse
18. Concluding Remarks

Chapter 4: Literature

A. The Formative Years: Literature Examinations Prior to CSE

1. Introduction
2. The Prevailing Philosophy of Literature Teaching in the 1960's
3. The Emphasis on Factual Detail
4. Reasons for the Emphasis on Factual Content
5. The Backwash Effect on Teaching
6. Separate Certification
7. O-Level: a Disincentive to Wide Reading
8. The Inappropriateness of the Prescribed Texts
9. Enjoyment of Reading
10. Alternative Forms of Examining
11. The Search for a New Style of Examining
12. The New Mode 1 Examinations
13. An Evaluation of these Examinations

B. The Findings

1. Continuous Assessment or External Examination?
2. The Internal Validity of Continuous Assessment
3. Pupils' Perceived Response to Course Work Assessment
4. Workload
5. The Effects of Excessive Course Work Requirements
6. Certification
7. The Preference for Separate Certification
8. Probable Consequences of Separate Certification
9. The Unification of English and Different Methods of Certification
Evidence of Dissatisfaction

Moderation: the Logistical Problems

Personality Factors

Concluding Remarks

Appendix 1: the Questionnaire

Appendix 2: Preliminary Information Sheet

Appendix 3: Introductory Letter to Respondents

Appendix 4: Sample Interview Questions

Bibliography
ACKNOWLEDGMENTS

Thanks are due to many people who have helped me in this study. I am indebted to the East Midland Regional Examinations Board for providing me with information and official documents relating to CSE, to the numerous teachers throughout the East Midland Region who completed the questionnaire, to the people who gave so freely of their time for personal interviews, to Mrs. Janet Broughton for her painstaking work in typing this thesis and to my parents and husband who have supported and encouraged me throughout the period of this research. However, my greatest debt is to Mrs. Margaret Mathieson, Senior Lecturer, and Professor Maurice Galton of the School of Education, University of Leicester who have been my supervisors. Their guidance and criticisms have been invaluable to me.
INTRODUCTION

1. EXAMINING ENGLISH: THE NATURE OF THE PROBLEM

English has always raised special problems for examining as is witnessed by the publication of books such as *English versus Examinations* (Jackson, 1965) and Whitehead's (1966) often-quoted comment:

... all past experience suggests the existence of a deep and inherent incompatibility between external examinations as we know them and the essential aims of good English teaching.

(L. p. 233)

Likewise, Dixon (1965) claimed that: "Many of the valuable experiences of English lessons are inevitably lost sight of in a mass examination" (p.27).

The reasons for this incompatibility are varied.

Firstly, it is important to recognise that English, as it has been defined in recent prescriptive writing, is not a subject in the normally accepted sense of the word:

> English is distinctive in the curriculum in that it is both all-pervasive and yet has relatively little subject matter of its own.

(Newsom Report, 1963, p.159)

Some argue that the grammar of the language constitutes its body of knowledge while others argue that literature represents the subject matter around which the study of English is built. Nevertheless, it is true to say that recent prescriptive writing perceives English to be concerned principally with the promotion of qualities and competences which are of broad personal and social importance rather than the transmission of specialized knowledge and skills. This is reflected in titles of books on the subject e.g. *English for Maturity* (Holbrook, 1961), *Growth through
English (Dixon, 1975), Sense and Sensitivity (Creber, 1968), The Preachers of Culture (Mathieson, 1975) and A Language for Life (the Bullock Report, 1975). Likewise, the Newsom Report (1965) claimed that: "The overriding aim of English teaching must be the personal development and social competence of the pupil" (p.153). As Shayer (1972) pointed out:

Few people stopped to consider the real nature of the subject, or came near to discovering that most paradoxical of all English facts, namely that it is not really a "subject" in the traditional sense at all. Perhaps this "content" myth is the fallacy behind all the others, that English, like certain other subjects, is little more than a body of factual knowledge to be transferred from teacher to pupil ... and lending itself naturally to examinations. (p.19)

This is at the heart of the assessment problem which English presents because whereas the factual nature of other subjects provides them with an objective basis, English is essentially subjective. Most recent prescriptive writing has urged teachers to encourage responses which are personal and original and to foster imagination, creativity and sensitivity in the use of and response to language. Such emphases present problems for assessment because mass examinations are most efficient at assessing the kind of factual material on which examiners can readily agree and which can be marked with a high level of objectivity. This has provided a persistent source of conflict between the aims of English teaching and the object of most examinations. Whitehead (1965) described the effect of this on literature, observing that the examination's quest for reliability:

... is achieved by concentrating on those tasks which examiners can agree about among themselves ... "Facts" are what makes standardising possible ... Consequently memory for facts, even insignificant facts, takes on an absurd importance. Many

* There are two related criteria which govern the quality of an examination: validity and reliability. A valid test actually measures the skills and/or knowledge it is intended to measure. The reliability of an individual score is judged in terms of its freedom from error i.e. it is a true measure of the things which the test purports to evaluate. There are three main sources of error: disagreement amongst markers; deficiencies in the test and variability in a candidate's performance from one occasion to another. A reliable score, therefore, is one which can be reproduced, either by the same person on a different occasion or by any number of independent assessors marking on the same occasion.
questions ask for nothing else ... this fosters a tendency to treat poems, novels and plays as though the valuable thing about them were their extractable content of event or argument - sometimes with the implication that it is important to remember quite small details ... since this is literature ... and not history, the facts which gain credit from the examiner are in reality fictions treated over-literally. What we observe in such an examination is indeed the creation of a unique class of mark-gaining items which we might call (borrowing a coinage from James Joyce) ficts. (pp.275 and 279)

The examiner's quest for reliability had overridden the English teacher's concern for perceptiveness and a sensitive, personal response to literature. Since an examination must first and foremost be reliable and since subjectivity is associated with a poor level of reliability, the attempt had been made to make literature examinations as objective as possible. Clearly, the essentially subjective nature of English was at the centre of the problem.

The problems which subjectivity raises for assessment may be illustrated by referring to a study involving a comparison of marker reliability in nine Irish Leaving Certificate subjects, including Mathematics and English (Macnamara and Madaus, 1969). An objective discipline such as Mathematics lends itself readily to objective testing and consequently a high level of reliability is usually achieved but even Mathematics is capable of yielding unreliable results once an element of subjectivity has been introduced into the assessment. The Mathematics results in this study were only slightly more reliable than the English results. The poor reliability of the English examination came as no surprise since the English essay was: "... notoriously troublesome to markers" (Macnamara and Madaus, 1969, p.12). The real surprise was that: "... marker unreliability in Mathematics is scarcely less pronounced" (Ibid, p.13). The source of this unreliability can be traced directly to the mark scheme employed. Examiners were instructed to subtract:

- Full marks for "A very serious blunder"
- 10 marks for "Blunders or serious omissions"
- 5 marks for "Less serious blunders or omissions"
- 3 marks for "slips"
Thus, an examiner had to distinguish between three different types of blunder: "a very serious blunder", "a blunder" and a "less serious blunder". He also had to distinguish between two types of omission: "serious omissions" and "omissions". When an examiner is required to make this sort of distinction, he is being asked to make a subjective assessment and reliability tends to deteriorate. Generally speaking, objectivity is associated with a high level of reliability and subjectivity is associated with a poor level of reliability.

In English, the problem is much more intractable because subjectivity is inherent in the very nature of the test*. Secondary school Mathematics is a largely objective discipline and thereby amenable to objective testing procedures. Although the assessment procedure described above was subjective, it is relatively easy to construct objective mark schemes for Mathematics. In the case of English, the testing instrument itself is highly subjective and predetermines the nature of the assessment which can take place. For example, an essay title randomly selected from a GCE or CSE English Language Composition Paper could be expected to produce an infinite variety of responses. Indeed, the more original the response, the more likely it is to be valued. It would be impossible to devise an objective mark scheme to encompass and account for the multiplicity of responses an essay title may be expected to produce. As Ballard (1923) pointed out:

One of the defects of the essay as a measuring device ... is the impossibility of making the essay amenable to rigid objective measurement ... An essay is an intricate mental product which can be analysed in a variety of ways and yet can never be analysed completely ... it fails through its very wealth and complexity. (pp.61-62)

* With the exception of multiple choice or other fixed-response or semi-objective procedures. However, these methods are much less popular and less widely used in English than the free response/continuous prose answer.
Admittedly, analytical mark schemes* are devised in an attempt to make the
assessment more objective. Nevertheless, these do not define or quantify
the actual content of the essays; they are limited to describing the
qualities which an essay is expected to exhibit. Thus, although
analytical schemes may facilitate agreement amongst examiners as to the
range of qualities that are to be evaluated, the amount of subjectivity
involved in the assessment is reduced very little. As soon as an examiner
is required to judge the extent to which a particular quality is present in
an essay, he is being asked to make a personal judgement similar to that
required of the Mathematics examiners in the study quoted above. This
judgement is coloured by all sorts of idiosyncracies: personal preferences,
previous experiences, values etc. This has been confirmed by a controlled
experiment which evaluated the reliability of analytical and general
impression mark schemes for assessing O-Level English Language composition
scripts:

... analytical marking was somewhat flexible in interpretation ...
The analytical marking of compositions is not completely objective
and gives room for a certain amount of latitude by the marker so
that general impression comes into play.

(Morris, 1968, pp.9 and 17)

Clearly, the subjective nature of the English essay and its assessment puts
the English Language composition in a special position amongst school
subjects. It goes far towards explaining why the assessment of English has
earned a reputation for being especially unreliable and why, when a number
of essay-based examinations are compared, English is likely to yield the
least reliable results (see, for example, Murphy, 1978). It was problems
such as these that the Regional English Panels for CSE had to face.

In fact, by 1965, when CSE was introduced, hostility towards

* In analytical marking, each criterion in the mark scheme is awarded a
separate mark and the total mark is arrived at by adding these together.
In general impression marking, the mark is awarded on the basis of the
examiner's overall impression of an essay. Although specific criteria
may be borne in mind, these are not assessed separately.
conventional examinations was mounting:

Criticisms of the General Certificate tests in English have been mounting for years ... A glance at back numbers of the quarterly "The Use of English", for example, will reveal scores of hostile comments. (English Examined, 1966, p.1)

In view of the difficulty of reconciling examinations with good English teaching, it may seem somewhat surprising that the Beloe Report (1960) had recommended the creation of yet another examination in English. However, its intention was that CSE should not become just another examination. The importance of CSE was that it was designed to be quite different from any other examination:

... it is our considered view that many of the defects and dangers we found in the existing examinations arise not from the intrinsic nature of examinations at this level but from the particular circumstances in which these examinations have grown up ... We are convinced that, if certain conditions are fulfilled, external examinations can make a constructive contribution to the educational process at this level. (Beloe Report, 1960, pp.27-28)

As Whalley (1969) observed:

The significance of the CSE examination lies not, therefore, in its presence as yet another examination which will aggravate the problem presented by the harmful effects of examinations, but as an attempt to solve it. (p.7)

It is in this light that CSE should be judged.

Although CSE represented one of the most far-reaching revolutions in examining that this country has experienced, almost two decades have passed since this great experiment was launched and:

... no authoritative survey has yet been attempted of the achievements and lessons of this recent work. (Dixon, 1977, p.8)

The present enquiry is an attempt to evaluate some of these "achievements and lessons" in the field of English in the East Midland Region.

2. THE PRESENT ENQUIRY : ITS NATURE AND PURPOSE

Since its beginnings in the 1850's, school examining had evolved as an external activity. Examinations were conducted by authorities which
were quite separate from the schools which used them. CSE sought to revolutionize this by giving the teacher a position of real power and influence in the new examination:

Effective teacher control of syllabus content, examination papers and examining techniques is the rock on which the CSE system will stand. (Examinations Bulletin, 1963, p.3)

Since the teacher was made the pivotal point in the new system, it was felt that the best way of gaining insight into the workings of CSE English in the East Midland Region was to concentrate on the views of practising teachers. Therefore, all of the data collected in this enquiry came from practising teachers. Their perspectives form the basis for an evaluation of CSE English in this region.

The initial stage of the investigation involved interviews with people who had held positions of responsibility in connection with the examination at various stages of its development. The questions revolved around the examining procedures employed in CSE and their appropriateness for English. Although some standard questions were used in all of the interviews, others varied according to the experience of the interviewee (see Appendix 4 for a sample of the type of questions used). Some of the questions were structured but many were open and, once an issue had been raised, the interviewee commented freely on any aspect of it which was of particular significance to him or her. Initially, the areas for investigation were suggested by background reading on the aims of CSE and by discussions with board officials including the Secretary of the Board and the officer responsible for English.

The interviews not only provided a sense of the history and development of the examination but also gave an insight into the varying perspectives of people who had fulfilled different functions. The interviewees fell into two broad and sometimes overlapping categories: those concerned with the operation of the examination and those involved in formulating policy and devising syllabuses. The first category included two Mode I Chief
Examiners, a Chief Examiner for a large Mode 3 Consortium, an Assistant
Chief Examiner, an LEA Adviser for External Examinations, a Mode 3
Assessor and Mode I Literature Moderator. The second category included a
Chairman of the Regional English Panel, two Vice Chairmen and two other
members of the Regional Panel. Since all of these people had, at some
stage, been closely involved in the running of the examination, it was
decided to redress the balance by interviewing a number of teachers whose
experience of CSE was limited to the classroom. Nine teachers, representing
Modes I and 3, all of whom had considerable experience of teaching CSE but
had not held any special responsibility in connection with it, took part in
this. These teachers came from schools throughout the region. The
interviews, which lasted between one and three hours, were detailed and
searching. Each interview was tape-recorded and then transcribed so that
the quotations used in the study are an exact record of what interviewees
said. In order to preserve their anonymity, each interviewee has been
assigned an identity number with the prefix "I" to show that the comment
was made during an interview.

At the same time as these interviews were being carried out, a detailed
literature search was undertaken. On the basis of the interviews and the
background reading, a questionnaire was devised dealing with the major
issues they had identified. Since the assessment of English is such a
broad topic, it was decided to limit the field of enquiry by concentrating
on those areas where EMREB had made important innovations: the assessment
of literature, spoken English and continuous assessment. The questionnaire
was divided into four sections. The first dealt briefly with the
respondents' personal details: teaching experience, experience of CSE,
mode of examination used etc. A second, longer section was general and
dealt with various issues including the value of different assessment
procedures, the desirability of retaining some of them in the light of the
proposed revision of the examinations' system at 16+ and the demands
associated with CSE. Continuous assessment was also included in this section. The third section focussed on the assessment of oracy and the final section upon literature (see Appendix 1). It was felt that it was important for the research to concentrate on those areas of the assessment which represented fresh departures rather than on those which used existing methods about which a considerable amount of information was already available. Therefore, the Mode I written language papers received comparatively little attention since the testing of language by a final examination was a well-established practice by 1965. This is not to suggest that there were no new or significant features to this assessment. Nevertheless, any changes which did take place were made within the conventional examination framework. Thus, the questions were designed to probe the attitudes of the sample towards the more novel aspects of the English assessment and the questionnaire went through numerous draft versions before a pilot study was conducted.

The pilot study took place in Leicestershire and four schools were selected to represent the variety of school type, size, geographical location and mode of examination to be encountered in the region. A total of twenty teachers took part in the pilot study and, after they had completed the questionnaire, meetings were arranged in order that they might comment on it. The final version of the questionnaire incorporated a number of modifications to the wording and format of the questions made as a result of these meetings.

Most of the questions took the form of fixed-response items and respondents simply had to tick the box or boxes opposite to the appropriate response(s). This question format was chosen for two reasons. Firstly, because the area of enquiry was broad, the questionnaire was long (twelve pages in all). It was appreciated that if busy teachers were to be expected to complete such a detailed questionnaire, some method must be found to facilitate completion, making it quick and easy. However, it was
not simply considerations of expediency which prompted the choice of fixed-
response items. If the questions had been open-ended, the data collected
would have been idiosyncratic and incomplete and it would have been
difficult to establish any common ground between the different viewpoints
or to reach definite conclusions on particular issues. The value of fixed-
response items is that they ensure that respondents address themselves to
the issues which have been identified as important by the researcher and
that their responses take a form which can be meaningfully analysed and
compared. However, it was recognised that fixed-response items do have
drawbacks. They force respondents to place their own views into categories
into which they may not comfortably fit and they allow them no freedom to
amplify their own points of view or to qualify the responses they have
made. For these reasons, spaces were provided throughout the questionnaire
where respondents could comment freely. This blending of fixed-response
and free-response items in the questionnaire was an attempt to combine the
best of both procedures.*

It was for similar reasons that two different methods of enquiry, the
personal interview and the large-scale, postal questionnaire, were employed.
The one-to-one interview can provide a detailed insight into an
individual's views but limitations of time and manpower mean that such a
procedure can only be used on a small scale. A large-scale study,
representative of the views of teachers throughout the region, necessitated
using a method whereby a large amount of data could be collected. Thus,
this study combines large-scale, quantitative data with qualitative data in
the belief that the two can be mutually enlightening. It was also felt to

---

* At the outset, complete confidentiality was guaranteed to both the
individual and his school. Therefore, each subject was assigned an
identity number. Where their comments have been quoted, these identity
numbers are prefixed with the letter "Q". This is to distinguish them
from the interviewees whose numbers are prefixed with the letter "I".
The column of boxes in the righthand margin of each page of the
questionnaire was for the purposes of coding and computation.
be important to collect some large-scale quantitative data since so many of the issues raised by the assessment of English have only been studied on a small scale. It is difficult, therefore, to know how far the views expressed are representative and how far they are the idiosyncratic views of the individuals concerned. The uncertainty is increased by the degree of conflict amongst the views expressed. Therefore, in an area characterized by small-scale studies and conflicting personal opinions, it has been useful to collect some large-scale quantitative data.

The third stage in the enquiry was to select a sample of subjects who would receive the questionnaire. Since the enquiry was intended to be representative of the entire region, considerable care had to be taken with the design of this part of the investigation. The East Midland Region is a large area covering the counties of Leicestershire, Nottinghamshire, Derbyshire, Northamptonshire, Lincolnshire and South Humberside. Not only do these counties vary in size and type but their educational provision varies enormously ranging from large, non-selective community colleges to selective grammar schools and secondary modern schools. If the sample was to be representative of the region as a whole, the inter- and intra-county variations in size had to be maintained in the sample.

The EMREB Register of Centres was used to select the sample. Prior to selection, colleges of further education and special schools were eliminated as unrepresentative but independent schools were retained. A total of 149 schools was selected using random number tables and the sample was stratified in accordance with regional variations in size. This ensured adequate coverage of the region as a whole and kept the proportions within and between areas constant (e.g. Derbyshire accounted for 90 useable centres whereas Lincolnshire had only 45. Therefore, the initial sample included 38 Derbyshire schools and 19 Lincolnshire schools.)

Initial contact was made with the Head of English Department in each school who received an introductory letter explaining the nature and
purpose of the enquiry. He/she was asked to complete a brief information sheet (see Appendix 2) which requested the following details:

(a) did the English Department use CSE?
(b) if yes, which mode(s) was/were used?
(c) the names of teachers whose responsibilities included CSE English courses.

A stamped addressed envelope was enclosed with each letter. 89% of the schools contacted (i.e. 133) replied. Eighteen of these schools did not use CSE for English and so the final sample included 115 schools. The information thus provided enabled the researcher to contact each teacher personally and request his/her assistance in the research (see Appendix 3).

640 teachers received a questionnaire together with a covering letter and a stamped addressed envelope. The questionnaires were dispatched in September 1979 and a 69% response rate was achieved. The resulting data was analysed using the Statistical Package for Social Scientists (Nie et al., 1975).

This study opens with two chapters which trace the history of examinations from 1850 until the introduction of CSE in 1965. CSE was both a response to and a reaction against a history fraught with problems. As Whalley (1969) pointed out:

In order to appreciate the significance and purpose of the CSE, and the point behind many of the details of its system, it is necessary to consider the background which this sequence of events provides ... although the CSE examination is a relatively new examination it is closely concerned with issues which have been of great importance for many years. Its introduction is not an isolated incident. (p.8)

Many of the problems which CSE was designed to overcome had persisted for more than a century which meant that CSE was tackling the same problems in 1965 that had beset examinations in the 1850's. As Ebel (1972) remarked:

... the problems of educational measurement are persistently perennial ... the problems ... that bothered our predecessors still bother us. (p.27)

This is not to suggest that CSE was part of a smooth evolutionary process: the next phase in a logical process of development. CSE is
important to the history of examining because it represented a turning point, a radical break with precedent. Whereas previous reforms had displayed a reluctance to revise radically the existing system, the most striking feature of CSE was its experimental nature. This was appreciated by the Minister of Education, Boyle (1963), who praised CSE for its "...willingness to break new ground":

The establishment of the Certificate of Secondary Education presents the teachers with a unique opportunity to create, for the first time in our educational history, an examinations' system which is wholly the servant of the schools.  
(Examinations Bulletin 1, 1963, p.iii)

It is in this light that the present study evaluates the development of CSE English in the East Midland Region.

Although Chapters 3, 4 and 5 are concerned with analysing and interpreting data from the interviews and questionnaire, the historical perspective is borne in mind throughout and colours the evaluation. Chapter 3, which is concerned with continuous assessment, opens by looking briefly at the history of this procedure and what it was hoped could be achieved by a greater use of it in CSE. It then examines the use which has been made of this technique in the East Midland Region and evaluates its usefulness, commenting on its pitfalls and problems as well as its successes. Chapter 4 is concerned with literature. A review of the state of literature examining prior to 1965 is followed by an appraisal of how far EMREB's revolutionary approach has succeeded in overcoming the characteristic problems of literature assessment. Examining spoken English is the subject of Chapter 5 which traces the history of oral examinations and the way in which this has influenced the standing of oracy in recent years. This is followed by a consideration of the assessment problems which the oral medium raises and the respondents' attitudes to these problems. Finally, each of the procedures which has been used by EMREB in its Mode I test battery (reading aloud, the talk, one-to-one conversation, group discussion and the continuous assessment of course contribution) is
analysed separately. A final chapter is devoted to drawing together
briefly the main findings of the survey and trying to reach some overall
conclusions. The bibliography is organised in alphabetical order by
sections. Works which are only referred to in a particular chapter are
referenced under the appropriate section. Those which are referred to in
more than one place appear in the general section at the beginning as do
works mentioned in the Introduction.

CSE was an attempt to revolutionize not only techniques of examining
but also the existing power structure within examinations. Almost twenty
years have passed since the introduction of CSE and, because so many of
its procedures were novel in 1965, it is now timely to consider what
teachers think has been achieved. This investigation was encouraged by
John Dixon's regret, expressed in 1977, that:

In general, over the past twelve years there has been more
flexibility and experiment in the public examining of English
than over the previous half century ... Unfortunately no
authoritative survey has yet been attempted of the achievements
and lessons of this recent work. (Dixon, 1977, p.8)
CHAPTER ONE

THE EARLY YEARS (1850-1947)

1. THE EARLIEST PUBLIC EXAMINATIONS

It is customary for histories of examining to open with an account of the ancient Chinese system of public examinations used to recruit civil servants (see Judges, 1969). However, between this ancient Chinese system and the sudden growth of examining in England, in the middle of the nineteenth century, there is little evidence of any civilisation producing a well-developed system of external examinations. Public examining was only established on a significant scale in the mid 1800's when, "Examining became a universal industry, Britain leading the civilised world" (Judges, 1969, p.29). In fact, the Oxford and Cambridge "locals", established in 1858, have been described as, "... the real beginning of the school-leaving examination" (Ibid, p.30). Thus, although the history of education stretches back over many centuries, public examining is a relatively new phenomenon, the roots of which are in the middle of the preceding century.

The number of public examinations instituted in the 1850's is quite remarkable. The following list provides some indication of the mushroom-like growth which took place in that decade:

1853    College of Preceptors
1853    Royal Military Academy at Woolwich
1855    Home Civil Service
1855 Indian Civil Service
1858 Oxford Delegacy "locals"
1858 Cambridge Syndicate "locals"
1858 University of Durham
1858 University of London

Even this does not convey the full extent of the development because it only refers to examinations established in the 1850's and those which were used by schools. There were numerous other examinations created later in the nineteenth century as well as those devised for institutions other than schools.* Why did public examining experience such a sudden expansion during this period?

Although it is impossible to provide a comprehensive answer to this question within the context of this chapter, it appears that this growth reflected major social changes then overtaking Britain. Patronage had been the traditional method of recruitment to positions of rank and importance but, as society became increasingly complex, there was a shift away from the traditional method of recruitment towards proven merit and ability. Some of the changes in the fabric of British society which led to the adoption of competitive examinations are outlined by Montgomery (1978):

Great Britain flourished at the centre of a mighty empire. Large social organisations of a new order were developed to cope with responsibilities at home and abroad, and these required competent officials if they were to operate smoothly. There was pressure in the middle of the nineteenth century for the introduction of tests of efficiency or even for open competition for recruits to important public posts... The middle of the nineteenth century saw a vast extension of examining... Examinations suited the spirit of the age and certificates came to be produced for candidates in their thousands... So all-embracing was the whole apparatus of examinations, and such was the demand for qualified persons that the very nature of society shifted. It was apparent by 1958 that a revolution had taken place clearly described by Michael Young in "The Rise of Meritocracy". (pp.2, 3 and 5).

* For example, the examinations devised by the further education institutes for the working classes (see the Beloe Report, 1960, p.4) at a period when secondary education was largely the privilege of the fee-paying middle and upper classes.
The composition of society was also changing:

At this time the middle classes were increasing fast, and the endowed and private schools increased in size and number to provide for them. Examinations came into being to meet the needs of these schools for standards. (Beloe Report, 1960, p.3)

Thus, in the early stages, this growth complemented fundamental social changes as well as expansion in the field of secondary education. The 1911 Report of the Consultative Committee on Examinations in Secondary Schools drew attention to the valuable role which examinations had played in this expansion:

... we wish to record at an early stage in our report our firm conviction that an immense debt of gratitude is owing to these bodies (i.e. the examining authorities) for the services they have rendered to national education... during all these years when the State has been unwilling or unable to do anything, these various bodies saw what was needed in the schools and set themselves to supply it. We have no doubt whatever that they should have much of the credit for the improvements in Secondary Schools during the last 50 years. (pp. 4-5)

However, since there was no central body to oversee and co-ordinate these developments, examinations tended to spring up on an ad hoc basis in response to particular needs:

Headmasters and mistresses, influenced either by patriotism to their own university or by the belief that one particular examination was more suited to their own curriculum, showed anxiety to be examined by one particular body, and so helped in many cases to lead to the establishment of more than one examination centre in the neighbourhood. (Ibid, p.30)

In many ways, this freedom from state control, and the ability to set up examinations as and when they were needed, was useful in the early days. Despite these advantages, the situation was a dangerous one as the Schools Inquiry Commission (1868) found.

2. THE SITUATION BY 1868: THE REPORT OF THE SCHOOLS INQUIRY COMMISSION

Unchecked by any form of state direction, secondary education quickly found itself flooded by, "... an uncontrolled proliferation of
examinations" (Beloe Report, 1960, p.3). The disastrous consequences of this inside the schools were well-documented by the Schools Inquiry Commission (1868):

When a school has to prepare boys for several different examinations, an adaptation of the school course to suit them all becomes impossible. One boy, who is reading for the army, has to be taught one set of subjects; another, who is to be a medical student, has to be taught another. It is easy, if the examinations are very stringent, to push this divergence between the different studies required, so far, as to make effective organisation of the school, as a place of general education impossible. (Volume I, p.324)

These conflicting requirements were seriously disrupting school curricula and the problem was compounded by professional bodies eager to jump on to the "qualifications bandwagon".

In addition to these examinations conducted by academic bodies, a large number of professional bodies organised general entrance examinations of their own for admission to each avocation. This had the effect of creating a large number of external authorities each of which had liberty in a sense to make its own rules for general school education. The bewildering variety of standard and requirement imposed by these various bodies... interfered very considerably with the systematic organisation of the curriculum.

(Report of the Consultative Committee on Secondary Education, 1939, p.38)

Another disturbing effect noted by the Commission (1868) was that examinations distorted a school's educational goals:

It may well be admitted, that the authorities in charge of the schools will always find it necessary to be on their guard against allowing the examinations to override everything else... Examinations... are liable to abuse; and if boys at school are induced to view them as the be-all and the end-all of school life, it is probable, that the good which they do in stimulating study, will be very dearly purchased... there is a further mischief... Such examinations have a tendency to dislocate the school work by rewarding highly what the school values low, and disregarding what the school makes of great importance. (Vol.1, pp.322-323)

That which is referred to nowadays as the "backwash effect" of examinations is by no means a new phenomenon. External examinations were undermining the wider educational ideals of schools less than twenty years after their introduction.
One of the most pernicious aspects of this backwash effect was that the work of the schools became biased in favour of an intellectual élite at the expense of the majority:

They tend to direct the attention of the master, not to his school, but to the cleverest of his scholars, and are therefore a temptation to neglect all those whom he believes it to be impossible to prepare for passing the examinations... Mr Green... reports that he constantly found the classes under the head master of a grammar school reading a book, which was plainly too hard for the majority, because it was prescribed for the next local examination, for which only one or two were going in. (p.334)

Examinations were also used for purposes they were not designed to fulfil. Thus, the Commission found that the complaint that the Oxford and Cambridge locals were too difficult was largely due to pupils below the age of sixteen attempting them:

The severity of the examination is often a subject of complaint... The requirements are probably not much, if at all, too severe for the boys who are on the point of leaving school; but... the examinations are unsuited to boys who have not yet reached the same point. (p.333)

The real crux of the problem was the absence of state involvement which left the field wide open to private enterprise. Without "... the directive power of central government" (Roach, 1979, p.58), it was possible for any number of external authorities, professional as well as academic, to establish examinations and demand their own peculiar requirements. This produced a situation which has been described as a "nightmare" (Spens Report, 1939, p.254). Although with hindsight, it is clear that state intervention was going to be vital to any attempt to rationalize the "system", in the age of the entrepreneur, state involvement was unpopular. The government, which followed a laissez-faire policy in many areas of public life, showed itself unwilling to become involved in the field of examining. This is illustrated by the origins of the Oxford and Cambridge locals which evolved out of an
experimental examination devised by a committee in Exeter in June 1857 (see the Report of the Consultative Committee on Examinations in Secondary Schools, 1911, pp. 11-12). In the first year, the promoters turned to the state, requesting the assistance of HMI's in validating the results of the examination. Although the request was granted, a marked reluctance to make state involvement a permanent feature of the examination was expressed and the committee subsequently approached Oxford and Cambridge for assistance in making the scheme permanent. This helped to establish examining as a non-state, non-centralized activity in which all future growth would be haphazard and uncontrolled. Thus, by 1911, the Report of the Consultative Committee observed that existing examinations were:

... the outcome of a varied succession of unconnected accidents, in which owing to the disinclination or inability of the State to superintend the work at its beginning, there has been no single co-ordinating force at work. (p.30)

Naturally, individual examining bodies were unlikely to initiate reform; there was little incentive for them to change so long as they were able to satisfy the need for external standards. Their vested interest in the continuation of the existing state of affairs meant that any attempt to check this wasteful situation must come from the state.

3. THE "PERSISTENTLY PERENNIAL" NATURE OF THESE PROBLEMS

A striking feature of these problems is the speed with which they had emerged. They testify to the powerful influence external examinations were exerting over schools less than twenty years after their inception. Clearly, the newness of the examinations and the small scale of secondary schooling provided no safeguard against difficulties which were to dog examinations throughout their history. In fact, almost every problem identified by the Commission would be instantly recognised
by teachers of any decade during the following century. For instance, more than half of that period was characterized by a glut of external examinations whose growth was totally unchecked by a central, coordinating body. It was the severity of this problem in the 1950's which made the creation of CSE a matter of such urgency. Facts such as this endorse Ebel's (1972) claim that, "... the problems of educational measurement are persistently perennial... the problems... that bothered our predecessors still bother us" (p.27). Likewise, the backwash effect of external examinations has been a tenacious problem. Teachers have persistently complained about the way in which their teaching is subjugated by rigid examination requirements. This problem has been especially acute in the case of English where there has been a fundamental incompatibility between the aims of enlightened English teachers and the nature of conventional examinations (see pp.i-vi). This conflict has proved to be one of the most difficult to resolve. Thus, the Schools Inquiry Commission's finding that examinations, "... dislocate the school work by rewarding highly what the school values low and disregarding what the school makes of great importance" (p.323) has a special pertinence to English and is a problem of central importance to the present enquiry.

The tendency of prestigious examinations to encourage a preoccupation with the most able pupils at the expense of the rest has also been a "persistently perennial" problem. The description of grammar schools in the 1860's (p.5) could, in fact, have been written ninety years later so accurately does it reflect the position in many secondary modern schools during the 1950's. Some modern schools became so obsessed with competing with grammar schools that entire schools were geared to cater for a small élite capable of taking O-Levels. The Commissioners' stricture that the locals were being used for purposes they were not
intended to fulfil also has a familiar ring to it. Between 1917 and
1951, the matriculation level of the School Certificate was also widely
misused. Although it was originally intended to exempt that small group
of pupils who would go forward to universities from taking matriculation
examinations, the matriculation level came to be regarded as a superior
general school-leaving qualification. Increasingly, employers requested
a matriculation certificate, instead of an ordinary pass, as a
condition of entry to certain occupations and the "matric" became the
aim of all ambitious pupils. Subsequent revisions of the School
Certificate, which led to the creation of GCE (1951), suffered a
similar fate. GCE was designed for the small group of children educated
in the grammar schools while the majority, who were educated in the
modern schools, were to be catered for by school-based assessment
procedures. However, modern schools used GCE at an ever-accelerating
rate despite the obstacles put in their way and ministerial discouragement.
In fact, by 1959, a third of all GCE entries were from institutions
other than grammar schools (Beloe Report, 1960, p.9). The commissioners'
warning that examination requirements were militating against the efforts
of schools to provide their pupils with a broad education up to the age
of sixteen would have also struck a chord with teachers during the
century which followed.

Examinations, moreover, have frequently failed to keep pace with
the rate of progress in secondary education. Throughout this period,
secondary education was expanding and changing rapidly. Examinations,
on the other hand, displayed a characteristic tendency to become fixed
and stereotyped. This meant that educational circumstances repeatedly
outgrew the examinations which catered for them and a mismatch between
the rate of development in secondary education and the retarded rate of
growth in examinations is at the root of many of the problems encountered.
Again this has been particularly acute in English which has been one of the most important growth areas in secondary school curricula. From a position of insignificance at the turn of the century (see Shayer, 1972), English has achieved a position of central importance in the modern curriculum. This rise in the status of English has been accompanied by important changes in its philosophy. Thus, English has suffered particularly from the insensitivity of examinations to changes in thinking about the nature of a subject.

4. THE RECOMMENDATIONS OF THE SCHOOLS INQUIRY COMMISSION

Contemporary attitudes towards state intervention are illustrated by the way in which the recommendations of the Commission were received. Realising that reform was unlikely to come from any other source, the Commission proposed, "... the creation of a Council of Examinations to consist of twelve members, two to be elected by each of the Universities of Oxford, Cambridge, and London, and six to be appointed by the Crown ...
... To this Council should be assigned the duty of drawing up the general rules for the examination of schools, and of appointing the examiners" (Volume I, pp.649-50). Considering the equal representation of academic and state interests on the proposed Council, one could hardly interpret this as an attempt to impose excessive state control.

Even so, the idea proved to be premature and quite alien to the prevailing mood of the period. Thus, although the original draft of the Endowed Schools Bill (1869) provided for the establishment of the Council, the idea aroused such vigorous opposition, especially amongst public school headmasters, that it was abandoned.

Although this early attempt at reform was abortive, it did have some important repercussions. Firstly, the formation of the Headmasters' Conference (1870) arose out of the concerted opposition of the public
school headmasters to the proposed council and secondly, the Oxford and Cambridge Joint Examination Committee was formed (1873) at the invitation of the Headmasters' Conference. Although this board originally offered only a Higher Certificate for sixth formers requiring a university entrance examination, in 1884 it added to the welter of examinations at the 16+ level by introducing a Lower Certificate. This tangle of examinations persisted into the twentieth century.

5. THE POSITIVE EFFECTS OF EXTERNAL EXAMINATIONS

This account has, so far, paid exclusive attention to the adverse effects of examinations. Although they were in urgent need of reform, it would be difficult to explain their continued use without acknowledging their positive attributes. Therefore, before passing on to the twentieth century, it is necessary to redress the balance by mentioning the beneficial effects which external examinations were perceived to have.

The evidence presented to the Schools Inquiry Commission (1868) revealed a consensus of opinion: in spite of the deficiencies which individual examinations might display, the overall effect was propitious:

There is almost universal agreement that some sort of... examination cannot be dispensed with if the schools are to be maintained in thorough efficiency. Even the best masters will not do so well without this aid as with it. (p.619)

This "tonic" effect was noted frequently by the commissioners:

While pointing out several serious faults, our Assistant Commissioners generally reported very highly of the good results which these examinations have already produced. Mr Stanton speaks of the quickening effect exercised on both teachers and scholars. Mr Fitch... remarks that all the best and most vigorous schools in his district made use of them, and that the good influence was perceptible, not only in the candidates sent in, but even in the lower classes... The steadily increasing number... of candidates that received certificates, is strong proof... of the good effect of thus constantly presenting a definite aim to the minds of teachers and scholars. (pp.330-331)
In fact, the Commission was sufficiently impressed by the benefits to be derived from examinations that they made them the hub of their suggested reforms: "... the examination... is the pivot of all the improvements that we have recommended" (p.648). Thus, although existing examinations were recognised as being deficient, it would be wrong to suggest that examinations in principle were regarded as necessarily evil or incapable of reform.

6. THE REPORT OF THE BRITISH ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE (1903)

This report provides a particularly revealing insight into the situation at the turn of the century because it records the views of practising teachers and university tutors. The Association had compiled a questionnaire, dealing with the impact of examinations, which was despatched to members of the Headmasters' Conference and university tutors. Because it produced a verbatim report of their replies, this survey provides a unique insight into contemporary opinion on the examinations' system. The Report did not make any startling discoveries nor did it suggest that a more established system of examinations was throwing up new problems. Instead it highlighted the persistence of problems originally identified by the Schools Inquiry Commission suggesting that these were causing anxiety amongst an ever-increasing body of teachers.

As might be expected, a major cause for concern was the web of conflicting demands made by different examining authorities:

At the present time external examinations are carried to such wasteful and mischievous excess that they are doing more harm than good to the advancement of education... Every professional body appears to hold that it is forwarding education, or perhaps rather satisfying self-respect and rising to its
position, by instituting schemes of examination and insisting on those particular tests as alone valid. Every British university frames its own scheme of subjects and books, to the exclusion of all others. Colleges, commissioners, boards, and committees follow the same course in respect of scholarship or admission tests. The schemes are arbitrary, conflicting, and needlessly inelastic, and together make havoc of all unity in school curricula. No one familiar with the whole field can place much faith in the opposed and contradictory conclusions enforced by different authorities. (826,* p.435)

It is hardly surprising to find that the plethora of examinations was as much a subject of concern in 1903 as it had been in 1868 since the one attempt to rationalize the system (1869) had failed. Moreover, despite the already crowded nature of the examinations' system, new authorities were still coming into existence. By 1903 the London University Extension Board and the Northern Universities' Joint Matriculation Board had joined the field and Bristol University established its School Examinations' Board in 1911.

When questioned about, "The need of unifying examinations with the object... that certain examinations may serve a common purpose" (p.444), respondents were "practically unanimous" (p.444) and their comments convey a sense of urgency:

... the crying need is unity and simplification... to supersede or replace the multifarious preliminary examinations (academic or professional) that now exist. (826, p.445)

This is the most urgent question in connection with secondary education at the present time. There ought to be a single qualifying examination for boys leaving school between sixteen and seventeen. (S16, p.445)

The need is paramount, especially for the smaller schools. (S17, p.446)

The need for this is very great. (S10, p.446)

If this report is to be accepted as a measure of the climate of opinion at the turn of the century, it provides strong evidence of the discontent

* All respondents were assigned identity numbers to preserve their anonymity. An 'S' preceding the number indicates that the statement was made by a schoolteacher and a 'U' denotes a university tutor.
which the glut of examinations had provoked.

The 1902 Education Act magnified this problem by greatly increasing the number and size of secondary schools. These schools turned to examinations to provide them with public credibility and to give their curricula form and purpose. In this way, the hold of examinations over the schools was strengthened and the 1903 Report provides firm evidence of this:

For less able boys the... encroachment becomes much more serious. ... for one, two, or (in extreme cases) three terms of the final school year boys must be withdrawn from parts or even the whole of the school curriculum, forfeiting the stimulus, the emulation, and the interest that attaches to collective learning, and must be set in ones or twos to prepare the particular subjects or authors imposed by the authority to whose regulations he must conform each detail. (p.436)

Particularly serious difficulties were created by "group" examinations in which the candidate had to achieve a pass in a number of subjects in order to gain a certificate:

It may happen that a boy has to get up the same subjects over and over again, to pass some examination in which he has been ploughed in some other subject. For instance, I have known a boy's whole education at a standstill for a year while he is getting up some one subject for the London Matriculation; and I have known boys go in four or five times for the L.C.C. Intermediate Scholarships, getting up the same 'elementary experimental science' year after year, their scientific education meanwhile being at a standstill. (p.437)

This situation was increasingly anachronistic now that secondary education was evolving into a national system. The 1902 Education Act had, "... marked a very important stage in the development of a national system of education" (Spens Report, 1939, p.64) and a system of examinations with national currency was needed to complement this development. Furthermore, the increasing rate of mobility added to this need for rationalization.
The Board of Education had warned that, "... 'external examinations were having unfavourable reactions on the work of many schools, often leading to cramming and over-pressure; that they frequently set wrong ideas before schools and pupils'." (quoted by the Spens Report, 1939, p.79) and there was plenty of first-hand evidence of this in the 1903 Report.

As to the effect of examinations upon study and teaching... (they)... certainly tend to narrow, not to widen, the range... from the examination point of view what pays is close adherence to the standard commentary or text book on the subject. (S26, p.436)

General examinations in all subjects are wholly pernicious in their effects, not only in checking individuality and progress among teachers, but in tending to substitute facility of reproduction for originality of thought amongst the taught. (S30, p.436)

... the whole question has weighed on me like a nightmare. I believe that examinations as they are... are giving a totally wrong trend to education. They are subsidising the receptive and discouraging the training of the instructive powers of the mind; they are encouraging a sort of cut-and-dried mode of teaching and learning which would have driven Arnold wild... cramming disgusts its victims. (S1, p.439)

... teaching is narrowed and everything neglected which does not pay. (U18, p.440)

... the final year of school preparation is broken up by distracting and discordant examination calls. (S26, p.436)

S25 described the results of external examinations thus:

(1) confine teacher to old ruts and discourage all attempts at improvement. (2) cause the teacher to devote disproportionate amount of time to written work. (3) foster 'cram'. The teacher considers not what is best for the boys but what the examiner will ask. (p.438)

S26 identified the crux of this problem when he observed that unless teaching and examining are closely co-ordinated, the educational goals of the school are liable to be sacrificed in the cause of examination successes. Enlightened teachers increasingly came to feel that the price paid for examination success was too high in educational terms. This problem has been particularly pertinent to English and will be a
central issue in the main enquiry.

The situation was aggravated by the tendency of the examinations to encourage early specialisation. In this respect also, they were in direct conflict with the schools which aimed to provide pupils with a general education up to the age of sixteen. They were aptly described as an, "... 'inducement to prefer premature successes to sound education'" (p.444). S26 regarded the situation as quite absurd: "'I have no faith in first forcing specialisation by scholarship tests and then attempting to redress the balance by enforcing supplemental subjects through subsequent requirements. This will but aggravate the evil and produce successive bouts of cram'" (p.441). Ull, however, sympathised with the teacher's dilemma:

The schoolmaster is compelled (a) by the natural desire to advertise his school, (b) by the absolute necessity of meeting the reasonable wishes of parents, to prepare his boys for open college scholarships, obtainable only by candidates under nineteen years of age, and therefore to allow them to specialise as soon as they show any special aptitude. This seems to me a misfortune. (pp.443-444)

The products of this system were described thus by Ull0:

... for their knowledge of chemistry and physics, scholarships are awarded to boys of eighteen who have in far too many cases a very inadequate grounding in mathematics, are ignorant of history and of modern languages, possess a smattering of Latin and cram up subsequently enough Greek to carry them through... Worse still is the condition of the mathematician who as regards general education is more poorly equipped than the rest. (p.443)

The case for a school as a place of general education had to be argued in 1903: "'Much will be gained when it is clearly recognised that school work must be general... that specialisation is the work of the universities,'" argued S13 (p.442). Examinations had become a strait jacket to many teachers, confounding their broader educational aims and frustrating enlightenment in teaching.
The comments were not limited to criticism of the existing system; suggestions for improvements were offered and it is interesting to note the prophetic quality to some of the recommendations. For instance, S25 advocated that "'Examiners should be drawn only from experienced and enlightened school masters... In the setting of papers the teacher should have a voice and a right of veto..." (p.438). This proposal was revolutionary in the first decade of the twentieth century and, although the following statement may sound like a truism now, its truth was less obvious in 1903: "'Curricula should not be controlled by examinations but directly, and the examinations... arranged to fit the curricula'" (p.439). The proposal by S26 that, "'... schools should, so far as possible, have examinations individual to themselves moulded upon the lines of their own teaching'" (p.445) harks forward to the proposals of the Norwood Committee (1943) and the eventual development of Mode 3 examinations under CSE (1965). The reform suggested by S14 to eliminate the existing tangle of examinations was, in fact, the solution which was finally implemented in 1917: "'We want two certificates based on as wide a freedom of teaching as possible - one for boys sixteen to seventeen, one for boys eighteen to nineteen'" (p.446). U10 also presaged future developments when he urged that, "'It is very desirable for teachers to be represented on examining bodies'" (p.450).

Perhaps the most striking feature of these criticisms is the resemblance they bear to those made forty years earlier in the evidence to the Schools Inquiry Commission (1868). However, the similarity between the circumstances prevailing in 1868 and those in 1903 is hardly surprising since the one attempt to overhaul the system (1869) had failed. A more vivid illustration of the "persistently perennial" nature of these problems is provided by more recent events. Modern examining reached a...
criterion point in the 1950's and the dissatisfactions voiced in that
decade could have been expressed fifty years or even ninety years
earlier so similar were the problems. However, circumstances in the
1950's were quite different from those in 1903 or in 1868. In fact,
between 1903 and the 1960's, examinations underwent major revision on
two separate occasions (1917 and 1951) and yet, despite these reforms,
the problems which beset examining at the beginning of the century re-
emerged in the 1950's.

Another similarity between 1868 and 1903 was the continued belief
that examinations were worthwhile. Despite enumerating the evils of the
existing system, the general tenor of opinion was still optimistic about
the benefits to be derived from a reformed system. Just as they had in
1868, contemporaries saw the way ahead in improvements to the system, not
its total abolition. The prevailing mood of the Report is conveyed by
Sl's remark that "'The true remedy seems... to be to reform and not to
abolish examinations'" (p.439). Likewise, the introduction stressed
that, "While pointing out the many evils which attend examinations, the
majority take the view that in some form they are necessary" (p.439).
The predominant mood was perhaps best summarized by U7 who "'While
recognising the evils of the system is forced to recognise its very great
merits as a most useful instrument in proper hands'" (p.440).

7. THE CREATION OF THE BOARD OF EDUCATION (1899)

External examinations had been increasing their sphere of influence
since 1850 thus demonstrating the need that they were fulfilling. As
long as this demand continued, and the government refused to intervene,
there was little incentive for the examining authorities to rationalize
their activities. Clearly, the onus must lie with the state and this is
why the creation of the Board of Education represented an important landmark. Its significance has been described thus:

... the creation of the Board of Education with greater powers of central direction had brought about important changes. The discussions about school examinations before 1900 had lacked any central focus, which the state... was now able to provide... The Committee itself remarked that 'the state is now organised as to be able to supply both motive power and unifying influence'. State activity indeed was to be crucial in carrying through the changes which followed. (Roach, 1979, p.48)

However, at the outset, the Board of Education was unwilling to tackle the problem as its reaction to the recommendations of its own Consultative Committee (1904-1905) revealed. At the request of the Board, the Consultative Committee had investigated examination standards and had drawn up proposals for a new school-leaving certificate. These proposals were, however, rejected by the Board on the grounds that: "... it was likely to elicit 'a fear that the state as a rival and competing agency in the conduct of examinations may ultimately command a monopoly'" (quoted by Roach, 1979, p.47). This reaction suggests that the traditional antipathy to state intervention was unabated.

Although the Board appeared reluctant to accept responsibility for examinations, these recommendations may have been badly timed. The 1902 Education Act was a major undertaking which had only recently been enforced and, as Roach (1979) pointed out, "In the early years of getting the new system of secondary education to work, the Board perhaps felt that it had enough on its hands without tackling secondary school examinations as well" (p.47). Nevertheless, state involvement was to be crucial in carrying through reform. Indeed, Roach (1979) has described the recent history of examinations as, "... one chapter in that general theme of twentieth century educational history - the steady growth in the directive power of central government" (p.58).
8. THE REPORT OF THE CONSULTATIVE COMMITTEE ON EXAMINATIONS IN SECONDARY SCHOOLS (1911)

The Board's increasing sense of urgency about examining is indicated by the fact that although it rejected the proposals of its Consultative Committee in 1905, only six years later the Committee returned to the topic, at the request of the Board, and produced a major report. This represented a turning point in the evolution of examinations because it was the first stage in a series of events which led to the state taking action to end the existing state of affairs.

The Committee described the prevailing situation as, "... a state of chaos" (p.1) and they denounced the existing range of examinations as "simply bewildering" (p.1). They regarded the need to streamline examinations as their prime objective and recommended that all existing examinations should be replaced by a single, two-tier system of certificates with national currency:

The first external examination to be taken normally by a pupil in a secondary school, save for what may be found necessary in the form of admission or scholarship examinations, should be one which would be a suitable test of the general attainments of an average pupil of 16 years of age... The only other external examination to be taken... should be one which would be suitable to the attainments of pupils of an average age of 18 or 19. (pp.106-107)

There was a sense of urgency in their treatment of the subject:

It would be difficult to exaggerate the importance of the question of examinations in secondary schools... The crux of the whole problem is the external examination which... is the point in the pupil's school career which cannot fail to influence to a large extent his previous school training... so vital is the question of these external examinations that we have devoted the greater part of our Report to its consideration. The further questions of examinations at entrance and during school life, important as they are in themselves, are of... comparatively minor urgency. (pp.1 and 3)

A feature of the existing system which caused the Committee deep concern was the pressure for early specialisation. In order to combat this
tendency, they suggested that the new certificate should be group-based rather than subject-based:

... (candidates) should be examined in each of the subjects which they have studied in the last year, or perhaps two years, of their approved course, and should be expected to obtain a fair aggregate of marks on the examination as a whole. Unless this is done there must always be a danger that young pupils will be allowed to drop useful but uncongenial subjects with a view to over-specialisation... at too early an age...

Further, we think that candidates should be expected to reach a given standard in a stated number of the principal subjects of their curriculum. (p.117)

This "group" requirement was regarded as a particularly happy solution to the problem of early specialisation because it would also ensure that a richer curriculum was introduced into those schools where the existing curriculum was narrow or incomplete:

... the Board of Education has attempted a double task. Where undue specialisation was practised, it has endeavoured to correct this defect; on the other hand, where the curriculum was scrappy, it has encouraged greater concentration of purpose upon the central needs of a liberal education. (p.112)

The Committee also believed that this group requirement would safeguard the freedom of schools to tailor their curriculum to suit their own needs:

The fact that they would have to be examined in all the subjects of their last year or so, and to reach a certain aggregate of marks on the whole examination, would probably make it possible to leave them considerable freedom in choosing the individual subjects in which a qualifying mark had to be obtained... (The Report) has refrained... from any precise definition of the range of instruction which a secondary school should provide and... this freedom is to be preserved. (pp.112 and 117)

One of the Committee's main aims was to shift the balance of power inside schools away from extraneous authorities and back to the schools themselves: "... the schools should be more or less supreme in deciding upon the pupil's course up to that age (i.e. 16 years)" (p.91).

Having laid down its principles for reform, the Committee was faced with the question: which type of administrative machinery would be best
fitted to bring about these changes? The administrative model finally
decided upon was only accepted after serious consideration, and rejection,
of two other proposals. The three schemes were: a system of external
examinations controlled by provincial authorities; a system of external
examinations organised by the Board of Education and a system held under
a widely representative council with executive powers. The Committee
eventually chose the third option, a solution which had originally been
suggested in 1868. They felt that it would enable them to tap the wealth
of expertise already in existence as well as being the least antagonistic
choice in a situation fraught with vested interests. They described it
as a "... means equally efficacious and economical, but less drastic
and less wasteful of existing experience" (p.132) and one which combined
"... the least sacrifice of experience and of good-will with the best
security for wise and effective reform" (p.132). The proposed council
was to be "... widely representative in character... (including)
representatives of the Universities, of the Local Authorities, and of
the teachers... as well as of the official experience of the Board of
Education itself" (p.132). In addition, "... a limited number of persons
of practical experience especially of the requirements of professional,
industrial, and commercial life" were to be included. This council
represented the most crucial element in the Committee's recommendations
because the onus for implementing change lay with it:

The function of the Council would be the supervision of all
external examinations in recognised secondary schools. It
would lay down regulations as to the scope, time, and method
of these examinations. It would control their organisation,
fix the fees to be charged for admission to them, and approve
the examiners. In all secondary schools aided or recognised
by the Board no external examinations would be permitted
except those held under the authority, or with the approval,
of the Examinations Council. (p.133)

The traditional attitude towards examining had been one of optimism
about the potential good to be derived from a reformed system (e.g. in
the Report of the Schools Inquiry Commission (1868) and the Report of
the British Association for the Advancement of Science (1903)). This
mood of confidence was undiminished in the 1911 Report as its opening
remarks illustrate:

... though we have criticised many of the existing examination
arrangements, we must not be taken as implying any antagonism
to external examinations as such. On the contrary, we consider
that under proper conditions, they are a necessary and a
valuable part of the educational machinery of a good school.
(p.4)

Thus, the faith in what might be achieved through a revised system of
examinations persisted.

9. THE INTRODUCTION OF THE SCHOOL CERTIFICATE EXAMINATION (1917)

In spite of the Consultative Committee's sense of urgency, "The
Board moved with somewhat deliberate speed after the appearance of the
1911 Report and, after a series of conferences with the interested
parties, did not issue its own proposals until July 1914" (Roach, 1979,
p.49). Progress was further interrupted by the outbreak of the First
World War (1914) and the reforms were eventually implemented in 1917.
Following the recommendations of the 1911 Report, the new examination
provided a single, two-tier system of assessment. The School Certificate,
designed for 16+ candidates, and the Higher School Certificate, for
pupils of eighteen or nineteen, became the only recognised examinations
for use in secondary schools. Grant-aided schools were forbidden to
use preliminary examinations designed for pupils younger than sixteen
years and the Board undertook to pay a fee to all schools on the grant
list whose pupils entered for the new examination.

The School Certificate was put under the control of the seven
university schools' examining boards. In order to ensure the equivalence of standards between them and to oversee the functioning of the new examination, the Secondary School Examinations Council was created. This was, of course, the central council which the Consultative Committee had designated such an important position in the new administrative machinery. In the event, "... the committee's plans for a central council to co-ordinate the work of the existing examining bodies was rather more ambitious than what was eventually achieved in 1917" (Roach, 1979, p.49). Nevertheless, it did represent an important stage in state involvement in public examining and its duties included maintaining standards and comparability between the boards, negotiating with university and professional bodies about exemption from additional examinations, promoting improvements within the system, investigating complaints and organising conferences.

Since the aim of the Committee had been to rationalize examinations, an interesting innovation was the inclusion of two pass levels in the School Certificate thereby enabling it to serve a dual purpose. The pass level was intended as a certificate of general education for pupils going straight into employment while the credit level denoted a higher level of attainment and was designed to give university entrants exemption from matriculation and professional examinations. It is not difficult to understand why making a single examination serve a dual function seemed such an economic idea to generations that had suffered from a glut of examinations. In its original intention, this innovation was expected to have a beneficial influence on the work of the schools.

Another deficiency which the reformers set out to correct was the "strait jacket" effect of external examinations in schools. Firstly, the pressure for early specialisation had a restrictive effect on the curriculum and this was compounded by the inflexibility of most examination
requirements. The suggested solution to this was the "group" requirement. Individual subjects were divided into three main categories: English subjects; Foreign Languages and Mathematics and Science. To obtain a School Certificate, candidates had to obtain passes in at least five subjects and a minimum of one subject had to come from each group. A fourth group, including subjects such as Music and Art, was added later but this group was never made compulsory. The reformers saw this group requirement as the guarantee of a sound general education and the Boards supported this aim by offering "... a wide range of alternatives in their syllabuses for this examination" (Spens Report, 1939, p.80). Although the intention was to safeguard the freedom of the schools to devise flexible curricula, it is important to recognise that by introducing a group requirement and a minimum number of subjects requirement, the School Certificate immediately placed constraints on this freedom. This was at variance with the spirit of the 1911 Report which had urged that, "... the schools should be more or less supreme in deciding upon the pupil's course up to that age" (p.91).

There was a growing awareness of the need to involve teachers in public examining. The 1911 Report had argued that:

... the judgement of the teachers should be more systematically taken into account in the conduct of external examinations... (The examinations) should give credit to good work done prior to the examination, especially for work which cannot justly be estimated by examination methods... teachers should have the power of making formal representations to the examining authority as to the suitability of the examination. In most cases we think this could be most conveniently done by giving teachers a definite representation on the council of the authority... Secondly, the teacher should have the right to submit to the examiner his own estimate of the relative merit of the pupils in his class. (pp.83 and 120)

The principle of teacher involvement was given further support by Circular 849 (Board of Education, 1914) and when the reforms were finally initiated, Circular 1002 stated that "... adequate arrangements must be
made for involving teachers, for permitting schools to be examined on their own syllabuses if they so desired, and for requiring heads of schools to submit their own estimates of the probable performance of their pupils" (Roach, 1979, p.51).

The role allocated to the teacher in the reformed system did not constitute revolutionary change. If the Mode 3 style of examining is taken to represent an extreme form of change which provides the teacher with a much more powerful role in public examining, then the measures adopted by the School Certificate allowed for a modest amount of teacher participation. They included Mode 2 type examinations which may be described as a compromise between the conventional Mode 1 form of examining and the revolutionary Mode 3 style of assessment. Nearly thirty years were to pass before the use of Mode 3 was recommended by the Norwood Report (1943) which advised the abolition of external examining at 16+. However, the Norwood proposals were never implemented and it was more than twenty years before Mode 3 was used on a significant scale under CSE (1965).

The principal features of an examinations' system which was to last more than three decades were laid down in 1917. School Certificate represented a landmark in the history of examining because: (a) it marked the end of more than half a century of deliberate non-intervention by the state - an acknowledgement that the state had an important role to play in ensuring an efficient, waste-free examining system and (b) it eliminated "the nightmare of a multiplicity" (Spens Report, 1939, p.254) of external examinations and replaced it with a unified system possessing national

* In Mode 1, both the syllabus and the examination are externally set and marked. In Mode 2, the syllabus is devised internally but externally approved and examined. In Mode 3, the school devises its own syllabus and carries out its own assessment with external moderation by the examining authority.
currency. The period in which this examination held sway has been characterized thus:

Fed by the free place system at one end and aiming, for an important minority of its pupils, at college and university places at the other, the secondary school of the inter-war period was dominated by the examination system with its twin peaks of the School and Higher School Certificates.

(Roach, 1979, p.46)

10. **AN ASSESSMENT OF THE SCHOOL CERTIFICATE EXAMINATION (1917-1950)**

Certain aspects of the School Certificate were successful from the outset. For instance, the decision that: "... the universities should be recognised as the responsible bodies through whom the examinations... should normally be conducted" (Spens Report, 1939, p.254) made an important contribution to its status. The uncertain status of a new examination is a potentially serious obstacle to its acceptance; indeed, the low public esteem of CSE is a stigma which has marred this examination throughout its existence. However, this is a problem which the School Certificate never had to face because its prestige was immediately secured by the decision to assign its supervision to the universities. As Whalley (1969) observed:

... the close association of the universities with the new examination undoubtedly gave the examinations prestige and helped them to gain acceptance. (p.10)

There is another reason why inviting the co-operation of the universities proved to be expedient. Prior to 1917, the universities operated their own matriculation examinations but, when the reforms were initiated, the SSEC was able to negotiate terms of exemption, with the majority of universities, for pupils who obtained School and Higher School Certificates. If the new examining authorities had not been associated with the universities, it would have been difficult to gain
exemption from matriculation examinations. The School Certificate was therefore successful in eliminating one element of the "examinations glut" which would have been difficult to remove in any other way. It was also successful in removing excess at the 16+ level of examining. Even the Spens Report (1939), which was extremely critical of the School Certificate, conceded that: "... it has served a very useful purpose in freeing the secondary schools from the nightmare of a multiplicity of external examinations" (p.254). Thus, in one fell piece of reform, a unified examinations' system with national currency had been created. This was a great achievement.

Throughout their history, external examinations had been noted for their tonic effect on the work of schools and School Certificate was no exception to this rule:

... the School Certificate Examination has in fact performed important services for the education provided by Grammar Schools. During the past twenty years it has been one factor in raising the general standard of attainment in the ordinary school subjects to a level which could hardly have been reached without the incentive offered by the prospect of success in a public examination. (Spens Report, 1939, p.256)

Another area where the School Certificate enjoyed considerable success was in the prevention of early specialisation. The group and the minimum number of subjects requirements effectively eliminated this.

Clearly, there were several important areas in which the School Certificate had been highly successful. However, many of its novel features were specifically designed to meet deficiencies in the existing system and a new system, whose consequences it was impossible to predict, was thereby created. Ironically, it turned out that the solutions to the old problems threw up new difficulties to take their place.

In summarizing the inter-war period, Roach (1979) aptly described the schools as "dominated" by the School Certificate examination. One of
the most "persistently perennial" problems associated with external examining is captured in that word – domination. As early as 1868, the Schools Inquiry Commission was pointing to the very real danger of examinations becoming, "...the be-all and the end-all of school life!" (p.323) and in 1903 the problem was just as acute with respondents to the British Association questionnaire denouncing the way in which education was dictated by examinations: "'At present examination rules education'", declared S32 (p.438). Thus, by 1917, the backwash effect of external examining was well-known and the reforms were expressly designed to combat this. The "... needlessly inelastic" requirements of the existing examinations were abolished and the Boards conducting the School Certificate offered "... a wide range of alternatives in their syllabuses" (Spens Report, 1939, p.80). Likewise, the group and minimum number of subjects requirements were designed to ensure breadth and flexibility in school curricula and to deter early specialisation. In spite of these measures, School Certificate quickly became the dominant factor in secondary education. The extent to which it influenced grammar schools by the 1930's is revealed by the Spens Report (1939):

... Most of our witnesses seemed unable to think of the curriculum except in terms of the examination, while some defined the curriculum entirely in such terms. (p.142)

An illustration of what this could mean in practical terms was provided by Valentine (1932): "One of my students told me that in her own school the pupils studied the Hanoverian period of English history for four consecutive years... because that was to be the period selected for their School Certificate examination" (p.16). On the basis of such evidence, the Spens Report (1939) concluded that, "... one of the gravest results of the present system" was that:
... the examination should be regarded alike by parents and children as the main, or even the sole, objective of the education given in the Grammar School. "The modern mania for examination results," one of our witnesses said, "renders many recipients of the School Certificate less cultured and efficient than they might have been without it." (p.266)

The crux of this problem is that an examination syllabus is a skeleton which it is the teacher's duty to "flesh out" into an appropriate teaching syllabus. In practice, this distinction between the broader, richer teaching syllabus and the narrow, precise examining syllabus is frequently blurred or disappears altogether so that the examination syllabus becomes synonymous with the teaching syllabus. This has been one of the most serious objections to external examinations throughout their history and it has proved to be one of the most difficult problems to solve. Even the range of measures adopted by the School Certificate, to emancipate the schools, failed to achieve their aim.

It has already been noted that, in solving existing difficulties, the School Certificate created new problems to take their place and this issue provides a case in point. The School Certificate had succeeded in ridding the system of excess but since it became the only recognised examination for use in secondary schools, this boosted its importance. The more popular the examination became, the more likely it was to become the main preoccupation of parents, pupils and teachers. Ironically, its success in overcoming the, "... nightmare of a multiplicity of external examinations" (Spens Report, 1939, p.254) simultaneously increased the likelihood of it becoming "... the be-all and the end-all of school life" (Schools Inquiry Commission, 1868, p.322).

One of the most disturbing features of this preoccupation was the pressure it placed on pupils:
... the pressure placed upon individual children is often wholly excessive. One of our witnesses spoke of "that feature of secondary education which causes the gravest alarm at the present moment among those who are closely in touch with young people - the tendency... to overwork and overstrain the adolescent pupil". Another witness added that "those who know schools from the inside are only too familiar with the sense of rush, hurry and overstrain". (p.257)

Another aspect of School Certificate which was designed to combat existing deficiencies, but gave rise to new ones, was the group requirement.

Originally devised as a built-in safeguard of the freedom of schools to provide flexible curricula tailored to their own needs, in practice this controlled and limited school curricula. Even though the examining authorities, "... offered and continue to offer a wide range of alternatives in their syllabuses", the group requirement acted as a constraining factor which placed "... certain subjects or groups of subjects in so sharply preferential a position and has in fact permitted one subject, French, to become in the greater number of schools virtually compulsory" (Spens Report, 1939, p.263). Thus:

... the majority of pupils presented for it offer a comparatively restricted number of subjects, namely - English, French, Mathematics, Science, Geography, History, Latin, and to a less extent, Art. (Spens Report, 1939, p.80)

Thus, the group requirement illustrates the way in which examinations could override a school's values, producing an arbitrary system which determined the subjects which were taught most frequently and those which were given less attention.

... the requirements of the examination have put a heavy premium on certain subjects to the detriment of others, and have compelled schools, in the interests of pupils desiring to obtain the Certificate, to teach certain subjects to all pupils throughout their course, even when they might be deriving greater benefits from taking alternative subjects or from taking fewer subjects to a higher level. (Ibid, p.257)

This problem was compounded by the Board of Education's stipulation that,
"The form and not the pupil was to be the unit for examination" (Circular 549 quoted in Spens Report, 1939, p.255). The intention was clear; the School Certificate was to be a more truly democratic examination suitable for pupils in the fifth form of grammar schools rather than for an intellectual élite destined for universities. This was, however, another provision which became distorted in practice.

The supposed flexibility of the group requirement also proved to be illusory:

... the obligation to pass in both Groups II and III, has contributed to overpressure on candidates, and has caused the failure of a large number who have accomplished a generally meritorious performance in the examination, and who might properly be regarded as having concluded satisfactorily that stage in their education... it has often meant in practice the diversion of a child's energy from a subject in which he is beginning to find meaning and interest to one from which no lasting benefit can be expected...'the time expended is out of all proportion to the results achieved'.

(Spens Report, 1939, pp.262-63)

Only two decades after its inauguration, School Certificate was condemned for having produced effects the very opposite of those which had been intended:

In Circular 103... the Board of Education explained that it was a cardinal principle that this examination should follow the curriculum and not determine it. In the actual working of the examination this principle has been reversed, and there can be little doubt that in many Secondary Schools the Certificate examination is now the dominant factor. (Ibid, p.80)

It is clear that its prestige had reinforced this tendency:

... the fact that this examination soon came to be regarded as the terminus ad quem for pupils under the age of 16, had the effect of stereotyping and narrowing the curriculum. The examination unquestionably checked effectively any tendency to develop special courses in the main portions of Secondary Schools for pupils below the age of 16 on the lines indicated in the Board's Memorandum on Curricula. (Spens Report, 1939, p.80)

The dual pass system was another feature of School Certificate which
constituted an unforeseen weakness. In view of the pre-1917 proliferation of examinations, it is easy to understand why fitting an examination to fulfil a dual role had seemed such an attractive proposition. With hindsight, the Spens Committee (1939) concluded that "... the attempt to combine the two different objects in one examination has been disastrous" (p.258). The value of the "pass" level was undermined by the presence of the "credit" level and a matriculation certificate became the aim of most, if not all, pupils regardless of whether they intended to go to a university.

The practical convenience of this arrangement was obvious. The able pupils intended for the universities would take one examination instead of two, and the organisation and time-tables of schools would to that extent be simplified... however... a 'matriculation certificate' which should mean nothing more than a certificate entitling the holder to admission to a university, has come to mean a superior kind of school certificate with its own special value in the eyes of employers and the general public, and to be the 'aim of thousands of Secondary School pupils who neither intend nor desire to enter the doors of a university'. Indeed, many children... are oppressed by a mistaken sense of failure if they obtain a School Certificate but do not satisfy the regulations which make it a 'matriculation certificate'. (Spens Report, 1939, p.258)

This problem, of misusing examinations for purposes they were never intended to fulfil, has a long pedigree. As early as 1868, the Schools Inquiry Commission drew attention to the difficulties which were created when the Oxford and Cambridge "locals" were taken by pupils for whom they were ill-suited (see p.5). Similarly, the creation of CSE became a matter of some urgency in the 1950's because GCE and FE examinations were being used by the modern schools. In fact, the following statement reflects accurately the circumstances prevailing in many grammar schools in the 1920's and '30's and in many modern schools in the 1950's and '60's:

... (the examinations) tend to direct the attention of the master, not to his school, but to the cleverest of his scholars, and are therefore a temptation to neglect all those whom he believes it to be impossible to prepare for passing the examinations. (p.334)
It was, however, written in 1868 and described the situation in secondary schools of that period. University requirements have persistently dominated the work of the schools when only a tiny proportion of their pupils actually go on to higher education.

The misuse of the credit level aggravated some of the problems which derived from the group requirement. For instance, the constraining influence of the group requirement was strengthened by the matriculation requirements:

The choice of subjects already restricted by the "group" requirement is still further narrowed by the requirements for matriculation. Nor would this matter seriously, if only those pupils were concerned who proposed to proceed to a University. In fact, matriculation requirements dominate the situation. (Spens Report, 1939, p.259)

Likewise, the cramming in subjects where a candidate displayed no real talent, encouraged by the group requirement, was reinforced by the matriculation requirements which: "... led to waste of time by students repeating certificates in order to complete the necessary list of subject credits' " (Report of the School Certificate Examination (1931), quoted by Roach, 1979, p.53). Similarly, the arbitrary value placed on subjects by the group requirement was strengthened by matriculation requirements and subjects lacking academic prestige were automatically devalued:

... the conjunction of the Matriculation Certificate and the School Certificate has helped to upset the balance between what are conventionally known as academic and non-academic subjects. (Spens Report, 1939, p.259)

Another area in which the School Certificate never fully realised its original aim was in the attempt to encourage teachers to become more involved in examining. Although the regulations made provision for a substantial amount of teacher participation, this early attempt to involve teachers proved to be premature. As Roach (1979) observed:
... it seems that English secondary school teachers in general preferred that their pupils should be examined by external bodies, and little breach was made in this tradition until the creation of the Certificate of Secondary Education in the 1960's. (p.46)

Ultimately, School Certificate had failed to free examinations from some of the most "persistently perennial" problems but in spite of levelling harsh criticisms against the examination, the Spens Committee (1939) recommended only modest reforms. None of their proposals would have entailed radical alterations to the School Certificate and they clearly envisaged change taking place within the existing framework. For instance, it was suggested that the group requirement should be relaxed so that although a candidate must still gain five passes, English should be the only compulsory subject and candidates could make up the remaining number from either Group 2 or Group 3 subjects. In order to check, "... individuals eager to obtain 'Matriculation Certificates' as things valuable in themselves" (Report of the School Certificate Examination (1931) quoted by Roach, 1979, p.53), the Committee advocated dissociating matriculation from the School Certificate. Nevertheless, they felt that School Certificate ought to form a prerequisite to matriculation in order to avoid the re-emergence of early specialisation. They also argued that the maximum number of subjects offered by a candidate should be limited to seven and that, in order to counteract some of the importance attached to the examination, greater use should be made of school records. Clearly, the Committee did not favour radical revision and proposed to leave the existing framework largely intact.

For the existing grammar schools... they suggested modifying the 1917 system, but not radically changing it. (Roach, 1979, p.56)

However, any influence which this report might have had was interrupted by the outbreak of the Second World War (1939).
11. THE 1944 EDUCATION ACT

When examining once again became a subject of public concern, it had become enmeshed in the broader issue of rebuilding a nation depleted by war. Education was perceived as having an important role to play in this regenerative process:

Upon the education of the people of this country the fate of this country depends... The war has revealed afresh the resources and character of British people... In the youth of the nation we have our greatest national asset... we cannot afford not to develop this asset to the greatest advantage. (Educational Reconstruction, 1943, p.3)

This Board of Education pamphlet conveys the prevailing spirit towards the end of the War. Indeed, such was the confidence in what could be achieved through education that the most far-reaching education act ever devised was passed during this period. It represented the government's attempt to "... develop this asset to the greatest advantage."

The 1944 Education Act represented a landmark in the history of education which had profound implications for the examining system. It launched secondary education into a period of unprecedented growth which is illustrated by comparing the percentages of children receiving a secondary education prior to 1944 and after the Act. Only 9.5% of the secondary age group were receiving secondary schooling prior to 1944 while, "For 90 per cent of the children attending state schools an elementary education was all that was available" (Dent, 1944, p.15). The 1944 Education Act altered the whole complexion of secondary education by stipulating that every child who was eligible for secondary education should receive full-time schooling up to the age of fifteen years. Although the provision for secondary education had been expanding gradually over the past century, nothing could compare with the rapid acceleration of growth effected by the Act. Furthermore, it was stated
that this education should be suited to the "age, aptitude and abilities" of individual children and, to this end, a tripartite system of schooling was introduced. A minority of intellectually able pupils were to be educated in grammar schools while the technical schools were to cater for those who showed a technical bias. The majority, however, were to receive their education in secondary modern schools.

It would be difficult to over-emphasize the magnitude and complexity of the task embodied in this act; not only did it increase the size of the secondary school population to tenfold its previous size, it also aimed to provide each child with an education suited to his own particular needs. This would have been an ambitious enterprise for a nation enjoying stability and a thriving economy but Britain, newly emerged from the throes of war, did not possess these advantages. It is hardly surprising that with reconstruction necessary in so many areas of public life, Britain was hard-pushed to fulfil such an ambitious educational policy.

It is important to recognise that this new generation of school children was essentially an unknown quantity. Even as late as 1960, one commentator observed that:

We just do not know about the potentialities of the large majority of the pupils in the schools, and we will only begin to know about this as a result of experience - of actually attempting systematically to educate the average child.

(Forum, Vol.2(3), p.87)

Prior to 1944, secondary schools had catered for a small, select and relatively homogeneous group of pupils but after 1944 they had to provide for an unselected and heterogeneous generation. The provision of free secondary schooling for all was as much "a leap in the dark" as was the 1832 Reform Act and contemporaries greeted it as an educational revolution. For instance, Dent (1945) described the Act as "profoundly important" and
"revolutionary in conception" (p.79) arguing that: "... it makes possible as important and substantial an advance in public education as this country has ever known" (p.3). Likewise, The New Secondary Education (1947) argued that it marked, "... a revolutionary change in education" (p.8).

Changes of such magnitude in the educational system were bound to have repercussions on the examinations' system. With hindsight, the Beloe Committee (1960) concluded that: "The 1944 Act included provisions which were bound sooner or later to throw up major problems in the field of examinations." (p.6) However, even before the changes wrought by the 1944 Act, the examinations' system came under renewed and heavy fire.

12. MOUNTING HOSTILITY TOWARDS EXTERNAL EXAMINATIONS (1929-1943)

As Roach (1979) observed: "By the early 1930's the tide of criticism of external examinations was beginning to rise again" (p.54). However, there was a new dimension to the criticisms voiced in this decade. Earlier critics had usually confined their comments to the various backwash effects which external examinations were perceived to have but new evidence, uncovered in the 1930's, cast an even more disparaging light on conventional examinations. In this decade, the science of assessment finally became established as a serious field of study and pioneering figures such as Hartog, Rhodes and Burt (1936) conducted controlled experiments into methods of examining which produced disconcerting results. They uncovered "distressing uncertainties" in public examining and pointed out that, "... an examination certificate... may by chance be given to candidates who by rights should be rejected. There are many examinations in which the final results in a particular subject may depend solely on the marks allotted by two examiners or even by a single
examiner. Our results show how serious may be the element of chance in such cases" (p.xiii). Thus, increasing understanding of assessment as a science was exposing new deficiencies in conventional examinations which gave the new generation of critics a powerful, new weapon with which to discredit examinations. For instance, Valentine (1932) argued that there was:

... serious ground for lack of confidence in examination awards, namely personal variability and unreliability of marking in the examiners themselves... In the marking of essays especially, extraordinary variations occur between the marks of different examiners; and this is most important in view of the prominent place that the essay occupies in many examinations. (p.26)

Moreover, the backwash effects of examinations seemed all the more unacceptable once it was appreciated that the examinations were often sensationally unreliable: "... so far as the influence of examinations may be educationally harmful, it is all the more important to enquire how far they are reliable. A revelation of great unreliability should tend to lessen the weight attached to examinations" (Valentine, 1932, p.17).

Thus, the climate of opinion shifted significantly at this period. Ever since their beginnings in the 1850's, there had been immense optimism about external examinations and a continuing belief that they were a potential source of good once reform had eliminated their backwash effects. However, educationists increasingly felt that backwash effects were not the idiosyncratic products of a particular system but that they were inherent in the very process of examining. This changed mood is illustrated by the recommendation in the Spens Report (1939) that the examinations for the new Technical High Schools should be conducted on an internal basis because:

We would not seek to impose upon the Technical High Schools any system of external examination with the necessarily uniform and rigid syllabuses which would inevitably follow in its wake. (p.279)
Thus, official reports were finally beginning to adopt the view that external examinations were synonymous with rigidity and uniformity in school curricula and that this was an inherent feature of examining. This view was shared by a growing number of educationists. For instance, Norwood (1929) argued:

... the chief blot on that admirable document, the Report on the Education of the Adolescent, is that it lays down that it is desirable to make available a special examination of a type suitable for pupils leaving post-primary schools... The usual safeguarding clause is solemnly inserted 'provided of course that the examination syllabus is not allowed unduly to dominate the curriculum'. As if in an examination of this magnitude it could ever do anything else, as if there were not proof overflowing from the short career of the School Certificate that it must do so. (p.206)

Norwood's concluding sentences convey the mood of pessimism which increasingly enveloped external examinations:

... the examination system is no friend to whom one can trust in blind confidence: from its first guise as a servant it may at any moment reveal itself as a hard and unsympathetic master: and it is far the most dangerous of the mechanical instruments which are employed in the service of education. (pp. 208-9)

This reaction against external examining culminated in the Norwood Report (1943).

13. THE NORWOOD REPORT (1943)

If the Spens Report (1939) was characterized by its moderation, then its heir, the Norwood Report, was characterized by its radicalism. The Spens Report had suggested modest reforms within the existing framework of examinations whereas those advocated by the Norwood Report would have entailed sweeping away the existing framework. The reaction against conventional examinations reached a climax in the Norwood Report which recommended the abolition of external examinations at
16+ and their replacement by a network of internal, school-based assessments. The Norwood Report represented a turning point in the attitude towards examinations which had prevailed in official reports over the preceding century. Whereas their mood had been one of optimism, the Norwood Report flatly denounced the traditional belief in the ultimate good of examinations. It argued that external examinations necessarily enslaved schools:

... we find it difficult to accept the dictum that external examinations can follow curriculum; the time lag is too long and the ruts grow too deep... external examinations must necessarily determine curriculum. (pp. 32-33)

It is not difficult to understand this volte-face. The failure of the 1917 reforms to bring about the desired changes made it increasingly difficult for educationists to conceive of an examination which could, in practice, be the servant of the schools rather than their master.

"At present, the examination dictates the curriculum and cannot do otherwise; it confines experiment, limits free choice of subjects, hampers treatment of subjects, encourages wrong values in the classroom" (Norwood Report, 1943, p.31).

Besides wishing to free the schools from the negative effects of external examining, the Committee felt that there were positive reasons why examining should be school-based:

... the examination is best conducted by the teachers themselves as being those who should know their pupils' work and ought therefore to be those best able to form a judgement on it... a method of examining by the teacher, combined with school records... would furnish a certificate giving information of real importance... and yet would preserve intact the freedom of the school and would rid teacher and pupil of an artificial restraint imposed from without. (pp. 32-45)

A factor in the existing situation which the Committee felt justified major revision was the Education Bill which would effect major changes in secondary education (see section 11). One of the most
"persistently perennial" problems in the history of examining has been
a mismatch between the dynamic nature of secondary education and the
characteristically fixed nature of external examinations. Examinations
have consistently lagged behind education which has evolved rapidly
and the Committee realised that, at this time in particular, there was
a need for far-reaching change. They regarded external examinations
as one phase in the development of secondary education, playing a
crucial role at an earlier stage of growth but anachronistic by the
1940's:

In the last twenty years the examination has rendered useful
and valuable service... At a time when the rapid expansion
of secondary schools caused uncertainty about standards...
when newly recruited teachers bringing with them little
tradition and little experience were in doubt about...
syllabuses and curricula, the programme put before them in
carefully devised regulations... gave a sense of direction,
defined levels of achievement and helped in no small measure
to establish secondary education on a sure and sound basis...
In the early days a scaffolding was needed for the building
of secondary education; and there was some justification for
examinations dictating curriculum... But secondary education
has passed through its early phases; it can stand alone, and
already finds the framework which supported it to be a
barrier preventing growth. (pp. 32-33)

The Committee were conscious of their role in creating a system which
must operate in altered circumstances and they realised that if the
existing system was allowed to continue, the results might be disastrous:

If the present School Certificate examination is retained...
it will mark off the Secondary Grammar School from other forms
of secondary education. A system will then be established
under which parity in secondary education will become
impossible... the establishment of the Technical School and
the Modern School... will be prejudiced from the outset.
(p.46)

The Committee recognised that during this period of expansion and
experimentation, examinations would need to be flexible:

... it will be increasingly difficult to reconcile an external
examination at 16+ with the full realisation of the aims of
the schools and with the enjoyment of that freedom which will then be held to be a vital necessity. (p.45)

The proposed abolition of external examinations at 16+ was nothing short of revolutionary because, although far-reaching changes had been implemented before, the principle of external examining had always been preserved intact. The Committee were aware of the drastic nature of the changes they were advocating and they conceded that such changes would have to occur gradually:

We recognise that sudden and immediate change is inadvisable ... it may be the part of wisdom to take one step at a time ...
We should expect that sufficient progress would have been made in seven years from now to enable a wholly internal examination to be adopted. (pp. 45-46)

Therefore, they, "... set forth an ultimate objective" (p.45) while conceding that, in the transitional period, interim measures would have to be employed. Their interim plan entailed using the existing external system as a basis from which to propel the schools towards the new internal system. To this end, during the next seven years while teachers were gaining the skills necessary to conduct internal examinations, the School Certificate was to be retained under the control of the university authorities. However, each board was to be requested to appoint a Standing Committee for the conduct of the examination on which the largest single group of representatives was to be drawn from the teaching profession. Teachers were to account for 40% of the membership and this was just one step in the direction of providing teachers with the expertise necessary to take full responsibility for examinations in the future. The Committee also wished to abandon the group and minimum number of subjects requirements so that candidates may take any number and any combination of subjects that they wished. It is interesting to note that even in their interim measures, the Committee regarded
flexibility and freedom as qualities of great importance. At the same time, schools were to be encouraged "... to offer their own syllabuses" (p.46).

Since school records were intended to play a major role in the new assessments, the Committee recommended that research into record keeping methods should be undertaken as a means of facilitating this development. A new type of certificate, divided into two parts, was also suggested. Significantly, the first part, not the second, was to "... contain a record of the share which the pupil had taken in the general life of the school", indicating "... the way in which he had used the opportunities offered to him by his education, using the term in its widest sense" (p.48). The fact that examination results were to appear in the second section, while a record of the pupil's education "... in its widest sense" was to precede it, neatly symbolised the Committee's sense of priorities.

The above recommendations were designed to facilitate the move away from external examining towards the "... ultimate objective" of a national system of internal, school-based assessments.


The transition envisaged by the Norwood Report was never realised. No immediate action was taken following the publication of the Norwood Report (1943) and the Minister finally acted in 1946 by re-constituting the Secondary School Examinations Council. He requested that they should review the situation and their report, Examinations in Secondary Schools, was published in 1947. This is not to suggest that the period between 1943 and 1946 was one of inactivity since conflicting opinions built
up over these years. Initially, it appears that the Norwood Report was well-received in the press and the favourable response led Board officials to hope for an early start to the reforms.

In fact the difficulties proved to be considerable, and opposition built up in several quarters... in 1944 the Headmasters' Association voted against the proposal that the School Certificate should be conducted internally. The Journal of Education, which had supported the Norwood view, suggested that the headmasters' attitude stemmed from a desire to maintain a clear demarcation between the grammar schools and other secondary schools... In August 1944 the Cambridge Local Examinations Syndicate entered the ring with a statement recommending that the examining bodies should continue to assess the candidate's whole performance in School Certificate as well as recording passes in individual subjects... In the same year the Syndicate's Deputy Secretary, J.L. Brereton, published an investigation of the examining system which was very critical of the Norwood proposals. When all these cross-currents of opinion are remembered, it is not surprising that it was 1947 before the Secondary School Examinations Council presented a new set of proposals to the Minister of Education. (Roach, 1979, p.57)

The Norwood Committee had sought to sweep away the external examinations' "scaffolding" and to allow the edifice of secondary education to stand alone but this proved to be abortive. The SSEC's revised proposals abandoned Norwood's "ultimate objective" (i.e. a national system of internal assessments under teacher control), while its interim measure (an external examination on a subject basis controlled by the universities) was made a permanent feature of the new system. Flatly contradicting the Norwood Report, the 1947 Report claimed that, "An external examination is profitable and proper" (p.5) for certain secondary school pupils and it recommended the establishment of a new examination, the General Certificate of Education, which was to be offered at three levels: Ordinary, Advanced and Scholarship.

Although the SSEC abandoned Norwood's "ultimate objective", it would be misleading to imply that Norwood's proposals were rejected out of hand since many of them were incorporated into the new set of
recommendations. However, the manner in which they were embodied in the new system may be described as a total perversion of the Norwood Report's ultimate intentions. For instance, having recommended the establishment of a new external examination with all the power and prestige which university backing could provide, the SSEC's following proposals, which were entirely in keeping with the Norwood ideal, may sound somewhat superfluous:

We recommend that -
(a) Individual secondary schools should carry out systematic internal examinations based on and designed to suit the particular courses and the pupils following them.
(b) The Ministry and Local Education Authorities (singly or in groups) should promote and encourage experiments in the conduct and assessment of internal examinations. (p.5)

In addition, they advised that, "Objective tests... should be set periodically... and the results recorded in school records" (p.4). These records were also intended to play a major role in this school-based system of assessments: "Every pupil on leaving a secondary school should be provided with a comprehensive school report containing the fullest possible information about him" (p.4). These internal assessment procedures were entirely in keeping with the spirit of Norwood and yet their purpose may be questioned in the face of a new and potentially powerful external examination.

Certain features of the new examination explain these apparently conflicting proposals. For instance, the pass level was fixed at the credit level of the old School Certificate immediately marking off the GCE as an examination designed for an intellectual élite. It must be remembered that the credit level originally had been intended as a means of satisfying matriculation requirements for university entrants. In 1951, this stringent level was officially recognised as the acceptable
standard for a pass in a general school certificate. Another regulation which was designed to place O-Level beyond the reach of all but a small proportion of the total school population was the stipulation that, "... the minimum age for sitting any part of the examination should for the time being be sixteen years on Sept. 1st in the year of the examination" and that an "... increase in the minimum age at which the examination can be taken, together with any raising of standards which may be appropriate" (p.7) should also take place as soon as circumstances permitted. Thus, the rationale behind these seemingly conflicting proposals becomes increasingly clear. GCE was intended to cater for approximately the same limited group of pupils as had been catered for by the School Certificate. The Norwood methods had, in effect, been confined to use by those pupils who would receive their education in modern schools. As the Beloe Report (1960) later commented:

The Council hoped that... they had devised an examination which would in practice be beyond the reach of any but those in selective courses. They envisaged that the needs of children in other courses would be met by arrangements for systematic internal examinations... accompanied by the extensive use of school reports. (p.7)

The SSEC also attempted to play down the importance of GCE in the grammar school scheme of education:

... for a pupil who is clearly fitted for "Advanced" work... sitting for some of the "Ordinary" papers would be profitless. It would be contrary to the intentions of the system for any pupil to sit successively for "Ordinary", "Advanced" and "Scholarship" papers in the same subject. We hope, therefore, that schools will rely upon their internal examinations and other similar evidence for their internal organisation; that they will defer as long as possible the entry of their pupils for these external examinations and that they will suitably limit the subjects offered. (p.7)
15. **CONCLUDING REMARKS**

An examinations' system which was to last for less than twenty years was established in 1951. It was essentially a compromise between the vested interests of the "traditionalists" and the aims of the "progressives" which resulted in a curious blend of conventional external examinations and novel internal assessments. External examining was intended to play no part in the secondary modern scheme of education and only a limited role in secondary grammar school education. Wherever possible, assessment was to be school-based.

Since the creation of CSE stems directly from the working of the 1947 provisions, the following chapter will trace their implementation and evaluate their efficacy. It is important to bear in mind that less than a decade after the introduction of these provisions, the Beloe Committee was asked to review the question of examining at the 16+ level. It is important to determine why the 1947 provisions were found to be unsatisfactory so shortly after their inception. Their deficiencies provide the raison d'etre for CSE and it is against this background that the "achievements and lessons" (Dixon, 1977, p.8) of CSE must be judged.
1. THE PROVISIONS MADE BY THE SSEC (1947)

While the Spens Report (1939) had advocated modest reforms within the existing framework of examinations, the Norwood Report (1943) had recommended far-reaching changes which would have swept away the existing structure. The recommendations which were finally implemented may, with some justification, be described as preserving the School Certificate under a new guise. Admittedly, some modifications, such as the removal of the group and minimum number of subjects requirements, had been made. Nevertheless, the characteristic features of the School Certificate were preserved. For example, the examining authorities which had operated the School Certificate were responsible for the GCE, a decision which, perhaps more than any other, ensured the preservation of the status quo. Furthermore, the O-Level pass was fixed by the direct adoption of that part of the School Certificate grading system which had been so highly coveted. Perhaps the most telling indication of the affinity between School Certificate and GCE appeared in the 1947 Report:

One effect, therefore, of our proposals is to make available .... what is virtually "School Certificate" on a subject basis. (p.7)

Clearly, the SSEC had no pretensions to providing anything other than the old system on revised lines for the grammar schools. Alongside this, they expected a system of internal assessments to coexist and to be used in preference to it by the modern schools and even by grammar schools in
appropriate circumstances:

It would be contrary to the intentions of the system for any pupil to sit successively for "Ordinary", "Advanced" and "Scholarship" papers in the same subject. We hope, therefore, that schools will rely upon their internal examinations .... for their internal organisation. (p.7)

These proposals were unrealistic since experience had shown that a certificate awarded by a prestigious academic body was highly coveted. It had proved to be impossible to confine the credit level of the School Certificate to its originally intended function as a matriculation qualification and it is difficult to understand why the SSEC imagined that GCE would not be used extensively as a school-leaving qualification when the attempt to confine the use of its forbear to university entrants had proved to be so impracticable.

The real measure of this lack of realism is provided by the extent to which the SSEC's intentions were undermined once the proposals were implemented. As the following section shows, internal schemes failed to materialize on a noticeable scale while external examinations proliferated.

2. EVIDENCE OF THE FAILURE OF THE 1947 PROVISIONS

A number of provisions had been made to secure the freedom of modern schools from external examinations. The rigorous pass level and the minimum age of entry for GCE were intended to exclude them from 0-Level and, since no other recognised examinations of a less exacting standard were available, it was assumed that they would devise their own internal procedures for certification. This was ratified by the Schools' Grant Regulations which provided financial support exclusively for the purpose of GCE entries; no other form of examination below the 16+ level was recognised, approved or financed by the State (see the Beloe Report, 1960, pp.15-16). In spite of this, the Beloe Report (1960) found that within the short space of thirteen years, ministerial policy had been flouted on a massive scale.
O-Level was adopted almost immediately as the leaving certificate not only in grammar schools but also in the modern schools where it was used on an ever-increasing scale and rapidly became the hallmark of a "good" modern school. Moreover, the acquisition of as many passes as possible became the object not only of grammar school pupils but also of the small but growing number of children in modern schools who entered for GCE. In fact, by 1959, a third of all O-Level entries came from establishments other than grammar schools (Beloe Report, 1960, p.9) and, by the end of the 1950's, this trend was accelerating at a great pace. Whereas in 1958 10,500 modern school candidates had entered for O-Level, the following year this number had increased by 50% with 15,600 candidates from 1000 modern schools taking the examination (Beloe Report, 1960, p.10).

Moreover, an NUT survey (1955) which attempted to gauge the popularity of different fourth and fifth year examination courses found that even in modern schools, GCE topped the list (quoted by Morris, 1960, p.1035).

This trend was both illustrated and encouraged by the establishment, in 1953, of a ninth GCE Board, the Associated Examining Board, whose examinations were ".... designed to have a different approach more suited to the needs of technical students" (Beloe Report, 1960, p.9).

Examining authorities offering school-leaving certificates below the GCE O-Level had also proliferated. In 1953, the College of Preceptors made available their School Certificate and they were followed in 1954 by the Royal Society of Arts. In 1956, the Union of Lancashire and Cheshire Institutes entered the field of school examining and they were joined in 1957 by the Union of Educational Institutes. The Northern Counties Examinations' Council School Certificate became available in 1959 and the East Midland Educational Union started school examinations in 1960. A Certificate of Commercial Education for use in schools was provided by the London Chamber of Commerce and the City and Guilds of London Institute expressed their intention of entering the field of school examinations as
soon as possible. This list includes the larger regional or national examining authorities but it is by no means exhaustive for there were also specialist examinations in use in schools such as those conducted by the Pitman Examinations' Institute, the National Council for Domestic Subjects, the General Nursing Council, the Armed Forces and so on. In addition, small, local examinations sprang up all over the country.

Typical of these were the Nottinghamshire local examinations (see Burdett, 1960). With the aid of the LEA, North and South Nottinghamshire formed themselves into two separate examining areas with local F.E. colleges acting as examining boards.

Since this development had been so ad hoc and unco-ordinated, the examinations varied enormously. Some were offered on a "group" basis; others were single subject examinations. Some were designed to be taken at the end of a four-year course and others were for those who had completed five years of schooling. Different examinations catered for different bands of ability: some were vocational; some were specialised and others were certificates of general education. Together they formed such a motley group that the situation was becoming as bewildering as it had been at the turn of the century.

Perhaps the most striking feature of this development was the rate of growth and, more particularly, the way in which this was accelerating towards the end of the 1950's:

"... the most striking and significant fact that has emerged from our survey is the growth, and more especially the rate of growth, in the numbers taking the examinations of the regional and national Bodies in recent years ... of 150 modern schools in England replying to the inquiry, while 76 or just over 50% were entering pupils for external examinations by 1958, a further 49 had plans for doing so shortly, bringing the proportion of schools entering pupils for examinations to over 80% of the sample ... figures (for four of the large non-GCE examining authorities) show an overall increase between 1958 and 1959 of 70% in the number of candidates and of over 45% in the number of centres, with a further increase of over 40% in the number of centres between 1959 and 1960. (Beloe Report, 1960, p.15)"

The sheer speed with which this trend had established itself is remarkable
especially when it is remembered that it did not represent isolated pockets of revolt but growth on a national scale. Moreover, it had occurred in spite of ministerial policies devised to deter it. The Minister's initial response to this was to declare his disapproval in government circulars but ministerial disapproval proved to be powerless in curbing this development.

These events are an almost exact replica of what had happened a century earlier. In the very same decade of the previous century, examining had undergone the same sort of prolific and unco-ordinated growth illustrating the "persistently perennial" nature of this problem. In fact, during the years in question (i.e. 1850-1965), more than half of that period was characterized by a glut of external examinations. The re-emergence of this problem in the 1950's illustrates its tenacity. As one critic observed:

"... the phrase, "the nightmare of a multiplicity of external examinations" is quite as relevant today as when it referred to the position in grammar schools before 1912. (Davis, 1962, p.61)"

The unchecked growth of examinations in the period prior to 1917 was accounted for by the government's laissez-faire policy. However, twentieth century education had been characterized by the increasing involvement of the State which could hardly be accused of inactivity in the 1940's having just passed a major education act and devised a new system of examining. Since the State cannot be accused of failing to act at this stage, why did examinations experience another period of frantic expansion?

3. REASONS FOR THE FAILURE OF THE 1947 PROVISIONS

Although it is possible to argue that the recommendation of the Norwood Report (1943) for a national system of Mode 3 examinations was premature, at least it recognised the importance of "parity of esteem" (p.14) in secondary education. In a tripartite system, where three different types of secondary school were expected to co-exist as equal but different types
of institutions, the Norwood Committee realised that an examining system which had inequality built into it would undermine this goal. They regarded the examinations' system as a critical factor in securing equivalent status for different types of schools:

If the present School Certificate examination is retained without alteration ... it will mark off the secondary Grammar School from other forms of secondary education. A system will then be established under which parity in secondary education will become impossible ... the establishment of the Technical School and the Modern School ... will be prejudiced from the outset. (Norwood Report, 1943, p.46)

The Norwood Report had attempted to meet the needs of the secondary school population in its entirety whereas, with hindsight, it is easy to see that the SSEC's provisions were ill-equipped to satisfy the needs of all secondary schools. In effect, the SSEC had retained a prestigious examination for an intellectual élite whilst an untried and untrusted system of internal assessments was expected to co-exist with it for the newer secondary schools. Although this represented an attempt at compromise between conflicting interests, it neglected the principle of parity of esteem for while modern schools were denied access to external examinations, grammar schools had access to both internal and external assessments. Cynically, one could argue that the SSEC had safeguarded examinations for the schools which had always enjoyed this privilege while the untried and untrusted Norwood methods were confined to experiments in modern schools with the less able "guinea-pigs".

Parity of esteem with the grammar schools was crucial to the modern school's success as a different type of educational institution - different in character, aims and approaches. If this failed to happen, it would be difficult to avoid an equation between the points where modern schools differed from grammar schools and the points where they were inferior. Their differences were liable to be viewed as the mark of their inferiority. An academic, examination-orientated education had been the traditional route to personal and career success and, as such, it enjoyed
the absolute confidence of the public. It was perhaps inevitable that any extension in the sphere of secondary education should be expected to conform to this pattern and that any which did not should be spurned.

However, the Government had asserted that:

No child must be forced into an academic education ... merely because that type of grammar school education is considered more socially desirable by parents ... This ... is the answer to those who think that secondary education for all must mean a grammar school education for all. (The New Secondary Education, 1947, pp.4-5)

Thus, the modern school was launched on a course which was at variance with public opinion, with almost inevitable consequences:

... they are the schools which have been subjected to more calumny and vilification than any other type of school ever evolved in these islands. The use of the term Secondary to denote "second stage" in 1944 misled many of the general public into the illusion that somehow "Secondary Education for All" meant "Grammar School Education for All" ... This was not so, and because of the disillusionment which followed the secondary modern schools bore the brunt of the attack which came from all quarters. (Heaton, 1958, pp.93-94)

The SSEC's attempt to overhaul the system has been described as "clumsy" (Judges, 1969, p.31), a justifiable criticism because a system in which new schools, needing to build up their public esteem, were precluded from using prestigious examinations prejudiced their chances of success from the outset. It has to be remembered that whenever secondary education had passed through a period of expansion, external examinations had played an important role in this process. Successive reports had paid tribute to the part played by external examinations in endowing schools with a sense of purpose and providing public credibility by measuring their work against external standards. It seems ironic that at a time when secondary education was passing through a period of unprecedented growth, it was denied this traditional source of support.

With hindsight, the Beloe Report condemned the SSEC's provisions as having not, "... satisfactorily met the realities of the situation" (1960, p.7). Indeed, lack of realism is the key criticism which can be
levelled against the SSEC. Their system embodied a number of unrealistic assumptions: that prestigious external examinations could be confined to a strictly limited role in the overall scheme of education; that internal assessments would be regarded as having equivalent status to external ones and therefore used in preference wherever possible; that modern schools could be excluded from an examination which conferred status on the schools which used it; that the number of subject entries could be limited and that GCE should not be used as a school-leaving qualification.

In the event, modern school candidates entered on an ever-increasing scale not only for GCE but also for the crop of non-GCE examinations which sprang up to meet their needs. As one shrewd observer remarked:

In this the schools have displayed a good deal more social realism than was characteristic of much post-war educational thought. (Taylor, 1960, p.63)

This claim is endorsed by "... the work by educational sociologists, notably Dr. Olive Banks, (which) has suggested that the prestige of a school is largely based on the social level of the career openings it can offer its pupils, and that this depends upon the sorts of examinations for which the school can prepare them " (Morris, 1960, p.1035).

It is already becoming clear that the history of examinations during this period was so inextricably intermeshed with the evolution of modern schools that the fate of one was bound up with growth and development in the other. Therefore, this chapter will pay detailed attention to the evolution of modern schools comparing their theory with their practice and tracing the way in which this triggered off a chain reaction of events which led to a renewed explosion of examinations and eventually to the creation of CSE.

4. THE MODERN SCHOOL : THE ORIGINAL CONCEPT

Although the working of the 1947 provisions makes it easy to level serious criticism against the SSEC, this is not entirely fair to a
generation which was embarking on a course the consequences of which were impossible to predict. Although the criticisms are strictly tenable, they ought to be tempered by insight into the mood in which secondary education for all was conceived. Contemporaries were conscious of taking a massive step forward and yet unsure of precisely which direction that step should take. It was this which made them feel that the modern schools should have freedom to develop their own special character and traditions:

... it will take a generation to implement fully this great scheme. During that time alterations will be needed, improvements will suggest themselves. New problems will arise. Everything to do with children must have room to grow ... The schools must have freedom to experiment, room to grow, variety for the sake of freshness. (The New Secondary Education, 1947, p.5)

The belief that modern schools needed growing space, room to develop their own identity, explains why they were given a free rein in the development of their curricula. Indeed, one of the most striking features of contemporary writings on this topic is the repetition of that word - freedom. For instance, the Ministry of Education's guidelines asserted that modern schools should provide an atmosphere in which pupils may, "develop freely along their own lines" and that teachers should be "... free to act" arguing that, "Freedom and flexibility are of its essence and are indeed its greatest opportunity " (The New Secondary Education, 1947, pp.29-30). Likewise, a Headmaster summed up the qualities of the modern school curriculum as "... real life, vitality and freedom" (Heaton, 1958, p.106). This explains why this new institution was jealously guarded against the encroachment of external examinations which were regarded as the antithesis of freedom:

... it is impracticable to combined a system of external examinations, which presupposes a measure of uniformity, with the fundamental conception of modern school education, which insists on variety. (The New Secondary Education, 1947, p.46)

Experience had shown that external examinations constrained the schools which used them and this was felt to be unacceptable:

Any tendency towards uniformity would be deplorable. We look forward to bold experiments and a wide variety of courses in the newer secondary schools. (Ibid, pp.4-5)
Thus, although hindsight makes us critical of the decision not to provide an examination for use in modern schools, it is important to appreciate the climate of opinion in which this decision was borne.

Moreover, the work of psychologists had demonstrated the importance of interest and a sense of involvement in the learning process. This awareness that learning is most effective when the child is fully absorbed in what he is doing gradually filtered through into educational thought influencing the modern school ethos:

The keynote of the new system will be that the child is the centre of education and that, so far as is humanly possible, all children should receive the type of education for which they are best adapted. (Educational Reconstruction, 1943, p.9)

Thus, it was argued, modern school curricula should be child-centred rather than subject-centred:

... experiments of many kinds are to be welcomed. The only proviso that must always be observed is that the real interests of the children must come first. (The New Secondary Education, 1947, p.24)

Thus, the new modern schools started life in a spirit of social idealism which is conveyed by the opening to The New Secondary Education (1947):

This pamphlet tells the story of a great adventure ... This work is the boldest investment in manpower, money and materials that any war-crippled nation has ever made - an investment in the children, the citizens of the future. (p.3)

In the event, the theory of secondary modern education did not match up to its practice. There were pressures inherent in their circumstances which precipitated these schools towards external examinations.

5. **THE UNREALISED IDEAL OF PARITY OF ESTEEM**

One of the key factors which precipitated modern schools towards external examinations was the unrealised ideal of "... parity of esteem" (Norwood Report, 1943, p.14) between different types of secondary schools. With hindsight, it appears that the modern schools were doomed on this point from the outset. In particular, their confinement to a novel form
of internal assessments placed them at an immediate disadvantage.

It has to be remembered that even before the Norwood proposals were "implemented", considerable hostility had built up towards them (see Chap. 1, pp. 43-44). An indication of this is provided by one commentator's reference to the Norwood Report as, "... that notorious document" (Forum, 1960, p.85) while one member of the Norwood Committee later observed that the report had initiated a "... prolonged controversy" (Journal of the A.A.M., 1961, p.58). Clearly, in inheriting the Norwood method of assessment, the modern schools also inherited the stigma attached to it. Widespread mistrust was perhaps inevitable: the Norwood approach represented a break with a tradition which was familiar and respected. The Norwood Committee had recognised this and had stressed that their system would have to be adopted wholesale in order to succeed.

Another problem deriving from their "ancestry" (Dent, 1958, p.142) was that many modern schools were borne out of Public Elementary Schools, were staffed by former elementary school teachers and used their buildings and equipment. Consequently they inherited:

... the stigma of social and educational inferiority which had always attached to the Public Elementary School ... During the first few years after 1945 this stigma touched off a widespread hostility ... whose intensity and persistence surprised even those students of the educational scene who had anticipated that something of the sort might happen. (Dent, 1958, pp.154-155)

The inadequate buildings and facilities which they inherited provided one of the most serious deficiencies with which modern schools had to cope. It has to be remembered that:

Until 1944 a "secondary school" meant a particular sort of school to which only a small proportion of the population could aspire, one which had better qualified and better paid staff, smaller classes, and more attractive premises and amenities than most of the other schools in its neighbourhood.

(The New Secondary Education, 1947, p.7)

The Minister of Education recognised that the newer schools must "... have the advantages and the amenities hitherto exclusively associated with the limited number of schools called secondary schools up to 1944" and pledged
that, "They will have equally good buildings ... The maximum size of classes will be the same for all" (The New Secondary Education, 1947, p.7).

However, this was an ambitious educational policy for a country devastated by war and it is hardly surprising that, in the event, it was only partially realised.

There is considerable evidence, in the books, journals and newspapers of the period, that inequality between grammar and modern schools was rife. For instance, the following account describes the appalling lack of facilities found in one converted elementary school:

Unbelievable as it may seem there are for them (the teachers) no cloakroom facilities, no staffroom, no lavatory. Thus, when a master has a free period, he must spend it in a classroom where someone else is teaching, or, if he wants a smoke, betake himself across the school yard to sit in the senior master's car. The Headmaster, denied the luxury of a room of his own, has a desk in the corner of a classroom. It is, accordingly, no unusual occurrence for some thirty boys, a teacher, the Headmaster, a part-time clerk and an anxious parent to be in the room at the same time. To add to the excitement, a telephone intermittently makes itself heard. (Dent, 1958, p.23)

The frequency with which such descriptions feature in the writings of the period suggest that these conditions were far from atypical of modern schools during those years:

Some four years ago I entered a secondary modern school. Built in 1893 ... it was intended to hold 250 boys and even then its playground space must have been appallingly inadequate. Now, with double that number of boys, the playground would be a downright hazard were it not for the fact that the school had spread itself about the town in halls and private houses (18 boys in a back bedroom) until its total accommodation lay within a circle of nine miles circumference. For many classes it was a case of morning assembly followed by morning departure. (TES, Aug. 26th 1960, p.239)

Thus, the teachers in secondary modern education frequently had to manage as best they could in schools ill-adapted to their new role. The quality of the learning, not to mention the morale of the staff, must have been seriously undermined by inadequate facilities. Moreover, these problems were not confined to the older schools but were even apparent in those which were purpose-built. For instance, in an article entitled, "Old Habits in a
New Town: First Impressions of Secondary Modern Life", a probationary teacher described the situation he found in a newly-built modern school:

Our school is designed for 600 pupils; we have approximately 815. This necessarily leads to overcrowded rooms and the use of the Crush Hall for "impossible" lessons ... I am constantly taking geography lessons in the domestic science room and in the art room, which virtually excludes adequate preparation before the children arrive ... eight classes of 30 pupils must share one set of 33 books - 240 pupils each using the book a minimum of once a week ... the frustration of trying to teach well, and failing, shows itself and produces a surprising ... turn-over of staff. The school opened two and a half years ago. Of the original staff of approximately 25, only five remain. (TES, Jan. 29th 1960, p.160)

Another teacher reported that in his school, "... it was sometimes impossible to seat the children if no one was absent, although the school was only recently built. Classes which would have been difficult enough for experienced teachers to tackle in ideal circumstances thus became intolerable at times because of the lack of adequate equipment" (TES, Feb. 5th 1960, p.220).

It would be difficult to over-estimate the importance of these inadequate resources. Indeed, Dent (1958) claimed that:

... the most obvious obstacle to progress was the fact that the Secondary Modern School "inherited a mass of out-of-date elementary schools completely unsuited to modern educational needs" ... Too great emphasis can hardly be laid upon the handicap that obsolete and ill-equipped buildings has been and still is. (pp.142-43)

This legacy was a major factor in motivating the schools to achieve the status and respectability conferred by success in external examinations. As one critic observed:

... the very best (modern schools) ... can genuinely claim parity of esteem. But where that has happened or shows any prospect of happening, it is almost invariably because a school has prepared its ablest pupils for the G.C.E. It has become plain to more and more head teachers that the way to public esteem is through success in nationally recognised examinations. (TES, Leader, 12th Sept. 1958, p.407)
6. "THE DAYS OF WANDERING"

Freedom and flexibility are of its essence and are indeed its greatest opportunity. *(The New Secondary Education, 1947, p.30)*

This was the rallying cry with which the modern school was launched. It was assumed that securing the freedom of these schools from external constraints would axiomatically release an internal dynamic, an inner sense of purpose which would propel the schools onto proper courses facilitating the development of a distinct character and traditions. In the event, this freedom turned out to be a stumbling block for many schools. For their teachers, there were no fixed points of reference, no established goals, no definite sense of direction and no established criteria by which to judge the efficacy of the experiments they were encouraged to undertake. The danger was that in an atmosphere of uncertainty, freedom would degenerate into aimlessness:

"Their future is their own to make." Yes, indeed ... but what sort of future? For many teachers there were no landmarks in this new territory. There were no precise aims or objectives laid down for Secondary Modern education, no recognised standards of attainment, no established "yardsticks" by which to measure the value of anything they might attempt. No wonder they felt lost. *(Dent, 1958, pp.146-47)*

Thus, a quest for a sense of purpose is frequently found in the literature of the period. The following comment is typical of this yearning for a positive sense of identity among modern school teachers:

In a district like our own where the grammar, technical and independent schools cater adequately for all the pupils who could obtain even a modest number of GCE "O" level passes our secondary modern schools still follow a timetable which differs in no way from that of the lower forms of a grammar school with the exception of slightly more time for practical subjects. We continually deplore our pupils lack of interest and inattention and yet in the staffroom it is difficult to find anyone who really believes that radical changes in our methods are possible ... why cannot the secondary modern school which is not bound by an examination syllabus be progressive? Are any secondary modern schools running successfully on a modified project method and if so could we hear about them? *(Waldron, TES, Feb.5th 1960, p.220)*

Another contributor to the *Times Educational Supplement* argued that "...without some incentive pupils of these schools wander willy-nilly in a haze
of educational haphazard". (Murray, TES, July 8th 1960, p.59). This comment prompted the following response:

Your correspondent ... refers to pupils of Secondary Modern Schools wandering "willy-nilly in a haze of educational haphazard." May I suggest that it is we, the teachers, who wander thus ... somehow we must bring ourselves to face the awful truth; perhaps then we may begin to go forward to the discovery of "how to educate the 80 per cent".  

(Kirkpatrick, TES, 22nd July 1960, p.120)

These feelings of aimlessness seem to have been especially pervasive at the fourth and fifth form level, a stage at which external examinations traditionally had provided a powerful sense of direction. For instance, one teacher claimed that, despite his school being well-equipped by the LEA, "... it remained a fact that, from the third year on, there was a pervasive feeling of non-involvement among the boys which stultified the efforts of the staff" (TES, Aug. 26th 1960, p.239).

In spite of this aimlessness, schools were encouraged to experiment and innovate: "... we look forward to bold experiments", claimed the SSEC (1947, p.46). However, those who enthused over the virtues of innovation sometimes paid little attention to the consequences of freedom in a situation where there were no established criteria by which to evaluate the efficacy of experiments. Schools were sometimes launched on to a course of frantic experiment as an activity which was laudable in itself:

... there was, during the early years, a flurry of rather wild and woolly experiment ... It was, unfortunately, only too easy for people not working in Secondary Modern Schools to be critical of those who were ... when experiments ended in chaotic failure ... It would have been more charitable, and more just, to remember both the magnitude and the terrifying vagueness of the task which the teachers in the new secondary schools had to undertake. (Dent, 1958, pp.146-147)

Another critic warned against the danger of individuality and experimentation being glorified for their own sakes:

Let us beware that in the commendable desire for progress, we do not sacrifice the majority of our secondary modern pupils on the altar of ... the idea that no other school has yet conceived. (Pinkus, TES, March 11th 1960, p.490)
The combination of the pressure to experiment and the teaching profession's own uncertainty made the modern schools easy prey for those with untried theories to peddle:

The Secondary Modern School has been the target of every new-fangled theory, every half-digested, ill-assorted idea. If ever it has seen a settled objective, half the busybodies in the world of education have rushed in to save it from itself. Turn out citizens. Turn out literates. Turn out technicians. Be more vocational. Prepare for leisure. Stick to the wider view.

(Dent, 1958, pp.xi-xii)

The SSEC had been anxious to ensure that modern schools did not emulate the examination-orientated education provided by grammar schools but in the absence of a distinct purpose of their own, many functioned as pale imitations of grammar schools. Holbrook (1961) described a visit to a modern school which illustrated this tendency:

I remember once being shown round a huge new modern school ... The headmaster took me into his large wall-papered study. "Of course", he said, "we only get the duds here." (p.3)

This remark reveals that even those selected to take the helm in modern schools frequently lacked a positive sense of identity and purpose. The grammar school provided this headmaster's yardstick and, when measured against this, his intake was of inferior quality. If the value of a modern school education was to be measured against the grammar school, it would inevitably fall short since it could not hope to excel in this rigorous, academic tradition. This negative outlook was widespread. As the Times Educational Supplement observed, modern school teachers were encouraged "...to fashion a new sort of school, knowing all the while that it was being measured against the old ... to build a new Jerusalem with one hand and fend off the critics with the other" (quoted by Dent, 1958, p.xi).

The modern schools were in a demoralizing situation because although they lacked independent criteria by which to measure the value of their work, they were being measured constantly against the established grammar school tradition.

In this atmosphere of uncertainty, it is not difficult to understand
why modern schools turned increasingly to external examinations to replace that void where a sense of direction should have been. The evidence presented in the Beloe Report (1960) underlined this crucial function which external examinations were fulfilling:

Many teachers in these schools believe that external examinations can provide, as it were, a landmark by which they can take their bearings and measure their standards and progress, in relation both to what is being achieved by other schools of the same type and to what is demanded by the world around them ... This we have been told is a need which, in this exploratory phase of the development of secondary education ... is keenly felt by many teachers. (p.23)

External examinations traditionally had performed this role during periods of expansion and yet schools in the 1940's and '50's were denied this support at a time when they were passing through a period of unprecedented growth. Furthermore, it is especially ironic that it was the growth sectors of secondary education (i.e. the modern schools) which were denied this traditional "prop" whilst the established grammar schools retained its support. Viewed in this light, it is not surprising that modern schools turned to examinations to fill the vacuum which official policy had created.

There is no doubt that the main reason for the introduction of examinations into modern schools is ... to provide a target or an aim. Vague phrases like 'Our object is to provide a good general education', whilst expressing the right ideal need further definition. What constitutes a good general education and what evidence is there to prove it? (Howard, 1960, p.30)

These convictions are echoed many times in the literature of the period and it is interesting to note how often terms like "target" and "aim" feature in descriptions of the role played by examinations. For instance, Rapstoff (1960) asserted that: "The examination does represent a tangible aim, a boon in a type of school where academic aimlessness may be a danger". (p.98).

7. "'CONTRACTING IN' TO THE 'COMPETITION-EXAMINATION-SUCCESS' SYSTEM"

It is difficult to appreciate the "tonic" effect which external examinations had when they were introduced into modern schools unless one
understands something of the demoralizing situation which prevailed in many schools before their introduction. When the results of 11+ Examinations were known, education departments up and down the country were inundated by anxious parents whose children were destined for modern schools. Their aim was to obtain places for their children in other types of secondary schools (grammar or technical) which offered better future prospects than the modern school held out. One headmaster conceded that their reluctance to send their children to modern schools was justified:

They rest on two facts: (a) The Grammar school child gets a better deal than the secondary modern child. In our authority, the capitation grant will be higher, the staffing ratios more favourable, and parents know that teachers in grammar schools are better qualified academically ... (b) An ever-widening range of worthwhile employment uses GCE results. (James, TES, Oct. 21st 1960, p.532)

Thus, many of the parents, for whom there was no alternative but to send their children to modern schools, felt resentful and their children entered the schools with an inbuilt sense of failure and alienation:

Whether we like it or not, any children of above average intelligence felt a sense of failure, and to some extent shame, if they failed to obtain a grammar school place ... Almost the first words uttered by the parent of the new boy in 1950 were: "I want him to have another chance for the scholarship". (Heaton, 1958, p.96)

For many, the main object was to leave the modern school as quickly as possible and the conventional escape routes, for the more able pupils, were the 12+ and 13+ examinations designed to re-evaluate pupils who had been wrongly placed by the 11+ examination. The disastrous effects which repeated selection examinations could have on the morale and learning in modern schools is conveyed by the following account given by a modern school headmaster:

... nearly every year there is the upset of an external examination. In the first two years the upset is greatest, for it is then that the external examination is designed to get them out of the school, with the obvious inference that any school is better than the secondary modern school in which they have been unfortunate enough to find themselves. It is against that kind of background that the modern schools set to work ... we are still working under the shadow of the over-age, or twelve-plus,
examination. I know that a few parents have their boys privately coached for the twelve-plus. I also know that these same boys never manage to settle down to real work during their first year. How can they settle down in a place they intend to leave as soon as possible? ... The next obstacle ... is the thirteen-plus examination or the technical school selection test ... Again, there is the wretched business ... of having boys crammed for this examination. Again, there is the disappointment for those many who do not pass ... These two examinations ... were stumbling blocks to all of us in secondary modern schools. (Heaton, 1958, pp.95-98)

In the early days, modern schools had few incentives to offer those destined to remain there for the rest of their school careers and most left as soon as they were able. Many children, on reaching the statutory leaving age of fifteen, failed even to complete their fourth year of schooling and a fifth year was out of the question for the majority. The absence of a uniform leaving stage created an atmosphere of restlessness and disarray at the top end of the school similar to that created by external examinations in the lower part.

... even today ... the Elementary School tradition of leaving at the end of the term in which school leaving age is reached has persisted. This means that many children leave at Christmas in their fourth year and many more at Easter. Apart from the fact that these children have a truncated course, there is inevitably a sense of restlessness in any class in which there are children who know they will be leaving school in a few weeks alongside others who will be staying considerably longer. (Dent, 1958, p.151)

Clearly, the modern schools were on a course which was at variance with the values and expectations of society in general. Faced with conflicting expectations, there were two principal courses open to them. They could: (a) attempt to divorce themselves from the values operating in society at large and follow the course recommended by the SSEC or (b) reverse their course so that it was more closely attuned to the values of the society in which they must take their place. In the event, the schools almost universally chose the second option; they defied ministerial policy and "... took matters into their own hands." (Taylor, 1960, p.63). In expecting the schools to resist the values which governed society as a whole, the Government displayed the social unrealism which characterized
their vision of society in general:

... the British people are learning by the hard way how
dependent is a civilized community on its farmers, transporters
and miners, its manual and technical workers ... the change in
the public attitude to the craftsman and the technician is
nothing less than revolutionary ... as the organisation of
industry continues to become more flexible, it will offer the
product of the modern school the same hope of promotion to the
more responsible and better-paid jobs as is now in practice
largely confined to the products of the grammar and technical
schools. (The New Secondary Education, 1947, pp.4 and 47)

As Taylor (1960) observed, the post-war world was depicted as:

... one in which the stresses and strains of occupational and
social competition would be considerably lessened. Official and
unofficial sources encouraged the view that equality of
occupational opportunity between children from different types
of secondary school was to be obtained not by providing all
children with access to the same qualifications, but by changes
in the occupational field itself. (p.64)

This outlook proved to be fallacious because "... the increasing demand for
educational certificates by employers, professional associations and others,

as minimum qualifications for entry to certain occupations" (Morris, 1959,
p.35) was a potent force operating on society in the '40's and '50's.

Indeed, so pervasive was this tendency that it was filtering down from
white collar occupations to blue collar work: "It is only since the war
that educational certificates have come to be at all widely demanded as
minimum qualifications for entry to apprenticeships". (Ibid). With hind-
sight, it is easy to perceive that the attempt to insulate the modern
schools against these forces was an error of judgement against which they
were bound to revolt. Indeed, one headmaster described the makers of
educational policy as "... the starry-eyed faddists who have dominated the
years since the war" (McCarthy, TES, Oct. 7th 1960, p.445). Another
critic observed that:

The modern school's gain in status during the fifties has been
largely a result of the way in which it has "contracted in"
to the "competition-examination-success" system rather than by
allowing its aims to be determined along "purely educational
lines". In this the schools have displayed a good deal more
social realism than was characteristic of much post-war
educational thought. (Taylor, 1960, p.63)
By taking their fate into their own hands, the modern schools turned public opinion in their favour, sometimes with astonishing success. The state of affairs in schools not using examinations (see the beginning of this section) provides a stark contrast to schools in which the therapeutic effect of examinations had been discovered. For instance, the headmaster of the Knoll School for Boys recounted the consequences of "contracting in' to the 'competition-examination-success' system" in his school:

Just over two and a half years ago I came to a secondary modern school of some 300 boys, none of whom stayed on for a fifth year. The only goal which could be offered to the brightest boys was to pass the entrance examination for the technical institute of a neighbouring authority. The staff and I decided that we should provide an examination aim not only for the 'top flyers' but for the middle stream boys as well. (TES, 10th June 1960, p.1194)

To this end, a five year GCE course was established for "A" stream pupils while the "B" stream embarked on a five year RSA course and the "N" stream started a four year UE1 course. Although the school had decided that three years was the minimum of preparation any boy should receive before entering for 0-Level, "Last summer, the temptation to experiment proved irresistible: we were glad we did so. Six boys out of eight passed in Mathematics, four out of six in English Language and two out of three in Geography ... All except one ... are staying to take further subjects next summer and some have the prospect of leaving school with up to seven subjects at 'O' level". (Ibid). This is one example, from many cited during this period, of modern school pupils achieving considerable success in 0-Level examinations. The 11+ Examination was not an infallible selection instrument and it was increasingly recognised that there were considerable reserves of talent in the modern schools - children able to pass 0-Level in a number of subjects and willing to stay at school for a fifth year in order to do so.

Perhaps more important than examination success itself was the tonic effect it had on "... the morale and reputation of the school":

... over 80 per cent (of parents) give a written undertaking that their sons shall remain for five years. Whereas previously there was no school uniform, it is now worn by 95 per cent of the boys.
To these two factors more than anything else I attribute the
growth of the school both in numbers and in status. The 300
referred to at the beginning is now 460 due to three causes:
forty boys staying on for a fifth or sixth year, giving us two
fifth forms; the reluctance of the parents of 41 boys out of
60 to transfer their sons to a nearby brand-new modern school
some 18 months ago when they had been scheduled to do so; the
admission of about 24 boys from private schools in the last two
year ... In many cases ... the boys themselves are working at a
pressure which surprises even them. (Ibid)

These changes, remarkable in themselves, are made even more noteworthy by
the speed with which they occurred. Within the short space of two and a
half years, external examinations had already had a dramatic impact on the
status and morale of the school.

The depressing state of affairs described on page 65 refers to the
situation at the Belmont Secondary Modern School for Boys before it
resorted to external examinations. The school eventually devised a system
similar to that in operation at the Knoll School in that examinations
pitched at three separate levels enabled the school to provide an examina-
tion objective for children throughout their ability range. O-Level was
taken by the most able, an RSA course by those of lesser ability and the
Harrow School Leaving Certificate catered for the least able pupils. Once
again, this produced a volte-face amongst the pupils and their parents:

It is interesting to notice the effect on the boys and the
parents. My experience is not unique, as I know from many talks
with colleagues ... Whereas a few years ago every boy in the
first year wished to take the over-age twelve-plus test, the
candidates now are less than 4 per cent of the intake. A
similar small number ask to take the thirteen-plus ... The
average boy in the fourth year, no matter what his academic
quality, now stays on for the full four years; it is the
minority who leave at the end of the term in which they are
fifteen. The fifth year will include 20 to 30 per cent of their
year group. I get the impression that the children feel that the
door is open to them ... A big change has taken place in the
attitude of the parents. They are happier. I now get letters
from them saying how much they want their boy to stay on.
(Heaton, 1958, pp.101-102)

Perhaps the key phrase in this account is that "... the children feel that
the door is open to them." Likewise, another observer described examina-
tion successes as "... that coveted Open Sesame" (Dent, 1958, p.158), a
phrase which epitomizes the significance of external examinations in modern schools. This was in sharp contrast to the situation created in 1947 when the SSEC, in effect, closed the door against the modern schools. Secondary education was a highly competitive field of endeavour, its entire route lined and determined by performance in academic tests. Examinations at 11+, 12+, 13+ and 16+ meant that, at each stage, access to the next level was controlled by examinations. Therefore, modern schools had to function in an examinations-orientated environment while they themselves were denied access to them. Moreover, although modern schools were denied the right to use examinations for their own purposes, they were nevertheless subjected to them and lived under their shadow. In the event, examination successes represented the weapon which enabled modern schools to compete with grammar and technical schools for parity of esteem.

As far as Crofts (1958) was concerned, his school had been transformed from a blind alley which could offer its pupils none of the conventional routes to self-advancement to a place of purpose and self-respect by the introduction of external examinations. He claimed that examinations:

... are helpful in replacing the notion that the modern school is a "dead-end" for the "also-rans" with the knowledge that it is indeed a place of opportunity. In human terms I have found this to mean the difference between a child arriving in 1943 with an expressed feeling of disappointment ... and a similar one, one of six in 1953, rejecting a grammar school opportunity for a more appropriate one in terms of predominant interests. (pp.22 - 23)

This tonic effect was not limited to examination classes but permeated the entire school charging it with a sense of purpose:

... there is a new air breathing through the secondary schools today. It is an air of purpose resulting from the presence of older pupils who have shown, by staying, that the school has something it is worth their while to get ... It shows in improved standards of work. It is apparent in the quality of both the teaching and the learning. It shows in many ways in the corporate life of the school. (p.24)

Heaton (1958) had noticed similar changes in his school: "It lifts up the whole of the standards throughout the school ... With this better material many sides of the school can flourish, and the crafts, drama and debate take
on a greater liveliness" (pp.104 and 109).

In 1953, Bournemouth Education Committee started what it described as "a grammar school course" (i.e. an O-Level course) in each modern school within its borough. In 1959, the Chief Education Officer published, "A Progress Report on Bournemouth: G.C.E. in the Secondary Modern School" which reflected the results from other studies. Firstly, the pass rate compared not unfavourably with that in grammar schools especially in the practical subjects. A premium of arguably greater importance was the boost it gave to the self-esteem of pupils, teachers and parents:

There have been many encouraging signs during the five-year experiment period. First is the reaction of the pupils. There is now a sense of purpose in the secondary modern schools, and the older pupils have introduced a dignity and maturity which has a highly beneficial impact throughout the school ... At the end of the period over 40 per cent of the pupils stay on at school for a full five-year course ... which is the best testimony that parents can give to the scheme. There has been a noticeable lessening of tension about the outcome of the common entrance examination, and visits to the office by parents of children who have not qualified for a place in the grammar schools have shrunk to negligible proportions ... The staff, too, have appreciated the new atmosphere of the schools and are proud of the improved status and standing of their schools ... This year ... the headteachers unanimously voted the scheme an unqualified success.

(Smedley, 1959, p.1037)

Perhaps most important of all, the headteachers concurred that the scheme "...had had highly beneficial effects on both pupils and schools far transcending examination results." (Ibid)

Although it would be easy to continue to replicate these examples, the evidence produced so far is essentially subjective and is based on the personal views of the individuals concerned. Therefore, it is useful to come across a controlled experiment designed to "... test objectively the view that external examinations have an improving effect upon the attitude to school work of pupils in secondary modern schools" (Kirkpatrick, 1965, p.128). This enquiry involved the construction of an attitude scale designed to measure pupils' attitudes to school work. Three groups of
subjects were selected: Groups A and B were preparing for external examinations whilst Group C, selected from schools not using examinations, was used as a control group. As might have been anticipated, "... the abler children in all schools had developed more favourable attitudes to school work than their less able fellows" (p.129). In spite of this, Kirkpatrick (1965) found that:

... the most able children at the schools where external examinations were taken had significantly better attitudes to their work than similar children in the control group of schools. (p.131)

Furthermore:

... in all schools, the children who intended to take an external examination had a more favourable attitude to school work than those who did not ... the differences between the average scores were almost all significant at the first percentile. (p.132)

Realising that it was possible to argue that his findings merely demonstrated that "... the children who had the better attitudes to begin with were the children who had decided to take the examinations" (p.133), Kirkpatrick (1965) carried out an additional analysis which demonstrated that in most cases the examination had been "... an active agent, affecting attitudes, rather than a mere vehicle for the expression of existing favourable attitudes" (p.134). Thus, he was able to conclude that "... the 'tonic' effect of external examinations upon those who take them was clearly demonstrated " (p.135).

Another systematic piece of research (Morris, 1959) endorsed these findings. Morris's study of five-year examination courses in modern schools in Birmingham once again highlighted their value as an incentive to work:

About three-quarters of the Heads drew attention to it, and some added that it was by no means limited to the fifth year children. Heads' tributes to this incentive value were often warm ... Among the values of the courses most often mentioned was the greater poise and maturity which resulted ... Parents' increased confidence ... is indicated by the recent hints of a decline both in the strength of the passions aroused by the 11-plus and in the proportion of children who leave the schools as soon as they are allowed to. Nearly all the heads thought that the disappointment their children felt about the 11-plus had been lessened by the introduction of the courses. A few added that some children were
turning down the chance of transferring to a technical school at 13, in order to take a fifth-year course ... from the proportion of their staff who were graduates, it was possible to make inferences about their prestige in the eyes of graduate teachers; in this respect, too, their status is slowly rising. It seems, then, that the prestige of the schools has risen, in both the professional and the parental eye. (pp.40-42)

An amusing illustration of this tonic effect was provided by Rapstoff (1960):

Some time ago I was fortunate to hear a conversation among three intelligent children who, by coincidence, came from various secondary modern schools. They each insisted that their school was recognised as the best in Birmingham. They were obviously proud of their schools. But for me the interesting fact was that the discussion almost petered out when they all revealed possession of the trump card, "Our school has a G.C.E. stream!" This experience suggests that among secondary modern children where GCE work is done, it may become the criterion of a good school. (p.98)

It is hardly surprising that when external examinations were capable of producing such a powerful reaction, there was almost irresistible pressure on the schools to use them. As more schools advertised their successes, the pressure on their counterparts to use examinations became more and more compelling. It is this "self-propagation" (Morris, 1959, p.41) which probably accounts for the accelerating rate of growth in examination entries towards the end of the 1950's (see p.51). In fact, by 1960 this trend had gained such momentum that the Beloe Committee predicted that:

... within a relatively few years the tendency towards the taking of external examinations amongst pupils below the G.C.E. group will gather such momentum that, even with continued ministerial discouragement, almost all schools will find themselves willingly or unwillingly conforming. (p.27)

In 1947, the SSEC had attempted to treat examinations as a detached and isolated phenomenon which could be omitted from one part of the system but as one critic remarked: "One cannot consider examinations in vacuo" (Howard, 1960, p.30). Examinations were not an isolated phenomenon detached from the prevailing value system and, in recognising this, the schools displayed the kind of social realism which was lacking in much post-war educational policy. The interdependence of the system as a whole
was argued thus by one headteacher:

... as there is already a Leaving Examination in the selective schools, there naturally is a demand for the institution of an examination in the secondary modern schools ... it is wrong to talk of the secondary modern school as a unit apart in this matter. We are continually telling the public that this generation has been given secondary education for all, and it is in this context that the question of a School Leaving Examination should be considered. (Burdett, 1960, p.12)

Another head, who had decided to defy ministerial policy and use examinations, claimed that: "For all these (pupils) ... full stature at 16 has in no way been limited by the outlook at 11 plus " (Crofts, 1958, p.22). The equation of "full stature" with the right to use external examinations shows how closely parity of esteem was bound up with access to the examining system. Indeed, examinations appear to have been the single most important factor determining the prestige of the modern schools in the 1950's. This is not to deny that the problems facing modern schools were many and complex or to suggest that they could be solved by a single or a simple answer. Nevertheless, the evidence points overwhelmingly to the omission of external examinations from the secondary modern "package" as the most important missing ingredient. This is confirmed by the fact that schools which retained their former inadequacies nevertheless experienced an immense boost to their self-esteem once they "'contracted in' to the 'competition-examination-success' system " (Taylor, 1960, p.63).

8. SCHOOL-BASED ASSESSMENTS

Although this chapter has been critical of the SSEC's attempt to combine external examinations with internal assessment procedures, there was a certain justification for the SSEC's decision not to adopt internal assessments wholesale as the Norwood Report (1943) had urged. This would have been a massive undertaking for which secondary education was ill-equipped. Firstly, the provision of free secondary education for all was in itself a major undertaking and it is unlikely that teachers simultaneously could have
taken responsibility for school-leaving examinations. Moreover, there
is evidence to suggest that Mode 3 examinations were somewhat premature
in the 1940's. It must be remembered that at the inauguration of the
School Certificate, there had been an attempt to encourage teachers to take
a more active role in examining (see Chap. 1, p.24) but as Roach (1979,
p.46) pointed out, most teachers seemed to prefer that their pupils should
be assessed by an independent authority. The fact that teachers had shown
a reluctance to accept this kind of responsibility casts further doubt on
the feasibility of Norwood's proposals. It also meant that, during the
School Certificate era, the teaching profession had failed to accumulate the
expertise necessary to operate examinations and it is hard to believe that
this could have been acquired in the seven years suggested by the Norwood
Report. These difficulties would have been exacerbated by the massive
increase in the secondary school population brought about by the 1944
Education Act and the inevitable strain this placed upon the supply of
qualified, experienced teachers. There was a severe shortage of teachers
throughout this period although the Emergency Training Schemes did provide
a stop-gap.

If the SSEC's unwillingness to replace external examinations by internal
schemes did derive from an awareness of the magnitude and complexity of the
task, this only renders their lack of involvement in initiating school-based
procedures even more unjustified. It had recommended that internal examina-
tions should become the principal method of assessment, yet it made detailed
provisions for the establishment of GCE while the creation of internal
schemes was left to chance. The SSEC took no part in initiating these
schemes, fostering their development or monitoring their progress. Instead
it retreated into its traditional laissez-faire approach. Thus, although
the State had taken increasing responsibility for education in the twentieth
century, school-based assessments represented a critical new area of growth
where the SSEC failed to act. Consequently, the Beloe Report (1960) found
that the development of internal leaving certificates had been quite random:

Some examinations came into being through the initiative of local groups of teachers; others grew up under the leadership of a local education authority. Some are conducted by the local education authority itself or by divisional executives, others by groups of teachers. (p.12)

A more serious problem than the unsystematic nature of these schemes was that they had simply failed to materialise on a significant scale. Even when it started to become apparent that school-based examinations were failing to develop, the Minister responded by issuing a circular in which he re-urged the use of school records, objective tests etc. and expressed disapproval of the use of external examinations in modern schools (Circular 289). These early warning signals, suggesting that school-based schemes were not going to develop spontaneously, should have prompted the Minister to positive action but as the Times Educational Supplement observed: "The Minister expresses the national policy and looks the other way." (4th Nov. 1960, p.611). Another critic made a similar point when he argued, "Let us stop playing the ostrich about examinations for the secondary school" (TES, 4th Nov. 1960, p.611).

By 1960, the Beloe Report found that only 30 out of a total of 146 LEA's in England and Wales had schools conducting internal leaving certificates. In other words, less than a quarter of the LEA's contained schools which were following the official policy. Moreover, it has to be remembered that by no means all of the schools in those LEA's would have been using these schemes. The full force of these figures is underlined by the fact that by 1960 137 of the LEA's contained schools which entered candidates for external examinations other than GCE. One member of the Beloe Committee observed that:

It seemed to us that a most undesirable situation was developing in which the Ministerial policy was in fact being set aside as of no account, a position which is usually regarded as discrediting and officially unrealistic. 

(Journal of the A.A.M., Vol. 12, 1961, p.61)
Thus, although one must be sympathetic to that generation's consciousness of taking "a leap in the dark" and their conviction that modern schools needed room to grow, it is difficult to justify the complete absence of provisions for monitoring and aiding the development of internal assessment procedures. Indeed, as one member of the Beloe Committee noted:

At the point at which we began our work, the area (i.e. examinations for the modern schools) ... remained completely uncharted in official policy.


9. THE BELOE REPORT AND THE MODERN SCHOOLS

There is a curious discrepancy in the viewpoints expressed during this period. The mood of the Beloe Report was one of deep pessimism about the use of external examinations in modern schools. They claimed to "... have found much which caused us misgiving, and some features that caused us grave disquiet" (p.17) arguing that if the existing situation was allowed to persist, it "... would entail grave dangers for the development of secondary education in the next ten years" (p.27). This contrasted starkly with the mood prevailing in the schools themselves which lavished warm praise on the examinations. Why was it that the examinations which were so warmly welcomed by the schools which used them caused the Beloe Committee such acute concern?

It appears that this divergence was due to the different standpoints adopted by the Committee and the schools. The Committee were able to take an overview of the situation and in this way they were able to detect overall trends at the macrocosmic level which may have been imperceptible at the microcosmic level of individual schools.

It is clearly of first importance ... to see not only what the pattern is now but what it may be in 1965 or 1970, assuming no action is taken to affect the course of developments ... first ... it seems inevitable that within five ... years' time the schools will almost without exception be entering those of their pupils who are in the appropriate age and ability ranges for external examinations ... The probability is that examining below the GCE level will become increasingly concentrated in the hands of the
eight existing regional and "national" Bodies ... it is likely to prove increasingly difficult for the smaller local bodies or groups to remain active, or for new Bodies to come into being without assistance. It follows that the existing Bodies will grow both in size and influence ... This in its turn must increase their problems, and in particular that of keeping in close touch with the schools who use their examinations ... by the end of the decade, if not before, the business of external examining below the GCE level will have become largely if not entirely concentrated in the hands of a limited group of examining bodies, free to pursue their own policies without reference to the long-term needs of the schools or of the educational system as a whole, growing rapidly and in danger because of their small number and the vast field open to them of becoming increasingly remote from the schools and the teachers they serve. (pp.19-20)

Thus, the Beloe Report identified disturbing overall trends which had hitherto gone unrecognised. Indeed, many of these problems were at a nascent stage which meant that the situation was fraught with problems which were potential rather than actual.

The problem of backwash effects illustrates this point. This had been one of the most "persistently perennial" problems associated with external examining (see, for example, the Report of the Schools Inquiry Commission (1868), the Report of the British Association for the Advancement of Science (1903) or the Spens Report (1939)) and the Beloe Committee regarded it as latent in the circumstances prevailing in the 1950's: "... as the numbers taking their examinations increase ... their syllabuses and papers will, if earlier history is any guide, come increasingly to influence the curriculum and teaching in the schools" (p.20). Thus, one might have expected a renewed outcry against the cramping effects of examinations in the 1950's. However, one of the most striking features of this period is the infrequency with which this criticism is met. For instance, Morris's study (1959) of examination courses in Birmingham modern schools demonstrated the comparative insignificance of this problem. Morris enquired "... whether the Heads thought that their schools were being cramped by offering courses leading to several external examinations ... About half thought that the examinations were not at all cramping; a quarter were non-committal; and a quarter thought that in some ways they
were cramping" (p.37). The fact that so few of the heads believed that
the examinations cramped their school's work is noteworthy but an even
more telling indication of the lack of concern over this issue is that the
heads had to be asked before they mentioned it. In comparison, they
actually "... drew attention to" the tonic effect of the examinations and
Morris was struck by the warmth of their praise. This was typical of
writings of this period. Most teachers who wrote about external examina-
tions simply failed to mention their backwash effects thereby indicating
how little this bothered them. Of those who did mention it, their usual
purpose was not to endorse the criticism but to defend their examinations
against what they recognised as a common objection. For instance, Dellan
(1960) concluded his account of the successful introduction of external
examinations into the Knoll School by defending them thus:

... one can hardly allow the criticism that an examination
programme is educationally restrictive when it allows the choice
of (thirteen subjects) to "O" level. (TES, 10th June, 1960, p.1194)

Even those who did acknowledge the stultifying effect of examinations treated
the subject dismissively. For example, Darlow's (1958) largely favourable
appraisal of the use of examinations in his school contained a brief
paragraph devoted to their backwash effect on teaching:

... the existence of an 'examination requirement' tends to make a
man regard his classes as recipients of a regular series of doses
of instruction which must be administered willy nilly and against
the clock. (p.1125)

But the problem is not dwelt upon— instead it is rapidly qualified: "...when I try to strike a balance between the advantages and disadvantages, I
am sure that the advantages are much greater" (p.1125). This climate of
opinion contrasts starkly with that prevailing from the 1860's onward when
teachers were intensely dissatisfied with the backwash effects of examina-
tions. In both instances there had been a proliferation of external
examinations and yet the reactions were quite different. Why?

Between 1868 and 1903 teachers were complaining about a situation
which had persisted for many years. In contrast, those backwash effects
detected in the 1950's were at an incipient stage where they may not have been perceived as a problem by most teachers. Thus, an obvious source of difference lies in the speed with which the problem was attacked on each occasion. The Beloe Committee (1960) acted so swiftly that a potential difficulty was eliminated almost before it was recognised by most educationists. In contrast, most of the teachers who criticised the glut of examinations at the turn of the century could not remember what teaching had been like before their introduction which made it easy for them to criticise their deficiencies without appreciating their benefits. Clearly, the situation was quite different in the 1950's; the teachers who welcomed examinations were acutely aware of what modern schools were like without them. External examinations had not become sufficiently well-established in the modern schools for their flaws to have emerged fully or for teachers' memories of what life was like without them to have become blunted. Any loss of supposed freedom, any rigidity which the examinations introduced, must have seemed a small price to pay when examinations conferred a sense of purpose and self-esteem on schools which had originally lacked these qualities. The teachers' viewpoint was determined by their position, entrenched in the struggle of the modern schools to achieve parity of esteem but the Beloe Committee were detached, objective observers, able to provide a dispassionate overview of long-term trends. Thus, the praise lavished upon examinations in the 1950's must not be taken to imply that examinations had been miraculously freed of their traditional problems. As Norwood (1929) had warned:

... the examination system is no friend to whom one can trust in blind confidence: from its first guise as a servant it may at any moment reveal itself as a hard and unsympathetic master.

(p.208)

10. THE BELOE REPORT'S FEARS JUSTIFIED

The Beloe Report's (1960) fear that examinations would exert an
increasingly powerful influence over schools was already being vindicated. There was evidence that examinations were beginning to dominate the curricula in some schools while in others the educational balance was being severely disrupted. For instance, one teacher observed that his school was:

... geared to what may paradoxically turn out to be an iniquitous introduction - that of GCE for secondary modern children. Timetabling these 25 pupils has the effect of throwing out the whole of the rest of the school, since they require extra periods of specific lessons, and therefore extra rooms and extra teaching time. The timetable has never been in operation for more than one term; subject teachers are always complaining that not enough time is given to their GCE classes and attempts at adjustment have to be made. (Correspondent, TES, 29th Jan. 1960, p.160)

The main difficulty undoubtedly arose from the use of GCE in modern schools. Teachers gradually realised that, in an attempt to meet its rigorous demands, the educational balance inside their schools was being overthrown:

How many other teachers have experienced these conditions in secondary modern schools taking GCE? I walked round such a school and found all those teachers who received one or more hundreds of pounds in special allowances were regularly taking classes of from five - ten bright pupils while I, taking a non-GCE subject, was taking over-size classes (even up to 80, in the hall) of more backward pupils. Separate staff meetings were held for the former teachers. (TES, 22nd April 1960, p.810)

With hindsight, this development is entirely predictable since the history of examinations had shown that an examination which acts as a passport to higher education or better job opportunities induces teachers and their pupils to treat it as "the be-all and the end-all" (Schools Inquiry Commission, 1868, p.322) of education.

There was evidence of a groundswell of revolt against this amongst modern school teachers towards the end of the 1950's:

A young teacher ... complained bitterly at a conference for new teachers that one man at his school spent all day coaching one pupil for GCE subjects, while he himself took 38 "C" streamers. Such things we know full well are quite common ... I cannot help thinking that many of us heads must have very guilty consciences over the way we run our fifth ... form courses.

(Headmistress, TES, Aug. 5th 1960, p.178)

However, one head openly admitted that, "The GCE pupils will give the school
purpose and raise the tone: the rest can get on as best they can" (TES, 12th Feb. 1960, p.279). An interesting perspective on this was provided by a pupil who had attended several modern schools. She observed that, "In some of them they seem to push the backward ones backward, and the front ones they push forward" (TES, 5th Aug. 1960, p.178).

This remark neatly captured the imbalance which characterized the approach of some modern schools. It was summarized by Blishen (1960) as, "... an agonising type of war between the needs of the majority and the needs of the minority" (TES, 12th Feb. 1960, p.279). However, this concentration on a small band of "high flyers" was questioned increasingly. As one teacher warned:

We are justifiably proud of our successes, but how far have they turned us away from the fundamentals? Let us beware that in the commendable desire for progress, we do not sacrifice the majority of our secondary modern pupils on the altar of ... the GCE class. (TES, 11th March 1960, p.490)

The value of an education dominated by GCE was also questioned because, as more schools experimented with it, the pass rate was found to be modest:

... in 1957 nearly 9,000 children from modern schools attempted subjects in the GCE examination; of these 1,697 failed to get any passes and 5,500 passed in three subjects or under. A little over 900 passed in five subjects and over. There is a growing feeling that the gaining of a certificate in an odd subject or two after a prolonged course is a waste of time. (Burdett, 1960, p.13)

A school servile to the rigours of GCE was the price being paid for the acquisition of a few O-Levels by a minority of pupils and some teachers were beginning to feel that this price was too high.

The GCE forms are the Olympus to which many will vainly strive but few will reach. The climb towards it may leave many weary broken travellers on its slopes, and while their view is limited to the conquest of its formidable crest, the view of greener pastures below will be lost. (Rapstoff, 1960, p.98)

There were three different types of examination in use in modern schools at this period: GCE; other external examinations below the GCE level and internal school-leaving certificates. The Beloe Report evaluated all three.
11. THE BELOE REPORT ON INTERNAL EXAMINATIONS

By 1960, it was clear that most schools were exceedingly reluctant to use internal assessment procedures. One of the most serious obstacles to their acceptance was their strictly local currency:

... the certificates awarded on them are ... unlikely to have more than local currency. The value of a certificate awarded by a group of schools in, say, a north London borough may be well understood by an employer or a technical college principal in that borough. But it may have little meaning or value to an employer or principal in Manchester or even on the other side of London; and this lack of wider currency detracts from its value. (Beloe Report, 1960, p.12)

Whereas the views of the Beloe Committee were at variance with wider opinion on many issues, this was a view which was widely shared:

Like so many other secondary modern schools at this time of year, we are ploughing through the motions of a 'school leaving' examination. These examinations are carried through with scrupulous efficiency and fairness: pupils revise quite thoroughly and papers of fair difficulty are devised ... Yet when all the results are known and the printed certificates completed, we know, and the children realise as well, that the piece of paper is really worthless, except perhaps within the small environs of the local town.

(Murray, TES, 8th July 1960, p.59)

This letter prompted a number of responses and not one of those printed defended internal examinations against this charge. One headteacher claimed that,"... no school leaving examination ... carried through by the internal staff will ever demand public respect like an external examination does "

(Ansom Smith, TES, 22nd July 1960, p.120). This reservation about the value of internal certificates was echoed many times in the literature of the period:

The weakness here is all too clear - the examination and certificate have no more validity than the ordinary school examination and report, and because they cannot be related to an external standard, will have little prestige.

(Burgett, 1960, p.13)

Even those teachers who welcomed the introduction of internal schemes in their own schools recognised this as a major drawback. For instance, Darlow (1958), although pleased by the tonic effect of the "Governors' Certificate" in his school, conceded that: "... it is not a perfect
instrument. First, the standards which permit a 'pass' ... are not necessarily acceptable to anybody else. People locally ... may accept its standards, some people only as 'better than nothing'." (p.1125). The Beloe Report found that this attitude was part of "... a strong and growing desire, both amongst teachers and amongst parents whose children are thought to be unsuited to attempt the GCE examination, to enter them for some other examination with wider than local currency." (p.19). Public opinion favoured external examining regarding it as the safeguard of acceptable, recognisable standards. Furthermore, a survey undertaken during this period suggested that internal examinations were at variance with the prevailing socio-economic climate:

In general, the smaller the urban authority, the more acceptable it finds purely local examination ... the medium-sized ones have taken an intermediate position; while the large cities (say 400,000 and over) have generally favoured an examination of regional or national standing ... the large cities are most likely to be adversely affected by the Ministry's policy on examinations ... for their industries are more dependent upon an impersonal system of recruitment according to ability. (Morris, 1959, p.1036)

This study suggested that while internal examinations might operate successfully under certain favourable conditions, they were not applicable on a large scale, again demonstrating the lack of social realism implicit in the SSEC's 1947 provisions. There were, in fact, several forces at work in society which undermined the SSEC's attempt to make examining an increasingly internal activity. There was an ever-widening demand for paper qualifications which was filtering down from professional and white collar occupations to blue collar work. This combined with the growth of large towns and cities and an accelerated rate of mobility to make internal leaving certificates, with their limited local currency, impracticable.

The Beloe Committee were also opposed to them on the grounds that most were designed to be taken at the end of a four year course which, they believed, was detrimental to progress in secondary education:

... the development of secondary education in this country has now reached a point at which external examinations taken ... at
the end of a four year course of secondary education, are inappropriate, and ... any further extension of their use would tend to hinder rather than to promote educational advance. (p.29)

This was a view which many teachers strongly opposed believing that four year examination courses promoted rather than hindered progress. As Moss (1960) argued: "Important though persuading children to stay for a fifth year is, the urgent immediate need is to get children to complete their fourth year" (TES, 28th Oct. 1960, p.574). It was claimed that not only did these certificates tempt many who would have left as soon as they reached fifteen to complete their fourth year but that success at this stage encouraged pupils to embark on a fifth year course:

... children each year decide, because of their success in the local certificate, to undertake a fifth year's work - and what is perhaps more pertinent, their parents are persuaded thereby to allow them to do so. (Cutts, TES, 28th Oct. 1960, p.574)

Once again the discrepancy between the Committee's viewpoint and that of practising teachers appears to have been due to the Committee taking an objective overview of long-term trends. One of the most pernicious features of examining prior to 1917 had been the tendency to over-examine pupils and the Committee saw this as latent in the existing situation:

... where fourth year examinations are available alongside a fifth year examination, there is always a danger, as was found to be the case before the old "Junior Certificate" examination was given up, that pupils will be entered for one after the other and thus subjected to excessive examining. (p.29)

The Committee regarded the government's laissez-faire policy on examinations in modern schools as likely to accelerate this tendency and, in a society which relied increasingly on paper qualifications, the temptation to accumulate as many certificates as possible was likely to become irresistible. Evidence from the literature of the period suggests that these fears were not unfounded. For instance, one headteacher described how:

I have now for several years used successfully the certificate examination of the College of Preceptors for my fourth-year boys ... After obtaining these certificates many boys thrilled with their first educational success elect to return for a fifth year and even a sixth year. We then train and enter them for the senior examination of the College of Preceptors and for GCE (O) level. (Ansom Smith, TES, 22nd July 1960, p.120)
Another head treated fourth year certificates as a "... useful halfway
test for children going on to GCE in the fifth year" (TES, 28th Oct.
1960, p.574). From the teachers' point of view, the powerful tonic
effect of these examinations was cloaking the underlying tendency
towards excessive examining of pupils during their final years of school-
ing. Once again the problem appears to have been at an incipient stage
where teachers were more likely to regard repeated examination entries as
a boon. For instance, Darlow (1958) proudly described how, "A crop of
subsidiary examinations has sprung up, following the stimulus of our own
internal examination" (p.1126), clearly regarding this "self-propagation"
as an asset.

A final reason why the Beloe Committee disliked internal certificates
was that they lacked the public esteem, and consequent power to motivate
pupils, which external examinations possessed:

... it would seem that the (tonic) effects which have been
described stem from the fact that the results of the examinations
are to some degree regarded ... by employers, technical college
principals and others who use them, as having external validity
... the current policy of the Minister (holds) that a system of
examinations for modern schools could be made effective without
introducing an element of externality. We wish to state our
conviction that there is a close connection, which is liable to
become still closer as the pupil's age increases, between the
tonic effects of such examinations and their externality. (p.23)

12. THE BELOE REPORT ON GCE AND NON-GCE EXTERNAL EXAMINATIONS

Since the Beloe Committee had expressed a strong preference for
external examinations, it is tempting to assume that they would have
welcomed the increasing use of GCE in modern schools. The Beloe Committee
were, in fact, deeply disturbed by the accelerating use of GCE, believing
that its rigorous academic approach was ill-suited to the needs of most
modern school pupils. Since they believed that these pupils required an
education which was fundamentally different from that provided by grammar
schools, they naturally felt that examinations for modern schools should
also be essentially different in character and style from GCE. It was for this reason that they rejected the frequent suggestion that the pass level in GCE should be lowered or that the GCE Boards should devise a new examination at a lower level in addition to O-Level:

... what is wanted is not an examination which simply reproduces the GCE pattern at a lower level, but an examination with a different character and aims ... (It) should be specially designed to suit the needs and interests of children in the ability range. (pp.30-31)

They were equally dissatisfied with most of the non-GCE external examinations. A feature which caused them "grave disquiet" (p.17) was that "... with one or two exceptions, the examinations lack a distinctive aim of their own; that those planning them seem for the most part to have been content to borrow, at a lower level and with more restricted syllabuses, from the GCE pattern, or that of the old School Certificate" (p.17). This further undermined the attempts of modern schools to establish a separate identity and encouraged their aping of grammar school curricula. Thus, one of the most damaging effects of these "watered-down" versions of GCE is that they nurtured the tendency of modern schools to function as pale imitations of grammar schools.

The Beloe Committee also called the professionalism of these boards into question:

... the marking arrangements seemed to be far from satisfactory ... marking was decidedly uneven; detailed mark schemes often appeared to be lacking, and moderating arrangements to be defective. Moreover, the marking seemed to be unduly lenient ... we too often had the impression that the standard implied by a question was denied by the leniency of the marking. (p.17)

Clearly, many of these authorities lacked the necessary expertise and the atmosphere of intense competition which existed between them prevented the kind of professional exchange which might have improved the quality of their work:

... because of the present relationships, or lack of them, between the different bodies, there was little evidence of interchange and discussion between them such as might encourage the dissemination of useful ideas and experience. (p.19)
These deficiencies were compounded by the indifference which many of these boards displayed towards research designed to improve assessment techniques and to developments in the philosophy of individual subjects: "We found little evidence that the Examining Bodies were paying special attention to research and to experiments in new techniques of examining appropriate to changing methods in the schools." (p.19)

Lack of co-operation has also been a "persistently perennial" problem in the history of examining especially during periods when the government has adopted a non-interventionist policy, thereby encouraging examining authorities to act as free and rival agents. For instance, the Oxford and Cambridge "locals" were so closely related in aim and purpose that the universities were exhorted to co-operate in their conduct. However, the Schools Inquiry Commission (1868) reported that in spite of, "Repeated attempts ... to bring about a co-operation of the Universities ... the examinations are still quite independent of each other." (p.335).

Another problem of long standing was the backwash effects of external examinations. Their stultifying effect on teaching was as much a subject of concern in 1960 as it had been in the 1860's:

... the syllabuses are narrow ... (and) the questions set are too predominantly of a kind that call only for memorised facts and opinions, rather than eliciting the pupil's interest, imagination or comment from direct experience. While we accept that at this ... level a testing of the candidate's factual knowledge is essential, we believe that an examination which does no more than that is inadequate to its purpose and deadening in its effect on teaching. (p.17)

In the long term, the position was a dangerous one. Not only were many of these boards lacking in the fundamental skills necessary to conduct public examinations, their failure to assimilate advances in the field of assessment or to keep abreast of developments in individual subject ideologies would make them increasingly unsatisfactory. The Beloe Committee's outlook was gloom-laden. In the absence of radical changes, they predicted that more and more schools would mount the "examinations'
bandwagon" placing the examinations in a powerful position where their
defects could be more widely perpetrated:

... examining below the GCE level will become increasingly
concentrated in the hands of the eight existing regional and
"national" Bodies ... It follows that the existing Bodies will
grow in both size and influence... This in its turn must
increase their problems, and in particular that of keeping in
close touch with the schools who use their examinations ... their
syllabuses and papers will, if earlier history is any guide, come
increasingly to influence the curriculum and teaching in the
schools. Thus, by the end of the decade, if not before, the
business of external examining below the GCE level will have
become largely if not entirely concentrated in the hands of a
limited group of examining bodies, free to pursue their own
policies without reference to the long-term needs of the schools
or of the educational system as a whole. (pp.19-20)

A final problem noted by the Beloe Committee was that since each
authority was free to determine the scope and standards of its own examina-
tions, there was a bewildering variety of requirements. One commentator
described this as,"A veritable jungle of examinations" (Holmes, 1961, p.51)
while another observed that examinations had " ... snowballed into a vast
disordered practice" (TES, 4th Nov. 1960, p.611). This seriously under-
minded their usefulness to employers who were confused by the range and levels
of attainment implied by different school-leaving certificates. One member
of the Beloe Committee observed that:

We ... met repeatedly the demand for clear and easily understood
evidence by employers and the increasing need for qualifying
credentials for school leavers. (Journal of the AAM, 1961, p.62)

More recently, employers have found it difficult to grasp the meaning of
the two-tier GCE/CSE system (see Freedman, 1981). Since this stream-lined
system has been a source of much misunderstanding, the medley which preceded
it is likely to have caused great difficulties had it been allowed to
persist.

The Beloe Committee concluded that the existing situation was
unacceptable, that the range of examinations in use should be eliminated and
their place taken by a single new examination specifically designed for use
by modern schools. The Committee's demand for far-reaching reform is easy
to understand once it is appreciated that it was not just isolated examinations which gave them cause for concern but the whole spectrum of examinations currently in use in modern schools.

13. CSE: A NEW CONCEPT OF EXAMINING

Since the Beloe Committee were so critical of all existing examinations, it is somewhat surprising to find them recommending the creation of yet another examination. However, their intention was that CSE should not become just another examination. The importance of CSE was that it was designed to be quite different from any other examination. In this, the Beloe Report represented an important shift in attitudes towards examining. Whereas up until the 1930's, the prevailing mood had been one of great optimism about the potential good to be derived from a reformed system of examining, in the 1930's this was replaced by cynicism. Educationists increasingly felt that their defects were not simply the products of particular examinations which could be overcome by judicious reform but were inherent in the very process of examining. This led to a reaction against examinations which culminated in the Norwood Report (1943).

However, by 1960 the mood had changed again and the Beloe Report displayed the traditional faith in the good to be derived from a revised system:

... it is our considered view that many of the defects and dangers we found in the existing examinations arise not from the intrinsic nature of examinations at this level but from the particular circumstances in which these examinations have grown up ... We are convinced that, if certain conditions are fulfilled, external examinations can make a constructive contribution to the educational process at this level. (pp.27-28)

In spite of their awareness of all the problems associated with examining, the Beloe Committee were confident that they could devise an examination which was capable of overcoming these obstacles.

The decision to introduce the examination was taken with a full awareness of the effects of external examinations - both good and bad; the hope was that it would be possible to develop an examining system that would encourage the good and avoid the bad.
The significance of the CSE examination lies not, therefore, in its presence as yet another examination which will aggravate the problem presented by the harmful effects of examinations, but as an attempt to solve it. (Whalley, 1969, p.7)

It is in this light that CSE should be judged and one of the main aims of this enquiry is to determine how successful CSE has been in achieving this goal in the field of English. It is a question which is especially pertinent to English since English, by its very nature, is at variance with the "one-shot" style of examining.

English is primarily concerned with qualities that are highly subjective. It aims to foster the development of personal maturity and it values the uniqueness of imaginative experience. Such emphases are incompatible with external examinations which depend for their efficacy on test items which can be marked with a high level of reliability i.e. items which are objective, factual, quantifiable. Thus, English provided a particular challenge to CSE to create a more valid examination and this study will attempt to evaluate how far CSE has succeeded in achieving this end.

One of the prime objectives of CSE was to overcome the domination of schools by examinations, to make examinations flexible and accommodating so that they might respond sensitively to needs in schools. Several novel features of the examination facilitated this development. Firstly, the fact that it was a new examination should not be overlooked. Moreover, it was not new in the way that GCE had been a new examination, growing out of its predecessor and heavily influenced by it. CSE was a completely new examination, unfettered by vested interests and established practices, and this meant that, "Every detail of its system had to be planned, and the planning was helped because again unlike GCE, the certificate would not be the main entry qualification to the universities and professional bodies" (Whalley, 1969, p.7). These were factors of great importance in providing CSE with the freedom to experiment and innovate.
The examination incorporated innovations which were designed to ensure that it reflected and complemented curricula rather than dictating them. Firstly, the decision to offer three alternative modes of examination provided the schools with considerable freedom in choosing their method of assessment:

The schools are thus offered considerable freedom of choice in deciding how best to adapt the examination to the needs of the pupils. It will be the responsibility of the teachers themselves to ensure that what is examined is what they want to teach; they will not be obliged to teach what someone else has decided to examine. (Examinations Bulletin 1, 1963, pp.4-5)

In addition to the flexibility afforded by three alternative modes, the level of teacher involvement envisaged was another crucial feature of the new examination: "Effective teacher control of syllabus content, examination papers and examining techniques is the rock on which the CSE system will stand" (Examinations Bulletin 1, 1963, p.3). Since the sheer size of a national examination tends to divorce it from the teachers who use it, CSE was made a regional examination in order that the practising teacher might make his voice heard:

The teachers in the schools using the examinations must have a major role in operating them ... This means that the Examining Bodies must be neither so large that their administrative offices become, geographically and in other ways, remote from the schools nor so small in number of candidates ... that they fall under the more or less exclusive influence of a single local education authority. (Beloe Report, 1960, pp.31-32)

All of this facilitated the experiments with novel techniques of examining which were strongly recommended by Examinations Bulletin 1 (1963) which attached, "... great importance to the development ... of new techniques of examining" (p.6) and urged, "... a fresh, indeed a courageously fresh, approach to techniques of examining" (p.26).

CSE represented the most far-reaching revolution in examining that this country has ever seen. However, more than a decade has passed since this great experiment was launched and as Dixon observed in 1977: "Unfortunately no authoritative survey has yet been attempted of the achievements and
lessons of this recent work" (p.8). The present enquiry is an attempt to evaluate some of these "achievements and lessons" in the field of English.
CHAPTER THREE

CONTINUOUS ASSESSMENT

A. INTRODUCTION

1. THE RATIONALE FOR THIS INVESTIGATION

Dixon (1977) evaluated the achievements of CSE thus:

Finally, I believe the main new contribution from the experiment of the last decade will be the sampling of course work. (p.8)

Until 1965, examining had evolved as an essentially external activity. However, CSE was designed to promote a new style of examining, one which would involve teachers more closely in its work. Thus, the SSEC (1963) urged the newly-formed CSE Boards to, "... move towards an integration of the roles of teacher and examiner" (p.26). This recommendation gave an immense boost to the concept of continuous assessment because no other testing procedure fuses teaching and assessment more completely than continuous assessment.

Despite the enthusiasm generated by continuous assessment, it remains a method which has been much prescribed but little described. Much of the work which has been published on this topic is the expression of purely personal opinions and there is little information about its nature and function which is based on the large-scale, systematic collection of data. It was inevitable that there should be a limited understanding of continuous assessment at the inception of CSE since this technique had not been used widely prior to 1965. Thus, NATE (1965) conceded that, "Naturally it will
be some years before all the possibilities and limitations of this method of assessment become clear" (p.27). Eggleston (1967) agreed: "Of the three classes of assessment procedure discussed here (i.e. conventional examinations, objective tests and continuous assessment), continuous assessment by teachers is least implemented and least well-known" (p.60).

These professions of uncertainty were to be expected in the 1960's but writers a decade later were still describing the same basic lack of information. Thus, Examinations Bulletin 31 (1975) reported that:

Published writings on the subject... reveal a wide diversity of opinion and relatively little factual information... Chapter 2 is a review of material published on the subject of continuous assessment. Necessarily, much of this is opinion, because facts are few and far between in this field. (pp. 11 and 17)

Moreover, even though Dixon (1977) asserted that, "... the main new contribution... of the last decade will be the sampling of course work", he conceded that: "Unfortunately no authoritative survey has yet been attempted of the achievements and lessons of this recent work" (p.8).

Even more disconcerting than the limited understanding of continuous assessment, or the fact that most of what has been written is based upon personal opinions, is the basic lack of agreement on this subject. For instance, there are conflicting opinions about the dual role of teacher-assessor which the teacher adopts when working on a continuously assessed course. One early commentator expressed doubts about the wisdom of this dual role:

Is it really desirable that teachers should be continually seen to be in an assessing role? Will not a fairly complex assessment procedure place an undue emphasis on what should be an incidental rather than a primary function of the teacher?

(Chapman, 1967, p.90)

These doubts were refuted by a later study which claimed that: "There was no evidence of the expected tension from the conflict of the dual role - teacher and assessor" (Eggleston and Kerr, 1969, p.190). However, an even later study reported that:
Besides affecting pupils' and teachers' work continuous assessment may also affect their relationship. There is strong support for the opinion that the teacher-pupil interaction may be drastically changed when the teacher adopts the additional role of assessor. (Examinations Bulletin 31, 1975, pp.32-33)

Another issue which illustrates the diversity of viewpoints surrounding continuous assessment is the question of its power to motivate pupils. Examinations Bulletin 31 (1975) found that, "Teachers committed to continuous assessment value the extra motivation it gives pupils particularly those of lower ability who do not demonstrate their true performance in examinations" (p.13). However, Spencer (1976) uncovered conflicting evidence on this topic:

Many administering CSE courses felt that internal assessment improves pupils' will to work, especially if a folio is prepared ... This claim has obvious face validity, yet some teachers claim it is not necessarily true. Pupils, they report, are not more highly motivated to complete pieces of work when they know they are for assessment purposes. (p.5)

While some teachers regarded continuous assessment as a powerful motivating force, others reported little or no impact or even a detrimental effect on the will to work.

These two examples demonstrate the range of issues raised by the use of continuous assessment which remain surrounded by uncertainty. Indeed, Examinations Bulletin 31 (1975) found it impossible to establish any common ground amongst the teachers included in its survey. Unable to make any definitive statements or to reach any firm conclusions, the authors attributed this lack of consensus to the newness of continuous assessment:

It is indicative of the growing pains inherent in the introduction of a new method of awarding grades that such diversity of opinion is to be found among just fourteen teachers. Differing shades of opinion make it difficult to summarize the many facets of continuous assessment seen by these teachers. (pp.66-67)

It is hoped that the present enquiry may make a useful contribution to our understanding of continuous assessment by providing some large-scale,
quantitative data in an area where most of the existing material takes
the form of small-scale surveys or is based on the personal views of
individuals. Furthermore, it is particularly useful to study this
procedure in an English context because continuous assessment was heralded
as an important breakthrough in the quest for more appropriate methods of
examining English. Therefore, there is a good case for a large-scale
investigation into this technique and especially of its application to
English.

The following sections will identify those deficiencies in existing
examinations which it was hoped that continuous assessment might overcome.
This will provide a basis for comparing the theory of continuous assessment
with its practice.*

2. THE DIVISIVENESS OF EXTERNAL EXAMINATIONS

A major criticism of existing examinations was that they separated
teaching and examining. Enlightened educationists came to regard it as a
fundamental anomaly that assessment should be an essentially external

---

* It is wrong to think of the term continuous assessment as describing a
single procedure since it is really an umbrella term which covers a
"family of assessment techniques" (Examinations Bulletin 31, 1975, p.29)
and may refer to different approaches in different contexts. As
Bulletin 31 (1975) pointed out: "Continuous assessment embraces a wide
range of procedures all of which have only one thing in common - they
occur over an extended period of time" (p.19). For this reason, it is
important to clarify the use of the term in this enquiry. It focusses on
a particular type of continuous assessment for which Bulletin 31 (1975)
coined the name "learning experience assessment":

Marks (or grades) may be awarded to work which arises directly
out of the process of learning and is the product of this
experience (learning experience assessment)... learning
experience assessment can only take place within a scheme of
continuous assessment... the course and the assessment are
closely related. (pp.11 and 12)

The present study focusses on this type of continuous assessment because
it is the form used in the East Midland Region's Mode 1 English
examination.
activity conducted by agencies who were often insensitive to needs and practices inside the schools. Thus, the axiom that examining should be an external exercise came under increasing fire:

... it is vitally important to stress the need to break down the barriers that have grown up between assessment as a classroom activity and assessment as an external exercise conducted by some outside agency. The separation is entirely artificial and has caused teachers to neglect assessment as an essential aspect of their professional work to which they should be prepared to devote more of their time and energy. The reasons for this divorce lie within the organisation and structure of our secondary school examining system, at least until the coming of CSE.

(McIntosh, 1972, p.14)

Educationists were seeking to effect a revolution in the relationship between teaching and examining so that instead of dictating what teachers taught, examinations should reflect what they wished to teach. Thus, Examinations Bulletin 1 (1963) advocated an "... examining system which will be the servant of the schools, not their master... the outstanding characteristic of the examinations for the Certificate of Secondary Education will be... that they will reflect and not inhibit the originality of the work being done in the schools" (pp.1 and 3).

In this light, it is not difficult to understand why continuous assessment seemed such an attractive procedure. By taking normal classroom activities and incorporating them into the final assessment, continuous assessment was capable of fusing the hitherto separate activities of teaching and examining. This potential strength of continuous assessment was especially pertinent to English where opinion leaders increasingly felt that their subject was completely at variance with conventional examining techniques: "... all past experience suggests the existence of a deep and inherent incompatibility between external examinations as we know them and the essential aims of good English teaching". (Whitehead, 1965, p.233). There was a growing feeling that there was little common
ground between English which fosters personal maturity and values the uniqueness of imaginative experience, and external examinations which are ill-equipped to test such subjective qualities. Thus, NATE (1965) argued that: "Written papers in their present form cannot adequately sample the full range of English work. Many of the valuable experiences of English lessons are inevitably lost sight of in a mass examination" (p.27). This disenchantment was reflected in a series of recommendations published by NATE (1964) for the guidance of the teachers who would form the regional English Panels for CSE:

(a) We believe that as far as English is concerned, there is a very special case for an internal examination, irrespective of arrangements made for other subjects

...  

(c) Coursework must be considered as a substantial part of the examination

(d) A written paper under formal conditions is inadequate for testing pupils' response to literature. Therefore assessment of this response should be a continuous process throughout the course (Bulletin 1, 1964, pp.4-5)

It was argued that some of the most pernicious backwash effects of external examinations in English e.g. fact-grubbing, cramming, question-spotting and the mechanical absorption of prepared answers would be eliminated by the use of continuous assessment.

... the promise of CSE was that the area of examinable English would be extended rather than contracted so that English examining might become coterminous with English itself, English as widely and liberally interpreted as the progressive teacher wished it to be... Course work constitutes one of the principal means whereby the area of examinable English may be extended. (Hipkin, 1967, p.25)

Clearly, the hope was that continuous assessment might be used to integrate teaching and examining and to develop them as mutually enhancing activities.
3. EXAMINATIONS AS SAMPLING PROCEDURES

A basic tenet of assessment theory is that the briefer the test and the smaller the number of items it comprises, the less valid and reliable that test is likely to be. The converse is, of course, that an extended test, composed of a greater number of items, is potentially more valid and reliable. Ebel (1972) demonstrated the relationship between test length and reliability in an article entitled, "Why is a Longer Test Usually a More Reliable Test?"

One of the best known properties of the tests commonly used in educational and psychological measurement... is that the longer they are (i.e., the larger the sample of tasks), the more reliable are the scores they yield. (p.249)

When measured against this criterion, continuous assessment appeared greatly superior to the conventional examination which, at best, involved three three-hour papers whereas innumerable separate tests could contribute to a continuously assessed course.

A brief review of the literature reveals that one of the principal reasons why continuous assessment was regarded as superior to external examining was the inadequacy of the "one-shot sudden-death" (Rowlands, 1974, p.104) sampling procedure:

The one-shot test is administratively convenient, but this convenience must be set against the potential hazards of teaching a subject to a pupil for five years and then ignoring all the accumulated evidence of his ability, basing decisions on his performance during a 2½ hour examination. One-shot tests suffer from three major sampling defects. The proportion of the total syllabus which can be examined in the time available is small; the range of skills appropriate for a particular subject may be represented neither comprehensively nor proportionately; possible variations of a candidate's performance in those skills which are measured, are not adequately sampled.

(Eggleston and Holford, 1971, pp.40-41)

Dunn (1974) described this difference as being analogous to that between a "snapshot" and a "movie" (p.44), continuous assessment providing a much
more detailed and comprehensive picture of ability than an external examination.

Once again, the argument was felt to have a special pertinence to English which encompasses a wide range of skills, many of which it is difficult to sample adequately in a conventional examination. The capacity of continuous assessment for sampling the skills more comprehensively and for incorporating normal classroom activities into the examining procedure was regarded as giving it a special validity in the assessment of English:

> English is an extremely complex subject involving the whole process of verbal communication and understanding. To attempt to assess it in three hours by means of three questions is unlikely to produce very reliable results. (Mather, France and Sare, 1965)

Surely the use of language, which is an everyday use if it is anything, can best be measured by having a good teacher observe his students uses of English everyday, continuously assessing them. (Bryan, 1973, p.54)

Likewise, Dixon (1965) argued that the first question which an English teacher should ask himself is: "... does the material assessed represent a fair sample of a sound English course?" (p.1) and he answered this question thus:

> ...In the traditional "0" level paper, the sample is limited to writing and, within writing, to the production of one or two pieces of original prose and the understanding of one passage of written English. Thus the assumption has to be made that this sample represents the candidate's abilities across the whole range of written purposes, audiences and contexts. It is a false assumption, as every teacher knows. The sampling of course work makes it possible in principle both to extend and - more important - to specify the range of writing to be sampled. (1977, p.5)

Clearly, important claims have been made for continuous assessment, especially as a more effective means of assessing English. It is now timely to evaluate some of these claims.
B. THE FINDINGS

1. INTRODUCTION

The above claims made it important for the questionnaire to pay special attention to certain questions:

- Was continuous assessment perceived as having satisfactorily bridged the gulf between the teaching and the testing of English?
- Had it succeeded in overcoming some of the worst defects of conventional examinations?
- Was it perceived as a fairer and more accurate method of examining English?
- Was it subject to any special drawbacks of its own?

Furthermore, since the present study was large-scale, it was felt that it could also make a valuable contribution by dealing with those issues over which educationists have been most sharply divided, e.g., its capacity to motivate pupils and the dual role of teacher-examiner.

Before proceeding with the data analysis, it is important to stress that this will not be treated as a contest between external examinations and continuous assessment with the aim of demonstrating the superiority of one procedure over the other. Such an approach would be inappropriate because each test will have its own particular aims and circumstances which different testing procedures will be better-equipped to meet. As 106 argued:

> It depends what you're testing. If you want to know can a person ... exhibit skills in a set time against the clock on a certain date, then an exam's best and if you want to find out what sort of person he is over an extended period of time in a variety of situations then continuous assessment is the answer. Which is better or fairer is almost irrelevant because one technique would be totally unsuitable for one set of criteria.

105 agreed: "I think the best is perhaps the one which works in one set of circumstances". Spencer (1976) reached a similar conclusion, arguing that: "Incontrovertible proof of the superiority of an internal or an external
system will never be available, because, in the end, value judgements about the purpose of the test are involved" (p.4). Therefore, the aim of this study is to provide information about the functioning of these procedures as a basis for teachers to decide which method is best suited to their particular needs on a particular occasion.

2. MARK WEIGHTINGS

Question 6 was designed to give a broad indication of attitudes towards this method relative to those towards external examining. It presented respondents with a list of differently constituted examinations ranging from:- 100% continuous assessment - 0% external examination  to 0% continuous assessment - 100% external examination and asked them to specify which alternative they preferred (i) for language and (ii) for literature. The resulting responses are recorded in Table 1 below.

**TABLE 1: PREFERRED ASSESSMENT FOR LANGUAGE AND FOR LITERATURE**

<table>
<thead>
<tr>
<th>CONSTITUTION OF EXAMINATION</th>
<th>PERCENTAGE FOR LANGUAGE</th>
<th>PERCENTAGE FOR LITERATURE</th>
</tr>
</thead>
<tbody>
<tr>
<td>% CA % EXAM</td>
<td>18</td>
<td>51</td>
</tr>
<tr>
<td>100 - 0</td>
<td>13</td>
<td>14</td>
</tr>
<tr>
<td>75 - 25</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>60 - 40</td>
<td>19</td>
<td>15</td>
</tr>
<tr>
<td>50 - 50</td>
<td>12</td>
<td>3</td>
</tr>
<tr>
<td>40 - 60</td>
<td>12</td>
<td>1</td>
</tr>
<tr>
<td>25 - 75</td>
<td>13</td>
<td>2</td>
</tr>
<tr>
<td>0 - 100</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>INVALID RESPONSE</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>
Perhaps the most striking feature to emerge was the difference in attitudes towards the examining of language and of literature. As far as the examining of literature was concerned, the respondents were heavily committed to continuous assessment with as many as half of the entire sample (51%) regarding it as the only satisfactory means of examination. It was interesting to find that such a large proportion of the sample held such strong and uncompromising views on the literature assessment. Even amongst those who favoured "mixed economy" examining (Dixon, 1977, p.6), there was a marked preference for the continuously assessed element to be most heavily weighted. Thus, a quarter of the sample (25%) favoured 60-75% continuous assessment compared with 4% who wanted 60-75% external examination. Overall, the responses demonstrated the overwhelming support for continuous assessment with over three-quarters of the sample (77%) wishing to use it as the principal or only method of evaluating literature. This provided a sharp contrast to the external examination which proved to be extremely unpopular. Only 2% would have liked 100% external examination for literature and a further 4% favoured it as the principal method of assessment (i.e. accounting for 60% or more of the marks).

There was a less clear-cut pattern of responses for language with no definite consensus emerging. Thus, although 43% of the respondents favoured 60% or more continuous assessment, almost as many again (37%) preferred 40% or less. In fact, the responses were fairly evenly dispersed across all seven alternatives. Another distinguishing feature of these responses was that the external examination was a much more popular mode of examining. 13% of the sample wished to rely exclusively on external examining for language and a further 24% wanted to use it as the principal mode of assessment (i.e. 60% or more). This meant that over a
third of the sample (37%) favoured the conventional examination as a major component whereas the figure for literature was only 6%. Despite this greater level of satisfaction with external examining in language, continuous assessment was still the more popular method.

Another interesting feature of the language figures was the popularity of "mixed economy" examining (Dixon, 1977, p.6). Over two-thirds of the sample (68%) indicated a preference for some form of mixed economy examining compared with less than half (45%) for the literature assessment. Although the East Midland Region has made little use of mixed economy examining, a wider use of this style of examining may prove not only to be popular but also to be profitable since there is sound theoretical justification for its use:

... it can be shown that the reliability and discrimination of writing and literature assessments improve when different single-marked assessments are added together. A combination of internal and external marks for these areas of work would in most cases be a better assessment than either alone, provided that both are validly assessing some aspect of English.

(Spencer, 1979, p.93)

The responses to Question 6 revealed a sharp division in attitudes towards the examining of language and literature. The sample clearly regarded continuous assessment as the most appropriate method of evaluating literature; indeed, half of them regarded it as the only satisfactory method. Although no definite consensus of opinion emerged on the testing of language, the external examination and mixed economy examining were regarded as more satisfactory in a language context than in a literary one.

Although Question 6 provided a general guide to attitudes towards continuous assessment, Question 15 was used to probe attitudes on specific aspects in greater detail. It dealt with elements of continuous assessment which had been identified as areas of particular concern or
uncertainty in the hope that a large-scale survey of more than 400
practising teachers might shed new light on these topics. Question 15
comprised a list of ten separate statements about continuous assessment
and respondents were asked to indicate their level of agreement with each
statement on a five-point scale ranging from "Strongly Agree" through
"Uncertain" to "Strongly Disagree". It was felt that it would be useful
to allow for varying levels of intensity in the responses (e.g. "Agree"
or "Strongly Agree") in order to gain some insight into the strength of
feeling provoked by particular issues.

3. **QUESTION 15(i): MOTIVATION**

The first item in Question 15 dealt with the contentious question of
the impact of continuous assessment on motivation. Since studies taking
a qualitative approach to this issue have produced a mass of conflicting
opinions, it was felt that it would be useful to obtain some large-scale
quantitative data on this question. Therefore, item (i) stated:

Continuous assessment is capable of motivating pupils to work
conscientiously throughout the course of study.

and respondents were asked to indicate their level of agreement/disagreement.*

The response was one of overwhelming agreement with 79% of the sample
concurring that continuous assessment is a potential source of motivation.
The word "potential" was used deliberately because item (i) was worded to
suggest that continuous assessment "... is capable of motivating" pupils
rather than a categorical statement that it axiomatically motivates pupils.
Indeed, some of the respondents stressed that the statement was conditional
rather than absolute and that they agreed with it in that sense. For
instance, Q395 added: "... 'is capable' yes, but does not necessarily do so".

* Strongly Agree (22%), Agree (57%), Uncertain (15%), Disagree (5%),
Strongly Disagree (1%).
The way in which the responses were dispersed between the "Agree" and "Strongly Agree" categories may shed further light on this qualification. A brief glance at the overall pattern of responses on Question 15 (see Table 2, overleaf) reveals that a comparatively small proportion of those who agreed with item (i) were in strong agreement. In fact, item (i) accounted for the highest proportion of responses clustering in the "Agree" category over the whole of Question 15 and this may have been due to the conditional rather than absolute nature of their agreement with this statement. Thus, it is probably true to say that those who felt that continuous assessment invariably motivates pupils were among the 22% who ticked the "Strongly Agree" category. It was also interesting to note the relatively high proportion of the sample (15%) who ticked the "Uncertain" category compared with the very small percentage (6%) who were in definite disagreement which reflected the degree of uncertainty surrounding this issue. Nevertheless, it appears to have been the experience of most respondents that this mode of assessment can provide an added incentive to work.

Their comments illustrate the general tenor of feelings on this issue:

... it helps a child to work better in that the whole year's work is taken seriously and not just the revision before the exam... motivation is the thing for this group... and I think that a lot of them show themselves very willing to work hard over a long period - particularly with reference to the folder. (I12)

It acts as a motivator for pupils to work throughout the entire course. (I11)

This evidence supports that of other studies. For instance, Robinson (1977) reported that the introduction of a continuously assessed scheme into his school had had, "... a very favourable effect on the morale and productivity of our fifth year" (p.50). Another teacher felt that the on-going nature of continuous assessment provided a boost to morale and confidence:
### Table 2: Attitudes Towards Continuous Assessment

(expresed as a percentage of the total response)

<table>
<thead>
<tr>
<th>ITEM</th>
<th>STRONGLY AGREE</th>
<th>AGREE</th>
<th>UNCERTAIN</th>
<th>DISAGREE</th>
<th>STRONGLY DISAGREE</th>
</tr>
</thead>
<tbody>
<tr>
<td>15(i) Continuous assessment is capable of motivating pupils to work conscientiously throughout the course.</td>
<td>22</td>
<td>57</td>
<td>15</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>15(ii) Continuous assessment gives a more rounded and accurate picture of an individual's real ability.</td>
<td>38</td>
<td>45</td>
<td>12</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>15(iii) Continuous assessment redresses the balance for those pupils who are prepared to work hard over an extended period as opposed to those who are capable of doing well by working hard just before an examination.</td>
<td>37</td>
<td>52</td>
<td>8.5</td>
<td>2</td>
<td>.5</td>
</tr>
<tr>
<td>5(iv) The fact that so many CSE courses involve an element of continuous assessment exerts considerable pressure on some pupils who feel the need to produce a large quantity of work of a consistently high standard.</td>
<td>12</td>
<td>44</td>
<td>24</td>
<td>19</td>
<td>1</td>
</tr>
<tr>
<td>15(v) The duties and responsibilities involved in continuous assessment are particularly demanding of a teacher's time and energy.</td>
<td>38</td>
<td>49</td>
<td>4</td>
<td>9</td>
<td>-</td>
</tr>
<tr>
<td>15(vi) One of the functions of a teacher is constantly to encourage pupils to give of their best. The need to make realistic assessments during the course conflicts, at some points, with the attempt to motivate them.</td>
<td>20</td>
<td>44</td>
<td>13</td>
<td>20</td>
<td>3</td>
</tr>
<tr>
<td>15(vii) Whenever an individual acts as both teacher and examiner, there is always the possibility that a variety of personal factors may influence the assessment to some extent.</td>
<td>11</td>
<td>56</td>
<td>12</td>
<td>18</td>
<td>3</td>
</tr>
</tbody>
</table>
TABLE 2 CONTINUED *

<table>
<thead>
<tr>
<th></th>
<th>STRONGLY AGREE</th>
<th>AGREE</th>
<th>UNCERTAIN</th>
<th>DISAGREE</th>
<th>STRONGLY DISAGREE</th>
</tr>
</thead>
<tbody>
<tr>
<td>15(viii) Continuous assessment of the study of literature avoids certain inherent tendencies of literature examinations: the learning of notes and regurgitation of second-hand opinions.</td>
<td>28</td>
<td>47</td>
<td>11</td>
<td>11</td>
<td>3</td>
</tr>
<tr>
<td>15(ix) Continuous assessment is particularly well-suited to the study of literature because it evaluates a pupil's response to a text when it is made rather than asking him to duplicate that response during an examination which may be quite remote from the time when it was studied.</td>
<td>39</td>
<td>43</td>
<td>10</td>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td>15(x) Large-scale continuous assessment, operated over an entire region, makes it difficult to be confident that the marks awarded really do reflect accurately the regional standard.</td>
<td>13</td>
<td>36</td>
<td>31</td>
<td>17</td>
<td>3</td>
</tr>
</tbody>
</table>

* For ease of reading, these responses are also reproduced individually in footnotes to the appropriate sections.
The kids feel that every single piece of work they do at any stage is equally important and, for those who haven't much ability, to feel that you've really done something today that has helped towards your exam result has been a motive for keeping at it'" (Examinations Bulletin 31, 1975, p.51). Teachers in Spencer's study (1979) also drew attention to the increased motivation which had resulted from the introduction of continuous assessment:

Increase amount of work done by poorer pupils because of greater motivation (Fife 4)

... I feel there has been a greater application to class work since the first assessment. Class has learnt to apply itself with greater concentration (Strathclyde 5) (p.87)

Cumulatively, this qualitative and quantitative data suggests that teachers see continuous assessment as a potent source of motivation. However, this raises certain questions: what is there about this method of examining which provides the incentive to work and why is it that a minority of teachers have found the general principle to be untrue in their own cases?

The on-going nature of continuous assessment appears to play an important part in its capacity for motivating candidates. The ways in which this can be utilized to increase motivation were described by Examinations Bulletin 31 (1975):

The beneficial effect of informing the pupil of his current progress is frequently mentioned... If a candidate's work falls behind that of the rest of the class a teacher can take effective action and since intermediate results affect the final assessment, the candidate himself is more easily persuaded to take action. This suggests that candidates who are assessed continuously should obtain better results than those of equal ability taking "one-shot" examinations. Both teacher and pupil have the information and incentive to improve their performances. (p.31)

The usefulness of an on-going assessment was also mentioned by respondents in the present enquiry. For instance, Q450 argued that continuous assessment is "... worthwhile because the progress achieved and the
intermediate milestones passed are a good guide to monitoring ability and making teaching relevant. Thus, it seems that continuous assessment is particularly effective in motivating pupils when teachers use it as a diagnostic tool and for monitoring progress:

The knowledge that he must conduct the entire course, from conception to final grade, makes the teacher aware of the inter-relationship of everything he does... He can use continuous assessment to motivate recalcitrant and conscientious students alike; he can use it to diagnose individual strengths and weaknesses and the success or failure of his teaching. He can use it, that is, as a tool to assist his teaching.

(Bryan, 1973, p. 54)

Thus, it seems possible that the success enjoyed by continuous assessment is partly determined by the way in which the teacher handles it. The teacher needs to be aware that continuous assessment is a mode of examining with special features which may be exploited to increase motivation. If this suggestion is correct and the teacher's approach is important, this may explain why some teachers find continuous assessment disappointing as a source of motivation. Some respondents attributed this to other factors in the examining environment stressing that motivation may be influenced by factors which are quite extraneous to the mode of assessment. For instance, Q309 argued that the character of a particular group of pupils may be an important variable: "The value of continuous assessment often depends on the personality of a particular group. Pupils' attitudes towards continuous assessment can negate the ideals behind it". Q289 stressed the importance of home environment: "... it often depends on home background whether continuous assessment is effective for pupils or not - favours stable, interested home". He concluded that, "Some respond, some don't - lots of variables". This view was perhaps best summarized by Q395:
Continuous assessment "is capable" yes, but does not necessarily do so (i.e. motivate pupils). A pupil's motivation may be based on some quite separate factor: the lack of motivation may have nothing to do with the system of assessment and may be based on factors which heavily outweigh the possible motivation which can be created by continuous assessment.

Finally, there was a small minority who held the view that an external examination is a more effective incentive to work, that continuous assessment quickly loses the impact which the imminence of a final examination never fails to provide. This view was typified by Il4: "In no way do I think that continuous assessment is motivational if it's completely continuous assessment apart from the very bright ones... I think children expect an examination. I think it's got currency".

Spencer (1979) reported similar views amongst some of the teachers in his survey:

> The impact of exams was completely lost and much of the written work demanded was not done... An exam (with the less motivated pupil) has an impact and means much more than any other viable alternative I have yet seen... Folder-work soon lost its importance for poorer/less willing pupils. (Fife 3) (p.90)

The quantitative data of item (i) suggests that such views are shared by only a minority of teachers. Indeed, the very high rate of agreement elicited by item (i) suggested that its capacity for increasing motivation was one of the most attractive features of this mode of examining.

4. **QUESTION 15(vi): THE DUAL ROLE OF TEACHER-ASSESOR**

The evidence of the preceding section suggests that the respondents saw continuous assessment as a potentially powerful source of motivation but since there are many factors which may undermine this, Question 15(vi) took one of the most controversial of these threats to increased motivation, the dual role of teacher-asseressor, and examined it in greater detail.
Some educationists believe that to fulfil the roles of teacher and public assessor* simultaneously is a contradiction in terms, that this joint role is a paradoxical one. They argue that the aim of examining is to provide as precise and as objective an evaluation of a pupil's attainment as it is possible to give while the teacher's primary goal is to motivate him. Continuous assessment, by requiring the teacher to perform both roles simultaneously, may create tension. This view was typified by Chapman (1967):

Is it really desirable that teachers should be continually seen to be in an assessing role? Will not a fairly complex assessment procedure place an undue emphasis on what should be an incidental rather than a primary function of the teacher? (p.90)

Spencer's study (1979) illustrated the type of difficulty which could arise as a result of this: "... 'the advantages in motivating some were offset by the tantrums of others who for the first time actually expressed their frustration and sense of inadequacy with continuing failure'" (p.88).

Since most of the existing material on this topic consists of personal opinions, it was decided that it would be useful to gather some large-scale quantitative data. Therefore, item (vi) stated:

One of the functions of a teacher is constantly to encourage pupils to give of their best. The need to make realistic assessments of their performance during the course conflicts, at some points, with the attempt to motivate them.

Respondents were asked to indicate their level of agreement/disagreement with this statement.** One fifth of the sample (20%) opted for the

---

* It is important to distinguish between the role of the teacher as a "private" assessor i.e. the kind of evaluation which he carries out every day as an integral part of his teaching and as a "public" assessor marking for examination purposes. It is this second role which some educationists regard as problematic.

** Strongly Agree (20%), Agree (44%), Uncertain (13%), Disagree (20%), Strongly Disagree (3%).
"Strongly Agree" category suggesting that a sizeable proportion of the respondents felt strongly about the tension in this role. A further 44% agreed which meant that overall more than three-fifths of the sample (64%) felt that this dual role created conflict between teaching and assessing at some points. Ill argued the case thus:

One of one's functions as a teacher of English is to constantly encourage pupils to give of their best and to make invidious contrasts between them is bound to conflict with the attempt to encourage them. This is an unavoidable contradiction in continuous assessment.

IO2 agreed: "... in English, the function of examining and the function of teaching are not necessarily the same... The role of the examination is a clinical one... The purpose is to say what is good, bad or indifferent and that's not the function of teaching. The function of teaching is to enable children to do as well as they can".

In contrast with the 64% who agreed with item (vi), almost a quarter of the sample (23%) disagreed. They gave various reasons for refuting the statement. For instance, Il4 felt that children welcomed an honest assessment of their achievements and that this did not interfere with motivation:

I don't think its contradictory at all. I think teachers are constantly having to make assessments... the teacher is an assessor by definition... children want to be guided by knowing how well they're doing. I don't think the kids see any conflict in your role and I think that's what matters.

IO5 argued that the two roles were quite separate and that this eliminated the possibility of conflict: "I have no difficulty whatsoever. It is my job to teach the subject as well as I can and then I think of myself in a totally different role when I mark their work". Some teachers argued that the obligation to give an objective assessment could be useful:
Sometimes it is useful to have to make comparisons between kids because you can get so wrapped up in encouraging groups of kids who are less able that you lose perspective on the relative quality of their work... I can think of cases where I have over-estimated how well kids are doing until I've started to compare them with others. (I13)

It also emerged that the way in which the teacher handles the whole process of marking and handing back work may be a critical factor in determining whether this conflict arises. For instance, I13 explained that:

The comment I make on their piece of work is about the relative value of their work, relative to what they have done and what they are likely to do and because this is separate from the mark on the mark sheet, that is why there isn't this contradiction.

Like I13, many other teachers who regarded the teacher-assessor role as a potential source of conflict had adopted strategies to cope with this difficulty. Some gave a "public" mark for the candidate's purposes which referred to the "relative" value of the work: relative, that is, to what had been achieved in the past whereas the "private" mark which went into the mark book was a more realistic assessment of its worth in relation to the wider standards of the CSE examination. Others, like I13, gave no "public" marks and relied on written comments to guide and motivate pupils. Others still, seeing no conflict between the two roles, used a single mark for both public and private purposes arguing that pupils preferred an honest assessment of their work relative to that of their peers and that other means could be used to motivate them if the actual mark was disappointing:

I find that the average fourteen or fifteen-year-old, if you level with him or her, will take it. I'm sure they're aware that one day the crunch will come and you've got to assess them. Personally I've never adjusted marks falsely to encourage people. I hope that my relationship with them, and that my other attempts to urge them on, will not rely solely on the mark at the bottom of the page. (I06)
In spite of the strategies adopted by some teachers, the responses to item (vi) made it clear that the majority found the dual role created conflict at some points. However, it is important to put this item into perspective by comparing the pattern of responses with that on item (i). Considerably more respondents both agreed and strongly agreed with item (i) than with item (vi) (i.e. 22% and 57% compared with 20% and 44%) suggesting that although most teachers perceived a certain amount of tension in this dual role, they did not regard this as a major obstacle to the power of continuous assessment to motivate pupils. Overall, the responses seemed to suggest that continuous assessment is a potent motivating force but that the need for teachers to adopt the dual role can create tension at some points.

5. **INCREASED PROFESSIONAL DIALOGUE**

It is often complained that teaching is a profession which is practised in isolation, generating little opportunity for teachers to come together to exchange ideas or to tackle common problems collectively. Therefore, one consequence of the teacher-examiner role which did receive warm praise was that it had promoted contact between schools and individual subject specialists.

Teachers who carry out assessments for CSE need to meet regularly and this provides the opportunity for professional dialogue. New ideas are developed, innovations are disseminated more quickly and teachers can collaborate over common difficulties. In this way, CSE has become a seedbed for experimentation and progress and some teachers attached great importance to this:
... it has brought teachers together to exchange ideas and information far more than ever before in the history of education. In my view, these meetings of teachers have done more to encourage teachers to utilize other methods in their teaching... than any course which has ever been organised by the DES or any other body. It has made a tremendous amount of difference to in-service training because it is in-service training without a course leader... I think if I were asked what is the most important aspect of CSE, I would say teacher involvement in the examination process... CSE brought more people into the examination sphere and it is not just a matter of teachers going round and examining someone else's pupils. This getting together periodically, and exchanging ideas, has done the education service a power of good. (I08)

I see my commitment to EMREB as a vital part of my teaching because it makes me a better teacher... By coming into contact with perhaps some of the best English teachers in the area, and many of the people at Nottingham are very high quality people indeed, it has helped my teaching tremendously. (I05)

I'm Chairman of the local English Panel... and this has been ideal for finding out what's gone on in other schools. I think that the amount of dialogue that has gone on in CSE can only contribute to a more interesting syllabus... I think it's made us concerned about what we expect from children... It's given us the chance to discuss what children should be producing at sixteen and we've been able to put these things into practice. (I14)

Other studies agree that when teachers become involved in assessment, professional exchange is increased. For instance, Burchnell (1968) described the introduction of a continuously assessed course into his school thus:

The extra work involved was considerable but not intolerable ... one of the merits of the scheme lies in the opportunity afforded to the schools involved to share their ideas, discuss their problems and interests, and strengthen the links which naturally exist between enthusiastic and responsible staffs. (p.29)

Examinations Bulletin 31 (1975) agreed that: "It has benefits in that teachers see more of each other's schools and can compare standards, understand each other's problems and spread good ideas more quickly. All these should lead to a raising of standards and the adoption of better curricula" (pp.51 and 82). Likewise, one of their subjects reported that:
"Getting to know other teachers and co-operating with them has come out as an additional advantage" (p.13).

However, it was felt that these glowing tributes did not necessarily reflect the general level of benefit derived from involvement in CSE simply because they were made by teachers who were heavily involved in the examination. That such benefits do accrue when the teacher is involved to that extent is beyond dispute. However, it was decided that the present study should try to gauge how widespread these advantages were felt to be amongst this random sample of practising teachers - not just those who had involved themselves in CSE panels and committees where they were likely to meet like-minded, enlightened individuals. Therefore, Question 16(xii) stated:

The meetings, panels and committees etc. involved in the running of the CSE examination have brought additional benefits for practising teachers, in that they provide the opportunity for the sharing of ideas and the dissemination of expertise.

Respondents were asked to indicate their level of agreement with this statement.* The resulting pattern of responses was rather different from what might have been expected on the basis of the remarks made by those heavily involved in such activities. Only a small proportion of the sample (7%) were in strong agreement with this statement although almost half (48%) agreed. This meant that altogether just over half of the sample (55%) believed that they had derived the kind of additional benefits described in item (xii) from their involvement in CSE. Neither the overall level of agreement nor the amount of strong agreement were high. Nevertheless, more than half of the teachers thought that CSE had had useful "spin-offs" on their teaching and this is a consequence of CSE

* Strongly Agree (7%), Agree (48%), Uncertain (33%), Disagree (11%), Strongly Disagree (1%).
involvement which should not be undervalued. CSE has provided the
opportunity for many more teachers to become involved in examining their
subject than happens in conventional examinations and this has implications
for the quality of their work overall.

Although only a small proportion of the sample (12%) disagreed with
this statement, a strikingly high percentage (33%) held no definite opinion
either way. This may have been because they were not sufficiently involved
in the running of the examination to comment. This suggestion is supported
by I06's observation about the way in which teacher participation operates
in practice. I06 believed that there was a widespread lack of involvement
in Mode I and, although he regarded this as a problem, he still felt that
teacher involvement was a worthwhile feature of CSE:

... the teacher involvement is a key feature which is very good
although I've got reservations about its operation. One is that
teachers aren't as involved as they ought to be. That's their
own fault, I think, but a lot of teachers don't involve
themselves in CSE work and then complain about the kind of
decisions that are made as a result of the people who do...
That's a difficulty but at least more teachers are involved
than are at O-Level so by and large the teacher involvement is
a good thing.

When asked if he thought that CSE had brought the average, practising
teacher the kind of "cross-fertilisation" of ideas described in Question
16(xii), he replied:

Yes, to a limited extent but not the "average, practising
English teacher" because the average, practising English
teacher hasn't been at the meetings... It's true that I've met,
over the past fifteen years, a lot of English teachers and I've
learnt a lot from them but I'm not average in the sense that all
English teachers are not involved to the extent that I am.

He illustrated the point by referring to a local panel meeting which we
had both attended: "... when I went to a meeting of the Leicestershire
North Panel, I already knew several of them by their christian names
although I'd never been there before but those people are always at the meetings I attend".* 105 had similar reservations: "I suppose that even with teachers doing this, we have created an élite. That's alright so long as we in the élite don't start to pontificate and tell teachers what to do... I think that to make this kind of examination work properly, you've got to make everyone who is involved feel part of it".

This problem appears to be widespread. For instance, the Secretary of the Southern Board observed that the opportunity to involve themselves in examining:

... is not one which has been universally welcomed by the teaching profession and even among those who have welcomed it there has been a reluctance on occasions to accept its full implications. We thus have an ironic and potentially disturbing situation. There is an assessment pattern which is becoming increasingly complex in terms of the techniques being used and a teaching profession, some of whose members appear unwilling to involve themselves directly in the assessment procedure... It is only through the informed involvement of the teaching profession as a whole that teacher control of public examinations can be achieved. (MacIntosh, 1972, pp.14 and 16)

Clearly, CSE has provided both the opportunity and the framework for teachers to become more involved in the examining of their subject but there is still a need for greater involvement on the part of all teachers. However, this does not necessarily imply criticism of those teachers who have not participated fully in CSE.

6. OBSTACLES TO FULL INVOLVEMENT

Teaching is a demanding profession and many of the respondents felt that English, with its heavy marking load and the constant need to prepare fresh lesson materials relevant to particular groups of pupils, was

*I06 was referring specifically to Mode I. Some interviewees suggested that Mode 3 encouraged a greater degree of involvement. Indeed, those who wish to be closely involved in the examination will probably opt for Mode 3 anyway.
especially so. Thus, some of them regarded examination commitments as an additional burden in an already onerous profession. Some teachers were deeply concerned that these extra responsibilities received little or no acknowledgement in their teaching timetables, that insufficient time was available to perform out-of-school duties as conscientiously as they wished and that financial remuneration was not provided for the internal responsibilities that they were obliged to undertake. A selection of their comments will illustrate the feelings on these issues.

As far as the students themselves are concerned, I feel that 100% assessment of coursework is a much fairer and more accurate method... However, the burdens and strains it places upon the teacher are very great and this should be recognised by giving teachers marking, preparation and moderation time during the school day for subjects which are continuously assessed - especially English. (Q 285)

The volume of work and the responsibility that goes with it (are) frequently not recognised by Management in the distribution of salary points. The CSE tends to overload English departments particularly. (Q 532)

Coursework... adds to the English teacher's workload. The average, conscientious English teacher works very hard without this added labour... Exam boards create hostility among Heads of English... (by) increasing the administrative paper work. (Q 007)

The demands of teaching English are always high... The planning of work is always great because each group has different demands and the marking of English essays is always demanding, in part because of the sheer volume. The assessment of folders is a difficult task, particularly time-consuming. (Q 380)

The degree of effort is not related to a criticism of the examination structure but to the loss or shortage of marking periods in school. (Q 345)

The main problem here is the staggeringly low proportion of preparation time allowed by the average English secondary school timetable. (Q 496)

The mass of form-filling between the end of February and the end of April! The internal moderation of Literature folders - a huge task - and vital to the integrity of the whole examination - totally unpaid and vastly time-consuming. (Q 386)

The volume of work required for internal moderation in a large comprehensive is most onerous at Easter time. Close scrutiny by head of Department is essential for satisfactory moderation -
most time-consuming and no extra fee for this in Mode I - under
Mode 3 the rewards (financial) are considerable. This discrepancy
seems unjust. (Q 470)

As the Head of Department, I am responsible for internal
moderation of: (1) Mode I Oral (3 weeks); (2) Mode I Literature
folios (100-120 candidates); (3) Mode 3 course work... During
the busy Easter Term I have to be taken off some classes -
usually sixth form to give oral exams and to entertain visiting
moderators. The clerical work involved for two Mode I courses
and a Mode 3 is very demanding and too complicated to delegate.
(Q 310)

When considering these difficulties, it is important to bear in mind
the original ideal: "Effective teacher control of syllabus content,
examination papers and examining techniques is the rock on which the CSE
system will stand" (Examinations Bulletin 1, 1963, p.3). If the examination
is to benefit from greater teacher involvement and if more teachers are
to enjoy the advantages which this study suggests accrue from full
involvement, serious account must be taken of the extra demands a teacher
faces when he takes on the additional role of assessor. As Hipkin (1971)
asserted:

... internal examinations should not be permitted to occupy too
much of a teacher's limited time and energy... Where teachers
are expected to undertake additional duties they should be
adequately recompensed and assisted. (p.23)

7. QUESTION 15(ii): INTERNAL VALIDITY

At the inception of CSE, it was claimed that continuous assessment
could give a more valid measure of attainment than the one-shot test. It
was felt that the present study should ascertain whether the teachers'
experiences of this procedure endorsed or refuted this claim. Did teachers
believe continuous assessment to be a fairer and more accurate method of
examining and, if so, why? Which of its features did they perceive as
contributing to its heightened level of validity? There were several items
in Question 15 which dealt with this aspect of continuous assessment; the first was item (ii) which was designed to investigate the internal validity* of continuous assessment:

Continuous assessment gives a more rounded and accurate picture of an individual's real ability.

Respondents were asked to indicate their level of agreement/disagreement with this statement.**

The responses suggested that this early claim has been vindicated. More than a third of the responses (38%) were concentrated in the "Strongly Agree" category. Only one item in Question 15 yielded a higher level of strong agreement than this. A further 45% agreed which meant that the overwhelming majority (83%) regarded continuous assessment as a particularly satisfactory measure of attainment. This high level of consensus was matched by an almost negligible rate of disagreement (5%). Clearly, its ability to improve the validity of assessment was perceived as one of continuous assessment's most attractive features. However, item (ii) gave no indication of why respondents believed this to be so.

The qualitative evidence of the written and spoken comments suggested that there were two principal reasons why respondents regarded continuous assessment as having a good level of internal validity. Firstly, it was regarded as a more satisfactory sampling procedure than the one-off examination which is subject to various sampling errors (see pp.100-101). The prudence of taking a wider sample of performances was one of the most

---

* The concept of internal validity is based on the professional judgement of experts in the field in question: "'Internal' validity is concerned with the content and structure of the test. Do these appear to be reasonable and satisfactory? We 'size up' a test, noting its characteristics, and form a judgement concerning its validity" (Hitchman, 1966, p.62)

** Strongly Agree (38%), Agree (45%), Uncertain (12%), Disagree (4%), Strongly Disagree (1%).
frequent arguments forwarded in favour of continuous assessment:

... it tends to give you a wider picture. You're not relying on a one-off composition or a one-off comprehension... I think it's very unfair to assess someone on the basis of a one-off examination. (II4)

I think continuous assessment is the best method of obtaining an accurate picture of an individual's real ability. (Q 231)

I can't really see the point of final, external examinations. I think that in Assessment Theory terms they're very, very poor samples of kids' overall performance or ability. (II3)

As far as the students themselves are concerned, I feel that 100% assessment of coursework is a much fairer and more accurate method than external examination. (Q 285)

It is a more valid and accurate assessment of the real quality of a candidate's work. I am completely in favour of it. (Q 395)

The present study is not alone in finding that continuous assessment is regarded as a superior measure of attainment. The moderators surveyed in Examinations Bulletin 31 (1975) concurred that it was, "... the fairest method available" (p.82).

While some teachers emphasised the accuracy of continuous assessment, others were more concerned by the inadequacies of conventional tests. For instance, when asked whether he thought that a final examination could fulfil a function which other types of assessment could not perform, IOI answered thus:

Yes, it can make people write worse! I think all the effects tend to be negative. Under pressure of time they write less accurately, less vigorously, less creatively because you can't require it at the switch of a button... I think that the great achievement of CSE was to make the normal procedures of class teaching accessible to assessment... in so far as CSE strays from that, it strays from one of its most important functions.

Some teachers were especially concerned about the implications of this for areas of English such as response to literature* and creative composition:

* This topic is dealt with separately in the literature chapter.
What's worried me over the years is the quality of composition work in a one-off exam... with no opportunity for candidates to work on stimuli... it's surprising how often class teachers are disappointed in the performance of many good candidates. ... As far as I'm concerned, the whole purpose of continuous assessment is to make absolutely certain that we are able to obtain the candidate's best work. I've long been worried about the simple composition exam... that's what we'd like to see evidence of over the course of five terms in a folio of work.

A similar point was made by II0 in describing the value of the Creative Writing Folio which was offered as an alternative to part of the Mode I Language Paper for several years:

The Creative Writing Folio was an addition which was of value, believing, as I do, that the idea of creative composition being something which you do on a Tuesday morning at 9.30 a.m. when you are pushed into an examination situation and told, "Create!" is the kind of thing which no professional author would attempt to do and expect to get the best work. The idea of instant creativity just because an examination timetable demands it, I don't think is necessarily on. It therefore gave the opportunity for children to do work rather in the same way that the literature folio gives them the opportunity to do work: which is over a period of time; which is assessed over a period of time and eliminates some of the problems of the on-the-spot examination.

Other studies have also uncovered acute concern over the invalidity of performances given in "one-shot sudden-death" tests (Rowlands, 1974, p.104). Snowdon (1965) reported that:

Final examinations often give what appear to be unreliable grades. A recent Cambridge report expressed concern that some students, known to have done good work for three years, should unexpectedly obtain a final examination result much below what has appeared to be their real quality. (p.39)

Continuous assessment, he argued, would eliminate "... 'accidental' failures at the end of the course" (p.39). Likewise, Examinations Bulletin 31 (1975) found that: "When the results of end-of-term examinations arrive there are often surprises: one pupil has done much better than was thought
possible; another has disappointed" (p.19).

Clearly, the internal validity of continuous assessment is attributed primarily to its capacity for taking a wide range of soundings over an extended period of time and for incorporating work done under normal classroom conditions into the assessment thus avoiding the danger of uncharacteristic examination performances.

8. QUESTION 15(iii): A FAIRER METHOD OF ASSESSMENT?

Another of the principal arguments in favour of continuous assessment at the outset was that it provided a more just method of assessment. Whereas the external examination tends to favour quick wits and verbal agility, it was argued that continuous assessment would allow children who lacked these qualities to show their true ability. Teachers knew that children who learnt quickly often wasted much of the course and still did well by making a concentrated effort during the revision period. Whereas these candidates benefitted from the "one-shot" style of examining, it was seen as detrimental to other mental types:

... anything which actually builds into the eighteen months from the beginning of the fourth year onwards is well worth doing. This is the argument we used in the early days and pointed out that for CSE-type candidates we did not want everything to depend on ready, quick wits and quick verbal fluency... This is so important so that children are not tempted to waste the fourth year... you're measuring will to work and you're measuring determination. This is why course work has done so much... (I03)

Was this still regarded as an important attribute of continuous assessment? Had the importance attached to this at the outset been justified by experience? In order to answer this question, item (iii) stated:
Continuous assessment redresses the balance for those pupils who are prepared to work hard over an extended period of time, as opposed to those pupils who are capable of doing well by working hard just before an examination.

Respondents were asked to indicate their level of agreement/disagreement with this statement.* Of all the issues pertaining to continuous assessment raised by the questionnaire, this was the one which elicited the greatest level of agreement (89%). Moreover, it was something about which many of them felt particularly strongly with more than a third (37%) indicating strong agreement. In contrast, only 2.5% of the sample disagreed with this statement. If the level of agreement elicited by an item is taken as the criterion of its success, then the extremely high rating enjoyed by item (iii) suggests that continuous assessment's capacity for rewarding conscientiousness has proved to be one of its most attractive features.

The comments which were volunteered convey something of the value which this was perceived to have. For instance, 101 argued:

We are certainly conscious of having a substantial minority of students who have qualities of persistence and conscientiousness that are better rewarded through CSE than through an on-the-day, limited examination... of the two examinations, there is no doubt in my mind that continuous assessment is the superior method of examining.

113 was of a similar opinion: "There was a feeling that (continuous assessment) allowed students a better deal for a variety of reasons such as kids who clam up in an exam, kids who earnestly endeavour over a period of two years rather than kids who have a trick of memorisation and maybe work for a month...". These convictions are endorsed by other studies. For instance, Examinations Bulletin 31 (1975) found that: "Teachers have commented on the high standard, especially at the pass level. The scheme

* Strongly Agree (37%), Agree (52%), Uncertain (8.5%), Disagree (2%), Strongly Disagree (.5%).
seems to encourage the average or weak candidate to persevere throughout the year". (p.40)

Before leaving this item, it is important to clarify an issue raised by one or two respondents who objected to what they regarded as a value judgement implicit in item (iii). For instance, Q395 argued that, "15(iii) implies some sort of merit in terms of examination successes. The pupil who gains a Grade 1 by hard, concentrated effort just before the exam is probably going to be able and willing to apply that hard, concentrated effort as and when necessary in their working life." This comment raises the important question of whether an examination is designed primarily to predict future achievement or to describe attainment. It is clear that Q395 was referring to the predictive function of examining whereas CSE is primarily a descriptive examination relating to attainment. Clearly, most teachers felt that continuous assessment provides a fairer description of attainment than one-shot tests can give.

9. A THREAT TO THE VALIDITY OF CONTINUOUS ASSESSMENT

Learning experience assessment (see p.97) enables elements of the learning process to be assimilated into the assessment which are normally inaccessible to examination and this has been seen as one of the great strengths of this method. For instance, Examinations Bulletin 31 (1975) stressed that continuous assessment "... allows non-cognitive factors to be taken into account in the assessment" (p.13). Similarly, Wiseman pointed out that "...'you have got to have a certain amount of knowledge and in order to test that you do not need continuous assessment. When you come to the development of aesthetic attitudes and the development of skills, then continuous assessment comes into its own'" (quoted by Examinations Bulletin 31, 1975, p.36). This capacity for tapping skills
and experiences which are normally inaccessible to examination has been widely regarded as one of continuous assessment's great strengths. However, it has encouraged teachers to include in their mark schemes criteria which have no place in a CSE examination. For instance, although in a purely internal test a teacher may consider it useful to measure the development of attitudes, these may not be suitable for inclusion in a public test of attainment. As IO8 observed:

I've seen cases where marks have been awarded for personal characteristics such as perseverance, endeavour, effort and so on and then these have been blindly added to raw marks for a written paper and these have been blindly added to raw marks based upon a subjective judgement of a teacher. All the constituents of this mathematical pudding have been materials which don't produce a homogeneous mixture at the finish... It isn't attainment anyway. Somebody may try extremely hard in a certain subject and still produce hopeless results and in such cases how can you use as one of your constituent marks in a score for attainment a mark for effort?

Ironically, one of the great strengths of continuous assessment is also a potential weakness. Because it facilitates a more comprehensive, and thereby more valid, assessment, it also facilitates the inclusion of inappropriate criteria which may undermine its increased validity.

This section and the preceding one may appear to contradict each other in that the preceding section emphasised the value of an assessment procedure which rewards candidates who are prepared to work hard throughout the course while the present section is critical of those schemes which give marks for qualities such as perseverance. However, there is an important distinction between favouring schemes which allow these qualities to play a part in demonstrating attainment and between introducing them directly into the assessment as measurable components of the mark scheme. In other words, while it is quite acceptable to use an assessment procedure which allows the candidate's industry to be reflected in his attainment, it is a different matter to introduce this
quality directly into the examination by making it a constituent of the mark scheme.

10. **QUESTION 15(iv): PUPILS' WORKLOAD**

Since continuous assessment was perceived as rewarding the conscientious worker, this raised the question of the workload characteristic of this mode of examining. If continuous assessment does favour the industrious candidate, does it follow that it is a more demanding form of examination requiring a greater amount of effort than the conventional examination? There were two items in Question 15 which dealt with this issue - one from the pupils' point of view (item iv) and the other from the teachers' point of view (item v). Since the preceding section was concerned with the pupils' conscientiousness, it seems logical to deal with the question of workload from the pupils' point of view first.

Various studies have explored this issue but they are not particularly relevant to the present enquiry because they deal with continuous assessment in higher education institutions. It was felt that the issue could be made more relevant to the situation in schools by taking account of the fact that school children study a variety of subjects, several of which may involve an element of continuous assessment. Therefore, Question 15(iv) stated:

> The fact that so many CSE courses involve an element of continuous assessment exerts considerable pressure on some pupils, in that they feel the need to produce a large quantity of work of a consistently high level of achievement.

Respondents were invited to indicate their level of agreement/disagreement with this statement:

* Strongly Agree (12%), Agree (44%), Uncertain (24%), Disagree (19%), Strongly Disagree (1%).
Over half of the sample (56%) were in agreement, 12% strongly so. Those who commented on this issue identified a number of factors which may contribute to this pressure. For instance, Q162 drew attention to the influence of home environment arguing that, "Particularly conscientious, middle-class pupils of moderate ability, subject to parental pressure" were especially susceptible. However, the difficulty which is likely to affect the greatest number of candidates was raised by I14 who upbraided teachers for failing to co-ordinate their course work requirements or to distribute them evenly across courses. The crisis point, he argued, often came towards the end of the final year:

... this is a very difficult area because some of them have four or five subjects with continuous assessment and if the teachers have not been assiduous in spreading it out over the period of the course, then some of them are put under a great deal of pressure to get their coursework ready for April or May in the final year.

At the root of this problem lies the tendency of teachers to operate as isolated subject specialists in the provision of examinations, taking little account of what other subjects will be demanding. Attempts to formulate "across subjects" and/or "within school" policies covering such matters as the size of course work requirements, their co-ordination and fair distribution rarely occur. The result is that more than half of those sampled felt that the overall demand which individual candidates may face "... exerts considerable pressure on some pupils".

Nevertheless, it has to be acknowledged that in comparison with the level of agreement elicited by some items in Question 15, a relatively small proportion of the respondents were in agreement with item (iv). Only 56% agreed making item (iv) the second lowest scoring item in terms of overall agreement. Moreover, only 12% felt strongly enough to tick the "Strongly Agree" category making it the second lowest scoring item in
terms of strong agreement too. However, these low scores may, in themselves, illustrate the problem. This suggestion is supported by the unusually high rate of uncertainty. Almost a quarter of the sample (24%) had no definite opinion on this issue supporting the suggestion that many subject specialists work with only a limited understanding of the kind of overall pressures faced by candidates. Thus, the comparatively low rate of agreement and the unusually high rate of uncertainty were probably a simple reflection of the fact that the people whose problem this was were not consulted and the teachers, who were consulted, had only a limited awareness of the problem. Moreover, they could only know the problem indirectly and at second-hand if they had any insight at all. A different picture might emerge if this issue was studied from the standpoint of the candidates. In spite of this shortcoming, the available evidence suggested that a problem of some urgency for teacher-examiners is the need to collaborate in the provision of continuously assessed courses and to develop a greater sensitivity towards what is being required of candidates in other subjects.

It also has to be recognised that continuous assessment is an inherently more demanding form of examination involving pupils in a more protracted and detailed assessment than a conventional examination entails. The sheer volume of assessable material which candidates are normally required to produce is the key factor; as Ill observed: "O-Level takes two and a half hours whereas I dread to think of the number of hours which goes into the testing in CSE". This demand "... for constant visible results" (Snowdon, 1965, p.40) is the principal reason why continuous assessment is such a demanding method of examination. As Examinations Bulletin 31 (1975) observed:
There is no doubt that continuous assessment makes considerably greater demands on the time and resources of all concerned. Teachers operating a continuously assessed scheme and their students... commented upon the increased load which has to be borne. (pp.13 and 32)

This raises questions about the type of pressure associated with this method of assessment. The available evidence suggests that a particular type of psychological strain is induced by continuous assessment.

Commenting on her experiences, Cowell (1972) observed that:

... To me, comparing assessment with examinations is like comparing months of nagging toothache with the short, sharp pain of having a tooth removed. (p.15)

Snowdon (1965) agreed that, "So far from liberating the student, continuous assessment, which replaces one large leap by many small leaps, imposes upon him stultifying demands for constant, visible results" (p.40). A similar view was taken by Coleman (1971):

When continuous assessment was first taken seriously, its followers insisted that fewer students would break down; without exams to worry about, students would be free to work. Unfortunately, what seems to happen is that the stresses are just spread out over the academic year; more, not fewer, students suffer. (p.7)

These studies refer to the use of continuous assessment in higher education institutions but one study, which focussed on its use in schools, agreed that it exerts a particular type of pressure: "...'a good candidate may be adversely affected by the effort required to maintain a consistently high standard over the year". (J.M.B., quoted in Examinations Bulletin 31, 1975, p.40).

When the idea of continuous assessment was first mooted, its exponents argued that it would eliminate the kind of tension induced by the one-shot examination. However, the evidence suggests that continuous assessment, with its demand for regular results, creates its own psychological strain. It is especially onerous for certain personality types who cope better
with one large hurdle at the end of a course rather than numerous smaller
hurdles throughout.

Although continuous assessment is undoubtedly a more demanding style
of assessment, the evidence of the present study suggests that at the root
of this difficulty lies another problem. It seems that the real crux of
the matter lies in the magnitude of the requirements.

11. THE SIZE OF COURSE WORK REQUIREMENTS

In 1965, Dixon claimed that: "The first problem is how big a sample
of each candidate's work to assess?" recognising this as a question which
needed to be solved at the outset. Likewise, NATE (1964), in attempting
to identify areas where difficulties might emerge in the new examination,
enquired:

Is there any danger that if course work is required the shadow
of the examiner will loom over both teacher and pupil throughout
two years? What safeguards can be taken against this?

(Prestwich, 1964, p.15)

This enquiry suggests that there is a danger that the shadow of the
examiner will loom over both teacher and pupil throughout the course and
perhaps the best safeguard against this is to pitch the course work
requirements at an appropriate level.

In many cases, the problem of pitching course work requirements at
a satisfactory level has never been resolved. For instance, to fulfil
the requirements of the Mode 3 syllabus in Il3's school "... a CSE kid
needs at minimum 60 pieces of work... and that means a tremendous amount
of work. Very few kids had done the minimum by the end of the course".
Likewise, Q178 found that: "The folder work required... makes excessive
demands on candidates' time.* Clearly, there has been a tendency to overload continuously assessed syllabuses. Thus, the excessive workload associated with continuous assessment is not simply the result of the overall burden placed on candidates by a range of courses - it starts at the microcosmic level of individual subjects. Evidence from other studies suggests that this problem is widespread. For instance, one of the subjects in Examinations Bulletin 31 (1975) reported that:

I have to produce approximately twelve pieces of assessable written work per term... In fact, I haven't succeeded in doing that this year... It's just too much. (p.52)

The frequency with which such complaints are met suggests that this tendency to overload the syllabus is characteristic of the continuously assessed course. It is understandable that mistakes should have been made at the outset when experience of continuous assessment was limited but it is interesting to find that this problem has persisted especially since continuous assessment, more than any other form of examining, is controlled by teachers. Moreover, this problem would not be difficult to rectify for as Q198 pointed out: "Quality not quantity is the thing to be assessed and this can be accurately assessed on a smaller sample". It would be relatively easy to reduce course work requirements to a more manageable level and this raises the question of why this problem has been allowed to persist when, from a technical point of view, it is

* It would be possible to quote many more instances of this problem. However, much of the detail on this issue is to be found in the literature chapter. (The Mode I Literature award is based entirely on a folio of written work produced during the course). Since this question of pitching course work requirements at an appropriate level is pertinent to the Mode I literature folio, the issue of workload is dealt with in more detail in the literature chapter. In order not to duplicate information which is reported there, this chapter attempts to deal with the more general aspects of this topic.
easy to correct?

Part of the answer to this enigma probably lies in this method's unjustified reputation as a "soft option":

Moderators are worried that their subject, particularly when continuously assessed, may be seen as a soft option... The status of continuous assessment was alluded to on a number of occasions... Some feel that it is regarded by many both within and outside the profession as a soft option which gives the lazy and incompetent teacher opportunity to get by without the discipline of measuring up to acceptable external criteria... it should be made clear that continuous assessment makes considerably more demands on pupils, teachers and moderators. (Examinations Bulletin 31, 1975, pp.81-82)

It seems possible that in reacting to this criticism, teachers have in fact over-reacted making continuous assessment so rigorous that this accusation is palpably untrue. Although this is a natural reaction, from an educational point of view it has had some unfortunate repercussions.

The aim of continuous assessment is to provide a more accurate measurement of a candidate's true level of attainment than a one-shot test can elicit. Although there is a relationship between quantity and quality with a larger quantity of assessable material providing a truer reflection of attainment, this relationship only holds true up to a certain point after which an excessive preoccupation with quantity can impair the measurement of quality (see pp.191-193 for a more detailed explanation). In other words, there is a delicate balance to be struck if increased quantity is to mean increased accuracy in the measurement of quality. Many continuously assessed courses display an obsession with quantity which has undermined the relationship between it and quality.

12. QUESTION 15(v): TEACHERS' WORKLOAD

Since continuous assessment had been found to be an especially
onious form of assessment for the pupil, it was wondered whether it was also demanding for teachers? Therefore, item (v) stated:

The duties and responsibilities involved in continuous assessment are particularly demanding of a teacher's time and energy.

Respondents were asked to indicate their level of agreement/disagreement with this statement.*

There was almost total consensus with 87% of the sample agreeing with the statement. This meant that item (v) had yielded a greater level of agreement than any other item in Question 15. An equally striking feature of this item was the strength of feeling it aroused with over a third of the sample (38%) opting for the "Strongly Agree" category. Only one item elicited a higher level of strong agreement than this. Clearly, this was an element of continuous assessment about which there was both firm agreement and strong feeling. This item also yielded the lowest level of uncertainty (4%) which is perhaps best explained by Examinations Bulletin 31's observation that: "Teachers are heavily involved in all aspects of CSE examining... But it is their involvement in continuous assessment (which) is most intensive and extensive" (p.49). The respondents held definite opinions on this aspect of continuous assessment because it affected them personally. This contrasted with the responses on the preceding item (i.e. pupils' workload) where there was a comparatively low rate of agreement and a very high rate of uncertainty. Naturally, teachers were less sure about something which did not influence them directly. Likewise, few of them volunteered opinions about the effects of heavy coursework requirements

* Strongly Agree (38%), Agree (49%), Uncertain (4%), Disagree (9%), Strongly Disagree (-).
on pupils whereas many of them commented on their own workload.

Indeed, if the amount of comment prompted by an issue may be taken as a rough guide to the level of concern it aroused, it is clear that the extra workload placed upon teachers by continuous assessment was very near the top of their list of concerns.

Although there was almost universal agreement on this, respondents were divided into those who argued that the additional workload was acceptable because it was offset by the advantages of continuous assessment and those who found the demands excessive. If the attitudes of the former group are analysed first, it emerges that there were two principal reasons why they felt that the advantages of continuous assessment outweighed the extra workload. The first was the beneficial effect it was perceived to have on pupils:

The only disadvantage is that it is demanding on the teacher but it's not a demand which I think is unreasonable or that they shouldn't be quite happy to take on... I think that if you feel that this is worthwhile and good for the pupils, then you just do it. (II2)

There's a tremendous marking load... and it is a drain on the energies of everybody... But we choose to do continuous assessment because we think it gives the kids a better deal. It's a choice we make and we know that it means that we have to work harder. (II3)

It does involve extra work but I believe in what our course and assessment are trying to do and so am prepared to put in that extra effort. (Q187)

Organisation of pupils' folders etc. does involve additional work... but it is not irrelevant to the success of the course so an extra effort seems worthwhile. (Q438)

The other reason which was mentioned frequently as making continuous assessment worthwhile, despite the extra workload, was that it enhanced the teacher's sense of professional worth:

Continuous assessment means very hard work; it also means a regular system of moderation... The benefits of such moderation are great; aims, approaches and syllabus content are continually refined. (Q068)
103 felt the same about the final moderation meeting for Mode 3 assessors:

"I find it a very hard day's work, a most enjoyable day's work because I know it's doing me a lot of good as a teacher." 113 agreed:

I think it's a good thing for people to be forced to consider not their standards as standards but their reasons for their criteria and how they apply their criteria because I think that has an effect on how they teach and not just on how they examine.

103 summed up these feelings when he claimed:

The thing that really teaches teachers how to teach well is to get them involved in examining and moderation. There is nothing which clarifies the mind as wonderfully as that.

Other studies endorse this finding. For instance, Spencer (1976) reported that:

The teachers... unanimously declared that... continuous assessment has improved teaching. Much more professional interest is taken by more teachers in planning courses and making them and the methods of assessment relevant to the pupils. As a result, greater confidence in their professional worth is engendered... The value to the curriculum of teachers being responsible for their own assessment was a constantly recurring theme. (p.5)

This finding, that teachers attribute feelings of increased professional confidence to their work as assessors, is important in connection with the suggestion that there is conflict inherent in the dual role of teacher-assessor (see pp.112-116). Against the potential tension inherent in this role must be weighed the fact that many teachers perceived their work as assessors as enhancing the quality of their teaching.

13. THE IMPORTANCE OF EXTRANEOUS FACTORS

Distinct from those who found the additional burden acceptable were those for whom the workload was so heavy that it was creating severe difficulties. When the comments of these teachers were scrutinised, it
emerged that the key to their problems was often to be found in factors which were, strictly speaking, extraneous to the method of assessment. However, when these factors were compounded with this demanding assessment procedure, the burden often became intolerable. For instance, a teacher's status within his school could be a critical variable determining whether the teacher found continuous assessment difficult to cope with simply because it is the policy in some schools to allocate most of the CSE classes to junior members of staff while the more senior teachers take the O- and A-Level work. Thus, the demands of continuous assessment are magnified when a teacher is responsible for a high proportion of CSE classes:

The amount of time necessary for CSE classes varies in proportion to experience. Younger teachers tend to be allocated more CSE classes, while the more experienced tend to be allocated O- and A-Level forms. (Q130)

The demands are obviously heavier on young teachers with a large CSE commitment and who may have 60+ candidates in any year. (Q286)

School size emerged as another important variable but for a different group of teachers. Instead of affecting the least experienced, school size tended to affect the teachers in 'senior positions who were expected to carry out internal moderation and the administration and paperwork which CSE entails. The larger the school, the more onerous these tasks were perceived to be:

Heads of department responsible for internal moderation and establishing order of merit have very heavy workload in schools with 100+ candidates. (Q212)

When large numbers are involved (i.e. 100+ candidates), workload becomes much, much more difficult. (Q201)

A member of my staff is in charge of CSE while I am primarily responsible for O- and A-Level work... I feel our Mode 3 syllabus, with nearly 350 pupils, is far too onerous for the teacher in charge. (Q162)
School type also emerged as an important factor. Indeed, Leicestershire Upper Schools provided a good illustration of the account which needs to be taken of extraneous factors. Circumstances inside these schools were such that some teachers found it immensely difficult to cope with continuous assessment. Firstly, in an upper school, virtually every class is being prepared for an examination which means that very little of the work is free from examination pressures. Secondly, English is a compulsory examination subject in most schools which means that English classes tend to be larger than optional subject classes. Furthermore, English is, by its very nature, a "bulky" subject in terms of the marking load it produces. Leicestershire Upper Schools also have a history of strong commitment to CSE, many of them entering all of their pupils for CSE English, including the most able who are capable of producing reams of written work. In addition, these schools tend to favour the CSE style of examining and so opt for continuously assessed Mode 3 syllabuses at O- and A-Level as well. The cumulative effect of these factors was that some teachers felt themselves to be seriously overworked as assessors. There was a feeling that the balance between teaching and assessing had been upset and some teachers were bitter about being forced to wade through masses of assessment at the expense of developing their teaching techniques:

... the balance is wrong. The freedom it (continuous assessment) should give you is eliminated because the energy and enthusiasm which should go into evolving a way of teaching and preparing particular texts for presentation is sucked out of you by the 90-95% of assessment and administration... of the mountains of written work which they're forced to produce. I think on two counts it's maximized because you're in an upper school which multiplies and intensifies the amount of marking for a start because you come into contact with far more examination classes... You also have to recognise that it's a subject like English which actually produces such a mass of extensive marking. (I16)

... the syllabus does allow flexibility and that's marvellous... but at the same time we have to get through so much written work
that limits us because we spend far too much time marking and we don't have enough time for preparing materials. (I17)

Although continuous assessment is best for the students, it is hell for the teachers because there is so much marking to do as well as grading meetings etc. I know of schools where the idealism has worn off in the welter of work. (Q192)

In an ideal world, with unlimited time, I think that 100% coursework is more interesting to teach and better for pupils. However, to reduce stress and workload on individual teachers, some externally assessed work seems to be necessary. (Q288)

... it's necessary to distinguish between the burden on the pupil and on the teacher. The latter often is so great that the "satisfactory" nature of the assessment suffers... Feeding work into the folder "god" and fattening up classes to produce "acceptable", stereotyped pieces for the folder is an endless and often unrewarding process. The focus turns to ASSESSING and ranking and away from meaningful teaching. Ironically, the teaching of straight "O" may give teachers more time and energy to teach more interestingly than with CSE continuous assessment which is so demanding on time... The bitterness over the work created is aggravated by the sheer volume many (indeed most) of the students produce when cut loose under the flexibility of Mode 3 (especially as over 40% are double entered for 0-Level Language too). (Q289)

The situation in some Leicestershire Upper Schools is a prime example of how extraneous factors may undermine the efficacy of a continuously assessed course. When the use of continuous assessment is being considered, external factors must be taken into account and a sensitive adjustment of assessment to circumstances made if the course is to be a success.

14. QUESTION 15(x): COMPARABILITY

On the whole, continuous assessment had been judged to be highly successful when measured against the validity criterion. The other key criterion of an assessment technique is reliability and, on this count, continuous assessment appears to have been less successful. Admittedly, a longer test is, in theory, a more reliable test (see Ebel, 1972) but in the practical setting there are other factors which may influence reliability:
In test theory, you make a test more reliable by making it longer... Assessment of say, a year's work, coupled with the teacher's knowledge of his pupils should therefore be more accurate... Despite these claims, it has to be recognised that there are other factors affecting reliability, and it cannot be too readily assumed that a continuous assessment procedure will produce more reliable results. (Spencer, 1976, pp.4-5)

As Spencer pointed out, the accuracy of an assessment procedure does not depend solely on taking an adequate sample of performances. Prominent among the threats to reliability is the difficulty of achieving comparable standards in any mass examination (see pp.164-165). Since course work assessment is class-based, it was wondered whether this made it particularly difficult to establish a regional/national standard and to ensure that it is applied at the microcosmic level of individual classrooms? Therefore, item 15(x) stated:

Large-scale continuous assessment, operated over an entire region, makes it difficult to be confident that the marks awarded really do reflect accurately the regional standard.

Respondents were asked to indicate their level of agreement/disagreement with this statement.*

The resulting pattern of responses suggested that comparability may be the main weakness in continuous assessment. Whereas a strong, positive attitude had emerged on most other items, the prevailing attitude here was either negative or uncertain. To take a positive attitude on this item, the respondents would have had to disagree with the statement. Only a fifth (20%) actually did disagree thereby indicating confidence in the large-scale comparability of continuous assessment whereas half (49%) agreed, indicating a lack of confidence. In addition to those who indicated real doubts, an unusually high proportion (31%) were uncertain. This meant that altogether 80% of the sample expressed either a definite lack

* Strongly Agree (13%), Agree (36%), Uncertain (31%), Disagree (17%), Strongly Disagree (3%)
of confidence or uncertainty about the comparability of continuous assessment.

Other studies have also reported deep concern amongst teachers about the influence of large-scale continuous assessment on comparability. For instance, Lee (1976) found that:

... the issue of internal assessment was the one which caused the greatest concern when we analysed responses to the Report... Problems which centred around the validation and moderation of internal assessment procedures within the CSE examination... have been a source of some contention since the early days of the examination and have represented a particularly sensitive issue for the many teachers whose commitment has been rooted in the traditionally oriented form of the GCE examinations. (p.17)

Examinations Bulletin 31 (1975) confirmed this fear: "... teachers and moderators feel that the difficulty of consistently adhering to a standard is the greatest problem" (p.12). Likewise, Hipkin (1971) asserted that:

Many teachers, even those well-disposed to the principle of internal assessment, worry that the proliferation of idiosyncratic, school-based examinations would make it impossible to maintain comparability of standards between schools. Any new proposal must seek to answer this objection. (p.23)

This problem is not confined to establishing consistent marking standards but is complicated by variations in the work tackled by different schools, the conditions under which the work is produced and the different levels of difficulty of different types of work. There is evidence to suggest that these issues remain unresolved. For instance, Spencer (1976) observed that, "... continuous assessment is constantly beset with difficulties arising from lack of comparability among the conditions under which work is done in different schools and among pieces of work of different levels of difficulty" (p.5). Likewise, Dixon (1977) urged that, "... with new approaches to assessment, including coursework, an explicit code is essential... practice varies a great deal between and within Boards and I think a national code for English needs to be worked out". (p.4)
Although continuous assessment makes comparability difficult to achieve, it would be a mistake to assume that an external examination axiomatically possesses more exact standards. It has already been noted that continuous assessment has been hampered throughout its history by its popular image as a soft option. There is an equally erroneous assumption that external examining axiomatically possesses stricter standards, possibly because of the greater degree of central control that it entails. This view was typified by IO2:

I do think continuous assessment has a role to play... (but) I'd certainly keep the final examination as the major factor... with continuous assessment you do your best... but I don't think that in a pure sense it can be as good or as reliable as a properly conducted, properly standardised examination procedure. Although it's nice to do continuous assessment, there are so many factors that influence it and I'm talking in a purist sense now...

This view was widely shared. For instance, IO8 argued that an external examination should form the major component of any assessment because, "... that is the only time when all the children are faced with the same work. It's a sort of moderating instrument". Likewise, IO4 claimed that:

... a final examination is a desirable element of any examination... I should certainly want a substantial proportion of the marks, no less than 50%, to be dependent on this final external examination... I believe that the genuine external element in exams introduces an objectivity into what would otherwise be simply a subjective attempt at appraisal.

This prejudice does not, however, stand up to close scrutiny:

... the measuring process is not something wholly accurate. Whatever an examination tries to measure, it cannot measure it with the accuracy with which we can measure the temperature of water in a jar or the amount of sugar in a solution. No examining board denies that there is a margin of error which is inseparable from the examining process, so that the results are an approximation, however close, rather than something precise. Examining is not an exact science... level of performance, or of attainment displayed, in work submitted on a number of occasions throughout the course is every bit as weighty a piece of evidence as an examination result... There are those who will grant that such methods have great virtues... but will say that
the price paid for this is that the grades thus obtained are not reliable or uniform, unlike the grades derived from a wholly external examination, which are. It is of the utmost importance to emphasize that this distinction is false on both sides... by various methods of moderation they produce a final assessment which is as reliable, valid and uniform as that produced by any existing external examination conducted on a large scale.
(Jennings, 1971, pp.6, 7 and 9)*

Admittedly, one of the principal duties of a public examination is to maintain consistent standards and failure to do this may be a weakness in continuous assessment. Nevertheless, it would be erroneous to assume that the alternative, external examining, is not subject to similar weaknesses. This is not to acquiesce in these difficulties but merely to suggest that it would be misguided to dismiss continuous assessment on the grounds of poor comparability when examining is such an inexact science anyway.

At the root of the comparability problem lies the dilemma created by teachers acting as public examiners. The problems which can arise when outsiders conduct school examinations were described by Chapters 1 and 2. However, simply because the teacher-examiner innovation has brought many benefits, it should not be forgotten that it creates its own special difficulties.

15. MODERATION

As Section 7 showed, running a public examination makes additional demands on a teacher's limited time, energy and expertise. These pressures are most acute in out-of-school duties like moderation, especially when these must be performed during the school day.

* See also the evidence of incomparability in external examinations provided by Chapter 4, pp.164-165.
The only thing that's always worried me about it (i.e. continuous assessment) is the business of moderation. Nobody has ever come up with a perfect moderating system simply because you can't afford the time these people need to spend in examination centres... You've got to be deeply immersed in all aspects of it before you can successfully moderate otherwise you're going to make a superficial judgement. This is the only worrying aspect of coursework assessment. (103)

Examinations Bulletin 31 (1975) agreed that the time factor is critical:

The problem lies in assessing in isolation in a school... Both teachers and moderators see a need for... more time to be committed to the moderation of continuous assessment so that there is a better understanding between teacher and moderator... The issue uppermost in the moderators' minds is the sheer physical difficulty of making the number of contacts they feel they should be making with the schools. It is made clear in their replies to the questionnaire, time and time again, that even with the present relatively light use of continuous assessment, they are not able to see their schools as often as they would like in order to satisfy themselves that standards are comparable. Some moderators are adamant that the level of contact is inadequate. (pp.13 and 80)

The dilemma is clear: if the teacher is released from school for more time to carry out his moderating duties more thoroughly, his school and his pupils suffer. The alternative of professional moderators arouses intense hostility because teachers are extremely reluctant to externalize such an important part of the examining process. Moreover, many teachers regard examining as excellent in-service training which they would be loathe to relinquish.

Another difficulty lies in the specialized examining skills required. Several studies have suggested that areas of technical competence such as moderation are ones in which the teacher's expertise is limited. For instance, Examinations Bulletin 31 (1975) reported that:

Both teachers and moderators see a need for training in assessment techniques... A number feel they require training in basic techniques... It was surprising how few of those interviewed had been on any sort of course of training or could recommend a particularly useful book on assessment techniques. (p.65)
In connection with this, it is interesting to note the unusually high proportion of the sample (31%) who did not hold a definite opinion on item 15(x). This item yielded a higher level of uncertainty than any other item in Question 15 and this may have been because this was an area of technical competence in which many teachers felt ill-equipped to comment. Thus, although the teacher-examiner role has many strengths, it has also created tensions which are difficult to resolve.

The difficulty of establishing comparable standards is not the only factor which may influence the reliability of continuous assessment.

16. THE INTRUSION OF PERSONAL FACTORS

While a close relationship between teaching and examining has been shown to have some merits, attention has also been drawn to the possible drawbacks of involving the teacher so intimately in the final assessment. One fear, expressed by some critics, is that continuous assessment is open to abuse or, at the very least, to the unwitting intrusion of personal feelings such as favouritism. For instance, Snowdon (1965) argued that continuous assessment, "... opens the door to 'non-objective' criteria in evaluating students' work" (p.40). Here again was an issue on which it was felt to be important to determine the views of practising teachers. Therefore, item 15(vii) stated:

> Whenever an individual acts as both teacher and examiner, there is always the possibility that a variety of personal factors may influence the assessment to some extent.*

Two-thirds of the respondents (67%) agreed with this statement while only a fifth (21%) disagreed but although most teachers agreed that personal

* Strongly Agree (11%), Agree (56%), Uncertain (12%), Disagree (18%), Strongly Disagree (3%)
factors do sometimes influence assessment, this did not appear to have
aroused either strong feeling or particularly widespread concern. In fact,
item (vii) yielded the lowest level of strong agreement across the whole
of Question 15 suggesting that this difficulty had not provoked a strong
reaction even though most teachers conceded that it existed. Moreover,
it did not command the high level of consensus which had emerged on certain
items. Of the ten items comprising Question 15, only three had elicited
a lower level of consensus. Thus, it seemed that although most teachers
accepted the accuracy of this statement, it was not something about which
there was intense feeling or especially wide concern.

Reliability was a subject of deep concern for some respondents and
so it was somewhat surprising to find that this threat to it had not
provoked a stronger reaction. Perhaps the key to this was provided by those
teachers who drew attention to the role of moderation:

In some ways I don't feel that is entirely my job though I think
I have got a responsibility to the kid I might be prejudiced
against and to the kid I might be prejudiced towards because of
the kind of feedback I give them in the two years. But as far
as getting the grade right at the end is concerned, I think that's
the moderator's business not mine. (113)

Likewise, when asked whether he thought the personal relationship between
assessor and assessed might distort the evaluation, 107 replied: "I'm
sure it does and the answer is that it's removed by moderation". Q633
agreed: "All assessment at this level is subject to error. That's what
internal moderation is for!" Clearly, some teachers did not regard the
influence of personal factors as a serious threat to reliability because
they perceived moderation as a kind of "catch-all" process which would
eliminate inconsistencies in the end.

The moderating system used in the Mode I English examination is
limited in its power to correct inconsistencies because it accepts the
inalienable right of the schools to establish the rank order of candidates.
This means that although an overall bias towards severity or leniency in
marking may be corrected, if the moderator regards the placing of certain
candidates as unsatisfactory, the school is not obliged to alter its rank
order. The assumption is made that the internal assessor knows the worth
of individual candidates relative to one another better than an external
moderator can know them. Thus, although this type of moderation may bring
individual markers into line with a regional standard, it is not empowered
to alter the actual sequencing of candidates. Although this principle is,
on the whole, a sound one it does not answer the objection that the
teacher may be biased in favour of certain candidates and prejudiced
against others who may be misplaced as a result of this. When asked
whether he had been satisfied with this type of moderation, 101 replied:

Not always but I think it was a correct principle so I was
satisfied with it as a principle but I wasn't always satisfied
with the way it was carried out - but that's the human factor,
 isn't it?

Those who expressed concern about the intrusion of personal factors
were particularly worried about the "halo" effect of a candidate's personal
qualities and about teachers developing fixed notions about certain
candidates. For instance, 108 observed that:

There is always a personal relationship between the teacher and
the child and that personal relationship can colour the
teacher's evaluation of the child's work. All teachers have said,
"My word! This piece of work from X is far better than I ever
anticipated it would be," and when a good child does a bad piece
of work say, "Well, this is a long way below his normal standard".
You tend to lean more towards the child who is good and give him
the benefit of the doubt and to penalize the child who has been
a nuisance in class etc. There is a very large subjective element
which creeps into continuous assessment... There is always this
subjective relationship between teacher and taught and it can
modify a teacher's judgement one way or the other. It is very,
very difficult to be completely impartial when you not only
teach but also examine your own children.

IO9 agreed:

If I, at the end of a two year course, am ranking the literature
folios, I might give one child a mark of 22 and another 20 out
of 30. I've made a decision that one child is better than
another but that decision is based on a purely subjective
assessment by me. There's nothing objective... It's my knowledge
of the kid which may be influenced by all kinds of things, not
just what's in that folio e.g. whether they're nice kids who've
worked hard. Now I would set out not to do this but I'm going
to do this because I'm a teacher and the teacher as an examiner
is a very tricky role and one I don't think we yet understand.
This doesn't mean that I think all written examinations are
reliable either particularly in English where, on the whole, the
examiner is going through a very subjective exercise.

Likewise Il4 confessed that: "... it is difficult to put to one side one's
own prejudices." Thus, it is important to recognise that a moderating
system which is designed to correct marking bias but ignores incorrect
placement cannot meet this drawback to continuous assessment.

Other studies agree that school-based assessments may be particularly
susceptible to the intrusion of personal factors. For instance, Examinations
Bulletin 1 (1963) warned that, "With the best will in the world, some
teachers and some schools are bound to develop fixed ideas about particular
pupils, who may as a result be over- or under-estimated" (p.25). Spencer
(1976) also claimed that:

The individuality of the marker may have a greater effect on
the grade a pupil scores than the type of work set. The
advantage of the teacher's knowing the pupils well may be offset
by day to day fluctuations in mood or in opinions... or he may
be influenced by the "halo" effect of a pupil's personal
qualities. (p.5)

Although the respondents conceded that continuous assessment is
susceptible to this type of influence, some argued that external examining
is not necessarily superior in this respect. For instance, Il2 agreed
that an assessor's judgement may be distorted by the halo effect of a
candidate's personal qualities but added: "It's the same sort of risk, anyway, as that faced by an examiner who's marking a script that's incredibly neatly written. You've still got to make the same sort of sensitive judgements about how you're arriving at your assessment... It requires sensitivity". Likewise, I09 felt that, "... the teacher as an examiner is a very tricky role... This doesn't mean that I think written examinations are reliable either — particularly in English". I06 claimed that he would be happy to use one hundred per cent continuous assessment, "... but if people are going to worry about whether I'm swayed by personal factors then I suppose we've got to have written examinations but I don't think they are all that reliable".

In addition to this kind of unwitting abuse of the system, there is also the question of blatant misconduct.

17. POTENTIAL FOR ABUSE

The belief that continuous assessment is, by its very nature, more open to abuse than the external examination is widely held. For instance, I04 argued that it is, "... open, rather too easily, to abuse. Any satisfactory examination system has got to monitor that possibility". I10 agreed: "Continuous assessment, in itself, is more open to error and abuse" while Q201 claimed that:

Moderation is the X factor in the CSE examination procedure. It is frighteningly capable of being abused and provides the area of uncertainty in assessment which, in turn, can lead to a lack of respect for the examination.

Clearly, the personal involvement of assessor with assessed undermines the public credibility of continuous assessment in the eyes of some teachers. Examinations Bulletin 31 (1975) supports this finding:
Continuous assessment generates vigorous comment both for and against it. On the one hand, it is seen as the most valid form of assessment... On the other, it is seen as cumbersome and open to abuse by staff and pupils... Feelings range from overriding enthusiasm... to deep suspicion. (pp.38 and 85)

A subject in Spencer's study (1979) illustrated this suspicion: "I have a dislike of any scheme of internal assessment, as opposed to an outside exam, since it involves too much of the personal... Examiners are faceless" (p.91). However, just as the belief that continuous assessment is a "soft option" is unjustified, so too the argument that external examinations are less open to abuse is easily exploded.

The type of abuse which the external examination facilitates was neatly summarized by Dixon (1977) when he observed that, "... outside the exam room no holds are barred" (p.4). This comment aptly describes the "us and them" situation a teacher confronts when he prepares his pupils for an external examination. He is free to employ any tactics, no matter how dubious they may be on educational grounds, in order to secure a pass for his candidates. He may provide his pupils with model answers to be absorbed mechanically and regurgitated in the examination and he may spot likely questions so that his pupils may revise these areas most thoroughly. In this sense, the conventional examination is more conducive to blatant misconduct than continuous assessment because it sets the teacher against the examiner encouraging him to pit his wiles against those of the examiner. Continuous assessment, on the other hand, places the teacher in a position where such conduct becomes unacceptable. The "us and them" situation in which the teacher unites with his class against the examiner is eliminated because the teacher is the examiner. This point was underlined by IO3 when he compared his work as an O-Level marker with his work as an assessor for EMRED's literature course:
The big difference that we found was that as O-Level markers we were used to marking the best considered and best remembered thoughts of the teacher. What we were marking in the folios was the individuality of the child having touched upon the theme... so we were looking at far more originality... in O-Level you still had to give the same marks even if you could hear the teacher's words coming through in every essay that the child wrote.

Spencer's enquiry (1976) supports this view that continuous assessment is not more susceptible to abuse:

> It is often argued that all internal assessments are susceptible to cheating and that parents might place teachers under pressure... Every person I consulted in England considered these arguments to be of no importance at all. (p.7)

This is not to suggest that all teachers are beyond foul play but provided that moderation is carried out conscientiously, malpractice invariably will be detected.

18. CONCLUDING REMARKS

It is clear that continuous assessment has raised important new problems which remain unresolved and there is a need for further work in this area. Nevertheless, the general tenor of opinion amongst the respondents endorsed Dixon's (1977) view that continuous assessment is one of the most important innovations to emerge from CSE. IO8's belief that, "... some account must be taken of the work which the child has done over the period of the course. The days of the set examinations on a particular day are very much out-of-date" is certainly supported by the evidence of this chapter.

One of continuous assessment's most novel and important applications is considered in the following chapter.
CHAPTER FOUR

LITERATURE

A. THE FORMATIVE YEARS: LITERATURE EXAMINATIONS PRIOR TO CSE

1. INTRODUCTION

The formative years of CSE coincided with a period in which
dissatisfaction with the GCE O-Level was mounting. Nowhere was this
dissatisfaction more acute than in English:

Criticisms of the General Certificate tests in English have been mounting for years among a vocal minority of teachers of
English. A glance at back numbers of the quarterly "The Use of English", for example, will reveal scores of hostile comments.
(NATE, 1966, p.1)

This was especially true of literature where the O-Level examination was
widely regarded as, "... a complete antithesis to enlightened teaching of
the subject" (Birmingham NATE, 1966, p.2). This was a critical factor in
the development of CSE literature which gave it a "Janus-like" quality:
on the one hand, looking back at established conventions and reacting
against them and, on the other, looking forward to a fresh start, an
opportunity to devise new techniques for examining literature: "... simply
because it is new, the CSE offers the opportunity for major change"
(Dixon, 1965, p.1). As Whalley (1969) pointed out:

The significance of the CSE examination lies not in its presence
as yet another examination which will aggravate the problem
presented by the harmful effects of examinations, but as an
attempt to solve it. (p.7)

This was particularly pertinent to literature which was fraught with
assessment problems which existing examinations had failed to solve.
Therefore, literature presented a special challenge to CSE to devise new and better techniques of assessment.

Since the existing examinations provide a yardstick by which to judge CSE, the aim of the following sections is to describe the state of literature examining prior to 1965 and to use this as a basis for evaluating "... the achievements and lessons" (Dixon, 1977, p.8) of CSE.

2. THE PREVAILING PHILOSOPHY OF LITERATURE TEACHING IN THE 1960's

Most GCE examinations encouraged an approach to literature which fell far short of the ideals of enlightened educationists. To understand the extent of this discrepancy, it is necessary to know something about the prevailing philosophy of literature teaching.

Great importance was attached to literature as a means of personal growth. It was claimed that through its study one's capacity for life was extended. Typical of these claims was that made by Whitehead (1966) who argued that through literature we gain:

... an extension of the capacity for experience ... the quality of the imaginative experience we derive from our books has a close bearing upon the quality of experience we are capable of in our everyday living. (p.48)

Similar convictions were expressed by Holbrook (1961):

... the best novel has the power to change our whole being, to refine our feelings and understanding, to develop our hold on life ... we read to become more aware of life ... and to exercise in imagination our capacities to deal with it. (pp.160 and 189)

Since it was regarded as contributing so much to personal development, literature was held to be of central importance in the school curriculum:

The reading of literature is perhaps one of the two or three most important activities that children practise in their school days ... (It will) prepare children more adequately for life. (Butts, 1967, pp.313 - 317)

It was not only English specialists who made such claims for literature: government reports agreed. For instance, the Eighth Report of the SSEC (1964) argued that:
... the reading of literature needs no justification. Its significance for personal values, for the width and depth of an individual's mind, and for his growth as a thoughtful member of society is self-evident. (p.20)

Likewise, the Newsom Report (1963) stressed that:

All pupils, including those of very limited attainments, need the civilising experience of contact with great literature ... literature can stretch the minds and imaginations of the pupils and help to illumine for them, in wider human terms, their own problems of living. (para. 473)

It is against these claims that the validity of existing examinations must be judged.

3. THE EMPHASIS ON FACTUAL DETAIL

The general opinion of O-Level is neatly summed up by Norris's (1962) complaint that, "... the present O-Level literature examination does not do the job as it should" (p.95). There were several reasons for dissatisfaction with existing examinations and each one will be discussed separately although in practice they did not occur in isolation.

A common criticism was that response to literature had been distorted into a factual or historical exercise characterized by "fact-grubbing". For instance, a correspondent in the Times Educational Supplement admonished O-Level for its treatment of poetry:

... not as poetry but as something else - fiction or history, for example, or material for academic exercises of the comprehension type. Too many questions ... seem to treat poetry in curiously inappropriate terms ... "Describe Excalibur and its history" only confirms the impressions of the candidate who regards a poem as not essentially different from a short story in the experience it offers. (March 18th 1960, p.553)

It was feared that the special nature of the literary experience was being lost in the face of the examination's preoccupation with factual minutiae: "... the conventional English examination has offered no positive encouragement to schools ... to regard literature as more than the factual content of three, four or five books" (Dixon, 1965, p.1). Likewise, the Birmingham NATE (1966) noted, "... with despair ... the sheer quantity of
trivial detail that the candidates are expected to know" (p.5).

Journals sometimes published accounts of the strategies adopted by teachers to cope with the conflict between their own aims and the approach demanded by the examination. The following is interesting because of what it reveals about attitudes towards O-Level:

I should ask all Heads to forbid absolutely the spending of more than three terms on set books. That at least gives the fourth year for reasonably serious reading.

(Use of English, Vol.13, 1961, p.103)

Clearly, O-Level was regarded as incompatible with "serious reading" and separate from the proper business of literature teaching. It was treated as an unavoidable evil which should be ignored for as long as possible.

Likewise the Birmingham NATE (1966) agreed that: "... there is no doubt that literature examining in its present form is a complete antithesis to enlightened teaching of the subject" (p.2).

The question this raises is: why did conventional examinations display an obsession with factual minutiae?

4. REASONS FOR THE EMPHASIS ON FACTUAL CONTENT

At the root of this problem lay the misconception of English as a "subject" in the normally accepted sense of the word:

Few people stopped to consider the real nature of the subject, or came near to discovering that most paradoxical of all English facts, namely that it is not really a "subject" in the traditional sense at all. Perhaps this "content" myth is the fallacy behind all the others, that English, like certain other subjects, is little more than a body of factual knowledge to be transferred from teacher to pupil ... and lending itself naturally to examinations. (Shayer, 1972, p.19)

The second, and more serious, difficulty is that examinations are most efficient at assessing the kind of factual material on which examiners can readily agree and which can be marked with a high level of objectivity.

As Whitehead (1965) observed, the examination's quest for reliability:

... is achieved by concentrating on those tasks which examiners can agree about among themselves and can mark with self-consistency ... "Facts" are what makes standardising possible ...
Consequently memory for facts, even insignificant facts, takes on an absurd importance. Many questions ask for nothing else ... this emphasis has a peculiarly unfortunate distorting influence upon teaching; it fosters a tendency to treat poems, novels and plays as though the valuable thing about them were their extractable content of event or argument - sometimes with the implication that it is important to remember quite small details. Yet after all the term "facts" is a misnomer, since this is literature ... and not history, the facts which gain credit from the examiner are in reality fictions treated over-literally. What we observe in such an examination is indeed the creation of a unique class of mark-gaining items which we might call (borrowing a coinage from James Joyce) ficts. (pp.275 and 279)

Thus, the literature examination aped the objectivity of other subjects by relying on "ficts" and although this gave it a semblance of objectivity, it also made it the antithesis of what literature teachers valued: "... these fragments of knowledge and skill which are common to thousands are not what we as English teachers value most; we value rather the qualities of observation, imagination, perception and judgement which are individual" (Whitehead, 1965, p.275). This was compounded by the misconception of English as a subject in the normally accepted sense of the word and together they encouraged the treatment of literature as a factual discipline.

5. THE BACKWASH EFFECT ON TEACHING

The problem, as Whitehead (1965) observed, is that, "... English teachers value ... the qualities of observation, imagination, perception and judgement which are individual" (p.275). These highly subjective qualities are difficult to assess and this created a disjunction between the ideals of the teachers and the demands of the examination. This point was underlined by a correspondent in the Times Educational Supplement:

... "the reading of a poem in class must not be aimed at imparting information - whether factual or narrative - it must be an experience, an active experience of poetry" ... this is what makes the poetry lesson essentially different from the kind of lesson that is concerned with the acquisition of facts or techniques ... while information or techniques can very easily be tested, it is extremely difficult to test the validity of an artistic experience, especially at a comparatively immature level.

(March 18th 1960, p.553)
The assumption that something as idiosyncratic as response to literature could be measured with any meaning or accuracy by such a blunt instrument as the mass examination was increasingly questioned: "... all past experience suggests the existence of a deep and inherent incompatibility between external examinations as we know them and the essential aims of good English teaching" (Whitehead, 1966, p.233).

Even the conditions under which examinations are conducted were seen as incompatible with responding to literature:

Suppose a teacher instructed his students when writing literary critical essays:
(a) never to have the text by them;
(b) never to consider the topic beforehand;
(c) never to spend more than 45 or 60 minutes overall.

It would seem a very poor method of eliciting a literary response. ... It is constraints such as these that inevitably reduce the value of the evidence collected in the traditional English examination papers. (Dixon, 1977, pp.5-6)

That which made the examination fair and objective was perceived as imposing absurdly artificial constraints on the study of literature.

However much teachers objected to the approach taken by the examinations, their backwash effects were so powerful that teaching invariably was influenced. Indeed, it was claimed that even when an examination syllabus reflected a teacher's aims, the backwash effect persisted:

... even when in the main the examination reflects what you want to teach, the probability is that the claims of the examination will somewhere reduce what you want to teach, and would like to explore in depth, to what is relevant to the examination ... any external examination, however carefully conceived ... must somewhere intrude upon these personal preferences and damage the personal approach to the subject in the classroom. (Clarke, 1970, p.24)

NATE (1966) agreed:

If ... the examination comes from outside the school and largely decides for the teacher what he will teach ... then there are clear dangers ... no syllabus can possibly satisfy anyone save the man or woman whose creation it is ... An external examination inevitably interposes itself to some degree between the teacher and his class. At the best, there must be compromises; there must be choices by the examiner which cannot command the full assent of the teacher. (p.35)

The problem was exacerbated by GCE's failure to assimilate developments in
the philosophy of English. While the concept of English had evolved rapidly since the beginning of the century (see Shayer, 1972), examinations had become stereotyped and largely unresponsive to changing needs:

... the syllabus should reflect developments in taste, critical opinion and teaching method ... A glance at recent syllabuses does not suggest that these developments are always taken into account. Indeed, there is often a striking similarity between the structure of the present examination and the papers of 10 or even 20 years ago, before the GCE came into existence. 

(Correspondent, TES, 1960, p.553)

6. SEPARATE CERTIFICATION

Another source of dissatisfaction with GCE was its method of awarding separate certificates for language and literature. Here too, GCE had failed to keep pace with developments in the field of English. At the beginning of this century, "English teaching at all levels ... was in a somewhat sorry plight" (Shayer, 1972, p.2) and literature was especially liable to be neglected. Most English teaching took the form of language work with a heavy grammatical bias. However, this century, there has been immense growth in the stature of English and, more especially, in the role and status of literature. Instead of regarding language and literature as separate activities, progressive educationists increasingly saw English as a unified whole with literature playing a central part:

... a justly conceived English examination ... does not (leave) literature out of account, but insists, rather, on giving it within our concept of "English" the central place which rightly belongs to it as the keystone which holds the arch together. 

(Whitehead, 1965, p.276)

However, GCE persisted in awarding separate certificates for language and literature. This had a twofold backwash effect on teaching. Firstly, it undermined the attempt to treat English as a unified whole:

We have failed to carry over into ... examinations the unified conception of English which nowadays underlies our best teaching of it; instead we allow examinations to segment this unity. 

(Whitehead, 1965, p.275)

Secondly, and perhaps more seriously, separate certification devalued
literature. Separate and single certification imply different valuations of literature with separate certification actively reinforcing the traditional inferiority of literature by facilitating its neglect:

... literature papers are taken by only a fraction of the language candidates ... It is thus possible for the vast majority of candidates in English to spend a year without any literature teaching at all. This is the problem on which to concentrate. (NATE, 1966, p.43)

Demands for single certification started to mount. For instance, the Eighth Report of the SSEC (1964) argued that:

The main defect of the present practice is that it provides a language examination largely separated from literature. We believe that (this implies) ... a misleading, unnecessary and, at this level, undesirable division. (p.20)

NATE (1966) agreed: "English is one subject not two ... we would place high on the list of our objectives the ending of the artificial distinction between language and literature" (pp.9 and 56). Despite these demands, the "lang-lit" division persisted, demonstrating GCE's inflexibility and resistance to change.

7. O-LEVEL: A DISINCENTIVE TO WIDE READING

The prevailing philosophy of English emphasized the importance of wide reading. Scores of reading schemes which were claimed to promote this end were printed at this period and NATE (1966) pointed out that the examinations should support this aim: "... syllabuses should encourage wide reading ... pupils' reading should not be limited to a small number of set books" (p.9).

As an inducement to wide reading, the conventional O-Level syllabus failed miserably. It was primarily a "set book" examination in which teachers selected, from a list of prescribed texts, a small number of books (usually three) for intensive study. Teachers rarely covered more than the minimum requirement; indeed, one of the best-known backwash effects of examinations is that the teaching syllabus tends to become synonymous with
the examining syllabus discouraging teachers from venturing outside the examination requirements. As Kennedy (1957) observed: "... it is necessary to consider the minimum practical requirements ... for there is little doubt that most schools confine themselves to the minimum" (p.92).

This was exacerbated by the questions set at O-Level which demanded that candidates should know the prescribed texts in great detail. This preoccupation with minutiae discouraged teachers from ranging outside the set texts lest their candidates should confuse details from one text with those from another. The teacher's dilemma was that if he followed the recommendations for extensive reading, he risked jeopardising his pupils' chances of success in the examination. In the event, most succumbed to the powerful backwash effect of the examination and concentrated on the minutiae they knew it would reward. As Whitehead (1965) observed: "... so important at all stages is wide reading - and so shamefully is it ... discouraged by the English examinations" (p.283). Moreover, the GCE Boards persisted in these narrow requirements from decade to decade. Thus, in 1957 Kennedy found, "... a discernible 'norm'. Five of the boards require a minimum of three books and three require four while one insists on only two books" (p.92). The position was virtually unchanged by the mid '60's as the survey undertaken by the Birmingham NATE (1966) revealed: "Often the 'three or four books' demanded show an astonishingly narrow range of choice both from year to year and from Board to Board" (p.3).

Perhaps the most damning aspect of these criticisms is that they were being replicated from decade to decade demonstrating the inflexibility of the boards and their failure to respond either to needs inside the schools or to criticism.

8. THE INAPPROPRIATENESS OF THE PRESCRIBED TEXTS

Many educationists were dissatisfied with the quality of the texts which the boards deemed suitable for close study. Here again the examina-
tions were in direct conflict with the prevailing English ethos. The work of psychologists had suggested that effective learning only takes place when a pupil's interest is engaged and this had influenced thinking about English. Thus, NATE (1965) urged that an English examination, "... should be judged by its ability to encourage in the candidate an active response, a sense of involvement in the work to be done" (p.2).

However, many of the texts set at O-Level were felt to be remote from the interests and experiences of the average sixteen-year old and the reliance on traditional "classics" meant that the idiom and period of the texts provided an additional barrier to involvement. Thus, the Birmingham NATE (1966) described the choice of set books as, "... an urgent matter ... the same texts, often totally unsuitable, recur with monotonous regularity" (p.3). The SSEC (1964) concurred that: "... pupils have frequently been put off literature by being made to study works barrenly treated and beyond their range of interest and capacity" (p.20). An illustration of what this could mean in the classroom was provided by Norris (1962):

Why should Shakespeare's be the only drama that O-Level students are allowed to read? Again, why should the choice of novel be so limited? If one considers that Jane Austen is not quite the most appealing material for candidates of 15 - 16 years old, who live in mid-twentieth century industrial England, one may very likely be left with Trollope ... (but) it is the poetry which is least satisfactory. Who can think Milton's Minor Poems suitable for an O-Level candidate? (p.96)

Again, the problem was one of long-standing with the range of texts varying little from board to board and from year to year.

The problem was compounded by a lack of comparability between the different texts prescribed at O-Level. Critics complained frequently about the uneven standards of the set books:

What is depressing ... as one looks through the lists is the uncertainty of taste and discrimination shown. Some books are well outside the range of the sixteen year old ... others may be manageable but are certainly of little interest ... yet others may have little discernible literary merit. (NATE, 1966, p.36)
Whitehead (1965) was also struck by this lack of parity:

On the one hand there are the established classics .... (and although these) are serious and complex enough to repay close study ... the qualities we value in them are to the sixteen year old both adult and difficult ... In the second category of set book, on the other hand, we find recent or contemporary fiction or autobiography ... at best they are worthy but slightly second-rate, at worst they hardly amount to more than effective journalism. (pp.280-281)

Once again, the same criticisms were being replicated from decade to decade. Thus, in 1957, Kennedy argued the need for:

... greater uniformity in the requirements for "O" level literature. I have considered only the 'standard' papers ... but some of the (alternative papers) could well be mistaken for "A" level syllabuses. Consequently an "O" in literature on a certificate can mean almost anything from uncritical absorption of three mediocre books to a considerable acquaintance with many aspects of five hundred years of English literature. (p.93)

A decade later the situation had not changed:

One thing that immediately emerges from this survey is the wide range of knowledge and expertise demanded in what is supposedly a standard national examination. Thus the difference between the requirements for the Cambridge Local and the Southern Universities JMB Paper B is so wide ranging that it is difficult to see in what way a pass in these two separate examinations can be equated ... the variety of scope and difficulty that exists between the Boards (makes) any standardization of a national examination surely impossible. (Birmingham NATE, 1967, pp.6 and 40)

9. ENJOYMENT OF READING

English specialists were concerned that reading should be, first and foremost, a pleasurable activity: "... every child should be made aware of the existence of literature as a source of enjoyment", urged the SSEC (1964, p.20). NATE (1966) agreed:

We want our pupils to enjoy reading good books, plays and poems ... and to attain the beginnings of perception, discrimination and sensitive intelligence that we hope will lead to an enduring love of good literature. (p.35)

These recommendations were made at a time when the set book "grind" was virtually guaranteed to eliminate any gratification which might be derived from reading:

... the most powerful argument for change is the sad result of our "O" level literature course: that in attempting to thrust
some heritage of literature on our pupils hastily and perforce, before (as we think) it is too late, we in most cases ensure that our examinable ration is all our pupils ever care to taste. What a pity we do not remember that in literature at least our compulsory schooling can only be (and should only be) a tempting start to a lifetime of self instruction.

(Norris, 1962, p.97)

Thompson (1966) agreed, arguing that, "Literature is for pleasure. That in our school context we've tended to forget because of examinations and the grim atmosphere they engender" (p.4).

The conventional question format contributed to this stultifying effect on enjoyment. The questions had become stereotyped and this encouraged cramming, rote-learning and the mindless regurgitation of second-hand opinions:

The trouble lies principally in the questions. Far too many of these questions are set to a formula which permits the regurgitation of prepared answers ... Chief examiners are aware that ... candidates should be encouraged to indicate their response to the book examined; only too often this leads directly to the reproduction of second-hand opinions given to the candidates by teachers. (NATE, 1966, pp.36, 37 and 40)

Thus, the conventional question format "... set a premium on hypocrisy and encourage pupils to repeat mechanically and without reflection other people's judgements" (Whitehead, 1965, p.282). Once again the fixed nature of the examinations was at the root of the problem. The "... highly conservative trend in examining from year to year" (Birmingham NATE, 1966, p.5), enabled teachers to become skilled in question-spotting and this encouraged the expression of second-hand opinions and prepared responses. Clearly, dissatisfaction with conventional examining techniques was widespread and it led to a search for alternative forms of assessment.

10. ALTERNATIVE FORMS OF EXAMINING

Since the set book was the butt of so much of the criticism of O-Level, it is hardly surprising that some critics regarded escape from the set book regime as the most promising way forward. This led to experiments with general literature examinations in which prescription was replaced by a
recommended reading list. The elimination of prescription seemed an obvious step to take— it provided the freedom for teachers to select materials according to their own enthusiasms and the interests of their pupils and to teach texts in whichever way seemed most appropriate. Furthermore, since questions based on general courses of reading were necessarily open-ended and non-specific, this precluded a preoccupation with minutiae. In theory, this approach offered an attractive alternative to the conventional examination. However, even before the introduction of CSE, doubts were being expressed about the wisdom of this approach.

As an approach to teaching, the general syllabus possessed some merit but as a technique of examining it was ill-equipped to elicit a genuine response to literature:

When no texts have been prescribed for study, it has been difficult to set examination questions which are not so general as to be vague, or so obvious as to be anticipated by the text book and teacher's note. (Examinations Bulletin 1, 1963, p.36)

The difficulty posed by this type of examination was how to design questions which were general enough to apply to a range of books and, at the same time, tested the candidates' understanding in depth. Most examinations failed to do both. In making the questions sufficiently open to accommodate a range of literature, examinations frequently failed to elicit depth of response. Thus, the Birmingham NATE (1966) observed: "We do not welcome the alternative of general questions on literature ... as at this level these become so vague as to be of little value" (p.3). Moreover, general literature examinations tended to become stereotyped and this encouraged the kind of cramming for which the set-book examination had been so heavily criticised: "Inevitably 'open' questions soon become stereotyped and lead to prepared answers" (Clarke, 1970, pp.22-23).

Since the general examination proved to be a disappointing alternative to the set book approach, some educationists came to regard the examining of literature as a choice between equally unsatisfactory alternatives:
The usual objection to set books is that they are restrictive and encourage second-hand opinions. To go to the other extreme... and set papers with "open" questions seems likely to invite disaster. (Clarke, 1970, p.22)

Examinations Bulletin 1 (1963) saw the problem in similar terms:

When set books have been used to limit the area of the examination they have narrowed the pupils' experience of literature and have led to the prepared answer ... When no texts have been prescribed for study, it has been difficult to set examination questions which are not so general as to be vague, or so obvious as to be anticipated by the text book and teacher's note. (p.36)

In fact, Bulletin 1 (1963) saw these difficulties as so insuperable that it recommended that the new examination for CSE, "... should be, in the main, a language rather than a literature paper" (p.36).

Another form of external examination in use prior to 1965 was the paper based on unseen extracts. However, this mode of assessment appears never to have been used as the principal method of assessment and has received comparatively little comment. This makes it difficult to form a firm impression of attitudes towards it. Nevertheless, it was an attractive alternative to the two more common methods of examining.

Providing that neither the passages nor the questions became stereotyped, the unseen offered the teacher the same freedom in the classroom as the general literature examination without entailing some of its worst drawbacks. While the general literature paper was found to be a poor test of understanding, it was relatively easy to devise a penetrating test of insight based on a passage which the candidate would have in front of him. If this method could avoid the stereotyping which is such a strong feature of external examinations generally, it was argued that it could eliminate some of the evils of the other forms of assessment: dictated notes; the rote-learning of second-hand opinions and fact-grubbing. It would also place a proper emphasis on the ability to respond sensitively since the best way of preparing for it was to expose candidates to a range of literature. The principal fear expressed by its critics was that this
method would encourage a tendency to fragment literature and treat it as "snippets" thus depriving pupils of the experience of full-length works which made sustained demands on them. As far as the writer is aware, the unseen extract has never been used as the principal method of assessment and therefore this potential drawback is offset by using the unseen passage as one element in the assessment rather than the entire assessment.

By 1965, none of the GCE Boards had managed to devise a satisfactory examination of literature making this a major challenge to CSE which was expressly designed to promote innovation and to improve techniques of examining. Could CSE provide an examination which would avoid the backwash effects of conventional examinations and at the same time provide an effective means of sampling response to literature?

11. THE SEARCH FOR A NEW STYLE OF EXAMINING

Clearly, the goal of progressive educationists was an English examination which unified the separate components of language, literature and oracy. Therefore, single certification was the only approved form of award. Moreover, literature was to be elevated to a position of central importance within this new concept of English. Reading was to be undertaken primarily for pleasure with pupils studying only those books which were relevant to their interests and experiences and which could contribute to their intellectual and emotional growth. They were to be encouraged to read widely and the existing preoccupation with factual minutiae and the mindless absorption of second-hand opinions was to be eschewed in favour of an honest response.

Since the external examination had proved to be so problematical, there was a general consensus that the solution ought to be sought in some form of school-based assessment:

... all past experience suggests the existence of a deep and inherent incompatibility between external examinations as we
know them and the essential aims of good English teaching ...
For English the only really satisfactory mode of examining is
a system of internal examinations with external moderation.
(Whitehead, 1966, p.233)

NATE (1966) agreed arguing that, "... as soon as the examining is taken
out of the hands of the teacher, serious problems arise" (p.3). The
continuous assessment of course work was the form of school-based assess-
ment which appealed most to English specialists. They regarded response to
literature as the kind of "transitory activity" (p.22) which Examinations
Bulletin 1 (1963) had recommended should be examined by course work. Thus,
Dixon (1965) argued that:

Wherever possible, normal school work should form part of the
assessment particularly when the stimulus to speaking and
writing is of a kind that cannot conveniently be used in an
examination - as, for example, when a class reads together a play
or a novel, and discuss it, with the experience still fresh. (p.2)

The problem with external examinations, he argued, was the need to:

... find ways of reaching and recalling the living response to
the text ... the living experience of books has become ... dim
and tarnished ... (and the examiner must) elicit from
candidates a once-warm response, now dormant or cold. (pp.21 and 24)

Continuous assessment was welcomed as a breakthrough in the quest for
more effective techniques of evaluating literature because it captured the
immediate response and incorporated this into the final assessment rather
than requiring the candidate to replicate it for an examination when the
freshness and immediacy of the original response had been lost. Some
educationists would deny that the spontaneous response is necessarily the
best response to literature in the belief that the process of returning to
a book and reflecting upon it can deepen and enhance the experience.
Nevertheless, the emphasis on spontaneity was perhaps of special value for
pupils in the CSE ability range many of whom find abstraction and repeated
reflection difficult. Thus, one of the most attractive features of
continuous assessment was its capacity for integrating teaching and
examining. Indeed, one of the great hopes for the CSE examination
generally was that it might bridge the gulf which existed between teaching
and examining. Thus, the SSEC (1963) expressed its hope that the new boards would, "... move towards an integration of the roles of teacher and examiner" (p.26).

Continuous assessment was also regarded as complementary to the study of literature because it avoided imposing the artificial constraints of external assessment onto response:

Ideally, this work of studies response is best done in the classroom. How can anyone write satisfactory answers within examination time limits? If the book is not available for consultation, as it naturally is in the classroom, then all there can be time and energy for is reproduction of memorized material. (NATE, 1966, p.37)

Another novel feature of CSE was its approach to the backwash effects of examining. Instead of regarding these as necessarily evil and therefore attempting to eliminate them, educationists increasingly accepted their inevitability and endeavoured to control them. It was argued that the backwash effect of a procedure should become a prime consideration when evaluating its usefulness. "For English, the most important consideration is not reliability but the nature of the backwash effect", claimed Whitehead (1965, p.275). NATE (1966) agreed: "... the yardstick of a good examination is to be found in the 'backwash' effect it has upon teaching in the classroom" (p.3).

This section has attempted to summarise the fermentation of ideas which contributed to the creation of CSE literature examinations. NATE (1965) was heartened by the fact that revolutionary principles had been established at the outset:

... most important of all, sound basic principles have already been laid down: the examination should "reflect school work" rather than condition it ... wherever possible, normal school work should form part of the assessment. (p.2)

How was this theory translated into practice by the regional panels when they devised their English syllabuses? The following section will attempt to compare the theory of CSE with its practice.
12. **THE NEW MODE I EXAMINATIONS**

Since the continuous assessment of course work was one of the most promising ideas in this difficult area, it is useful to start by considering its use by the regional panels. Only one of the thirteen panels* adopted continuous assessment wholesale. The East Midland Region abandoned conventional examining techniques in favour of an examination based entirely on a folio of work produced during the course. NATE (1965) was greatly encouraged by this initiative presenting it as an example of the "... boldness in breaking new ground" (Examinations Bulletin 1, 1963, p.3) which CSE had been designed to promote. "We heartily welcome this initiative", they announced (p.19). It is hardly surprising that NATE attached importance to this syllabus since there was little in the schemes proposed by the other panels which they could draw attention to as models of boldness and originality. An indication of their essential conservatism is that only two other panels made any use of continuous assessment. Candidates in the South Western Region were allowed to submit course work in place of the prose and drama sections of the examination, poetry being the only compulsory examination item. NATE (1965) described this as, "... an excellent compromise" (p.19). The South East Panel allocated one third of the marks to course work and the remaining two-thirds to a set text examination. NATE offered no comment on this syllabus. The majority of panels (i.e. the remaining ten) confined themselves to the familiar external examination. However, within that conventional framework there were signs of change, of attempts to make examination conditions more like those to be found in the classroom. For

---

* The figures quoted in this section are taken from the NATE document, *Criteria of Success in English* (Dixon, 1965). They take into account the North Region's exclusion of literature from its English examination and the sub-division of certain boards into separate areas with their own regional panels. This meant that altogether there were thirteen separate Mode I literature examinations in use.
instance, the emphasis on memory and rote-learning in GCE was countered by those panels (four in all) which decided to allow candidates to use unmarked texts in the examination. (A fifth panel allowed candidates to use poetry texts only.) The fact that nearly half of the twelve panels employing external examinations were allowing unmarked texts into the examination room was an encouraging sign. Likewise, five panels incorporated the unseen approach into their examinations while seven opted for the general literature examination. While neither of these procedures were new, they certainly would have represented a fresh departure for many teachers whose experience of literature examining was limited to the set book examination. Thus, although the panels displayed an essentially conservative attitude towards actual techniques of assessment, within the conventional framework, there was some experimentation. As Hipkin (1967) observed: "Certain CSE categories are conventional ... What has changed often is what those categories imply" (p.33).

Against these gains must be weighed the continuing popularity of the set book approach. In fact, nine of the twelve panels offering external examinations proposed to use set texts. It must be remembered that the set text was a feature of the existing system which had attracted vociferous criticism and it had seemed that the most immediate challenge to the new examination was that of eschewing prescription. The adoption of the set text approach made these new syllabuses a travesty of the, "... fresh, indeed courageously fresh, approach to techniques of examining" (p.26) recommended by Examinations Bulletin 1(1963). Moreover, the imprudence of adhering to the conventional approach was immediately apparent in the specimen papers of those boards intending to use it:

... the dangers of the set book examination are far from being overcome ... the backlash in the classroom of many of the specimen questions we have seen is bound, in our view, to kill an honest response to literature. To date we submit that the case for the set book is not proven. (NATE, 1965, pp.23-24)

Thus, even before the examination proper was launched, the dangers of the
set book stereotype were already apparent. Conventional procedures were surprisingly popular and nowhere was this more strikingly illustrated than in Mode 3. Although Mode 3 is recognised generally as the most experimental and progressive mode, Hipkin (1967) found that more than half of the one hundred syllabuses included in his survey had adopted the set book approach.

The conservatism displayed by this new generation of examiners is perhaps not as surprising as it appears at first sight. The panels consisted of practising teachers who would have derived most of their examining experience from GCE and its imitators. Thus, few of them would have had expertise or insight into novel techniques of examining making this an area where they were ill-equipped to make the kind of far-reaching contribution which was expected of them. Even though many of them were dissatisfied with existing methods, these did have the virtue of familiarity and, for teachers faced for the first time with the awesome task of responsibility for a public examination, that which was established may have seemed preferable to the untried and untrusted.

Over half of the panels using external examinations (i.e. seven out of twelve) opted for the general literature approach. This was more novel than the set book procedure and would have represented a fresh departure for many teachers. Nevertheless, its popularity (i.e. more widely used than the "unseen" and second only to the set book method) is somewhat surprising considering that even before CSE was introduced there was evidence suggesting that this was an unsatisfactory form of examining. This was confirmed by the specimen papers of those boards intending to use general examinations. NATE (1965) found these even more deficient than the set book papers:

... general questions offer even less help than questions on specific books ... (They) preclude anything but the prepared question or the barest elementary response. We find, with regret, that many questions are still an open invitation to
cramming. Stereotypes are already well-established ... we must conclude from the available evidence that the intractable problem of setting 16 year olds general questions on literature has not yet been solved. (pp.24 - 25)

Thus, as far as techniques of examining were concerned, the innovations were piecemeal, most panels displaying a preference for the established and the familiar.

One of the few achievements in terms of "... breaking new ground" (Examinations Bulletin 1, 1963, p.3) was in the range and variety of texts introduced into the examining of literature under CSE. In his survey of books recommended for study in Mode I, Butts (1966) found that:

... a real attempt has been made to break away from the idea of a limited number of set books, frequently of a remote relevance, and to substitute a greater freedom of choice in both quantity and quality ... a great deal of ingenuity has been exercised in suggesting the amount and kind of literature to be studied. (p.296)

For instance, whereas the drama prescribed at O-Level tended to be almost exclusively Shakespeare, Butts (1966) found that the regional panels had recommended fifty plays by thirty-four different authors (p.297). Likewise, O-Level fiction was selected from an exceedingly narrow range which relied heavily on such nineteenth century authors as Scott, Dickens and Jane Austen. In contrast, Butts reported that the CSE panels had suggested 124 titles by seventy-five writers. There was similar variety amongst Mode 3 syllabuses where Hipkin (1967) found that 302 prose texts had been recommended by almost one hundred separate syllabuses (p.309).

In addition, CSE literature had a distinctly modern flavour. Of the 302 texts included in Hipkin's survey (1967) over three-quarters of them were written after 1914 and over half of the twelve most popular modern authors were still living (p.309). Hipkin concluded that:

... all three modes encourage a wider view of literature so that CSE candidates are reading extensively, especially in the fields of the modern novel and non-fiction. (p.25)

Clearly, this was an area in which a trend in sharp contrast to conventional practice had emerged and it is useful to speculate on the reasons why
innovation was abundant here when panels had been so reluctant to experiment with novel assessment techniques. It may well be that while the average teacher had little expertise in the area of assessment, course content was something to which he could make a much more positive contribution because it drew on his skill as an English specialist. CSE seems to have failed to take real account of the fact that until 1965 teaching and examining had evolved as essentially separate activities which meant that, as far as examining was concerned, most teachers were novices. This, perhaps more than any other factor, explains why English panels experimented freely with course content whilst adhering to established examining techniques.

Another area in which the regional panels responded positively to enlightened opinion was in their approach to certification. GCE's fragmentation of English into language and literature had been denounced by leading educationists and the panels responded with an almost total ban on separate certification. All but one of the fourteen panels offering Mode I examinations in English opted for single certification. This commitment to single certification was regarded as so fundamental to CSE that panels even resisted attempts to set up Mode 3 syllabuses employing separate certification: "Panels sharply discouraged schools proposing to issue certificates with 'fragmentary endorsements' of literature or oral English" (Hipkin, 1967, p.306). This had far-reaching consequences for the study of literature because it safeguarded against the neglect of reading which separate certification facilitated. NATE (1966) pointed out that under GCE only about 25% of the English Language candidates also entered for literature: "It is thus possible for the vast majority of candidates in English to spend a year without any literature teaching at all. This is the problem on which to concentrate" (p.43). It was a problem to which almost every regional panel did address itself. By insisting on single certification, they ensured that literature was treated as a
compulsory component of English under CSE.

Although single certification ensured that literature was not neglected, it is erroneous to assume that it necessarily led to the treatment of English as an integrated whole. In some cases, the adoption of single certification masked the perpetuation of conventional divisions between language, literature and oracy. As Dixon (1965) observed:

... an agreement to add together percentages for "language" and "literature" as some panels propose, does not represent a unification of English work, though it is a necessary step in that direction. (p.4)

The real task was to evolve a whole new concept of "examination English" in which conventional divisions were eliminated altogether. Single certification was only one step in that direction but some panels never went beyond it.

Fresh attention also was paid to the language in which examination papers were couched. Conventional examination language was criticized for discouraging the candidate by employing a formal, foreboding and "frozen" style (Dixon, 1965, p.11). CSE examiners were urged to modify the tone of examination language:

Although we recognise that there may be a difference between a teaching and an examining question, we hope that the form of the examination questions will follow good classroom practice ... we should like to see a more humane approach to examining the wording of questions so that candidates are encouraged to do their best. (Birmingham NATE, 1967, pp.4 - 5)

Dixon (1965) agreed that: "Within the examination proper, conditions should be similar to those of the classroom ... the tone and the material ... should be judged by its ability to encourage in the candidate an active response" (p.2). Evidence that these recommendations were filtering through to practising teachers was provided by Hipkin (1967) who was struck by the new style of language apparent in the Mode 3 papers he evaluated:

Candidates have been encouraged, involved, "led-in", briefed and prepared so that the activity of examining, normally cold and neutral, has become an educative activity in itself. (pp.25 and 33)
Nevertheless, examining was a new skill for many teachers and therefore it is hardly surprising that Dixon (1965) found a mixed picture among the Mode I specimen papers. Questions which were helpful appeared alongside others with a "... frozen or formal style" which precluded "... direct, honest discussion" (p.11). As with so many features of the examination, question style needed time to develop.

13. AN EVALUATION OF THESE EXAMINATIONS

The mood of cautious conservatism displayed by most of the panels was at variance with that of Examinations Bulletin I (1963) which favoured innovation: "... there is a strong case for experimenting with new methods of examining," it urged (p.25). Despite the fact that literature was fraught with assessment problems and that there was mounting dissatisfaction with conventional procedures, the external examination and the set book emerged as the most prominent features of the new examination. Dixon (1965) observed that, "... enlightened syllabuses do not necessarily produce enlightened questions", arguing that it is therefore necessary to compare "theory with practice" (p.5). In the case of literature it is equally true that enlightened recommendations did not automatically lead to progressive syllabuses. In fact, when the theory of CSE literature is compared with its practice, it is clear that many of the early achievements were piecemeal and limited. As Dixon (1965) observed: "... the regional response is patchy as yet" (p.46). EMREB provided the only clear-cut example of "... boldness in breaking new ground" (Examinations Bulletin I, 1963, p.3) by eschewing both the set book and the final examination in favour of 100% continuous assessment of course work and freedom for the teacher in the choice of texts. EMREB received warm praise for this initiative from NATE (1965, p.19). However, it is important to recognise that it was the willingness to experiment rather than a proven method of assessment which was being praised. Although course work assessment seemed to offer
considerable promise, little was known about the way in which it would work in practice.

Since continuous assessment was such a novel examining procedure in the 1960's, it now seems timely to evaluate its achievements. How successful has it been in realising the hopes invested in it? Has it proved to be as well-adapted to the special needs of literature as its advocates claimed that it would be? The following section represents an attempt to evaluate some of the "... lessons and achievements" (Dixon, 1977, p.8) of the most far-reaching experiment to emerge amongst the Mode I literature syllabuses.

B. THE FINDINGS

1. CONTINUOUS ASSESSMENT OR EXTERNAL EXAMINATION?

Since continuous assessment is one of the most innovative and least well-documented aspects of literature examining, it was felt that its usefulness merited special attention in the questionnaire.

Initially, it was decided to gain a general impression of the level of support for continuous assessment relative to that for conventional examinations. Question 6 provided a list of differently composed tests ranging from: 100% continuous assessment - 0% external examination to 0% continuous assessment - 100% external examination and respondents were asked to indicate which alternative provided the most satisfactory means of assessing (i) language and (ii) literature. The literature results are recorded in the final column of Table 1, overleaf.
Although seven alternatives were listed, by far the largest category, accounting for half of the entire sample, was 100% continuous assessment. This suggested that not only were many teachers satisfied with continuous assessment but that they were satisfied with it as the only method of evaluating literature. Even amongst those who favoured "mixed economy" (Dixon, 1977, p.6) examining, there was a preference for the continuously assessed element to be most heavily weighted with a quarter of the sample preferring 60 - 75% continuous assessment compared with 4% who favoured 60 - 75% external examination. The overall pattern of responses demonstrated the immense popularity of continuous assessment and the very limited support for external examining; only 6% of the respondents wanted 60% or more external examination compared with 77% who favoured 60% or more continuous assessment. Clearly, this innovation has gained wide support amongst the teachers who use it with the majority wishing to use it as the principal or the only technique of evaluating literature. 103 conveyed something of the enthusiasm generated by course work assessment in the 1960's:

One of the great glories of the 1960's was the number of teachers who actually became excited about the folio approach. They had
not had the opportunity to do this before ... There really was a number of very fine ideas at that time.

104 also had been struck by the impetus which the folio approach had released noting the, "... great output of effort on the part of pupils" which had resulted. 101 summed up these feelings when he observed:

It seems to me that the great achievement of the East Midland Panel was the concept of the literature folio. Unlike other CSE Boards, which got trapped into the old set books problem, the East Midland Board developed this concept of the literature folio.

A comparison of the above figures with those for language (i.e. Question 6(i), see Table 1, middle column) reveals that it was in the examination of literature that the respondents considered continuous assessment to be particularly satisfactory. Whereas over three-quarters of the sample favoured 60% or more continuous assessment for literature, the figure dropped to less than half (43%) for language. Furthermore, over a third of the sample (37%) preferred 60% or more external examination in the language assessment compared with 6% who favoured this for literature. Clearly, it was in the testing of literature that continuous assessment enjoyed widest support.

Although Question 6(ii) revealed the enormous popularity of continuous assessment, it provided no indication of why it was held in such high esteem. In order to investigate the reasons for this, Question 6 was followed up by some more detailed items.

2. THE INTERNAL VALIDITY OF CONTINUOUS ASSESSMENT*

The principal complaint against the external examination was its invalidity as a method of testing response to literature. Thus, the key question raised by the use of continuous assessment is whether it has heightened the validity of literature examining?

One of the original attractions of continuous assessment was its

* For a definition of internal validity, see the footnote on p.123.
capacity for fusing normal classroom activities into the final examination rather than attempting to revive "... a once-warm response, now dormant or cold" (Dixon, 1965, p.24). In order to determine how valuable this attribute was perceived to be in practice, Question 15(ix) stated:

Continuous assessment is particularly well-suited to the study of literature because it evaluates a pupil's response to a text at the time when it is made, rather than asking him to duplicate that response during an examination which may be quite remote from the time when the literature was studied.

Respondents were asked to indicate their level of agreement/disagreement with this statement.*

The response was one of overwhelming agreement with 83% of the sample indicating that they favoured course work assessment for this reason. An equally striking feature of the response was the high level of strong agreement. 39% of the sample strongly agreed with this statement providing some indication of the level of feeling it aroused. In contrast, less than 10% disagreed. Clearly, this capacity for tapping the immediate response and incorporating it into the final assessment was still perceived as one of the strengths of course work assessment.

The other key criticism of conventional procedures was that they encouraged rote-learning and the mechanical absorption of second-hand opinions about literature. Since continuous assessment had successfully overcome one of the main defects of external examining, how far did respondents feel that it had been capable of eliminating this other key deficiency? Question 15(viii) stated:

Continuous assessment of the study of literature avoids certain inherent tendencies of literature examinations: the learning of notes and regurgitation of second-hand opinions.**

46% of the sample agreed with this statement and a further 28% strongly agreed which meant that altogether three-quarters of the entire sample

---

* Strongly Agree (39%), Agree (44%), Uncertain (9%), Disagree (7%), Strongly Disagree (1%)

** Strongly Agree (28%), Agree (46%), Uncertain (11%), Disagree (11%), Strongly Disagree (3%), Invalid response (1%)
believed that continuous assessment avoided these backwash effects. Although this was an impressive rate of agreement, it was nevertheless lower than that achieved by the preceding item (15(ix)). The rate of strong agreement was lower (28% as against 39%) and the level of disagreement was higher (14% as compared with 8%). This suggested that continuous assessment was perceived to be marginally less successful in eliminating these backwash effects than it had been in overcoming the drawbacks of a delayed response. The written comments provided the key to why this was so. For instance, Q289 pointed out that:

Regurgitation can take other forms in the preparation of classes to produce literature essays ... It is still remarkably easy for the student to regurgitate the teacher's opinions and comments made when a novel was read and discussed, for example. After all, it is by doing so that the student will feel confident of a satisfactory mark.

Q557 agreed: "Regurgitation doesn't require an examination situation."

Perhaps the last word on this issue should go to Q418 whose comment best reflects the general consensus of opinion in that it recognised that continuous assessment represented an important step forward in surmounting this difficulty whilst acknowledging that it could not entirely eliminate it:

There has to be some guidance (not necessarily in note form) as to how to tackle a particular question ... Therefore, some regurgitation is unavoidable. I believe it is minimised, however.

3. PUPILS' PERCEIVED RESPONSE TO COURSE WORK ASSESSMENT

In addition to this heightened internal validity, continuous assessment was perceived as having a beneficial effect on pupils. This was revealed by Question 25(x) which stated:

Many CSE pupils benefit considerably from the Mode I approach to examining literature which allows the work produced during the course to determine the final grade.*

This was another statement which commanded huge support with 85% of the

* Strongly Agree (17%), Agree (68%), Uncertain (9%), Disagree (3%), Strongly Disagree (-), Invalid response(3%)}
sample agreeing. There was, however, a comparatively low rate of strong agreement with only 17% of those in agreement opting for the "Strongly Agree" category. Nevertheless, only 3% of the sample disagreed with this statement.

A number of respondents volunteered comments on this issue revealing that reasons for adopting this view varied. Some teachers were most impressed by the influence of continuous assessment on attitudes to work. For instance, IL2 remarked:

It overcomes the unfairness of a one-off, overkill examination ... it helps a child to work better in that the whole year's work is taken seriously.

Others emphasized its beneficial impact on enjoyment. Comparing external examining with course work assessment, IO5 claimed:

I don't really see the value of making people learn, for the sake of it, facts that they're just going to forget. I'd much rather people write an essay on it and get the enjoyment out of the literature ... I think that being made to learn facts and regurgitate, as far as literature is concerned, tends to spoil the enjoyment ... I think there is quite a definite place for an external examination; what I said only applies to literature.

IO3 made a similar point in comparing his work as an O-Level marker with assessing literature folios:

... as O-Level markers, we were used to marking the best considered and best remembered thoughts of the teachers but what we were marking in the folios ... was the individuality of the child having touched upon a theme. We were looking at far more originality ... The idea was to expose the children to something that would catch their imaginations ... It's much more creative and imaginative on the part of the child; the child has got to respond.

Thus, the two principal reasons for adopting the view that continuous assessment has a positive effect on pupils were its influence on motivation and enjoyment.

4. WORKLOAD

Clearly, continuous assessment has been very successful in several important areas. However, since the external examination possesses certain
inherent flaws, it seemed possible that continuous assessment might also display weaknesses particular to this mode of examining.

Examinations Bulletin 1 (1963) provided the key to a problem which is characteristic of continuous assessment when it warned that, "... teachers will need to be constantly on their guard against the danger of including so much in the subject syllabuses that there is no time left for teaching which is not directly related to the examinations" (p.27).

The present study suggests that this tendency to overload the syllabus is particularly marked in the continuously assessed course and, although external examinations are sometimes accused of displaying this fault, it appears to be especially acute in the case of continuous assessment.

The first indication that this procedure was especially onerous came from a question which did not deal specifically with continuous assessment but asked generally about the workload associated with CSE courses.

Question 11 enquired:

How great are the demands made by CSE English courses on your time and energy as opposed to the demands made by the rest of your teaching commitments?
(a) extremely high
(b) higher than the demands made by the rest of your teaching commitments, but acceptable
(c) no different from the demands made by the rest of your teaching timetable
(d) less demanding than your other teaching commitments

The demands associated with CSE were perceived to be unusually high.

Almost a third of the respondents (30%) described the demands as, "extremely high" and almost half (48%) found them, "higher than those made by their other teaching commitments but acceptable". Thus, more than three-quarters of the sample found CSE particularly onerous and this contrasted sharply with the tiny proportion (2%) who regarded CSE as less demanding and 15% who described the demands as, "no different from those made by their other teaching commitments".*

* Some respondents in this category indicated that they worked in Leicestershire Upper Schools where most classes were examination classes and where other courses also employed continuous assessment and the Mode 3 style of examining. 5% of the responses to this question were invalid.
This finding did not necessarily suggest that the demands of CSE could be attributed to the use of continuous assessment. However, the following question established this beyond doubt. Question 12 was used to identify those elements of CSE which were regarded as most onerous. It comprised the following list of activities which were known from the initial interviews to be potentially demanding and those respondents who had indicated in Question 11 that the demands of CSE were high were asked to specify which of them applied in their own case.

<table>
<thead>
<tr>
<th>ACTIVITY</th>
<th>% RESPONSE RATE</th>
</tr>
</thead>
<tbody>
<tr>
<td>(i) The planning of appropriate work</td>
<td>67</td>
</tr>
<tr>
<td>(ii) The marking and ranking of work which contributes to continuously assessed course work folders</td>
<td>91</td>
</tr>
<tr>
<td>(iii) The chasing up of work which has not been handed in, but is necessary for the completion of the course work folder</td>
<td>81</td>
</tr>
<tr>
<td>(iv) The agreement trials, cross-moderation or standardisation meetings etc. which are necessary in order to establish a departmental or group standard</td>
<td>35</td>
</tr>
<tr>
<td>(v) The sheer volume of work required by the syllabus</td>
<td>39</td>
</tr>
<tr>
<td>(vi) The volume of work necessitated by final internal moderation procedures</td>
<td>54</td>
</tr>
<tr>
<td>(vii) Please specify any other contributory factors</td>
<td>18</td>
</tr>
</tbody>
</table>

The top-scoring item was (ii) which a strikingly high 91% singled out. The comments which they made about this aspect of their work were informative:

Most members of my department feel that the English literature essays are very time-consuming ... the mark load is very heavy for each teacher. (Q341)

There's a tremendous marking load ... it's a drain on the energies of everybody. (I13)

The freedom it (i.e. continuous assessment) should give you is eliminated because the energy and enthusiasm which should go into evolving a way of teaching and preparing particular texts for presentation ... is sucked out of you by the 90 – 95% of assessment and administration ... of the mountains of written work which they're forced to produce. (I16)

The second highest scoring item also concerned continuous assessment with 81% of the respondents indicating that item (iii) applied to them.

Referring to the Mode I literature requirements, Q334 argued that:
Much time is spent in chasing children up with folder work, particularly in helping those who have been absent or missed the occasional piece of work. There is no leeway as they need to produce so many pieces of work.

Thus, the two highest scoring items were both related to continuous assessment. This meant that almost all of those who found CSE especially demanding singled out procedures relating to continuous assessment as contributing to their difficulties.*

This was confirmed by Question 15(v) which dealt specifically with the demands of continuous assessment. It stated:

The duties and responsibilities involved in continuous assessment are particularly demanding of a teacher's time and energy.

Respondents were asked to indicate their level of agreement/disagreement with this statement.** The response was one of overwhelming support with 87% agreeing that it is particularly demanding of time and energy. 38% strongly agreed indicating the strength of feeling aroused by this issue. Nobody strongly disagreed and only 9% disagreed.

The demands of continuous assessment are underlined by comparing the literature folio with the language examination. Since the language component of the Mode I syllabus is tested by a final examination and the literature component is continuously assessed, it is possible to compare the two procedures. Respondents clearly found the language examination much less onerous than the literature folio as the following comments demonstrate:

I have always found it necessary to devote about 80% of my teaching time to literature in order to cover enough material and complete the coursework. The remaining 20% is not adequate to cover the language syllabus. (Q395)

* Only 18% of the sample mentioned other contributing factors under item (vi) which suggested that the list had been fairly comprehensive. Furthermore, each additional reason which was mentioned accounted for a negligible proportion of the response rate.

** Strongly Agree (38%), Agree (49%), Uncertain (4%), Disagree (9%), Strongly Disagree (-).
I spend a great deal of time on the literature folio - probably as much as 70%. (Q084)

In practice, we find that we spend far more time on the literature ... it does tend to dominate the fourth and fifth year work. (I12)

Q056 also found that if the literature requirements were to be fulfilled, "... little lesson time is left for language work."

These comments suggested that it was not simply that the language element made lighter demands than the literature syllabus. In practice, the literature requirements were so onerous that many teachers felt obliged to neglect language in order to produce the sheer volume required by the literature folio. This suggested that the language component was more susceptible to neglect than the literature because continuous assessment demands tangible evidence of work completed throughout the course in order to secure a pass. An external examination, which does not demand, "... constant visible results" (Snowdon, 1965, p. 40) is much more prone to neglect, especially when it is vying for time with a continuously assessed element as is the case with EMREB's English syllabus. This not only explains why the respondents tended to concentrate on literature at the expense of language, it also explains why continuous assessment is an inherently more demanding style of examination. This point was underlined by I15 when she observed that whereas a conventional examination usually lasts between two and three hours, "I dread to think of the number of hours which goes into testing in CSE."

Since course work assessment is, by its very nature, a more demanding mode of examination, it is especially important for the course requirements to be pitched at an appropriate level. The evidence of the present study suggests that course work requirements are frequently mispitched. This statement is endorsed not only by a close study of the Mode I syllabus but also by the respondents' descriptions of their own Mode 3 syllabuses. For instance, I13 explained that the Mode 3 syllabus followed in his school:
... demanded twenty book reviews and twenty-four other pieces of written work for a folder ... that means a tremendous amount of work. Very few kids had done the minimum by the end of the course.

It is hardly surprising that this school was finding its own syllabus difficult to complete since there were also language requirements to be met. Commenting on another Mode 3 syllabus, IL7 remarked:

The syllabus does allow flexibility and that's marvellous ... but at the same time we have to get through so much written work that that limits us.

Syllabuses such as these led Q604 to confess that although he found the Mode 3 style of examining attractive, "... those in existence are slightly off-putting because of course requirements."

The problem was not confined to Mode 3 syllabuses. There was also widespread concern about Mode I's demand for sixteen pieces of prose of approximately 500 words each. There was general agreement that this requirement was well within the capabilities of the more able pupils even if it did take up most of the time which was allocated to English. However, difficulties arose when it came to fulfilling these requirements with less able candidates:

The less able are not really capable of churning out 8000 words ... Many of them are not really capable of producing 500 words per piece of work ... so although there needs to be some sort of minimum requirement, it is extremely difficult to fulfil. (I05)

I find it very difficult with some children to get sixteen literary-type essays and given that at the best one only gets five forty-minute periods each week, two subjects intertwined in one, plus the examination requirements for a pass, make it very difficult. (I14)

The literature folio for the less able has always been an area which has caused some schools difficulties about the amount of time which is devoted ... in order to achieve the volume of work required. (I10)

In order to determine whether these feelings were widely shared, Question 25(ii) stated:

The effort necessary by some candidates to produce the required volume of work in the Mode I literature folio (16 separate pieces of work of about 500 words each) is extremely high.
Mode I teachers were asked to indicate their level of agreement/disagreement with this statement.* A third of them (32%) strongly agreed and a further 40% agreed which meant that almost three-quarters of this group felt that the literature syllabus was making "extremely high" demands on some candidates. In comparison less than a fifth of them (19%) disagreed with the statement.

A comparison of these results with those for Question 25(ix) made it clear that it was the written course work which was the problem - not the reading requirement. Question 25(ix) stated:

The Mode I requirement that candidates should have studied seven books or an equivalent amount of material in shorter works makes appropriate demands on pupils in terms of the amount of reading they should accomplish during the course.**

Over two-thirds of the Mode I respondents (69%) agreed with this statement and a further 9% strongly agreed which meant that over three-quarters of the sample were satisfied with the reading requirement while less than a tenth (9%) were dissatisfied. The general consensus of opinion was perhaps best summed up by Q029 when he argued, "It is not the number of books which I dispute but the volume of essay work."

It seems that *Examination Bulletin 1's* warning about the dangers of overloading the examination syllabus (p.27) has frequently gone unheeded. The question this raises is: why is the workload so frequently mispitched? It seems probable that continuous assessment's image as a "soft option" is a factor contributing to this tendency to overload the syllabus. One study reported that:

---

* Strongly Agree (32%), Agree (40%), Uncertain (9%), Disagree (19%), Strongly Disagree (-).

** Strongly Agree (9%), Agree (69%), Uncertain (11%), Disagree (7%), Strongly Disagree (2%), Invalid Response (2%).
The status of continuous assessment was alluded to on a number of occasions ... Some feel that it is regarded by many both within and outside the profession as a soft option which gives the lazy and incompetent teacher opportunity to get by without the discipline of measuring up to acceptable external criteria. (Examinations Bulletin 31, 1975, p.81)

Thus, it seems probable that rigorous course requirements may be seen in part as a defence mechanism, an attempt to counter these criticisms.

IO2 regarded the teacher-examiner as another important factor in this:

... when you have a teacher-controlled examination, they will want to put into the exam everything that is desirable from an educational point of view ... an examination is a sampling process and very often if mistakes are made in teacher-controlled exams, it is because they want to push everything in.

IO2 drew attention to an important distinction which is frequently blurred in the practical setting between the teaching syllabus and the examining syllabus. An examination syllabus is necessarily selective whereas a teaching syllabus should be richer and broader. When this distinction is blurred, it appears that the effects are likely to vary according to the examining procedure. With an external examination, the danger is that the teaching syllabus will contract to become synonymous with an examining syllabus which is necessarily limited. In a continuously assessed course, there is a different danger: the teaching syllabus is liable to become overloaded by an over-ambitious examination syllabus.

5. THE EFFECTS OF EXCESSIVE COURSE WORK REQUIREMENTS

Since overloading is such a characteristic feature of the continuously assessed course, it is important to determine its effects. As far as the candidate is concerned, one probable outcome is that the enjoyment of reading is marred:

The present requirement of 16 essays for a literature folder is a penance for most middle-range and certainly almost all low-ability candidates ... which militates strongly against one's attempts to make reading enjoyable. (Q471)

... sheer volume of 16 essays is far too demanding for most pupils and frequently kills all possible enthusiasm for literature. (Q217)
Thus, even when a course has qualities which ought to make it enjoyable i.e. freedom for the teacher to choose his own texts and to devise his own assessments, this is easily undermined when the course work requirements are wrongly pitched.

Another difficulty is that it induces an unbalanced approach to the subject as a whole:

With a below average CSE group, I find it extremely difficult to fit in anything other than literature work. (Q335)

The literature course is so overbearing that it stops the children from being able to concentrate on language work. (Q334)

Mode I literature is a particularly excessive workload - a folder of 8,000 words is an effective way of losing most Grade 4/5 language candidates. (Q155)

Neglect and an unbalanced approach are the probable consequences of overloading one element of the syllabus.

The internal validity of the course is also likely to suffer. Il3, who worked in a school where the course requirements were extremely heavy, noted various ways in which this undermined the internal validity of the course:

One of the pernicious effects of the course is that teachers felt themselves to be forced into giving kids written work which was busy work. It was work that fulfilled the necessity of the course rather than fulfilled their educational needs at that stage. For example, if a kid grasped the principle of "X" he might be asked to write about it when the writing added probably nothing at all to his or her knowledge of the principle. That's one of the bad effects of having such a large demand of literacy in CSE ... (I would like) fewer pieces of work so that kids can actually go into subjects in greater depth so that instead of short, shallow arguments ... they might actually do some research on the topic and develop different kinds of writing skills ... Nothing is done that is ungraded in this place ... I'm more interested in the problems that are caused for the kids' perception of the situation. For instance, I think that because we mark every single thing that the kid does, there's far too little exploratory writing by them ... there's a pressure on them to write formed, confident solutions to things rather than to honestly express doubts or misunderstandings. (Il3)

Il3 isolated three important ways in which excessive course work requirements can undermine the internal validity of a course. Firstly, it causes written work to be undertaken not because it has an inherent educational value but as an end in itself because it fulfils the requirements of the syllabus. There
is pressure on candidates to "churn out" work in quantity regardless of its contribution to learning. This, in turn, can shift the emphasis away from quality and onto quantity thereby encouraging candidates to produce superficial work rather than exploring issues in depth. Finally, it can exclude certain types of writing.

It is not only the pupil who is liable to suffer. Evidence from other parts of the questionnaire suggested that it can become a heavy drain on the teacher's time and energy: resources which are finite. Question 13 asked those who had indicated that they found CSE courses particularly demanding: "... how often do you find it difficult to give as much extra or special preparation to the rest of your teaching commitments as you wish to give?" A small proportion of the group (14%) said this happened "frequently" while a much more sizeable group (66%) claimed that it happened "sometimes". This meant that 80% of this group found that their CSE responsibilities occasionally prevented them from giving the attention to their other teaching commitments that they wished to give. This compared with the 16% who found that this happened "rarely" and 4% who said it "never" happened.

Clearly, pitching course work requirements at an appropriate level is a critical factor determining the success or failure of a continuously assessed course. It has to be remembered that in the 1960's continuous assessment was widely regarded as a breakthrough which would release teaching from the stranglehold of external examinations. However, the evidence of the present study suggests that the tyranny of the external examination may be replaced by another kind of tyranny: that of the course work folio.

6. CERTIFICATION

The regional panels made a sharp break with existing practice in the area of certification which had become a major issue in the years preceding
the creation of CSE. Educationists had increasingly urged the need for a new form of single certification and, in the event, all but one of the fourteen regional panels decided to issue a single certificate for English. Here, then, was an area in which the regional panels had genuinely attempted to meet the demands being made by educationists. In spite of this, twelve years after its introduction into the East Midland Region, single certification was abandoned in favour of dual certification* which quickly gave way to a system of fully separated certificates in language and literature. Clearly, an important shift in attitudes had taken place during these years since by 1979 single certification had been discarded in favour of the very system which it had originally succeeded! Since separate certification was an anathema to enlightened English teachers in the early '60's, it was felt that this change was an important topic for the questionnaire to explore.

The questionnaire dealt with teaching and certification separately but in each case enquired whether respondents preferred a unified or a separate approach. Question 16(i) probed their attitudes towards teaching:

> English is one complex and indivisible discipline and therefore an integrated approach to the teaching of language and literature is most appropriate.

Respondents were asked to indicate their level of agreement/disagreement with this statement.** Altogether over two-thirds of the sample (68%) favoured the unified approach with 26% of them opting for the "Strongly Agree" category indicating strong commitment to this approach amongst a sizeable proportion of the sample. Although only a fifth of the sample

---

* This is a hybrid form of certification in which separate certificates are awarded for language and literature (as in separate certification) but entry in both areas is compulsory (as in single certification).

** Strongly Agree (26%), Agree (42%), Uncertain (13%), Disagree (17%), Strongly Disagree (2%)
(19%) disagreed with this statement, the unified approach to teaching did not command the strong support which certain items had enjoyed. Nevertheless, most teachers were still committed to the approach which was strongly recommended in the 1960's. For instance, 104 claimed:

No approach to English can be considered satisfactory if it disregards one element of English. The language and literature aspects of English must complement and inform each other.

110 agreed:

English is one subject and indivisible ... To consider doing language purely and simply on its own ... is almost the equivalent of teaching a child Metalwork Theory and never letting him near a lathe ... the values of literature in the teaching of language are enormous so that I would suggest that it is almost impossible to see how you can teach language except through the medium of literature.

This view, that good English teaching fuses the separate elements, was widely held:

... the best language teaching arises from a study of literature. Only by showing children good writing and getting them to appreciate good writing ... can you hope their own writing will develop. (109)

This desire to integrate English was also evident in the responses to Question 25(i):

A CSE literature syllabus should stipulate the inclusion of all the main branches of English literature (prose, poetry and drama).*

An impressively wide consensus emerged on this issue with almost half of the entire sample (47%) opting for the "Strongly Agree" category and almost as many again (42%) agreeing. This meant that the overwhelming majority opposed the type of syllabus which fragments literature and facilitates the neglect of any part of it. Only 4% favoured such a syllabus. The comments on this topic convey something of the strength of feeling it aroused:

... it would be a great mistake to allow teachers to omit any major branch of literature. (Q239)

Surely no literature syllabus can be called that without drama and poetry as part. (Q275)

* Strongly Agree (47%), Agree (42%), Uncertain (4%), Disagree (3%), Strongly Disagree (1%), Invalid response (3%)
... the spirit of the literature folio is ... an integrated approach to literature - not a set book approach - but a consideration of all forms of literature. (I01)

Clearly, the desire to unify English was not confined to opposing the "lang-lit" division. Teachers applied this precept to the subject generally and were not satisfied with a treatment of literature which would undermine the general principle of wholeness.

When it came to certification, a distinctly different attitude emerged. Three methods of certification were specified (single, dual and separate) and respondents were asked to indicate which they regarded as "most appropriate". Separate certification was by far the most popular method with over half of the sample (56%) indicating a preference for it.* The second most popular option was dual certification although only just over a fifth of the respondents (22%) opted for it and single certification rated as the least popular procedure with less than a fifth of the sample (19%) favouring it. Thus, although most teachers preferred an integrated approach to teaching, they nevertheless wanted separate certificates. Although these attitudes might appear to be incompatible, the teachers did not regard them as such. For instance, I05 asserted: "English is one subject and indivisible ... but as far as certification is concerned, I do want two." I01 agreed: "I think you teach it as one subject and you certificate it as three" (i.e. language, literature and oral). Thus, while attitudes towards an integrated teaching approach have changed little, there has been a noticeable shift in attitudes towards certification.

7. THE PREFERENCE FOR SEPARATE CERTIFICATION

Single certification does not appear to have been abandoned on educational grounds. There was no evidence of widespread dissatisfaction with the philosophy on which single certification was based nor a feeling

* separate certification (56%), dual certification (22%), single certification (19%), uncertain (3%)
that it was impracticable. The real considerations appear to have been political and utilitarian.

Although the CSE Boards threw almost their entire weight behind single certification, the GCE Boards continued to resist the pressure for this which ultimately worked against the interests of CSE. The fact that CSE provided only one certificate in English where GCE offered two reinforced the "poor relation" image which had afflicted CSE from the outset.

I backed separate certification ... because ... I believed that our students were not getting their due deserts when measured against people who were taking O-Level and obtaining two certificates as against their one ... This made CSE a poor relation. (I05)

I maintained from the outset ... that the best interests of pupils would be served by there being two separate gradings, if only because that was the procedure adopted by the GCE Boards. (I04)

English teachers felt that they were spending an extremely large amount of time on teaching a subject which only gave one certificate. You could spend the same amount of time teaching English for GCE and CSE but for GCE there were two certificates ... Furthermore, you also got two certificates if you spent the same amount of time on two other CSE examination subjects. (I06)

The desire for parity with GCE was the principal reason why teachers began to press for separate certification. Just as the modern schools had never really succeeded because of their failure to gain parity of esteem with grammar schools so too an examination which bore obvious inequalities when measured against GCE was not entirely accepted. As I09 pointed out:

Meeting after meeting ... people said to me: "We're doing our children a disservice because their parents want them to get two qualifications in English which they would get if they were taking O-Level but by taking CSE, even if they get a Grade I, they're only getting one."

Although it was originally argued that CSE should be free to develop its own character and approaches independent of GCE, and that under no circumstances should it become a pale imitation of O-Level, it was unrealistic to imagine that it could escape the influence of the more prestigious examination, especially those features which gave GCE the
edge over CSE in the employment market.

This desire for equality with GCE was intensified by the impending reform of examinations at 16+:

With the prospect of reorganisation of 16+ examinations and the power struggle that's likely to ensue ... it soon became apparent that one way for English to at least get some degree of comparability was to have separate examinations. (106)

Thus, the reasons why teachers became dissatisfied with single certification mostly fall under the heading of "... commercial viability" (110).

There were educational grounds for the adoption of separate certification although these were by no means as pressing. Firstly, there was the obvious anomaly of the literature syllabus carrying only 30% of the total mark when it clearly dominated the time allocated to English:

I was always opposed to the single certificate on the grounds that far more time had to be devoted to the literature element than was reflected in the distribution of marks. (Q395)

Some teachers also favoured separate certification on the grounds that language and literature demanded different skills and capacities:

It seems to me that there is a consumer demand for a certificate of competence in written English which we call English language ... I think that the study of English literature requires different skills and responses and should be certificated separately. (101)

Some teachers felt that this could actually prejudice their pupils' chances of success:

I did feel strongly that we made a mistake with insisting that English was one subject and one, therefore, would only have one certificate. I felt that a lot of our students had abilities in one direction and not in another and that this probably brought them down to a lower grade. (102)

In order to determine how widely held this view of English was, Question 16(x) stated:

English language and literature demand different skills and capacities

and respondents were asked to indicate their level of agreement/disagreement
with this statement*. Although only 7% strongly agreed with this statement, a further 61% agreed which meant that more than two-thirds of the sample believed that language and literature do demand different skills and capacities. There was, however, a substantial minority (21%) who disagreed with this statement.

8. PROBABLE CONSEQUENCES OF SEPARATE CERTIFICATION

Against the advantages of separate certification must be weighed the possible disadvantage of the neglect of literature. One of the main reasons why single certification was introduced was to ensure that all children received some literary education as part of their English studies during their final years of schooling (see p.162). It is possible that this may be lost now that the safeguard of single certification has been removed. Many interviewees feared that this would be the consequence. For instance, 112 saw the, "... devaluation of literature" as the most likely outcome of separate certification. 106 agreed: "... all it will do is undermine the position of literature."

These fears are not unfounded; even though separate certification has only been in operation for a short while, the predicted trends are already evident. There has been a decline in the entries for literature relative to those for language as the following figures demonstrate.

<table>
<thead>
<tr>
<th>YEAR</th>
<th>EMREB FIGURES FOR LANGUAGE (AS A % OF ALL CSE ENTRIES)</th>
<th>EMREB FIGURES FOR LITERATURE (AS A % OF ALL CSE ENTRIES)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1977</td>
<td>43.7</td>
<td>42.3</td>
</tr>
<tr>
<td>1978</td>
<td>45.5</td>
<td>43.1</td>
</tr>
<tr>
<td>1979</td>
<td>49.1</td>
<td>41.6</td>
</tr>
<tr>
<td>1980</td>
<td>50.3</td>
<td>39.8</td>
</tr>
<tr>
<td>1981</td>
<td>41.3</td>
<td>32.3</td>
</tr>
<tr>
<td>1982</td>
<td>42.3</td>
<td>31.7</td>
</tr>
</tbody>
</table>

* Strongly Agree (7%), Agree (61%), Uncertain (9%), Disagree (19%), Strongly Disagree (2%), Invalid response (2%).
Although the proportion of candidates entering for language rose steadily between 1977 and 1980, the opposite trend was apparent in the literature entries. Thus, it was not simply a matter of the literature entries failing to keep pace with the rising figures for language - they actually dropped. Even the most recent figures for 1981 and 1982, where there has been a decline in the language entries, show that the pace of reduction in the literature entries still outstrips the decline in language entries. Less than a third of all candidates took literature in 1981 and 1982 and this raises the question of how far the study of literature is likely to decline overall? Although it is only possible to speculate on this issue, it is worth bearing in mind that over two-thirds of the sample (68%) indicated their conviction that the proper study of English fuses language and literature into a unified whole. Teachers who adhere to this philosophy are unlikely to neglect literature as an educational goal even though they discard it as an examination goal with some pupils. Nevertheless, as Chapter One showed, the worth attributed to a subject is closely linked with its examination status and since the examining of literature is already declining, an ebb in both the status of literature and in the amount which is studied at CSE seem inevitable.

109 predicted that this decline was likely to be sharpest amongst the least able candidates:

Fewer and fewer children at the Grade 3 level and below in CSE will do any literature in their final two years at school.

This claim was supported by the respondents' suggestion that the less able pupils found both the amount and type of work required by the literature folio exceedingly difficult (see pp.189-192). Moreover, since these children are usually weak linguistically, the teacher may be tempted to concentrate on language work at the expense of literature. As 109 warned:

If people are concerned with doing "bread and butter" CSE English language, cultural deprivation could be the result. Fewer and fewer children at the Grade 3 level and below will do any literature in their final two years at school.
Thus, children at or below the Grade 3 level form the group who are most at risk from this decline. In this context, it is worth remembering that the Newsom Report (1963) urged that, "All pupils, including those of very limited attainments, need the civilising experience of contact with great literature" (p.155).

This provides a further reason why course work requirements must be pitched at an appropriate level. This is not to imply that fixing requirements is an easy task in CSE which is designed to cater for a particularly wide ability band (40th - 80th percentile). The type and amount of work which pupils at either extremity of this ability band are capable of producing is extraordinarily wide making the task of fixing satisfactory course work requirements peculiarly difficult. Nevertheless, this question demands attention if the continuously assessed course is not to act as a disincentive to the study of literature rather than a complement to it.

9. THE UNIFICATION OF ENGLISH AND DIFFERENT METHODS OF CERTIFICATION

Although EMREB used single certification for twelve years, ultimately there are questions about its success in carrying the unified concept of English through to its logical conclusion. This was a principle of the utmost importance during the formative period of CSE. For instance, Whitehead (1965) argued that a, "... unified concept of English ... nowadays underlies our best teaching of it ... the place of reading and literature (is) within a unified examination in English" (p.275). Likewise, the SSEC (1964) recommended that O-Level language examinations should be replaced by "English" examinations which tested, "... a use of language having links with the best types of English writing and encouraging wide reading of these"(p.20). However, EMREB never really effected this fusion of language and literature which was the goal of educationists. There was a written language examination accounting for 50% of the total mark, a literature
folio accounting for 30% and the remaining 20% was allocated to an oral
test. The marks for each component were simply added together to give a
final grade in English. NATE (1965) was critical of this approach arguing
that:

... an agreement to add together percentages for "language" and
"literature", as some panels propose, does not represent a
unification of English work, though it is a necessary step in
that direction. (p. 4)

NATE regarded this as an interim measure while the panels experimented with
new approaches to examining. For the future, it urged all panels to make
a "... conscious effort to develop talking, writing and reading 'as an
organic whole'" (p. 4). Clearly, NATE was not recommending the type of
cosmetic surgery which would plaster over the divisions between language
and literature. Instead it favoured the kind of radical innovations which
would eliminate them altogether. These developments have never taken place
in EMREB's syllabus which preserved the segregation of English into
conventional content areas under the umbrella of single certification.
Although this allowed teachers to treat their subject as a unified whole,
the actual form of the examination was divisive.

This has become even more prominent under separate certification which
actually strengthens divisions. Under single certification, the "lang-lit"
boundary was, to a certain extent, relaxed. For instance, until 1977, it
was possible to include a wide range of written work in the literature folio:

... a final assessment ... must be based upon the original
writing that the candidate has done ... but an allowance should
be made in the assessment for private reading, voluntary written
work and research carried out by the candidate, provided that
there is evidence of this. It should also allow for voluntary
creative writing in prose, poetry and drama as a result of the
course ... In addition, original work arising from literature
quite independent of the areas being studied (e.g. as a result of
a broadcast series, a theatre visit etc.) may be included in the
folio. (Mode 1 Syllabus, 1976, p. 5)

The form which these pieces could take was not closely defined, assignments
being described simply as "pieces of work". However, with the introduction
of separate certification came an end to this flexible approach. The
suggestion that work might be included which was not strictly interpretation of texts was removed. The form which assignments might take also was more rigidly defined after 1977 when "pieces of work" was replaced by "continuous prose" thus prohibiting all responses which did not take this form. Clearly, separate certification has brought a more precise specification of a narrower range of written activities. IO6 explained why it had crystallised the division between language and literature thus:

... one of the Schools Council and CSE regulations is that you can't receive two certificates for the same kind of work. If creative work is going to appear in Mode I language papers, then we can't really give it weight in a literature examination ... not grudgingly, we accepted that you couldn't really have creative work in a literature folio.

Question 25(iii) was designed to discover how respondents felt about the removal of the opportunity to include a wider range of responses in the literature folio. It stated:

It is regrettable that only pieces of continuous prose are considered suitable for inclusion in the Mode I literature folio, leaving out other types of response (e.g. a creative response or comprehension questions requiring extended answers).

Mode I respondents were invited to indicate their level of agreement/disagreement with this statement*. There was a marked tendency to regret this development with a quarter of the group opting for the "Strongly Agree" category and a further 42% agreeing. This meant that overall two-thirds of the group (67%) opposed this development compared with less than a fifth (18%) who did not. The majority of respondents would have preferred to use a more catholic range of written activities than recent syllabuses have allowed.

The comments which were volunteered are useful for elucidating respondents' views on this topic. Some questioned the assumption that response to literature should necessarily take the form of conventional

* Strongly Agree (25%), Agree (42%), Uncertain (15%), Disagree (14%), Strongly Disagree (4%).
The nature of the response to literature demanded by the present Mode I syllabus is too academic and restricted. Candidates who are really capable of responding in such an analytic way ... will probably be doing O-Level any way. (Q388)

II3 agreed:

The critical evaluation of literature is a terrifically high level skill. It asks first that, taking for example a story, the kids should have read the story, that they should have understood the plot and grasped something of the characterisation and then, on top of that, it asks that they should be able to identify within the text reasons for their own response ... That's an incredibly difficult distancing operation ... It's evaluative, it's certainly synthetic and it involves a lot of judgmental skills and the use of evidence as well. I think that they are incredibly difficult skills and yet that, as I understand it, is what standard CSE "lit. crit." asks of kids ... Is that a good goal to have and, secondly, are there better goals that one might have that are more suited to the age and ability group?

A better goal, as far as II3 was concerned, was a more creative approach:

... there is a very strong argument for including such things as creative responses which are particularly related to a given text and are rewarded in so far as they are seen to be a response to a particular text, so they don't just happen to be good poems or good stories but have obviously assimilated something about the nature of one text and tried, in some sense, to recreate something of that text in the new piece of writing. I think that is one of the very important ways of learning about books and relating it to one's own life ... One is hidebound in a literature course by this critical emphasis. I want kids to experiment writing their own poetry partly because in the long term I think it will develop their own response and their own critical ability ... the literary critical approach, when it is divorced from the creative side, is harmed; they both lose out.

Q073 agreed:

I fear that the exclusion of a "creative response" ... has done great harm in encouraging an inappropriately "academic" approach to literature. A genuine, personal, creative response can indicate accurately the pupil's understanding of the text.

Clearly, many of the respondents believed that the creative approach has an important role to play in evaluating response to literature.

However, as IO6 pointed out, certification is an important consideration which may be an obstacle to including this type of work. Indeed, although many respondents favoured combining an integrated approach to teaching with separate certification, the two are not compatible. Separate
certification demands that different examinations test different skills. Thus, it is difficult to see how separate certification, with its emphasis on segregation, can be combined with a unified approach to teaching which insists on integration. Ultimately, these two approaches are incompatible and the attempt to combine them in a continuously assessed course is bound to create inconsistencies.

Separate certification highlights the irrationality of treating English as an aggregation of separate components. For instance, when a recognised writer creates prose, poetry or drama, his finished products are treated as literature but when a child produces similar work, it is categorised as language. Moreover, this creative use of language is invariably required on a language paper but is rarely tested in a literature examination. Since literature is language used well, it is difficult to see what justification there can be for compartmentalising the creative use of language as "language" or "literature" or for the decision that one type falls on the language side of the barrier while another falls on the literature side. Admittedly, there is a difference between using language oneself and responding to its use by others but if this is the real distinction, I06 points to another absurdity inherent in this fragmentation:

As soon as you put an extract from a novel as a piece of comprehension on a language paper, you're really evaluating literature aren't you? That's why the division is so ludicrous. If someone has understood an extract from a novel sufficiently well to score a high mark on a comprehension, would it be unfair to say that he is sensitive to literature? It is such an artificial thing really.

This division flies in the face of our insights into the true nature of English: that English concerns both response to the way in which language has been used by others and the attempt to use it well in one's own speech and writing; that these two activities are interdependent and the decision to hive them off into separate subject areas belies the organic nature of usage/response which are compounded in a proper study of English.
If English teachers genuinely regard their subject as an organic whole, and the evidence of this survey suggests that this conviction has remained firm over the years, this must be carried through to its logical conclusion in the examining of English. Even though single certification was abandoned, philosophically it remained a sound principle and it failed only in the practical setting where it had to vie for acceptability with a less satisfactory system adopted by a more prestigious examination. Since the principle was sacrificed for reasons of expediency, single certification failed only in a limited sense and remains a worthwhile goal for the future - but to succeed, it must be adopted by the entire examinations' system, not just part.

10. POETRY AT THE CSE LEVEL: ENJOYMENT

Perhaps the most controversial and emotive area of literature teaching is poetry. There are some who adhere to the view expressed rather bluntly by Q383: "Poetry is appreciated by a small percentage of the population. Don't force it on the others" and others who agree with Q613 that: "Verse is an essential component of all literature teaching." The response to Question 25(i) suggested that this latter view was shared by the majority (i.e. 89% agreed that the examination syllabus should not permit the neglect of any major branch of literature). Nevertheless, examiners report that there is considerably less evidence of work with poetry in the literature folios than of work on other branches of literature. Although the questionnaire did not deal with the separate branches of literature individually, it was decided that since there was evidence that poetry raises particular problems (see Mathieson, 1980), it would be useful to analyse attitudes towards the teaching of poetry in greater depth. Items 25(vii) and 25(viii) dealt with two aspects of studying poetry: the difficulties created by the special nature of poetic language and the ability of pupils in the CSE ability range to derive real enjoyment from
its study.

Whereas a consensus had emerged on most issues, respondents were sharply divided on these two items underlining the controversial nature of the subject. Item 25(viii) stated:

It is difficult for the majority of pupils in the CSE ability range to derive real enjoyment from most poetry.

Respondents were asked to indicate their level of agreement/disagreement with this statement*. Although almost half of the sample (47%) disagreed with this statement, over a third (36%) agreed suggesting that the respondents were sharply divided on this issue. Moreover, although the majority disagreed, the contents of the literature folios suggest that poetry is taught less frequently than prose or drama.

Some respondents commented on the difficulty of teaching poetry regarding it as the least successful branch of literature in terms of enjoyment:

Teaching CSE poetry is a weary, hard business. (Q332)

Poetry is always the most difficult of the three forms. (Q418)

However, most of the respondents who mentioned difficulties attributed these to extraneous factors and not to the inherent nature of poetry or the inability of CSE pupils to respond. For instance, Q289 observed that, in teaching poetry, the teacher is often, "... fighting against cultural and peer group conditioning." This problem was illustrated by Q596 when he remarked: "I find many CSE pupils dislike the idea of 'poetry' but do like individual poems!" Resistance to poetry was felt to be marked amongst particular groups of pupils:

Many boys feel that poetry is "cissyish". (Q170)

It is often more difficult to teach poetry ... because so many pupils, particularly those of lower ability, are already determined to dislike it. (Q473)

* Strongly Agree (5%), Agree (31%), Uncertain (13%), Disagree (39%), Strongly Disagree (8%), Invalid Response (4%).
Q247 felt that the root of the problem lay elsewhere: "... it is difficult to teach poetry to CSE students only because during their early school years they have had inadequate exposure to poetry." These comments suggested that teachers perceived the image and status of poetry both within and outside schools as the key factor in the difficulties they experienced. It was also evident that when a teacher presents poetry to a class, he frequently has to contend with blatant prejudice. Nevertheless, some teachers were anxious to emphasize that carefully selected poetry could be successful once these prejudices had been tackled:

Poetry may well be difficult but there is much excellent modern poetry which can "open doors" for CSE pupils if carefully chosen and taught.

Q008 agreed, explaining that, "My answer is based on the body of poetry as a whole but carefully selected poetry can evoke a response." Having noted that boys regarded poetry as "cissyish", Q170 pointed out that, "... although this is their initial response, some quite enjoy poetry".

A criticism which is sometimes levelled against CSE literature is that the main criterion in the choice of texts is their superficial entertainment value. However, the respondents were almost unanimous in their agreement with the Newsom Report (1963) that: "... without it (i.e. poetry) English will never be complete" (para. 477) in spite of the obvious difficulties some of them faced in teaching it.

11. POETIC LANGUAGE

Question 25(vii) stated:

Compact and condensed language is a characteristic feature of most poetry. Thus, poetry is often difficult to teach to the CSE ability range because the language is not easily accessible*.

In responding to this item, the sample was once again deeply divided. Although over half (54%) agreed with the statement, a third (35%) disagreed.

* Strongly Agree (8%), Agree (46%), Uncertain (8%), Disagree (30%), Strongly Disagree (5%), Invalid Response (3%)
Admittedly, there was a greater degree of consensus on the difficulties created by poetic language than there had been on the issue of enjoyment but there was still a marked division of opinion. It was interesting to note that in amplifying their views on this issue, most respondents raised the same point which was not specifically concerned with the nature of poetic language but which raised another important aspect of language usage:

In discussing poetry they show insight and emotional reaction, although they find it hard, sometimes, to recapture their thoughts when trying to commit them to paper. Many pupils can express their ideas excellently but not in the written word. (Q170)

Written response difficult. (Q022)

They can respond but verbalising this is often difficult. (Q162)

Rather it is difficult for the pupil to produce a written response. (Q019)

These views were perhaps best summed up by Q607 who argued: "I feel strongly that there should be greater emphasis laid on oral work within any literature as we continually discriminate against children who have difficulties with manipulation of the written mode of expression."

This item had raised an important issue which is not only relevant to poetry: the validity of confining response to the written medium. As Q133 pointed out:

... too often we restrict CSE pupils to a medium they are not totally at home with. Other art forms could perhaps increase their enjoyment of aesthetic experiences without making them sweat out their thoughts. Spoken/taped reactions would also channel thought without making a penance out of the whole thing. Children have "English" lessons from the age of 4 - if at fourteen they are still not at home with traditionally accepted "good English" then perhaps we could make their response the more natural, acceptable one.

Their comments suggested that many teachers regarded the oral medium as particularly useful for examining literary response, especially amongst the CSE ability range, and yet most panels have never included an oral element in their literature assessments. In contrast, some do acknowledge the complementary relationship between literary and oracy in their oral examinations by including a reading test based on literary extracts. This
test of a candidate's ability to interpret the meaning of a piece of
literature is often followed by questions and discussion of the passage.
Others use the candidate's literature course as the basis for one-to-one
conversation with an examiner. These procedures implicitly acknowledge the
connection between literacy and oracy and it is interesting to speculate
about the reasons why literature examinations have failed to exploit this
relationship. By 1965, literature was firmly established as a GCE
subject which was tested exclusively by written examination. Oral
examining, on the other hand, was relatively new and this facilitated
experiments with fresh approaches. This probably made it easier to
incorporate a literary element into oral examinations than it would have
been to include an oral component in literature examinations.

The exclusion of oracy from the literature examination seems even less
tenable when it is viewed in the wider context of developing English as an
organic whole. The practice of dividing up language, literature and oracy
and, more especially, the convention that oracy is confined to the
language examination does not appear to be based on any well-founded
principle. Just as the use of literary extracts in English language
comprehension exercises underlines the absurdity of fragmenting English so
too does the practice of restricting spoken English to the language
examination.

A large Mode III consortium in the East Midland Region operates a
syllabus which is unusual in that, for many years, literature was tested
exclusively in the spoken medium and only recently has this been
supplemented by written course work. It thus provides a working model of
the uses which spoken English may serve in the evaluation of literature.
107's description of this syllabus is worth quoting in full because it
suggests several ways in which English has profited from this coalescence
of literacy and oracy:
(It was felt that) oral work in general was such a strong feature of every child's learning process that, however difficult it was to assess and moderate, nonetheless it should be attempted ... I think it's a lot easier in the discussion of literature and I think this is still one of its great strengths in that it has a specific content area which makes oral assessment so much easier ... It (also) goes back to the idea of weaker candidates being able to perform orally things which they wouldn't be able to do in written work. But as it evolved, we found that the good candidates, not just the weaker ones, were able to show genuine oral skills of a very high quality. One of the things I really like about it is that it avoids John Dixon's fears that written work can be completed without knowledge of the texts. It's very easy to know whether a candidate's copied something from the board or from a crib or from another candidate when it comes to an oral examination. It's you and the candidate and there's no dissembling, no cloaking of abilities ... it's also much easier to assess the kind of enjoyment that candidates have gained from literature over two years. It's very hard to be enthusiastic and not to seem trite in written work whereas it's so much easier because oral expression lends itself to that kind of communication.

This suggests that both literature and oracy may gain from integration. While the absence of an obvious subject matter has marred some oral examinations, literature provides a natural topic for discussion thus eliminating the air of falseness which is apt to surround an oral test.

Response to literature is also enhanced by the naturalness of oral communication which discourages the hypocrisy and regurgitation which conventional examinations foster. This evidence suggests that the removal of barriers between language, literature and oracy may bring all-round gains to the study of English.

12. TEACHER FREEDOM IN THE CHOICE OF TEXTS

CSE was designed to promote the freedom of individual teachers from external constraints:

It will be the responsibility of the teachers themselves to ensure that what is examined is what they want to teach; they will not be obliged to teach what someone else has decided to examine. (Examinations Bulletin 1, 1963, p.4)

The Mode I literature syllabus which best captured this spirit of freedom was that devised by EMREEB:
Teachers are to devise courses on areas of literature of their own choosing: the range of a particular area and the breadth or depth in which it is studied depending upon the interests and abilities of their pupils. (Mode I English Syllabus, 1965, p.45)

EMREB made the two main areas where examinations influence teaching (i.e. the teaching syllabus and the final assessment), the responsibility of class teachers. Teachers were free to select their own course content and to tailor the assessment to it. Indeed, such was the freedom afforded by this syllabus that IO1 described it as, "... a Mode 3 element in a Mode I examination."

Since teachers in the East Midland Region have enjoyed an unprecedented degree of freedom, it was felt that this was an important area for evaluation. Teachers' attitudes towards freedom in the choice of texts were probed by Question 25(iv):

The EMREB approach to literature is particularly appropriate to the CSE ability range because it allows the teacher freedom to choose literature which is appropriate to his pupils' needs.

Respondents were asked to indicate their level of agreement/disagreement with this statement*. Their responses suggested that the overwhelming majority of teachers welcomed this freedom. Over half of them (56%) agreed and a further third (31%) strongly agreed indicating both the degree of consensus and the strength of feeling on this issue. A further indication that this freedom was highly prized was provided by the warm tributes volunteered by many respondents:

I think the last fourteen years have been very valuable. It's given the East Midland teachers a freedom that I don't think any other region has had and I think that is very good from the classroom point of view. (IO2)

The joy of CSE literature was that teachers could select reading material which they believed to be appropriate to the children they were working with or indeed which was their own strength. (IO9)

We have plenty of freedom to choose what we want to do ... Therefore, the staff do things that they enjoy and that the children enjoy and there's a motivation within that to read a lot. (IO2)

* Strongly Agree (31%), Agree (56%), Uncertain (6%), Disagree (2%), Strongly Disagree (1%), Invalid Response (4%).
One advantage of this flexibility, pinpointed by Q207, was that the ability to choose his own literature, "... gives the teacher the opportunity to assess individual needs rather than 'class' needs ... and to keep interest from flagging." Whereas the conventional examination encourages teachers to treat their classes as "units", a teacher working under this type of syllabus is able to take individual needs into consideration. Perhaps the warmest tribute of all came from I01 who assessed the wider significance of this freedom thus:

Schools began to build up stocks of the kind of books that enlightened English teachers wanted to teach and CSE was a factor in enabling them to do this. That was one of the spin-offs onto the curriculum that the examination had. I think CSE, as a concept and as an actual examination, did more to raise teaching standards in the '60's and '70's than any other single factor in education.

The freedom which this syllabus has conferred upon practising teachers is clearly a major factor in its popularity. Nevertheless, this should not be taken to imply that teachers valued this freedom without question or qualification. The value they attached to freedom in the choice of texts was tempered by a consciousness of the need for constraints in other areas. This delicate balancing of freedom and constraint is illustrated by comparing the responses on Question 25(iv) above and Question 28. Question 28 listed various types of literature syllabus and respondents were asked to indicate which one they preferred. The response rates are given in brackets after each option.

(i) Complete freedom for the teacher to choose the literature most appropriate to his pupils' needs (13%)

(ii) Freedom for the teacher to choose the literature most appropriate to his pupils' needs providing that he must include all the main branches of literature (prose, poetry and drama) (67%)

(iii) A number of set texts to be chosen from a syllabus of prescribed texts (3%)

(iv) A syllabus of prescribed authors which provides freedom concerning the actual works studied (2%)

(v) A syllabus which takes a thematic approach outlining a number of themes and suggesting relevant material (9%)

Invalid response (6%)
Despite the high valuation of freedom in the choice of texts, item 28(i) was surprisingly unpopular with only 13% of the respondents opting for it. By far the most popular alternative was item 28(ii) which was similar to item (i) except that it imposed one important constraint on the freedom of the teacher ("... providing that he must include all the main branches of literature"). Five times as many respondents (67%) opted for this item as had opted for item (i) demonstrating that although teachers prized the freedom to choose texts appropriate to their own enthusiasms and their pupils' interests, they did not approve of the kind of unbridled freedom which would allow teachers to neglect whole areas of literary experience. Thus, it seemed that they favoured a kind of structured freedom which prevented the neglect of any major branch of literature.

Question 28 also highlighted the extreme unpopularity of any form of prescription. Two types of prescription were specified, one (item iv) offering distinctly more freedom than the other (item iii). Both proved to be exceedingly unpopular with 3% opting for item (iii) and 2% opting for item (iv). The final alternative attracted a slightly better response although it still only accounted for 9% of the sample. The lack of support for this approach was interesting bearing in mind that the original syllabus had been strongly committed to the thematic approach recommending it from 1965 to 1976. However, in the 1977 syllabus, this emphasis was greatly weakened and, in the following year, all mention of it was removed.*

It seems that teachers who had enjoyed the freedom to present literature in whichever way seemed most appropriate to them disliked the imposition of a thematic approach and that this was reflected in the panel's decision to discontinue recommending it.

* In the 1976 syllabus, over 50% of the space devoted to literature was concerned with suggesting themes and appropriate texts. In 1977, this material was removed although the thematic approach was still referred to. In 1978, all reference to the thematic approach was removed.
13. **TYPE OF SYLLABUS PREFERRED**

In the pre-CSE setting, GCE was criticized for failing to make the study of literature relevant to the interests and experiences of its adolescent candidates. The syllabuses displayed a rigid concept of "literature" relying almost exclusively on pre-twentieth century texts where the language and concerns were inaccessible for many pupils.

With the introduction of CSE, a different approach to reading emerged. In fact, this was the area where practising teachers had perhaps the greatest impact on CSE literature. Teachers rushed forward with suggestions for more appropriate reading materials - suggestions which greatly extended the range of examinable literature and shifted the emphasis away from traditional and on to modern literature. When the original syllabuses were published, the sheer variety of recommended texts and, more particularly, the reaction against traditional literature, caused consternation in some quarters (see, for instance, Butts (1966, 1967), Hipkin (1967) and Holbrook and O'Malley (1965)).

Since this was a controversial area at the outset, it was felt that the questionnaire ought to ascertain recent opinion on this issue now that practices have become established. Therefore, Question 26 enquired:

The majority of pupils in the CSE ability range benefit most from studying:

(i) mostly modern literature written during the twentieth century  
(77%)

(ii) mostly traditional literature written prior to the twentieth century  
(-)

(iii) equal proportions of modern and traditional literature  
(19%)

The figures in brackets denote the response each option received. The responses revealed overwhelming support for the modern approach with over three-quarters of the respondents (77%) opting for (i). This contrasted with the extreme unpopularity of the "mostly traditional" approach which nobody preferred. Although nobody favoured a course with a traditional
bias, a substantial minority (19%) preferred equal proportions of modern and traditional literature*.

The importance attributed to modern literature at the outset has been vindicated as far as the teachers in the sample were concerned. There appear to have been two principal reasons for their preference for modern literature: its relevance to the experiences and concerns of adolescents and the accessibility of its language:

Accessibility of language is of great importance ... For this reason ... literature of the twentieth century is generally most successful. Subject matter is crucial; the candidate needs to feel immediate personal response which can be hindered if language is a barrier. (Q415)

Some traditional works may be mixed with modern ones, but pupils sometimes are reluctant to make an effort to cope with unfamiliar material. (Q59)

Q289 also argued the importance of reading "... which can be made to seem relevant to their own lives" claiming that, "... this is usually more easily done with modern literature." Thus, most teachers were satisfied to place the emphasis on modern literature and although traditional literature had not been rejected entirely, most teachers clearly believed it had a limited role to play in CSE. This is confirmed by a brief perusal of the Chief Examiner's Reports which use the contents of the literature folios to document the popularity of different texts. These popularity ratings reveal a strong bias in favour of modern literature. The reasons for this were perhaps best summed up by Q607:

I believe our primary aim must be to develop an interest in reading and a love of literature in all pupils and I feel this is best achieved through twentieth century literature.

Until 1977, the Mode I literature syllabus allowed teachers to decide whether to follow an intensive or an extensive course of reading:

Teachers are to devise courses on areas of literature of their own choosing, the range of a particular area and the breadth or depth in which it is studied depending on the interests and abilities of their pupils. A detailed study of a few texts

* 4% of the responses to this question were invalid.
having some common ground or dealing with some central theme is as acceptable as one which draws together a wide range of material. (Mode I Literature syllabus, 1965, p.45)

However, in 1978, the syllabus was revised:

Candidates will be expected to have studied seven books or an equivalent amount of material in shorter literary works.

(Mode I Literature syllabus, 1978, p.9)

Since teachers had formerly had greater freedom in this matter, Question 27 enquired about their preferences:

The majority of pupils in the CSE ability range benefit most from studying:

(i) a small number of texts (e.g. 4 or less) in great detail (3%)
(ii) a larger number of texts (e.g. 5 or more) with less emphasis on detail (67%)
(iii) one or two texts in depth plus a number of other texts read less closely (26%)

Only 3% believed that CSE candidates benefitted most from (i), the intensive approach, while the majority (67%) agreed that candidates should be required to study at least five texts. A further 26% found it valuable to combine the two approaches*.

14. CONCLUDING REMARKS

Taken as a whole, this analysis provides a clear picture of the type of literature syllabus which was favoured. Generally speaking, respondents preferred a syllabus which gave them freedom in the choice of texts with the proviso that all the main branches of literature (prose, poetry and drama) should be included. This was the only constraint on their freedom to devise their own courses which the teachers approved and thus there was little sympathy for the thematic approach and even less for prescription. There was general agreement that literature should be treated as an integral component of a unified English syllabus and that separate certification was preferrable to single certification. As far as the

* 4% of the responses to this question were invalid.
course content was concerned, respondents liked to read extensively rather than intensively choosing from a varied range of modern literature and giving less attention to traditional works. When it came to assessment, there was immense enthusiasm for the course work approach and little support for the external examination. However, there was widespread concern about the difficulties of pitching course work requirements at an appropriate level and, in particular, about the consequences of an excessive demand for written work.

If the sorry state of literature examining up to 1965 is compared with what has been achieved through the folio approach, EMREB's experiment undoubtedly has been a great success. This claim is endorsed by the overwhelming support for continuous assessment amongst the respondents. The principal complaint against conventional examinations was their poor level of internal validity as a means of testing response to literature. The folio approach has enhanced the validity of literature examining in a way which no other procedure has proved capable of doing. Moreover, whereas the conventional examination detracts from the pleasure of reading, there was a strong sense of the gratification which could be derived from this approach by pupil and teacher alike.

Although learning experience assessment is highly successful, it is not an unqualified success. The principal difficulty has been in pitching course work requirements at a satisfactory level. The consequences of inappropriate requirements can be serious: the validity of the assessment may be undermined and what might have been a pleasure becomes a chore. It may also exclude less able candidates whose teachers are tempted to concentrate on language rather than tackling an over-demanding literature syllabus. This enquiry suggests that it is dangerously easy to replace the tyranny of the external examination with the tyranny of the course work folder. Despite these disadvantages, the EMREB approach incorporates features which represent a huge step forward in examining literature.
IO1 conveys something of the special achievements of this unique experiment and so perhaps the final word should go to him.

It seems to me that the great achievement of the East Midland English Panel was the concept of the literature folio. Unlike other CSE Boards which got trapped into the old set books problem, the East Midland Board developed this concept of the literature folio which gave teachers maximum freedom in the literature that they chose to do with their pupils ... Schools began to build up stocks of the kind of books that enlightened English teachers wanted to teach and CSE was a factor in enabling them to do this ... I think CSE, as a concept and as an actual examination, did more to raise teaching standards in the '60's and '70's than any other single factor in education ... the great achievement of CSE was to make the normal procedures of class teaching accessible to assessment ... in so far as CSE strays from that, it strays from one of its most important functions.
CHAPTER FIVE

ORACY

A. INTRODUCTION

1. THE HISTORICAL CONTEXT

To understand the current state of oral examining, its status and perceived role, it is vital to know something about the history of neglect which has undermined its standing in recent years. There are many signs of this neglect. For instance, as recently as 1965, Wilkinson found himself hampered by the lack of a vocabulary with which to describe oral skills. He attributed this "lexical poverty" to the neglect of the subject:

The neglect of speech in our training of young people is only too obvious when one attempts to get a vocabulary in which to describe its various aspects. For instance, there exists no term for the ability to use the oral skills of speaking and listening; and conversely the fact that no term exists for the concept has meant that people have been unaware of its importance. (p.13)

Even the term "oracy", which is now in common use, was coined by Wilkinson in an attempt to provide a word which describes, "... general ability in the oral skills." (p.14)

Another sign that, "The spoken language in England has been shamefully neglected" (Ibid, p.11) was provided by Hitchman (1964) who reviewed the research on testing spoken English. He found that most of the work undertaken was limited to Public Speaking and, even here, he concluded that, "... the reported research is meagre" (p.56). Moreover, most of the work was of American origin. If the amount of research devoted to a topic may be regarded as one measure of the level of concern it has aroused, the absence
of British work indicated a serious lack of attention.

One of the important achievements of CSE, therefore, was the boost it gave to oral examining which became established on a national scale under CSE. There was an inevitable sense of venturing into uncharted territory about this development. As Examinations Bulletin 11 (1966) confessed: "There is very little material available for consultation on oral examining" (p.32). This raises the question: why is it that spoken English had been so "shamefully neglected"?

2. THE TRADITIONALLY LOW STATUS OF THE VERNACULAR

Contempt for the vernacular stretches back a long way in English education. In fact, Compton (1941) traced its roots back to the pre-Renaissance era:

We have never in our Secondary Schools succeeded in casting off the burden of the narrow, grammatical, classical curriculum which... became more and more confined, as time went on, to the study of the grammar of the classical languages and the reading of a few representative texts... there had never been established a tradition of the importance of sound teaching in the mother tongue - even in the oldest grammar schools before the time of the Renascence... English had suffered an eclipse of some three hundred years after the Norman Conquest... (and) it is clear that the prestige of the English Language was not well enough established by the time of the Renascence in England to enable it to stand up against the claims of the classicists at the time of the endowment and re-endowment of the many schools founded or modernized during the late fifteenth and sixteenth centuries. The effect of this disparagement... was to stress the importance of the literary and written language over the spoken word. (pp.8-10)

Although secondary education was available for only a small minority of the population until the middle of this century, even when the lower classes were provided with an elementary education, during the nineteenth century, the low valuation of spoken English persisted:

In the "Manual to the System of Teaching... in the Elementary Schools of the British and Foreign Schools Society", published in 1825, we find the advocacy of a minimum of conversation between teacher and taught, and an attempt to mechanize, as far as possible, the whole teaching situation. There was no attempt to provide good models of speech, or to promote fluency in
"A small handbell is used to direct the movements of the children, in order as much as possible to avoid verbal commands." (Compton, 1941, p.12)

The elementary schools were, of course, extremely overcrowded and this made it impractical to allow individuals to speak at length. Indeed, Compton cited a prescribed reading lesson as an example of the expedients which had to be employed in the face of extreme overcrowding. It recommended that teachers should write a sentence on the blackboard and, in order that each child may take part, each pupil should read a word. Clearly, oral skills were undervalued throughout the education system.* By the twentieth century, the vernacular had accumulated a long history of neglect and contempt which has influenced recent attitudes towards it.

3. EVIDENCE OF CHANGING ATTITUDES

In the present century, a significant shift in attitudes has occurred as educationists increasingly came to recognise the need for good spoken English. Official reports throughout this century have urged a larger place for oracy in school curricula:

... the first and chief duty of the Elementary School is to give its pupils speech. (The Teaching of English in England, 1921, p.60)

Undue emphasis has hitherto, in many instances, been laid on written exercises ... One of the chief aims of the course should be to secure clear and correct speech. (Hadow Report, 1926, pp.190-192)

Admittedly, the concept of oracy prevalent at the beginning of this century smacked of elocution training, but at least oral skills of some kind were being given greater recognition. Again, in the 1940's, the Norwood Report (1943) argued that:

It happens too often that little stress is laid upon oral expression ... we would invite the earnest attention of all

* In contrast, oral assessment had always played a part in higher education. In fact, it was not until the eighteenth century that written examinations started to replace viva voce examinations for honours degrees.
teachers to the importance of giving full opportunity to their pupils to hear English spoken ... and to utter English themselves. (pp.94-95)

Likewise, the Ministry of Education Pamphlet 26, Language: Some Suggestions for Teachers of English and Others (1954) recommended a "... more systematic" (p.75) training in oracy.

By 1963 the concept of oracy had changed:

The overriding aim of English teaching must be the personal development and social competence of the pupil. And of all the different aspects of English, speech has by far the most significant contribution to make towards that development. Inability to speak fluently is a worse handicap than inability to read or write ... This is not essentially a matter of accent or pronunciation ... far more important is the need to ensure that they can speak easily, clearly and with interest, and have something to talk about. Personal and social adequacy depend on being articulate. (p.153)

Thus, by the 1960's, the emphasis on elocution had been replaced by a concept of oracy as the tool for intellectual development and a pre-requisite for social and personal competence. Tributes to the importance of oracy grew in strength and number at this period as the decision to create CSE stimulated attention to innovative techniques of assessment.

Official reports, throughout this century, urged a more important role for spoken English. Although, in one sense, this illustrated the growing importance attached to oracy, in another, it underlined its persistent neglect. Since report after report found it necessary to recommend a greater use of oral work, this suggested that the recommendations were not being heeded. Why was it that despite the increasing recognition of the importance of oral competence, spoken English was still neglected in schools?

4. A MISMATCH BETWEEN EDUCATIONAL GOALS AND EXAMINATION PRACTICE

To answer this question, it is necessary to bear in mind the relationship between a subject's examination status and its curricular status:

In education to test a subject officially is to recognise its importance and to give it prestige. (Hitchman, 1963, p.53)
However, this relationship is not a simple, one-way process because that which is examined axiomatically achieves status. Thus, there tends to be a relationship between the examination "weight" carried by a subject and the time devoted to it in schools. The prestige of a subject is inextricably associated with its examination status, examined subjects achieving an exaggerated importance at the expense of non-examination subjects. Even when educational values shift, as in the case of spoken English, if examination practice remains unchanged, the impact which an innovation has on school curricula tends to be limited.

It is this discrepancy between practice and principle which has hindered the progress of spoken English for most of this century. There has been a mismatch between the aims of educationists and the practices of the examinations' boards. For the first half of this century, public examinations concerned themselves exclusively with written competence. Oracy finally broke through the public examinations' "barrier" in the 1950's when there was a surge of oral examinations. In 1953, the English Speaking Board and the Northern Universities' Joint Matriculation Board pioneered some of the earliest tests in spoken English and these were followed a year later by the Associated Board of the Royal Schools of Music. Several other authorities also established oral examinations during this decade (e.g. the Union of Educational Institutes, the Guildhall School of Music and Drama, the Royal Society of Arts and Durham University Schools' Examining Board). Although, in one sense, this period of sudden growth represented a turning point, in another, these examinations made only a limited contribution to the rising status of oracy because they were mostly offered by non-GCE authorities and, therefore, lacked the academic lustre of O-Level. Admittedly, two of the afore-mentioned bodies were GCE boards but the way in which they treated oracy did little to enhance its academic respectability. For instance, the examination offered by the NUJMB grew out of an initiative pioneered by a Liverpool grammar school. Since grammar
school pupils normally sit GCE examinations, the Headmaster naturally requested that his spoken English examination should be incorporated into GCE:

An approach was made to the Northern Universities' Joint Matriculation Board, inviting them to conduct an examination in oral English as part of the General Certificate of Education. The request was turned down but on submission of details of the scheme used internally an offer was made to conduct an experimental examination with the issue of a form of certificate, but not as part of the GCE. (Baxter, 1952, p.651)

Thus, the NUJMB agreed to examine oracy but not to accord it GCE status.

The NUJMB was not alone in denying spoken English O-Level status. From 1956 until its closure in 1964, Durham University Schools' Examining Board operated a successful oral examination which experienced a steady growth in candidature. Indeed, when the university announced its intention to close the board, the schools which had been using the oral test arranged for its continuation under the Oxford Delegacy and the Cambridge Syndicate. In spite of the support for this examination from schools, "The Durham test was never a GCE test, nor was it tied to any written paper" (Wise, 1964, p.71). Thus, although oral assessment was becoming more widely available in the 1950's, there was a dangerous trend to its development. It was becoming established as an inferior discipline denied GCE status. In 1965, Wilkinson felt it necessary to assert that oracy "... is not a frill" (p.58). Wilkinson's statement is understandable in view of the treatment of spoken English by the GCE boards.

Spoken English was finally introduced into O-Level English in 1964 (as part of the London Board's Language examination). However, it was not treated as an integral element of English but as an optional extra. A candidate who took this additional element received a certificate entitled, "English Language with Spoken English" with the implication that oracy was an appendage which could be "tacked on" to the end of an award. Furthermore, although the written and oral components were independent, candidates were entitled to a certificate on the strength of a pass in the
written papers regardless of their performance in the oral tests but those who passed the spoken English test only received a certificate if they also passed the written examination. Clearly, oracy still had a long way to go to achieve real academic respectability.

5. THE CONTRIBUTION OF CSE

The commitment to spoken language which had been growing throughout this century finally found expression through the creation of CSE. In fact, it is no exaggeration to describe CSE as the real turning point in oral assessment. CSE was the first national examination freed from the vested interests of the university boards. Operated by practising teachers, it was free to serve their interests and this produced a climate favourable to innovation:

The introduction of the CSE with its massive confrontation of teachers with problems of assessment has produced a climate for reform. (Eggleston, 1967, p.58)

This stimulated fresh thinking about techniques of assessment:

CSE ... burst as a tremendous, great searchlight upon the darkness of the examining world. It has made people more aware of what they are trying to do in an examination, what they're trying to test and measure. I don't think it was ever really considered at all properly in the old O-Level days. CSE started right from the very outset by trying to improve examining techniques and I think it has made a tremendous contribution in that respect. (108)

Examinations Bulletin 1 (1963) presented it as a "challenge" (p.26) to teachers to create: "... new and better means of testing" (p.3). This was particularly pertinent to English:

Previous examinations in English are not much help ... First, they test only a limited range of skills. They offer for assessment ... the writer of one type of essay, the comprehender of one type of passage, the reproducer of a few set books. Even this narrow range of skills is not fully assessed ... Why assess oral ability? First, because we have agreed in principle to assess performance over the whole range of language.

(Dixon, 1965, pp.1 and 29)

Thus, the pent-up enthusiasm for spoken English was unleashed. Indeed, as Allen (1977) observed: "In the early days of CSE one of the most fiery
enthusiasms was for the inclusion of an oral assessment" (p.15). This was reflected in the syllabuses of the regional English Panels.

Examinations Bulletin 1 (1963) had expressed the hope that, "... no examination which is concerned solely with written work will be considered adequate for the Certificate of Secondary Education" (pp.40-41). In the event, there was universal agreement amongst the sixteen Mode I panels that CSE English should include an oral component.* It is remarkable that there was not a dissenting voice amongst the panels on this question. However, as the syllabus of the London Board had demonstrated, merely to include an oral test did not guarantee its status – a more accurate indication of this was provided by its standing within the examination. Would CSE perpetuate the treatment of oracy as a "frill" or an optional extra to English? Again there was an impressive degree of consensus amongst the panels with all sixteen agreeing that spoken English should be compulsory.

The commitment to spoken English did not stop short at this point but was underlined by the mark weightings allocated to it. Sometimes compulsory items are allocated such small mark weightings that their contribution to the examination as a whole is negligible. Recognising this, Examinations Bulletin 1 (1963) stressed that oral tests, "... should form a substantial portion of the English examination, accounting for between one-third and one-half of the total marks" (pp.36-37). In the event, almost all of the panels opted for a smaller mark weighting. Hitchman (1968) reported that fourteen of the panels quoted the percentage of marks they allocated to oracy and that this ranged between 20% and 30% with an average of 23%. Thus, no panel had awarded less than a fifth of the marks to spoken English. Although this fell short of Examinations Bulletin 1's recommendation, it was

* These figures are taken from a survey carried out by Hitchman (1968). Certain regions were sub-divided into separate areas with their own English panels which meant that there was a total of sixteen, separate Mode 1 syllabuses in operation.
by no means an insubstantial proportion of the overall mark. Although this could be attributed to a continuing undervaluation of oracy, it is more realistic to attribute it to natural caution. Most teachers were novices to the business of examining while even experienced examiners were likely to have had little or no experience of oral assessment. In this light, the decision to allocate an average of 23% of the marks to spoken English suggests a commitment to oracy tempered by natural caution. This interpretation certainly seems to reflect the feelings of people who served on the original Regional Panel in the East Midland Region. Referring to the decision to include an oral examination carrying 20% of the marks, IO3 observed:

It would have been so easy to let that go because it was such a difficult thing to run in those days but we kept it in ... and this was an achievement.

IO9 agreed:

The fact that we had the oral component as part of English Language for CSE was a major breakthrough ... Because we weren't sure, we didn't run away from it.

Although in one sense, CSE represents an important milestone in the history of oral examinations, these achievements made little impact on GCE which, for the most part, continued to neglect oracy. Thus, oracy became even more firmly established as mostly beyond the pale of GCE, genuinely accepted only within the sphere of CSE. In one sense, CSE has done oral English an important service, in another, its unacademic image has been confirmed.

CSE has been in operation for nearly two decades and although its approach was revolutionary in 1965, oral tests are now an established feature with which practising teachers are fully acquainted. Therefore, it seems timely to enquire how teachers feel about oral examining in the light of this experience. After that initial surge of "fiery enthusiasm", how has oral examining worked out in practice? What are "... the lessons and achievements" (Dixon, 1977, p.8) to be drawn from this work? The
following section attempts to provide some answers to these questions.

B. THE FINDINGS

1. LEVEL OF COMMITMENT TO ORAL EXAMINING

In 1977, Allen observed that although:

In the early days of CSE one of the most fiery enthusiasms was for the inclusion of an oral assessment ... if my antennae are working aright, the mood has changed to something approaching either parochial indifference ("I can do my own thing") or doubt about the inclusion of any oral assessment. (p.15)

This observation was based on the JMB/TWYLREB 16+ scheme from which spoken English had been omitted:

Though there are some participants who are asking for its inclusion, there are many who prefer things as they are and others who seem not to care either way. If the point is made that the more enlightened schools tend to join this 100 per cent course work scheme, a large shift in opinion since early CSE days is indicated. (p.15)

Allen suggested a major shift in attitudes in the West Yorkshire and Lindsey Region making it important to enquire whether a similar reaction had taken place in the East Midland Region?

The evidence of the present study suggests that such a reaction has not occurred in the East Midland Region. Neither "parochial indifference" nor "doubt about the inclusion of any oral assessment" featured as prominent attitudes among the sample. On the contrary, there was firm commitment to oral examining which was perhaps most strikingly illustrated by the responses to Question 7 which listed various procedures employed by CSE and enquired which the respondents would wish to retain in a revised 16+ examinations system. The resulting responses are indicated in brackets.
7. If the examination system at 16+ is revised, which aspects of the present examination procedures would you like to see retained for the C.S.E. ability range? ... Please place ticks in the Language boxes of all of those items which you would like to see included in Language examinations at 16+.

<table>
<thead>
<tr>
<th>FOR LANGUAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>(i) The opportunity for teachers to submit their own Mode III syllabuses (70%)</td>
</tr>
<tr>
<td>(ii) Optional external syllabuses devised by the examining board (53%)</td>
</tr>
<tr>
<td>(iii) Democratic teacher control at all levels of the examination procedure (62%)</td>
</tr>
<tr>
<td>(iv) The assessment of oral ability (92%)</td>
</tr>
<tr>
<td>(v) An externally prescribed syllabus, but one which allows wide choice (52%)</td>
</tr>
<tr>
<td>(vi) A narrowly prescribed syllabus (7%)</td>
</tr>
<tr>
<td>(vii) A content-free syllabus involving guidelines as to the aims and objectives of the course (46%)</td>
</tr>
<tr>
<td>(viii) The assessment of aural ability (74%)</td>
</tr>
</tbody>
</table>

The most striking feature of these responses was the high level of commitment to oral assessment. Whereas most items spread the respondents more or less evenly between the desire to retain or abandon an element, there was almost total consensus (92%) about retaining oral assessment. Clearly, respondents were united in their continuing commitment to oral examining.

At the same time as Allen (1977) identified a change in attitude in the TWYLREB Region, EMREB was replacing single certification with a system of dual certification.** This meant that 30% of the marks, which had previously been allocated to literature, were unaccounted for and the way in

---

* The conditional verb tense was used in this question because the GCSE proposals had been shelved by the government at the time when the questionnaire was despatched.

** This is a hybrid form of certification in which separate certificates are awarded for language and literature (as in separate certification) but entry in both areas is compulsory (as in single certification).
which these "spare" marks were apportioned between remaining components (i.e. written language and oracy) provided an important indication of current attitudes towards oracy. Originally, oracy had accounted for 20% of the marks in a three-component examination. After dual certification, it accounted for 30% of the marks in a two-component examination. Had spoken English not received any of the spare marks, its weighting within the examination would have been effectively reduced lending support to Allen's claim that the popularity of oral assessment was diminishing. The way in which the spare marks actually were distributed appears to have been intended to preserve the original balance between written and spoken language rather than to alter it. Just as in the original syllabus written language had been the principal component accounting for 50% of the total mark, so too in the revised syllabus written language remained the major element accounting for 70% of the marks. EMREB has continued to strike a delicate balance in its mark weightings so that oracy has remained a substantial but never a principal component of the English assessment. The accumulated evidence suggests that the level of support for oracy, rather than faltering over the years, has remained constant.

When the respondents' comments were analysed, three principal reasons emerged for this commitment. Firstly, it was argued that spoken language plays such a vital role in a child's language development that any English assessment which ignores it must be inadequate. The belief that an English examination which omits oracy is invalid was widely held.

I feel strongly that oral ability is such an integral part of language - the way a person thinks and expresses himself - that greater attention should be given both to examining means by which this can be stimulated in the classroom and assessed accurately. (Q609)

The oral examination is not an easy problem but for an examination to be thorough in all aspects of English it is a "must". (Q585)

Oral examining is vital to a rounded assessment. (Q602)
... oral work is such a strong feature of every child's learning process that however difficult it is to assess and moderate, nonetheless it should be attempted. (I07)

I cannot see a language used purely and simply with a pen in your hand. (I10)

Another argument frequently forwarded was the importance of the spoken word for the CSE ability range:

CSE children are often articulate verbally: their lives are strongly based on verbal communication ... surely speech ought to be a greater criterion for assessing perception and awareness of language and literature - especially literature? (Q133)

I feel oral work and assessment are vital to CSE courses in English as so much communication for these students will be by word of mouth in their own lives and future employment. (Q315)

The majority of children, when they leave this school, will do very little writing. Many will do very little reading but they will do a great deal of listening and a great deal of talking; so one of the chief functions of the English teacher should be to train them for those functions in life. (I09)

Respondents also stressed the role which speech plays in the modern world arguing that examinations should reflect the importance attached to spoken communication. For instance, Q604 argued that more marks should be allocated to oracy because of:

... the ever increasing emphasis laid on spoken language in the modern world. A living language reflects such shifts in emphasis; an assessment procedure aimed at estimating ability to use such a language should make the appropriate changes for the sake of accuracy.

Oral ability seems more important in (the) present world - less writing needed. Exam boards can influence teachers - pressure them to give more teaching time to oral work - by including oral assessment as (an) important part of the examination. (Q007)

Clearly, many teachers wished to eliminate the "... gap between a school world, where writing is all-important, and the real world, where the greatest need is for skill in speech" (Dixon, 1965, p.29).

Despite their commitment to oral examining, it would be wrong to imply that the respondents were complacent. They were, in fact, acutely aware of the difficulties raised by oral assessment and this is perhaps best illustrated by their responses to Question 21 which enquired what proportion
of the total mark they believed should be allocated to oracy. Although this question referred to Mode I, it was addressed to the entire sample because it was important to know how all teachers felt about the weighting of spoken English. The figures in brackets indicate the response rate for each option.

21. The Mode I Language Examination allocates 30% of the total mark to the oral assessment. This allocation:

   (i) represents an adequate proportion of the total mark (59%)
   (ii) should be increased (5%)
   (iii) should be decreased (34%)
   (Invalid response (2%))

Although firmly committed to oral examining (i.e. 92% had indicated that it should be retained in a future 16+ examination), most respondents believed that its mark weighting ought to be carefully controlled. Only 5% felt that oracy should account for more than 30% of the total mark, and while the majority (59%) were happy for it to carry 30% of the marks, a substantial minority (34%) believed that this should be decreased. Clearly, most respondents felt that oracy should play only a limited role in the language assessment. If these preferred mark weightings may be regarded as a rough guide to attitudes towards oral examining, their essential caution suggested that most teachers had reservations about assessing spoken English.

2. **THE SPECIAL NATURE OF THE ORAL MEDIUM**

The reasons why teachers were reluctant to allocate too many marks to spoken English concerned the special problems raised by examining in an oral rather than a written medium. Nowhere are the problems inherent in English assessment (see pp. i-vi) more acutely present than in the examination of oracy which poses all of the characteristic problems of English in a particularly serious form:
"it suffers from all of the weaknesses and ills (... subjectivity, relative unreliability etc.) of the essay examination, but in a more acute and less curable form."
(Wood quoted by Trimble, 1934, p.550)

The special difficulties associated with oral examining may be illustrated by comparing the written composition with the oral composition. Hitchman (1968) describes as forms of oral composition, "... conversation, group discussion and the giving of a talk" (p.224). The term "composition" aptly describes such activities because regardless of whether the candidate is operating in the written or the verbal medium, he must construct language to convey meaning. Therefore, the essay and the conversation are examples of the same type of test carried out in a different medium and they provide a useful means of comparing spoken and written English.

Many of the problems posed by essay assessment also occur in the conversation test. Firstly, the nature of the response is the same. Regardless of whether a candidate is required to write an essay or to hold a discussion, it is his unique, personal response which is elicited. This predetermines the nature of the assessment which can take place. In an objective test, it is possible to specify the response required of a candidate but in a free-response test there is no such thing as a single, correct answer. A wide variety of responses may be acceptable and, although it is possible to give some sort of judgement on their degree of excellence and appropriateness, there is no such thing as an absolutely "right" or "wrong" answer. Thus, there is a large element of subjectivity in a free-response testing procedure such as written or oral composition. As Ballard (1923) observed:

One of the defects of the essay as a measuring device ... is the impossibility of making the essay amenable to rigid, objective measurement. (p.61)

In both cases, the test may be expected to produce a wide variety of responses. Indeed, the more original the response, the more likely it is
to be valued, thus making it impossible to devise an objective mark scheme to encompass and account for the multiplicity of responses which may be expected. Although analytical mark schemes are devised in an attempt to make the assessment more objective, these do not quantify the actual content of compositions; they are limited to describing the qualities which a composition ought to exhibit. Thus, although analytical mark schemes may facilitate agreement amongst examiners as to the qualities that are to be evaluated, the amount of subjectivity in the assessment is reduced very little. This is because a subjective mark scheme can only ever consist of approximations rather than precise definitions. Thus, the examiner's judgement is called into play - he has to interpret his mark scheme. As soon as an examiner is required to judge the extent to which a particular quality is present, he is being required to make a personal judgement which inevitably will be coloured by his predilections, previous experiences, values etc.* As Hitchman (1966) observed:

Why is it that two examiners of equal competence are quite likely to produce very different assessments in Spoken English for the same candidate? The answer seems to be that (a) judges do not perceive the same things, and (b) they evaluate what they do perceive in different ways. ... An American experimenter has spoken of the "performance-as-given" as distinct from the "performance-as-perceived" ... It is the observer who takes these things (i.e. the objective given observable facts of the oral performance) and by fusing them with what he himself can "give" creates the "performance" ... for each observer there is an individual "performance-as-perceived". There are exactly as many performances as there are observers. (pp.34-36)

Up to this point, the difficulties which face the oral and written examiner are similar. In both cases, subjectivity is the very essence of the test and it precludes an objective assessment. However, the difficulties faced by an oral examiner are infinitely more acute because the oral medium introduces a new dimension of subjectivity into the assessment.

* For a more detailed explanation of these arguments, see the author's report: "Improving the Reliability of Essay-Marking: A Survey of the Literature with Particular Reference to the English Language Composition", CSE Research Project Report 5, (June 1980), pp.4-8.
These additional complexities become apparent when one considers the nature of the testing instrument in each case. A written paper remains the same regardless of whether the candidate takes the examination in Liverpool or Brighton, regardless of whether he takes it at 9 o'clock in the morning or 3 o'clock in the afternoon and regardless of whether he is a friendly or a hostile character. This attribute of written tests is so obvious that it is rarely considered worth mentioning. It is only when the written test is compared with the oral test that the importance of this is appreciated. The testing instrument in an oral examination is not an inanimate piece of printed paper which remains the same regardless of when, where or with whom it is used but an individual who not only assesses the performance, he also takes part in it.

In order to clarify this point, it is useful to draw a distinction between an "examiner", an "assessor" and an "examiner-assessor". For the purposes of this argument, an examiner is defined as a person who administers a test, an assessor is someone who evaluates test performance while an examiner-assessor simultaneously conducts the test and rates the performance. From these definitions, it is clear that the person who marks written compositions is an assessor because he is not directly involved in carrying out the test; he simply marks a disembodied written performance. The person who conducts an oral examination has a much more complex role to play: that of examiner-assessor. Unlike the assessor, he is personally involved in that which he attempts to assess and this is the point at which all sorts of extra difficulties occur.

One observer summed up these special difficulties when he described oral examining as characterized by, "... heavy dependence ... upon a complex of personality and social factors" which make it, "exceedingly difficult to isolate assessable elements" (Allen, 1977, p.10). Another commentator referred to the, "... intensely human situation" of oral examining (Hitchman, 1966, p.41). Indeed, the whole crux of the problem lies in the
direct confrontation of personalities in oral assessment. In order to explore this issue, the author has found it necessary to devise a term to describe a particular type of oral test which is especially subject to the influence of personality and social factors. The term chosen is "personal interaction" because it conveys the distinctive features of tests which fall into this category e.g. conversation on a one-to-one basis with an examiner, group discussion and simulated situations. The important characteristic of a personal interaction test is that it depends on the candidate and the examiner interacting. Thus, the interplay of personalities is a strong feature of personal interaction tests. Clearly, not all oral tests fall into this category. Reading aloud and giving a talk are tests to which the examiner does not contribute directly and therefore the performance is not open to his influence. The best example of a personal interaction test is conversation on a one-to-one basis with an examiner which is a widely used examining technique. Although certain tests display the characteristics of personal interaction and others do not, it is worth remembering that the whole ambience of an oral examination, regardless of the individual test items, is an "... intensely human situation" involving the interplay of personalities and "a complex of social factors".

3. THE ROLE OF THE EXAMINER

Oral examining is a highly specialised and extremely sophisticated activity which requires special expertise on the part of the examiner. An oral examiner must display a flair for putting people at their ease, for listening sympathetically and for "drawing people out". He must strike that delicate balance between knowing when to speak, when to listen and when to sustain a pause in the hope that the silence indicates reflection and will lead to further utterance. A written examiner may be extremely capable and yet display none of these personal qualities whereas an oral examiner needs
not only to be a reliable marker but must also possess these sophisticated social skills. Thus, there is an extra dimension to oral examining and this was underlined by Hitchman (1966) when he included in his book on oral examining a whole chapter devoted to the special personal qualities which an oral examiner ought to possess. He asserted that:

Perhaps the greatest single factor deciding the success or failure of the test is the personality of the assessor. (p.41)

Since individual examiners will vary in the extent to which they display the necessary personal qualities, the examiner is a critical variable in the test situation - an unavoidable source of variability.

This difficulty is exacerbated because the examiner, "... almost always plays the dominant role" (Hitchman, 1966, p.42). Consequently the performance which he marks may be largely the product of his own behaviour.

Of course, the examiner may use his control of the dialogue to good effect in order to make it a more exacting trial of the candidate's linguistic competence. However, he may also have a detrimental effect on the candidate's performance: he may talk too much; fail to "draw out" the candidate; ask inappropriate questions etc. Thus, the examiner's conduct and personality are confounding variables in the measurement of oracy and individual tests cannot be standardised in the way that written tests are standardised.

This problem is exacerbated by the essentially spontaneous nature of the oral medium which makes it impossible for the test (i.e. the examiner's input) to be standardised in the way that a written paper may be standardised. A written test is the final product of much thought; it has been modified and redrafted until it is felt to be acceptable. In an oral examination, although the examiner may employ a prepared list of topics, the unpredictable nature of conversation means that he will be forced to make unplanned responses. An examiner's reaction, on the spur of the moment, cannot be carefully prepared and it is thus subject to all the
weaknesses of a spontaneous response. The comments which were volunteered by the respondents suggested that the role of the examiner was the subject of much concern:

- It is hard to know whether the discussion mark represents the pupil's contribution or the teacher/examiner's. (Q099)

- Discussion depends too much upon the skill of the teacher/examiner. (Q408)

- Oral examining depends very much on the abilities and personality of the examiner; one examiner might be far better able to communicate with the candidate and therefore achieve better results with the individual. (Q178)

- A great deal is dependent on the attitude and help given by the teachers. (Q016)

Some teachers placed this difficulty at the top of their list of concerns. For instance, Q089 argued that the aspect of oral examining which caused him greatest concern was, "... the reliance on teacher skill to ensure an effective discussion". Likewise, when asked to specify what he regarded as the main difficulty in oral assessment, 107 replied:

"... examiner participation. If the examiner doesn't show a genuine response to the candidate and enable him to take part in a genuine two-way dialogue, then clearly the examiner is working against the best interests of the candidate."

Some respondents presented this as their main reason for wishing to limit the marks allocated to oracy. For instance, Q052 felt that oracy is:

"... important but too subject to variables of personality etc. to merit 30% of the marks. 20% would be better in my estimation."

Q497 argued that:

"Whilst I would like to see further marks devoted to oral work, there is an even greater danger in this area that the teacher will be subjective and allow personal characteristics to influence assessment.

This belief that oracy was educationally valuable, and therefore ought to be heavily weighted, and the consciousness that, from an assessment point of view, it was unreliable, was a dilemma which was widely shared.

It is important to recognise that some of the respondents based their reservations about oral examining on their experiences as trained
moderators. For example, Q310 stated:

It has been my experience from both internal and external moderation of oral examinations that some teachers are incapable of conducting the discussion in a fair or relaxed way. They have not the proper personality traits to make suitable interviewers or sympathetic listeners. Most simply talk too much.

This tendency for examiners to talk too much was complained of frequently by moderators. For instance, Q209 reported that:

I have witnessed, as a moderator, too many (discussions) where the teacher/examiner, unwittingly or not, talks too much and asks questions which deny the candidate a full opportunity to express and explain opinions.

Q613 pointed out that:

Moderators are invariably experienced class teachers. The teacher-examiner may not always be. Unwittingly, he may ruin his own candidates' chances by failing to draw them out in conversation.

Q587 volunteered a confession which endorsed these complaints:

Teachers, I feel, should prepare for discussion along with pupils. I often find it difficult to ask questions "to order", especially the type of question which would draw out the pupil. I often find it easiest to ask the more superficial questions.

The whole problem was summed up neatly by 115 when she argued: "The test is only as good as the tester".

One of the respondents raised the question: "To what extent should personality be taken into account?" This question clearly misses the point. The real question is to what extent can these confounding variables in the assessment of oracy be eliminated? Ultimately, it would be neither possible nor desirable to eliminate the personal element from a personal interaction test because trying to eradicate personality from a personal interaction test would be like trying to remove the water from an ocean.

Moreover, although this section has focussed on the deficiencies of personal interaction tests, their greatest weakness is also their greatest strength. Because they involve the interplay of people in a conversational exchange, they are the most natural and the most valid form of oral test. In contrast, other widely used testing procedures such as reading aloud and giving a talk
occur neither naturally nor regularly in most people's everyday lives. Thus, it is ironic that the most valid form of oral test also raises the greatest problems from an assessment point of view.

This study supports Hitchman's (1966) concern that, "... the greatest single factor deciding the success or failure of the test is the personality of the assessor" (p.41). However, the examiner is not the only source of variability in an oral assessment; another variable in this "... complex of personality and social factors" (Allen, 1977, p.16) is the candidate.

4. THE CANDIDATE'S PERSONALITY

During an oral examination, a candidate may experience a wide, and possibly unexpected, range of emotions varying from nervousness and awkwardness to enjoyment and a sense of gratification. A strong emotional reaction is bound to influence performance. The candidate's performance may also be influenced by his perception of the examiner's personality and attitude towards himself. Thus, the assessment may be affected by factors which are, strictly speaking, quite unrelated to oral ability. This raises the important question of precisely what a personal interaction test measures? Ostensibly it measures linguistic competence but in reality there are confounding variables in the test situation. Certain personal characteristics, which are quite unrelated to his oral competence, may exert an undue influence on the candidate's ability to perform in an examination setting. This was pointed out time and time again by the respondents:

Oral assessment is often marking confidence.  (Q275)

... the assessment of oral contribution often turns out to be a reward for extroversion and confidence.  (Q447)

... a shy, nervous pupil is at a bad disadvantage in the oral examination, particularly with the talk which has to be sustained unaided.  (Q416)
Many candidates, normally able to converse easily, become nervous and tongue-tied as soon as they know they are taking an examination. This includes even the higher ability pupils.  
(Q122)

The talk tests introversion/extroversion rather than oral facility.  (Q195)

During the reading test, fluent readers often stumble ... or mispronounce a word because they know this is part of the examination.  (Q360)

Whatever methods of oral examination are used, the stress created for the student is bound to have an adverse effect on his result.  
(Q237)

Oral assessment must note the variable in pupils' personality, confidence etc. ... personality is bound to dictate performance in situations which make difficult demands.  (Q289)

As Q289 points out, personality traits may play an unusually large role in determining performance in situations which place the candidate under pressure. An oral examination appears to be just such a situation, generating tension in large numbers of candidates. As Q473 noted: "The oral examination is the one which more pupils appear to approach with trepidation". Likewise, Examinations Bulletin 11 (1966) reported that:

Most examiners and observers noted a good deal of nervousness, and only one pair thought that it diminished rapidly after the start of the interview.  (p.31)

As Hichman (1966) observed:

By its very nature, a testing situation tends to destroy the normal conditions under which a speech performance takes place. The knowledge that he is to be tested often changes the behaviour of the speaker.  (p.52)

Examinations Bulletin 11 (1966) pointed out the problems this may create for the assessor:

It may be that such a failure constitutes a radical criticism of the candidate, and he ought to be properly penalised, especially since many candidates overcome their nervousness, but the assessment would then come to look very like a moral judgement on personality, not an examination mark, and most teachers would rightly flinch from such an impertinence.  (p.24)

The importance which was attributed to personality in determining performance was also illustrated by the responses to Question 18 which asked respondents to identify the ability group for whom they felt that oral
assessment was most valuable. The response rates are given in brackets after each option.

For which candidates is the inclusion of an oral assessment most useful in aiding the assessment of their overall ability?

(i) The most able candidates (7%)
(ii) The average candidates (12%)
(iii) The below average candidates (13%)
(iv) All children benefit equally regardless of ability (20%)
(v) Personal characteristics are a more important factor than ability in determining how much an oral assessment aids the overall assessment of a candidate*

The category which dealt with personal characteristics was more than twice as large as any other category suggesting that level of ability was regarded as less important than personality in determining the usefulness of oral examination. Thus, the quantitative evidence of Question 18 supports the qualitative data of the teachers' comments in suggesting that respondents regarded the candidate's personality as a crucial factor determining the efficacy of oral assessment.

This was underlined by the responses to Question 19. The response rates are again given in brackets after each option.

Which of these methods of examining oral performance do you prefer?

(i) An oral examination on a one-to-one basis between the teacher/examiner and the candidate (41%)
(ii) The group discussion method whereby candidates are assessed on their contribution to a discussion (3%)
(iii) A combination of group discussion and the individual oral examination with the teacher/examiner (19%)
(iv) An examination in which the candidate performs individually but an audience is provided from the candidate's peer group (5%)
(v) The suitability of each method depends largely on the personalities of individual pupils.**

Apart from the popularity of the one-to-one method of examining, the other

* 5% of the responses to this question were invalid and 2% failed to reply.
** 5% of the responses to this question were invalid.
important feature to emerge from this question was the significance attributed to personality. The second largest category, accounting for over a quarter of the sample, was that which stated that personality was a more important variable than examining method.

Thus, personality was regarded as more important than level of ability in determining the efficacy of oral assessment and as a critical factor determining the usefulness of different examining techniques. This cumulative evidence suggests that the candidate's personality is another confounding variable in the measurement of oracy. As Allen (1977) observed:

The heavy dependence of discussion upon a complex of personality and social factors makes it exceedingly difficult to isolate assessable elements. (p.16)

5. THE INTERPLAY OF PERSONALITIES

So far this discussion of the role of personality has dealt with the contribution of examiner and examinee separately. However, their separate contributions represent only a part of the total role which personality plays in oral communication. An equally important consideration is the interplay of their personalities - each will react upon and modify the behaviour of the other as well as being reacted upon and modified himself. Although Hitchman (1966) claimed that, "... the greatest single factor deciding the success or failure of the test is the personality of the assessor" (p.41), he also added that: "... the potential force of this factor is enormously increased in the interplay of personality between assessor and assessed" (p.42). As an examination progresses, it is not just oral skills which are displayed; personality also unfolds:

In order to know a candidate's speech, the assessor must, to a certain extent, know the person. At least he knows him better at the end of the test than he did at the beginning ... Prejudice may flavour the relationship ... on either side there may be feelings of dislike or active liking, of repulsion or attraction. (Hitchman, 1966, p.42)
Admittedly, this difficulty occurs, to a lesser extent, in assessing written work. Examiners sometimes observe that, in marking essay scripts, they feel that they are making judgements about the writer's personality. For example, ME32 argued: "I find that often I am marking the maturity or personality of the candidate which shows through the medium of the essay". However, the problem is nowhere near as acute as it is in the oral medium because the written examiner deals with a script which is quite separate from the person who wrote it. Even though it may convey something of his personality, it is still disembodied written communication. Oral examination, on the other hand, involves direct, face-to-face confrontation and this is bound to provoke some sort of personal reaction. Thus, changing one of the participants is also likely to change the outcome in terms of speech performance.

There is a useful comparison with written composition here. The study of essay-marking by Britton, Martin and Rosen (1966) highlighted, "... the extent to which a candidate's final mark would be likely to vary in accordance with the particular 'marking channel' his script happened to pass through" (p.1). This problem is magnified in an oral examination because the "marking channel" has an impact on the assessment at two separate points rather than one. Whereas in the written medium only the mark is open to the influence of the "marking channel", in the oral medium both the mark and the performance it describes are affected. The examiner may, unwittingly or otherwise, give an impression ranging from genuine interest and sympathy to boredom and indifference. There are similar problems on the examiner's side. He may be confronted with "the dreaded monosyllabic type" (Q207), a hostile and aggressive character, a charming candidate who converses with ease or a shy and reflective candidate who reacts with less charm but with greater intelligence. How are such differences likely to affect his own conduct and his assessment? An indication of the kind of influence this may exert was provided by one of
the examiners in Examinations Bulletin 11 (1966):

"... the examiner has to be so preoccupied with the provision of open-ended questions and other leads that assessment becomes really difficult. The result was that I tended to reward co-operation and ready follow-up of leads more than the other qualities." (p.14)

In connection with this, Bulletin 11 (1966) warned that:

... the examiner should guard against over-estimating his response, naturally and justly a grateful one, when he encounters a congenial poise and set of social mannerisms and assumptions. It is difficult for everyone not to be flattered by a familiar kind of charm and not to be repelled by the apparent truculence of a hostility which has its roots in social and adolescent prejudice. (pp.24-25)

One critic was so concerned by the influence of these confounding variables on assessment that he urged the abandonment of oral examinations:

... speech itself, its effects and what it reveals of personality, changes with mood, with occasion and with the relationship to the listener. Understanding this, we are bound to ask how speech can be adequately tested in isolation and at one brief moment. To attempt to assess spoken English within the isolating and restricting setting of any test yet devised must be to involve oneself in an attempt to assess the whole personality and development of the speaker in quite impossible conditions. Surely it would be presumptuous to start trying. (Allchin, 1964, p.16)

Although only a small proportion of the sample (8%) agreed that oral assessment should be abandoned, many of them shared Allchin's concern over the effects of these confounding variables on reliability. For instance, Q214 claimed that: "There are so many different factors at play, it is nigh on impossible to be conclusive or definitive about oral assessment."

So far, this account of personal interaction tests has focussed on their deficiencies. However, simply to dismiss them, as fraught with pitfalls, would be misleading. If the personal interaction test failed to exhibit any positive features, it would be difficult to explain its immense popularity. Since it is a widely used testing procedure, it is important to attend to those features which account for its popularity. Conversation is the most natural and necessary form of oral activity. This point is underlined by comparing the conversation with other common forms of oral testing: reading aloud; prepared or impromptu talks;
debates and dramatic activities. All of these represent specialized forms of language usage. As Abercrombie (1964) observed:

Not everyone engages in them, or needs to ... together (they) would form a very small part of the sum total of instances of the use of spoken language which occur in the world. (p.11)

These exercises are largely academic and have only a limited application to the exigencies of real-life communication. However, everyone takes part in conversations:

Conversation ... contains the most natural, the most frequent, and the most widespread occurrences of spoken language. All humans indulge in conversation. (Abercrombie, 1964, p.11)

This gives the conversation test a high degree of internal validity. Thus, it is ironic that conversation's greatest strength is also the source of its greatest weakness.

Apart from the "... intensely human situation" of oral examining (Hitchman, 1966, p.41), there are other factors which make examining in an oral medium infinitely more problematical than examining in a written medium.

6. THE EVANESCENT NATURE OF THE ORAL MEDIUM

A written examination axiomatically produces a permanent record of performance which enables the examiner to re-read parts of the essay or check it against other essays. Cross-referencing of this kind is an important part of the examiner's work which enables him to establish, modify and confirm his marking standards. The fact that a written examination automatically produces a permanent record of performance is rarely considered worth mentioning. It is only when it is compared with examining in an oral medium that this appears to be an important attribute of the written examination. The oral examiner works in an infinitely more problematical medium because the performance he marks is evanescent. The act of writing naturally leaves behind it a record of performance whereas
the act of speaking leaves no trace. This has important implications for assessment; it means that the examiner must make his judgements very quickly in the fleeting moments at which the performance is given. Since the examiner has no tangible evidence, he cannot refer back to an earlier part of the performance or check it against others. Of course, he will attempt to make comparisons, but his evidence is not an objective record of the "performance-as-given" but his impressions of the "performance-as-remembered". Not only is his impression of an earlier performance highly subjective, his ability to use this as a basis for comparison relies upon the quality of his memory. How clearly does he remember earlier tests? How far have his impressions been distorted?

Methods of recording speech are available but none is entirely satisfactory. Tape-recording has the advantages of being cheap and easy to use. However, it captures only the verbal elements of speech and non-verbal components of communication are lost. Some teachers are opposed to tape-recording because it provides an incomplete record of performance while others argue that it destroys any semblance of normality which an oral examination might have. The other method of capturing performance is to use a video recorder. The obvious advantage of this method is that it provides a visual as well as a verbal record but the expensive equipment necessary renders this method financially impractical. Thus, at present, there is no method of recording oral performance which is cheap, easy to administer, provides a comprehensive record of performance and does not interfere with the naturalness of the situation.

The evanescence of oral performance was frequently mentioned as another reason for wishing to limit the marks allocated to it:

There are a great many problems in examining orals ... orals by their very nature are ephemeral. There isn't the same opportunity for a considered judgement to be made. It tends to be much more impressionistic however many notes you may take so to make the English examination heavily weighted in terms of the oral may lead to some questions as to the accuracy of the judgement of the subject as a whole. (110)
Concern about the influence of this component on the subject as a whole was widely shared. For instance, having claimed that he was confident of the reliability of other components of the English examination, I01 added:

... but the oral was another matter which tended to escape from us a bit ... The oral was only measurable, and the marks that it produced were only valuable, within limits ... it's not that I don't think oracy should be assessed. I think it's very difficult to assess, it's more evanescent in its nature and consequently could have had a disproportionate effect at and about borderlines than one feels that it should have done.

This led some respondents to argue that the marks allocated to oracy must be carefully controlled because the greater the proportion of marks it accounted for, the more liable it was to distort the final award. I02 summed up these fears thus:

I don't think that the oral can be examined reliably ... the reasons for that are, first of all, the ephemeral nature of oracy; it's here one minute and gone the next. If you were to tape-record it, you would only introduce more complicating factors. That means that you can't have any recall, you can't challenge it or appeal in the way that you can on written papers and then there are also non-verbal factors involved in orals which are almost incapable of being properly assessed ... It can't be reliably assessed ... Maybe it can be in a crude way but not with any finesse or precision.

Further light was shed on this issue by the responses to Question 17(ii). Question 17 consisted of six statements about oral assessment and respondents were asked to indicate their views on each using a five-point scale ranging from "Strongly Agree" through "Uncertain" to "Strongly Disagree". The responses to Question 17(ii) are indicated in brackets below.

Oral performance is, by its very nature, ephemeral and therefore difficult to assess.

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Strongly Agree</td>
<td>(9%)</td>
</tr>
<tr>
<td>Agree</td>
<td>(56%)</td>
</tr>
<tr>
<td>Uncertain</td>
<td>(8%)</td>
</tr>
<tr>
<td>Disagree</td>
<td>(25%)</td>
</tr>
<tr>
<td>Strongly Disagree</td>
<td>(2%)</td>
</tr>
</tbody>
</table>

By far the largest category, accounting for 56% of the sample, was the "Agree" category. A further 9% strongly agreed which meant that altogether two-thirds of the sample (65%) felt that spoken English was difficult to
assess because of its ephemeral nature. However, it has to be
acknowledged that a substantial minority (27%) disagreed with this state-
ment. Nevertheless, this figure must be balanced against the majority who
agreed that the transitory nature of oral performance does create assessment
problems. Thus, this quantitative data supports the qualitative data of
the teachers' written and verbal comments.

This section, in dealing with the ephemeral nature of oracy, has
inevitably raised the issue of reliability. Indeed, those teachers who
mentioned this often raised its implications for reliability. Therefore,
it seems appropriate at this point to consider Question 17(i) which dealt
with the reliability of oral examining. Again the responses are indicated
in brackets.

   Your present oral examination assesses oral ability as reliably
   and accurately as possible.

   Strongly Agree   (3%)
   Agree           (41%)
   Uncertain       (34%)
   Disagree        (18%)
   Strongly Disagree (3%)
   (Invalid Response (1%))

Throughout the questionnaire, the "Uncertain" category tended to attract a
small percentage of the responses. This had been anticipated because this
category was designed to cater for respondents who did not have a definite
view on a particular issue and it was predicted that this group would tend
to be small in most instances. Thus, it was interesting to find such a
substantial proportion of the sample (34%) expressing uncertainty about the
reliability of their own oral examination. A further 21% definitely
disagreed with the statement which meant that altogether over half of the
sample (55%) were either uncertain or had definite misgivings about the
reliability of their oral examination. However, this must be balanced
against the fact that the largest category, accounting for 41% of the
responses, was the "Agree" category and that a further 3% strongly agreed.
This meant that just under half of the sample (44%) agreed that their oral
examination was as reliable as it was possible for it to be. This quantitative data suggests that the respondents were not as dissatisfied with the reliability of oral assessment as they might have been anticipated to be from the tenor of their comments. This surprisingly high level of agreement is perhaps easier to understand when the careful wording of the statement is taken into consideration. It did not claim that the oral examination was reliable in absolute terms but used the qualifying statement, "... as reliably and accurately as possible". The conviction that their own test was as reliable and accurate as it was possible for oral examining to be was quite compatible with the belief that the oral medium raises especial difficulties for accurate examination. Teachers of such a highly subjective discipline as English are resigned, perhaps more than any other group of teachers, to the fact that examining for them will always be an inexact science. However, within the limitations imposed by the nature of the subject, they may still feel that a particular test is as reliable as it is possible for it to be. This may partly explain their continuing commitment to oral assessment inspite of their misgivings about its reliability in absolute terms. Nevertheless, there does appear to be an ambivalence in the sample's attitudes towards examining spoken English.

7. AMBIVALENT ATTITUDES

Although the respondents were firmly committed to testing spoken English, many of them were deeply concerned by the difficulties it raises. At first sight, these two attitudes may appear to be contradictory and incompatible. In fact, this illustrates an ambivalence which characterized their attitudes. On the one hand, there was a conviction that oracy was educationally important and, on the other hand, there was an awareness that it created extraordinary difficulties for assessment. This ambivalence was met time and time again in their comments:
The oral work is an important feature of the CSE examination but I know that it's fairly unreliable ... and this is where you come up against the problem of trying to teach what you think is worth teaching but examine what you know is efficient in terms of examining techniques and oral isn't. But I think that most teachers would agree that it ought to continue in their schools so you've got a conflict between what you feel in your heart is a good thing for the pupils and what you know in your mind is not very well examined because of factors beyond your control. (IO6)

... in its present form, the oral element in CSE lacks intellectual credibility ... This does not mean that I think oracy unimportant: it is important. We're just not very good at assessing it yet. (Q045)

I think it's a very important part of the syllabus but the least satisfactory part ... it's how to assess it which is the problem ... but I still think an oral part of the syllabus should be compulsory. (II4)

It's not something which is written down to be looked at by different examiners so I don't think we could ever be certain that we produced an examination which is reliable ... The fact that we have the oral component ... is very important, even if we don't really know how to do it. Because we're not sure, we haven't run away from it. (IO9)

I think oral ability is important but impossible to assess. (Q560)

... it was felt that oral work was such a strong feature of every child's learning process that however difficult it was to assess and moderate, it should be attempted. (IO7)

The oral examination is not an easy problem but for an examination to be thorough in all aspects of English, it is a "must". (Q585)

The allocation of marks for oral content should be decreased not because of its unimportance but more because of the unreliability of assessment. (Q605)

This sense of conflict raises an issue which is central to English.

English seeks to promote personal and social skills which are educationally important but exceedingly difficult to quantify or evaluate. In connection with this, IO2 drew a distinction between a "teaching syllabus" and an "examining syllabus". He argued that the teaching syllabus should have breadth but that an examining syllabus is a sampling device which is necessarily selective and which is governed by special considerations (i.e. validity, reliability and administrability). He saw the basic problem of English teaching/assessment thus:
... when you have a teacher-controlled examination, they will want to put into the examination everything that is desirable from an educational point of view but that need not be reliable from an examining point of view ... you have to bear all sorts of other criteria in mind in an examination. I don't think that in English all that it is desirable to teach is possible to examine. An examination is a sampling process and, very often, if mistakes are made in teacher-controlled examinations, it is because they want to push everything in.

I06 was making a similar point when he drew a distinction between "a reliable examination" and "an educationally desirable examination" claiming that, "... the two aren't compatible". In other words, validity and reliability do not always coincide in English.

Clearly, this distinction between a teaching syllabus and an examining syllabus is a useful one and yet it is not feasible, on this basis, to dismiss oracy from the examining syllabus and confine it to the teaching syllabus. Those teachers who claim that oracy can only be examined with limited reliability but nevertheless argue that it should be included, do make a great deal of sense. They recognise that to examine something is to assure it a place in the curriculum. If spoken English was omitted from the examination syllabus, this might take oracy back to its pre-1960's status when educationists argued its importance but it was liable to be neglected in the classroom. The dilemma faced by English panels is that if they advocate oral work in their syllabuses but omit it from their examinations, it is likely to be neglected in schools. If, on the other hand, they decide that its importance should be reflected in its mark weighting, they lay themselves open to the charge of setting unreliable examinations. Although, in theory, it is possible to distinguish between a broader teaching syllabus and a narrower examining syllabus, practical experience, spread over more than a century, shows that the examination syllabus tends to become synonymous with the teaching syllabus and that it is not feasible to make too sharp a distinction between the two.

Although spoken English is a highly valued classroom discipline, it is impossible to ignore its defects as a medium for examining. The
conflict which this creates is epitomised in these two contrasting quotations:

I think it's tremendously important ... I cannot see a language used purely and simply with a pen in your hand. (I10)

... it "suffers from all of the weaknesses and ills ... of the essay examination, but in a more acute and less curable form."

(Quoted by Trimble, 1934, p.550)

8. DIFFERENT TESTING PROCEDURES

An oral test is a particular type of test, distinct from any other, in that its medium is verbal communication. The term "oral examination" is, however, a deceptively simple one. It is really an umbrella term which spans a wide diversity of activities including prepared talks, impromptu talks, debating, prepared reading, sight reading, recitation from memory, conversation on a one-to-one basis, group discussion, role play in simulated situations, scripted and unscripted drama. Clearly, "oral examination" is an ambiguous term describing multifarious and unequal activities. The diversity does not end here because the circumstances of the examination may also vary. Candidates may perform individually, in pairs or in small groups and the performance may be private (i.e. taking place before the examiner alone) or it may be public (i.e. performed before a small audience taken from the candidate's peer group). Here too there is variation because the audience may be encouraged to participate or they may simply provide a silent group of listeners. The range of permutations based upon these different techniques and contexts is extremely wide. The kind of confusion which this has sometimes resulted in was outlined by Allen (1977):

It is puzzling in that skills in one kind of situation are compared with skills in another very different kind. Often lectureettes are used in the school's own assessments and group discussion used for moderation. But are the skills recognised in lectureettes those most appropriate to giving lectureettes or are they basic to all oral expression? If so, what are they? Are these skills revealed equally well by lectureettes, group
discussion, taped parody? If not equally well, which activity demonstrates the skill best? To be plain, do we know what we mean by "oral ability"? Is there a factor "0" that is present in oral expression (beyond the mere mode) ... If we act on the assumption that lectureettes, debates, conversation, discussion, reading aloud etc. are all comparable one to another ... are we not assuming that they manifest equal evidence of general oral ability? Has such diversity any meaningful common feature? ... If we are to develop oral assessment, we shall have to decide what kinds of oral activity/skills matter. What we have all too often at the moment is a consensus of connivance. The argument for flexibility and variety in examining is not a final one, but is conditional on the value of what is included. (p.17)

The practice of treating different procedures as interchangeable is unsatisfactory because different procedures demand different skills, have different strengths and weaknesses and raise different problems in connection with validity, reliability and administrability. Therefore, it would be misleading for this study to discuss the validity or the administrability of oral assessment in toto because what applies to one test used in a particular context may be quite untrue of another technique used in a different setting. Consequently, the amount of generalized statements which can be made about different techniques are limited in that few statements would apply equally to the gamut of testing procedures and situations outlined above. For this reason, any serious treatment of the subject must deal with the different procedures separately. Therefore, the items comprising the Mode I test battery are dealt with separately in the sections which follow.

The respondents' general attitudes towards different examining techniques were, however, probed by Question 19. The response rates are given in brackets after each option.
Which of these methods of examining oral performance do you prefer?

(i) An oral examination on a one-to-one basis between the teacher/examiner and the candidate (41%)
(ii) The group discussion method whereby candidates are assessed on their contribution to a discussion (3%)
(iii) A combination of group discussion and the individual oral examination with the teacher/examiner (19%)
(iv) An examination in which the candidate performs individually but an audience is provided from the candidate's peer group (5%)
(v) The suitability of each method depends largely on the personalities of individual pupils* (27%)

The most striking feature to emerge from Question 19 was the popularity of the one-to-one technique of examining (option i) which was by far the largest category accounting for 41% of the entire sample. The preference for "private" assessment was further underlined by the response to option (iv) which was similar to option (i) except that it provided for a peer group audience. Whereas this method was the most popular option when conducted on a private basis, when carried out publicly, it was one of the least popular procedures with only a handful of respondents (5%) favouring it. Thus, although almost half of the sample (46%) favoured assessment on a one-to-one basis, only a very small proportion (5%) preferred the public setting to the private one. Group discussion is one of the most innovative methods of oral examining and one which has never enjoyed widespread use. Its general lack of popularity was reflected in Question 19 where it emerged as the least popular procedure with only 3% indicating a preference for it. Nevertheless, even this distinctly unpopular method of examining experienced a dramatic increase in favour when it was combined with the one-to-one method (option (iii) which 19% ticked). The other noticeable feature of this question was the proportion of respondents who declined to select a procedure on the grounds that personality is a more important

* 5% of the responses to this question were invalid.
variable in the efficacy of oral measurement. The importance attached to personality was indicated by the fact that this was the second largest category accounting for over a quarter (27%) of the sample. Clearly, many teachers regarded the candidate's personality as more important than examining method in determining the usefulness of oral assessment.

Overall, this question demonstrated a number of interesting features: a marked preference for private over public methods of examining; the popularity of the one-to-one technique; the importance attributed to personality as a factor determining the success of oral assessment and the distinct lack of popularity of less conventional procedures such as group discussion and the provision of a peer group audience.

9. READING ALOUD

9:1 THE HISTORICAL CONTEXT

It is appropriate to open this evaluation of the various testing procedures by considering the reading test. From a chronological point of view, the reading test has a claim to be considered first because, whereas most other methods represent recent innovations, the reading test has quite a long pedigree. Although spoken English, "... has been shamefully neglected" (Wilkinson, 1965, p.11) in most secondary schools until recently:

Reading aloud ... has a long and honourable history in the training of teachers and in the training of pupils.
(Hitchman, 1968, p.223)

Reading tests were used extensively in teacher education throughout the second half of the nineteenth century and, since oracy generally was so neglected, the introduction of reading aloud into the training college curricula was important. However, the value attached to reading was somewhat different from what it is now since we value the ability to use articulation, intonation, stress and pace in order to interpret the meaning and mood of a passage:
The method (offers) a candidate the opportunity for imaginative involvement, and a chance to demonstrate his powers of understanding and interpretation. (Examinations Bulletin 11, 1966, p.7)

However, in the nineteenth century, the value of reading was largely utilitarian:

It is easy to see why reading aloud was considered so important ... A great part of education ... was carried on by reading aloud ... This was to use reading as a teaching technique. Reading aloud was a tool of learning. (Hitchman, 1963, p.53)

Nevertheless, the fact that reading aloud had been introduced into the examining of training college students gave it a prestige which was lacking in other areas of spoken English.

This has had important consequences for the evolution of oral assessment in the twentieth century. When the testing of spoken English finally began to establish itself in schools in the 1950's, reading aloud was already well-established and its historical precedence was reflected in the unanimous agreement of the new authorities to include a reading test. The present study has found detailed accounts of twelve separate oral examinations in operation between the 1950's and 1965*, and in every case a reading test was used. Although the other items which constituted these tests differed, reading aloud was a constant feature. As Wilkinson (1965) observed: "Very few of the examinations the writer has come across omit it" (p.78).

* The boards which instituted spoken English tests in the 1950's include: the Northern Universities' Joint Matriculation Board (1953), the English Speaking Board (1953), the Associated Board of the Royal Schools of Music (1954), Durham University Schools' Examining Board (1956), the Royal Society of Arts, the Guildhall School of Music and Drama and the Union of Educational Institutes. In 1964, spoken English was made available as part of O-Level English for the first time by London University Schools' Examination Board. Other initiatives included the Brighton Schools' Examination experimental oral assessment (1962–63), the Manchester Research Project trial oral examinations (1960's) and an oral examination designed specifically for use in secondary modern schools by the English Speaking Board (1963). The sparsity of spoken English tests also made the introduction of oracy into the Tasmanian School Leaving Certificate (1962) a subject of interest in England. This is not to suggest that these twelve were the only oral examinations in operation between 1950-1965. They were, however, the only ones of which this enquiry has been able to find detailed accounts.
Oral examining was given an immense boost by the inauguration of CSE which was openly committed to innovation. Indeed, it was presented to teachers as a challenge to institute major changes:

... simply because it is new, the CSE offers the opportunity for major change. The range, content and methods of English assessment need not be restricted by precedent.

(Dixon, 1965, pp.1-2)

It is important to bear this in mind when evaluating the approach to oracy adopted by the new regional panels.

In some ways, the regional panels were faithful to the spirit of CSE. For instance, their unanimous agreement that oracy should form a compulsory component of English represented a definite break with precedent. However, when their testing procedures are scrutinised, the picture is somewhat different. Hitchman (1968) conducted a survey of the testing procedures employed by the sixteen Mode I English panels and found that by far the most widely used method was conversation - available in fourteen of the regions. Conversation is the most natural form of spoken English, and it had long been undervalued, so it was fitting that it should be elevated to a position of importance by this new examination. Rather more surprising was the fact that the only other method which could match its popularity was the reading test - employed by thirteen areas. The third most popular test, the talk, was offered by nine areas only and each of the remaining procedures was used by seven areas or less. Admittedly, the reading test was a time-honoured and well-established testing technique and this made it an obvious candidate for inclusion, in one sense. In another sense, these same qualities made it an unlikely choice for inclusion in CSE. It could hardly be described as the type of exciting, new approach to examining which CSE was supposed to promote. Nor was its widespread adoption greeted with enthusiasm since, to many observers, it smacked of convention and conservatism:
Initially, the inclusion of this method was energetically opposed, for tenable reasons. It was clearly neither original nor exciting, and it smelled too much of stale examination air, and not sufficiently of the fresher atmosphere which was looked for in the CSE. (Examinations Bulletin 11, 1966, p.7)

Hostility to this development was widespread. For instance, the NATE critique of CSE, Criteria of Success in English (1965), stated:

... the conventional approach ... has been for an external examiner to have a set passage read to him ... it would be unnecessary and dangerous to continue this tradition ... reading aloud ... is a very specialized form of oral work ... both in form and content, we find the approach too limited. (p.30)

Similar views were held by Allchin (1964):

The methods of testing have been decided on, in many cases following the traditional procedure ... the procedure that was in operation in some schools as far back as 1952. The absence of fresh thinking ... is to be seen as all the more regrettable when one recalls the comment in Examinations Bulletin 1, that the traditional method is "abnormal by any standards". (pp.16-17)

The position was a paradoxical one. Reading aloud was neither new nor exciting; nor was it regarded as a particularly valid exercise. Despite this, and the opposition its use aroused, it was the second most popular testing procedure in the new examination. The most feasible explanation of this paradox is that reading's traditional examination status had secured it an important position within the new examination. This illustrates the complex relationship which exists between a subject's examination status and its educational importance and it demonstrates the importance which traditional examination status can confer on a procedure even when its educational worth has come into question. It has to be recognised that large-scale oral assessment was a new venture and it was perhaps inevitable that the new panels should resort to well-tried and familiar methods when tackling such a difficult and uncharted area of assessment.

9:3 THE PRACTICES OF THE DIFFERENT PANELS

Section 8 drew attention to the diversity of practice possible in an oral examination and the reading test illustrates this point. The reading
was a compulsory test in seven of the CSE regions but optional in the remaining six. Whilst all thirteen panels considered prose to be suitable material for reading, several allowed the candidate to opt for poetry and one included drama. The context of the examination also varied from area to area. In over half, the reading test was "private", conducted in the presence of the examiner alone, but in several areas an audience was provided from the candidate's peer group. The method of choosing the passage and the degree of preparation allowed also differed. In some regions, the passages were prescribed and the candidate was given a short period immediately prior to the examination for preparation. Other authorities allowed the candidate to select his own piece and to prepare it at leisure. A variation on this approach was to allow the candidate to choose the text while the examiner selected the extract. One area allowed the candidate to recite a piece from memory while, in another, the candidate gave his reading to a group, responded to questions on his performance and then re-read the piece. In some tests, the reading was used as a talking point and, in others, the candidate was required to introduce his reading.

In the East Midland Region, reading aloud, the talk, conversation with the examiner and group discussion have all featured in the Mode 1 test battery at some point. However, while other items have changed, been abandoned or modified, the reading has been a constant, unchanging feature of the test. In fact, it has been the single most durable element of the examination from 1965 to the present day. Its resilience is all the more interesting when one considers the hostility towards it amongst opinion leaders. Practising teachers also appeared to be far from satisfied with it. This was revealed by Question 23 which indicated the level of satisfaction not only with the reading test but also with the other two items which constituted the Mode 1 test battery.
23. The Mode 1 Oral Examination comprises three elements:— (a) reading from an extract; (b) a talk on a topic chosen by the candidate; (c) a discussion with the teacher/examiner.

(i) Which element of the assessment are you most satisfied with?

(a) the reading (23%)
(b) the talk (20%)
(c) the discussion (48%)

(ii) Which element of the assessment are you least satisfied with?

(a) the reading (39%)
(b) the talk (28%)
(c) the discussion (21%) *

The discussion rated as by far the most popular procedure with almost half of the sample (48%) indicating that they found it most satisfactory. There was little to choose between the levels of satisfaction recorded for the talk and the reading although the reading received a slightly higher score than the talk. Nevertheless, less than a quarter of the sample (23%) found the reading the most satisfactory test item. This, on its own, did not suggest that teachers were particularly dissatisfied with the reading test. It simply showed that only a small proportion of the sample regarded it as the most satisfactory component. However, when the evidence of Question 23(ii) is combined with that of Question 23(i), a more comprehensive picture emerges. The reading was rated as by far the least satisfactory item on Question 23(ii). Thus, the cumulative evidence of Question 23 revealed the reading to be the least satisfactory element. Not only did it receive the highest score for "least satisfactory item"; it also had a poor score on the "most satisfactory item". In this light, the enduring nature of the reading test is even more remarkable.

The information provided by Question 23 did not, on its own, suggest the reasons why teachers were dissatisfied with the reading test. However, free-response items appeared throughout the questionnaire and they provided

* 9% of the responses on Question 23(i) were invalid and 12% of the responses on Question 23(ii) were invalid.
the opportunity for respondents to amplify their own point of view. It was interesting to note that, of those who chose to comment on the oral examination, a large proportion of them commented on the reading test. If the number of respondents who chose to comment on a particular item may be taken as a rough guide to the level of concern it aroused, it is clear that the reading was the subject of much concern.

9:4 THE CHOICE OF READING EXTRACTS

A brief perusal of their comments made it clear that some respondents were dissatisfied not with the reading test in principle but with the particular tests which had been set by EMREB. The main subject of complaint was that tests had varied in intrinsic level of difficulty:

- Passages selected for reading vary from year to year and from passage to passage in difficulty of interpretation and expression. (Q385)
- Some of the reading passages seem much harder than the others with regard to style, vocab. etc. (Q397)
- The passages chosen are not always to my mind of a similar standard. (Q591)
- The passages are variable: demand very different responses and interpretation. Some call for very little expertise. (Q352)
- ... the provided extracts vary in complexity - a greater standardization is essential. (Q327)

It is standard practice to provide the examiner with a range of extracts to choose from - in the case of EMREB five passages are provided. However, it seems to have proved difficult to provide extracts which differ sufficiently in content and style to justify their inclusion but also make similar demands in terms of the level of difficulty of the vocabulary, syntax, style, mood etc.

The variability which the element of choice introduces into the reading test is all the more unfortunate because a particular strength of the reading test is that it can provide a fixed, standard element in the examination.
Examinations Bulletin 11 (1966) pointed out this special role which reading may play in an oral examination when, in searching for, "... some more specific frame of reference" it observed that:

In the reading aloud such a frame of reference is supplied by the passage chosen, which then serves in some sense as a standardized instrument for all candidates. (p.12)

Thus, the reading may serve as a fixed point of reference, a yardstick which remains constant across an infinite number of tests. This is important because in oral examining there is so much which is subjective, so much which is variable: so much, in short, which it is impossible to standardize. Thus, the reading provides the opportunity to incorporate into the examination a component which is easily standardized. Some of the respondents valued the reading test for this reason. For instance, Q570 claimed that it served as, "... a useful, fixed tool of comparison between pupils". However, this depends upon candidates being tested on extracts which are genuinely equivalent. This enquiry suggests that it has proved to be difficult to combine the principle of standardization with an element of choice. Clearly, the most effective way of standardizing the reading test is by providing that all candidates read the same material.

This raises the question of why it is standard practice to provide a range of extracts? It is difficult to find an answer to this question because the element of choice is a feature of reading tests which is neither questioned nor explained. For instance, Hitchman (1966) deals with reading aloud in meticulous detail touching on such minutiae as the precise position in which the candidate should stand to perform a reading, how many minutes he should be allowed for preparation, how many words the selected passages should contain etc. Despite the detail in which he deals with this topic, Hitchman fails to consider the question of whether it is desirable to provide a range of extracts. Instead he lays down that a satisfactory test should provide, "... several sets of passages for the assessor's choice" (p.85) without further explanation or justification. The scant attention
which has been paid to the question of choice implies that this is an
unimportant aspect of test design which does not demand careful considera-
tion. In fact, it is a vitally important consideration which may undermine
the validity of the test if neglected.

In practice, the right to vary the passages is limited to the examiner. Thus, the element of choice cannot be intended to aid the candidate since
he plays no part in the decision. As far as the writer is able to
ascertain, the only useful function which the choice of passages may serve
is in relieving tedium for the examiner. The examiner has to sit through
numerous readings and, as Q609's comment illustrates, the reading may
quickly become the most laborious part of the examination for him:

The reading is difficult to assess because, finally, the way in
which the majority of candidates execute this section of the
test is very similar. It is also the least interesting for the
examiner.

Thus, the examiner may find it more refreshing if there is some variety in
the passages. Against this positive aspect of the element of choice must be
weighed its potentially damaging effect on test validity. Clearly, a
selection of extracts is not a feature of reading test design which should
be employed without careful consideration.

There are some teachers who believe that the element of choice should
lay with the candidate not the examiner. For instance, Q333 argued: "I
think it essential that the reading is of the pupil's own choice". Teachers
who favour this approach usually argue that it will facilitate discussion.
For instance, Q156 remarked: "... a prepared reading of the candidate's own
choice might ... also provide additional material on which to base a
discussion". Examinations Bulletin 11 (1966) also felt that this approach
might allow the discussion to develop naturally out of the reading:

... he (the examiner) can invite the candidate to justify his
performance, to discuss the special merits of the passage in
question and ... what prompted (his choice) ... The possibilities
would be considerable. (p.?)
In the event, this method was found to be disappointing:

... it offered little as a subject of discussion, for most of the examinees had small interest in the books they brought for reading, nor were they able to comment fruitfully upon them. (p.8)

Furthermore, it raises difficulties for assessment since candidates will inevitably select materials which differ enormously in their level of difficulty. Not only does this complicate the examiner's task but the role of the reading as a fixed, standardizing instrument is completely lost. The examiner's task is further complicated by the fact that the chosen extract may be new to him. His assessment is likely to be more meaningful when it is based on a passage which he has studied beforehand rather than on his immediate impressions of an unknown passage. There is also the problem of candidates who make ill-informed choices and fail to do justice to their ability. Furthermore, one of the strongest arguments in favour of free choice does not stand up to scrutiny. Theoretically, it enables candidates to select materials which interest them but, in practice, candidates usually fail to exploit this opportunity. It is also difficult to see why a self-selected piece must necessarily provide a more useful talking point than a set passage. Most of the questions which an examiner may ask about self-selected extracts are just as applicable to prescribed extracts. Indeed, it seems just as reasonable to require a candidate to hold an intelligent conversation on a set piece as on one which he has chosen.

9:5 THE INTERNAL VALIDITY OF THE READING TEST

Another key criticism of the reading test concerned its internal validity*. There is a distinction between speaking and reading which raises the question: how valid are the oral skills measured in a test of reading?

* The concept of internal validity relies on the professional judgements of those with expertise in the field: "'Internal' validity is concerned with the content and the structure of the test. Do these appear to be reasonable and satisfactory? We 'size up' a test, noting its characteristics, and form a judgement concerning its validity."

(Hitchman, 1966, p.62)
Many respondents felt that it was not a valid exercise because the ability to read well is an artificial skill detached from the exigences of real-life communication:

> It is very artificial, testing skills which are largely irrelevant to the pupils' oral competence. (Q338)

> ... reading aloud is perhaps the least useful skill. (Q467)

Reading aloud is a most unnatural exercise and many CSE pupils will not require this skill in adult life. (Q570)

> I think this part of the test could be dispensed with - it's rather a false test. (I14)

> ... the problem is: are we actually testing what we set out to test in the first place? I'm not sure that we ever did because the syllabus demanded reading aloud ... (there were) always questions as to whether that is what we should have been testing. (I09)

> I'm not sure what it tests or what it proves. (I12)

Reading aloud to one other person (is a) very difficult and unrealistic task for the majority of pupils. (Q212)

I07's comments perhaps provide the key to this contradictory state of affairs. Referring to a large Mode 3 consortium, he argued: "Another area which is under discussion at the moment is: why have a mark for the reading? A lot of teachers are questioning the value of that aspect of the examination". But it was his own views on this issue which were most illuminating:

> Do you know, I haven't really got an opinion on it at the moment. I'm wondering about it ... and I've not come to any specific answer yet ... it's been there since the beginning.

The most telling part of this statement was the assertion: "... it's been there since the beginning." It lends support to the suggestion that the standing of the reading test derives largely from its traditional examination status.

The teachers sampled in this survey were not alone in expressing dissatisfaction with the internal validity of the reading test. Their views are supported by a number of studies. For instance, Abercrombie (1964) distinguished between three types of spoken language: "reading aloud", "monologue" and "conversation", arguing that:
there is something specialized about the first two categories of uses of spoken language, "reading aloud" and "monologue". Not everyone engages in them or needs to... These two categories together would form a very small part of the sum total of instances of the use of spoken language which occur in the world. (p.11)

Likewise the assessors involved in Examinations Bulletin 11 (1966) claimed:

"We gained little from the reading aloud. It told us nothing new about the power of communication" (p.8). Even Hitchman (1966), who advocated the inclusion of reading aloud, had reservations about its value:

Reading aloud, when the attempt is made to do it well ... has valuable indirect results; but the occasions on which in after life a speaker will have cause to read to an audience can be few. (p.68)

Allchin (1964) agreed:

It is the "abnormality" ... that seems so regrettable ... In no normal circumstances would a child be requested to read an isolated prose paragraph to someone who was interested, not in the content of the passage, but only in the manner of his reading. (p.17)

NATE (1965) was also vigorously opposed to the procedure:

... it would be unnecessary and dangerous to continue this tradition as CSE develops. To begin with, reading aloud ... is a very specialized form of oral work ... both in form and content we find the approach too limited. (p.30)

Clearly, opposition to the reading test is widespread because, although a reading is spoken, it is not speech in the normally accepted sense of the word. Whereas the normal purpose of speech is to construct language in order to create and convey meaning (i.e. oral composition), a reading tests the ability to use the voice to interpret someone else's meaning:

To read well requires an intelligent appreciation of the meaning and mood of the piece of prose. He will be presenting his interpretation of the print by means of skilful phrasing, fluent rhythm, expressive intonation, the flexible use of pace and pause. (Hitchman, 1966, p.83)

Although comprehension is involved in the act of reading aloud, it is misleading to regard it as a basic comprehension exercise. It is essentially a test of oral interpretation in which verbal tactics are employed to convey the meaning and mood of a piece. There is a potential difficulty in this
because although reading tests are not designed to test literal comprehension, comprehension is a vital element of the reading process. If the test is pitched at an appropriate level, the act of comprehending will be an integral part of the act of oral interpretation. If, on the other hand, the test is pitched at too high a level, comprehension, rather than being subsumed into the act of interpretation, may become the dominant concern. This difficulty is pertinent to the Mode I reading test which a number of respondents felt to be too difficult:

The reading favours the more able candidates since it appears designed to test them. Less able candidates with possible reading difficulties flounder unnecessarily. (Q002)

... the passages chosen have been markedly unsuitable, all requiring a high degree of reading ability. (Q407)

Reading passages are often difficult and help the more able disproportionately. Important element is comprehension not delivery. (Q128)

At this point, it is worth recalling that a test is judged valid in terms of its measurement of particular things. If the object of a test is to examine oral interpretation and if, at the same time, the literal comprehension is pitched at such a high level that most of the candidates' energies are channelled into making sense of the basic meaning, then that test may be described as having low internal validity in that it purports to measure one skill but exercises another.

This raises the difficult question of the level at which CSE reading tests should be pitched, bearing in mind that CSE caters for a wide band of ability? The target group is the 40th - 80th percentiles in the ability range and it is worth remembering that this is twice as wide as the band for which O-Level caters. The difficulty of constructing a single test which is appropriate to such a range of candidates is exacerbated by the policy of some schools for entering as many pupils as possible for public examinations (especially in a key subject like English). It seems that children below the level for which CSE was designed now enter for it - a
practice which creates problems, as the following quotation illustrates:

Being concerned with remedial children, I feel that the reading section of the oral examination is far too difficult therefore creating confusion and embarrassment for the child. (Q237)

The real question is: should remedial pupils be entered for an examination which was designed for entry at or above the 40th percentile in the ability range? This problem is not new; throughout the history of examinations, schools have entered pupils below the specified target groups for examinations. For instance, as early as 1868, the Schools Inquiry Commission found that boys below the specified age were struggling to pass the prestigious Oxford and Cambridge "locals". In the 1920's and '30's, pupils who should have been content to gain a School Certificate at the pass level strove for matriculation certificates and, in the 1950's, secondary modern school pupils entered for O-Level at an ever-accelerating rate.

With the creation of CSE, this problem emerged amongst a group of pupils who were formerly beyond the pale of public examinations: the least able.

In order to present a fair and balanced picture, it is necessary to acknowledge that a minority of teachers strongly approved of the reading test. Nearly a quarter of the sample (23%) found reading aloud the most satisfactory element in the Mode I test battery. Perhaps the most persuasive tribute to the value of reading aloud came from Q239:

I think that more importance should be given to reading aloud. The generally poor quality in my experience ... of reading aloud has repercussions in poor comprehension ability, lack of understanding of punctuation, and little feeling for words. It is, I think, a vital point at which literature and language teaching meet.

Q239 felt that a good reading accurately indicates a candidate's level of understanding but it is still doubtful whether this is an oral skill which is valid for inclusion in a CSE spoken English test.

One positive attribute of the reading test is that, unlike the personal interaction test, the examiner is free to devote all of his energies to assessing performance. In a personal interaction test, the examiner must administer the test as well as assessing performance and much of his
attention may be focussed on providing leads and responding to replies. Reading aloud, in contrast, releases the marker to focus upon assessment and this is an important merit of this testing procedure. In spite of such advantages, the general consensus of opinion suggested widespread dissatisfaction with the internal validity of the reading test.

9:6 CONCLUDING REMARKS

Reading aloud exemplifies a contradiction which is inherent in English: the discrepancy between that which is educationally desirable and that which is assessable. IO6 described the dilemma thus: "... you come up against the problem of trying to teach what you think is worth teaching but examine what you know is efficient in terms of examination techniques." This problem usually manifests itself in a demand for an educationally important item to be included in an examination despite its deficiencies from an assessment point of view. However, in the case of the reading, the dilemma is reversed! There is a good case for including reading aloud from an assessment point of view. First, it may serve as a fixed, standardizing element in an examination where so much else is incapable of standardization. Secondly, since it is not a personal interaction test, it is not subject to the weaknesses which afflict that type of test. Furthermore, it releases the examiner to devote all his attention to the assessment. However, from an educational point of view, it is difficult to defend because it is a highly specialized and largely superfluous oral skill. Although a good reading can indicate understanding and sensitivity towards literature, it is doubtful whether this is an oral skill in its own right.
10. THE TALK

10:1 INTRODUCTION

While reading aloud is a test of "oral interpretation", the talk is a test of "oral composition" in which the candidate must construct language in order to convey meaning. Another difference is that its importance seems to have become established in the classroom rather than in the examination room whereas reading tests were used initially in teacher training examinations. Articles in journals such as the Use of English suggested that the talk had become established as a useful classroom activity by the 1950's*. Indeed, Examinations Bulletin 11 (1966) described the talk as a "well tried" and "... traditional activity of the classroom" (p.9). It is hardly surprising that a method which had proved to be valuable in the classroom should filter through into examinations. However, it was not taken over wholesale by public examinations as was reading aloud. Of the twelve examinations in operation prior to 1965 included in this survey (see footnote, p.258), seven omitted the talk (i.e. the Northern Universities Joint Matriculation Board, the Durham Schools Examining Board, the Royal Society of Arts, the Union of Educational Institutes, the Brighton Schools Examination, the Manchester Research Project and London University Schools Examinations Board) and five included it (i.e. the English Speaking Board, the Associated Board of the Royal Schools of Music, the Guildhall School of Music and Drama, the Tasmanian School Leaving Certificate and the English Speaking Board's test specifically for use in secondary modern schools). Thus, for every examination which included the talk, there was another which omitted it. Similar trends were apparent in CSE. Of the sixteen Mode I syllabuses in operation by 1968, nine employed the talk and seven excluded it whereas

* See, for example, "Getting them talking", by W. Fry (Use of English 5(4), 1954) and "Speeches Lesson, Sir" by A. Hadfield (Use of English 8(4), 1957).
reading aloud was omitted from only three of the examinations (see
Hitchman, 1968). Thus, although the talk was an established teaching
technique, it made a more limited impact on examinations than did the
reading. However, reading's popularity as a testing technique does not
appear to have reflected its use in the classroom. Although reading aloud
was a common classroom practice, it was used primarily as a means of
approaching literature rather than as an oral skill in its own right.
Thus, the relationship between the teaching and the testing of oracy may
have been a paradoxical one where that which was taught most was examined
less and that which was most widely examined was less widely taught. In
fact, by 1977, Allen found, "... wide disagreement about what kind of
activities are best assessed, which are most valuable and how they are to
be compared" (p.15). He described the situation as "confused" and as
marked by "fundamental contradictions" (p.15).

10:2 ITS POPULARITY IN THE EAST MIDLAND REGION

The talk has had rather a chequered history in the Mode I syllabuses
of the East Midland Region. Between 1965 and 1973, the oral examination
consisted of two items: reading aloud and an item described as "informal
conversation". This was really a hybrid test displaying features of the
conversation and the talk. Although the subject of the conversation was,
"... chosen and introduced by the candidate, who may make use of visual
aids", the examiner was instructed to respond to this lead, "... with the
aim of fostering the natural growth of conversation" (Mode I syllabus,
1973, p.7). Only in 1974 were the two items separated out to form two
distinct tests and, between 1974 and 1980, the Mode I examination consisted
of three items: reading, talk and conversation. In 1981, the syllabus
reverted to its original format and the talk was once again subsumed into
the conversation element. Since the talk has had a rather uneven history
in the East Midland Region, it was important to discover the attitudes of
the sample towards it.

Once again, Question 23 provided a guide to the respondents' general attitudes (see p.262 for full details of this question).

<table>
<thead>
<tr>
<th>TEST ITEM</th>
<th>MOST SATISFACTORY ELEMENT (%)</th>
<th>LEAST SATISFACTORY ELEMENT (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reading</td>
<td>23</td>
<td>39</td>
</tr>
<tr>
<td>Talk</td>
<td>20</td>
<td>28</td>
</tr>
<tr>
<td>Discussion</td>
<td>48</td>
<td>21</td>
</tr>
<tr>
<td>Invalid Response</td>
<td>9</td>
<td>12</td>
</tr>
</tbody>
</table>

Judging from the combined evidence of 23(i) and 23(ii), the talk is the component which has provoked the least extreme feelings, either of satisfaction or of dissatisfaction. Reading aloud proved to be the least satisfactory item overall eliciting by far the highest rating for the least satisfactory item and only a low score on the most satisfactory item. The discussion, on the other hand, emerged as the most satisfactory item overall with the highest rating on part (i) and the lowest score on part (ii). If, in overall terms, the discussion had elicited the highest level of satisfaction and the reading the greatest level of dissatisfaction, between these two extremes lay the talk. Admittedly, the talk was the item which the smallest number of respondents found most satisfactory but this did not automatically indicate that it was a particularly unsatisfactory test. Its low rating on 23(i) was not related to a high rating on 23(ii) suggesting that even if the talk was not an especially satisfactory element in the examination, neither was it a particularly unsatisfactory one. Indeed, the figure of 28%, who found it the least satisfactory test item, compared favourably with the discussion score of 21%. Nevertheless, this should not be allowed to disguise the fact that there was a good deal of disquiet with this requirement.

10:3 INTERNAL VALIDITY

Just as the internal validity of reading aloud had been challenged, so
too the validity of giving a talk was questioned. It was felt that neither exercise involved the kind of oral activity that most people engage in during their everyday lives. "Unnatural" and "artificial" were words which featured time and time again in their criticisms suggesting that many teachers were dissatisfied with the nature of the task:

... the instruction to "give a talk" is an instruction to do something rather unnatural and artificial. To take the initiative and speak at length without encouragement or "cues" in public (even if to an audience of one) is not something that one is normally called upon to do, even in adult life. (Q156)

I feel that any formal "talk" component presents students with an artificial situation few will encounter post-school. (Q195)

The talk is very artificial, testing skills which are largely irrelevant to the pupil's oral competence after school. (Q338)

The discussion can show all that the examiner really needs to assess the candidate's oral capabilities, given that most people do not have to "give talks" in their general lifestyles. (Q217)

The respondents felt that both the talk and the reading placed candidates in a situation which made demands which were divorced from those of real-life oral communication. Thus, the internal validity of two out of the three items which constituted the Mode I test battery was questioned. This suggested that although the overwhelming majority of respondents were committed to the principle of oral examining, many were dissatisfied with the form which it took. Conversation was regarded as the most valid mode of assessment because it: "... contains the most natural, the most frequent and the most widespread occurrences of spoken language. All humans indulge in conversation", whereas activities such as reading aloud and giving a talk are, "... specialized ... Not everyone engages in them, or needs to ... together (they) would form a very small part of the sum total of instances of the use of spoken language which occur in the world" (Abercrombie, 1964, p.11).

Although the charge of unnaturalness cannot be denied, the talk can be defended on the grounds that it is a form of oral composition. Since the value of conversation is universally recognised, it follows that the talk,
as an alternative form of oral composition, also exercises important oral skills. Thus, although the form which the test takes may be regarded as artificial, it is nevertheless a useful device for testing skills which the candidate will require throughout his life and, although he may never use them in that specific form again, he will undoubtedly need the skills which oral composition entails. Thus, the talk may be justified as a means to this end.

It must also be acknowledged that the talk has certain advantages over conversation because it is a "purer" form of oral composition than the conversation. In the talk, the speech is "one-way", relying exclusively on the candidate's competence whereas in any form of personal interaction test, the candidate's performance may be influenced by the examiner's contribution. His performance may be enhanced or marred by the examiner's conduct; either way, it is not his pure, unadulterated oral competence which is being measured. Thus, the interplay of personalities between assessor and assessed is a source of variability in a personal interaction test and a testing technique which is not open to this kind of personal influence has some important advantages of its own:

The talk is the candidate's responsibility and less subject to outside influences. (Q209)

The talk tests the ability to talk freely and at length; and places the onus fairly on the candidate to perform well rather than relying upon the teacher's ability to sustain the examination. (Q405)

Thus, the talk is a better test of the candidate's ability to maintain a coherent flow of speech. There is an additional advantage implicit in this: the examiner, freed from the need to administer the test, can concentrate on assessment. Although from an educational point of view the talk is sometimes regarded as invalid, it clearly has some advantages from an assessment point of view.

Conversation has different strengths and weaknesses. It demonstrates
additional skills which derive from the "give and take" of dialogue:

... as in a boxing match, he has to improvise on the spur of the moment. No preparation of a position is possible - he has to take what comes and make the best of his opportunities.

(Hitchman, 1966, p.66)

In a talk, the candidate's speech may be well-rehearsed whereas a conversational exchange is spontaneous and reveals how well a candidate can use language on the spur of the moment. Although there will be occasions in life when he may voice a prepared point of view, the occasions on which he will need to improvise will be infinite.

Since two out of the three items in the Mode I test battery were felt to be artificial exercises divorced from the needs of real life communication, it seems appropriate at this point to consider item 17(v) which dealt with this issue. It stated:

A formal examination of oral ability makes the situation in which the candidate has to perform unreal and artificial.

Respondents were asked to indicate their level of agreement/disagreement with this statement.*

By far the largest category, accounting for 45% of the sample, consisted of those who agreed that a formal examination is, by its very nature, artificial. A further 8% strongly agreed which meant that there was widespread agreement that artificiality is an inevitable consequence of a formal assessment. However, a substantial proportion of the sample (17%) were uncertain about this issue and over a quarter (29%) disagreed. Nevertheless, this must be balanced against the fact that over half of the sample (53%) agreed with the statement. There was widespread agreement amongst the sample that artificiality is a characteristic of formal examining.

10:4 PUBLIC OR PRIVATE ASSESSMENT?

Although Question 19 demonstrated the popularity of the one-to-one

* Strongly Agree (8%), Agree (45%), Uncertain (17%), Disagree (27%), Strongly Disagree (2%), Invalid Response (1%).
examination in the East Midland Region (see p.256), it became established practice in many regions to carry out the talk before a peer group audience. In 1968, a total of nine Mode I syllabuses included the talk and eight of them required that the speaker should be provided with an audience. Indeed, five areas encouraged the audience to question the speaker as they would do in the classroom. The importance of providing a group of listeners was endorsed by Hitchman (1966) who drew a distinction between a "public" and a "private" test arguing that some procedures are ideally suited to the public mode while others are best treated privately:

... conversation is best fitted to the "private" type of test, the giving of a talk to the "public" type and prose reading and poetry speaking are suited to either. (p.61)

Clearly, the talk was regarded as an essentially public activity in which the audience was an important component.

In spite of this, when EMREB introduced the talk as a separate item in 1974, it gave teachers the choice of whether to conduct it publicly or privately. The only stipulation was that: "All children presented for the examination in any one school must be tested in the same situation (i.e. either all with or all without a participating audience)" (Mode I syllabus, 1974, p.10). Given this choice, teachers in the East Midland Region tended to favour the private mode of testing. Indeed, many were strongly opposed to the public mode of assessment regarding it as inhumane and inappropriate.

I am certain that the "one-to-one" basis is the least harrowing. (Q207)

Children show much concern at the prospect of having to take an oral examination before anyone but their teacher: their fears must be allayed: privacy is the order of the day. (Q132)

Conduct on a one-to-one basis seems to me to be the only "fair" method. (Q585)

Individuals are better examined away from their peer group. (Q008)

I find that the pupils themselves prefer this. (Q036)
... the important thing with the oral test is to do it on a one-to-one basis. I do not like the idea of examining any pupil with an audience of his peers. I don't think I could bring the best out of a pupil in that situation. (105)

Clearly, the public mode of examining adds a new dimension to the personality problem in oral assessment. The respondents argued that the public context may provide another variable affecting performance:

The provision of an audience from the candidate's peer group is likely to make all but the most confident extremely nervous, depending - as they do - on the approval of that peer group. (Q348)

One-to-one oral tests are better for the shy child. The well-adjusted child also benefits as they can be easily put off by interruptions and remarks from their peers. (Q597)

At my school we have tried examining pupils with a peer group audience, with the result that some (girls especially) simply refused point blank to "perform". (Q156)

I have found that many of the less forthcoming in class "open up" on a one-to-one basis. (Q587)

Q346 also made the point that the oral examination, "... is the only opportunity for contact on a one-to-one basis between the two". Throughout the course, a teacher's contact with his pupils is largely class-based making the examination a valuable opportunity for one-to-one contact. Indeed, respondents expressed surprise at the new sides to themselves which candidates sometimes revealed in a one-to-one situation.

It is not only the candidate's attitude towards his audience which may affect the performance; the audience itself provides another confounding variable:

I presumed that the peer group would be "friendly", whereas a number of them ... tend to be "anti" anything school-based and therefore not at all useful as an audience. (Q546)

The influence of the audience was also highlighted by Examinations Bulletin 11 (1966):

... the relative success of an individual was increased by the presence of an active, co-operating audience ... performance appeared to improve significantly when the audience was smallish and genuinely sympathetic and cheerful. (p.11)

Although the presence of an audience is intended to enhance performance by
making the situation more realistic, experience suggests that the
listeners may have an adverse effect either because of their own lack of
coop-eration or because of the candidate's anxiety in facing them. There
was a general consensus of opinion amongst the respondents that the one-
to-one test was the fairest method of examining and that an audience was
more likely to undermine the internal validity of the test than enhance it.

This qualitative data was supported by the quantitative evidence of
Question 19. The responses are indicated in brackets after each option:

Which of these methods of examining oral performance do you prefer?

(i) An oral examination on a one-to-one basis between
the teacher/examiner and the candidate (41%)
(ii) The group discussion method whereby candidates
are assessed on their contribution to a
discussion (3%)
(iii) A combination of the group discussion method and
the individual oral examination with the teacher/
examiner (19%)
(iv) An examination in which the candidate performs
individually but an audience is provided from
the candidate's peer group (5%)
(v) The suitability of each method depends largely on
the personalities of individual pupils (27%)
(Invalid response (5%))*

Those who favoured the one-to-one examination (41%) formed by far the
largest category. This contrasted sharply with the tiny group (5%) who,
although preferring the candidate to perform individually, thought that an
audience should be provided. It was also interesting to note that those who
felt that personality was more important than method of assessment (27%)
formed the second largest category. This question underlined the
disapproval of peer group audiences and the importance attributed to
personality as a factor in oral assessment.

A minority of respondents did believe that the talk was an artificial
exercise without an audience:

* For a general discussion of these results, see p.256.
The talk often seems artificial on a one-to-one basis. (Q223)

In a one-to-one situation a talk can be an artificial and painful experience for both examiner and candidate. (Q571)

One should have to give the talk to one's peers ... It's ridiculous to give a talk to one other person sitting there solemnly. It's a totally unreal situation. (I12)

An examination carried out with peers is a far more supportive situation than "isolation" with teachers. (Q333)

However, the teachers who shared this view represented only a small minority of those who commented on this issue.

The question of whether to examine the talk with or without an audience raises a problem which is central to English: there is frequently a mismatch between that which is educationally desirable and that which may be examined effectively. As I02 put it: "I don't think that in English teaching all that it's desirable to do is possible to examine". Although the audience may be a vital component of the talk in a classroom situation, in an examination, where the prime concern is an accurate assessment, an audience represents another confounding variable in a medium already fraught with problems.

10:5 THE NATURE OF THE TASK

In 1977, Allen described the current state of oral examining as "confused" arguing that, "... fundamental contradictions in the idea of oral assessment have lain unresolved from the start" (p.15). The talk illustrates this because there is a fundamental contradiction between the essential nature of the prepared talk and the way in which some examinations boards attempt to use it. The confusion centres on a distinction between conversation and spoken prose. In an article entitled, "Conversation and Spoken Prose", Abercrombie (1964) explored the differences between these two speech modes, observing that:

Most people believe that spoken prose, as I would like to call what we normally hear on the stage or screen, is at least not far removed, when well done, from the conversation of real life ... spoken prose is far more different from conversation than is
usually realized ... the whole structure of conversation is different from that of prose ... It comes as quite a surprise to find how different it is. (p.13)

This distinction is pertinent to the talk. It raises the question: is the talk more akin to spoken prose or to conversation? Clearly, the prepared talk has an obvious affinity with spoken prose. The very nature of the task (i.e. prepared) encourages an unspontaneous and well-rehearsed performance. Whereas a conversational exchange is unprepared and unpredictable, the success of a talk is closely related to the degree of preparation which has been given to it. Admittedly, speakers occasionally attempt a "spontaneous" presentation but it is difficult for a prepared talk to capture the characteristic features of a genuinely unprepared utterance and even more difficult for inexperienced speech-makers to display this skill. Problems arise because many examining authorities using the talk hope it will facilitate a conversational mode of speech. For instance, EMMEB's Mode I syllabus for 1974-1980 stated that: "Notes ought not to be necessary, though brief headings may be glanced at" (Mode I Syllabus, 1974, p.11). This regulation was clearly aimed at fostering conversation rather than spoken prose and the confusion was taken a stage further by the practice of merging talk and discussion in an activity entitled "informal conversation" which was really a hybrid test displaying features of both (see p.273). A similar confusion was apparent in Examination Bulletin 11 (1966). It warned that the instruction "to give a talk" might lead to a very formal mode of speech:

There seemed to be a danger that inevitable over-preparation and plagiarism would lead to a stilted affair without connection with the needs of normal communication. (pp.9-10)

Clearly, it was intended that the talk should elicit "normal communication" but this raises the question: is the presentation of a talk a normal mode of communication by most people's standards? Having stated that the talk is intended to meet, "the needs of normal communication", Examinations
Bulletin 11 (1966) quoted the following comment by a candidate:

> Of course, the circumstances of talking are very unnatural, they aren't ordinary. (p.11)

The real crux of this problem was identified by Q324 when he observed that:

> The talk presents difficulties in talking between formal and informal speech i.e. the distinction between "spoken" and "conversational" English is blurred here.

The type of problems which this "blurring" raises were outlined by Examinations Bulletin 11 (1966):

> The use of notes must be permitted, but it raises the most serious objections to the method, for it might be difficult to decide exactly at what point a talk is being read aloud from prepared notes, as opposed to a careful and necessary reference to written material. Notes copied more or less directly from encyclopaedias will probably be delivered tonelessly and boringly, yet they are sure to be tidy and accurate. If they have been learned by heart, then a good deal of effort and retention must be given the appropriate credit - but the display will be a forced product of the examination, and not a demonstration of oral attainment. (p.10)

They also encountered the type of talk which entailed, "... a very thoroughly prepared and sophisticated piece of work which yet leaned so heavily on visual aids that the speaker was able to add very little to his exhibition" (p.10). The dilemma is that a well-prepared and thorough talk encourages reliance on written notes and visual aids. However, if the aim of the test is to elicit the candidate's oral competence, then reading from notes, reliance on visual aids and the recitation of memorized material may undermine the validity of the test.

The issue of which mode of speech is most appropriate to the talk is charged with confusion. Examiners are frequently looking for naturalness and spontaneity and use the talk to elicit a conversational use of language. Although a skilled speech-maker may create an impression of unrehearsed utterance, this subtle skill may be difficult for candidates to display. Thus, there is a discrepancy between the intrinsic nature of the test and the way in which some examinations attempt to use it.
Nervousness and apprehension were noted more frequently in connection with the talk than with any other testing procedure:

... pupils who hope to do well find the talk most worrying. (Q114)

The talk inhibits a relaxed atmosphere. (Q207)

The talks seem to me to be very stilted. The kids ... don't express their ideas as well as they would if they were more relaxed and were given the opportunity to talk about something they knew about in conversation. (I12)

I feel that a shy, nervous pupil is at a bad disadvantage in the Oral particularly with the talk which has to be sustained unaided. (Q416)

This was confirmed by Examinations Bulletin 11 (1966) where, "All three markers noted very extensive nervousness" (p.11) in connection with this method. Why is it that the talk seems particularly liable to cause anxiety?

When the talk is compared with other items in the Mode I test battery, this is easy to understand. In the reading test, the passage acts as a "prop" to the candidate and he is not dependent on his own ability to sustain the flow of speech. Similarly in the conversation, the candidate takes part in an exchange in which the examiner's cues and contributions support him. Thus, a conversation and a reading may appear less intimidating than a talk in which the candidate relies entirely on his own capacity to sustain the flow of speech. Moreover, there is no opportunity for the examiner to relax the candidate in a talk whereas in a conversation the examiner may use various tactics to put the candidate at his ease; indeed, the conversation has been noted for the relaxed atmosphere which it promotes: "... both examiner and observer stated that any initial nervousness was soon discarded" (Examinations Bulletin 11, 1966, p.14).

Thus, although the talk does not raise the problem of personal interaction, it does raise the problem of nerves in a particularly acute form.

More positively, it is possible that candidates may find the opportunity
to give a talk on a specialist interest a valuable experience because it places them in a position where they are experts with something of interest to convey. This may be a novel and gratifying experience for the candidate as 105 observed:

\[
\text{Amongst my greatest joys in teaching over the last fifteen years have been some of the talks that we've had from very ordinary pupils. I think that the oral examination ought to be quite a momentous thing in a kid's life where they feel that they are an authority on something. And often they are and it really gives them a chance to shine. It's a very valuable part of English ... when you get into that kind of situation, the rapport you get is tremendous.}
\]

The personal satisfaction to be derived from giving a talk was also stressed by Examinations Bulletin 11 (1966):

\[
\text{The tape-recordings of this test cannot possibly convey the feelings of eventual enjoyment which the examiner, observer and candidates received ... Much to the candidates' surprise they found oral expression almost exciting ... In our opinion, the experiment was successful because the candidates left us happy and interested. Some told us this, others told their teachers subsequently. (p.12)}
\]

It was also felt that the opportunity to create an oral composition may be particularly valuable for pupils in the CSE ability range who often perform less fluently in the written medium:

\[
\text{In my experience, below average students (in terms of technical competence - spelling, punctuation, paragraphing etc.) can produce lively, interesting, well-prepared talks. (Q285)}
\]

\[
\text{Children of low ability often perform well orally. (Q352)}
\]

\[
\text{I am invariably surprised at how well (often better than expected) candidates do in the oral examination. They sometimes reveal sides of themselves previously hidden. (Q021)}
\]

However, it must be acknowledged that some candidates make little effort to prepare a satisfactory talk. In such cases, the talk can become an embarrassing and pointless exercise for candidate and examiner alike. As 112 observed:

\[
\text{... you do get some kids who read a football magazine the day before and walk in saying they're going to talk about Liverpool. To me that's not a very valid exercise.}
\]

\[
\text{The talk ... is often ill-prepared. (Q570)}
\]
Often candidates have no real interest in the subject selected for the talk - chosen for convenience. (Q528)

These findings are supported by Examinations Bulletin 11 (1966):

A considerable number spoke on topics they knew little or nothing about. The impression given in these cases was that the subject had been "mugged-up" for the occasion. (p.10)

Clearly, the success or failure of the talk is a reflection not only of oral competence but also of the candidate's conscientiousness in preparing and practising his material prior to the examination. Thus, the attempt to measure oral competence by the talk can be confounded by variables which are, strictly speaking, extraneous to oral ability. In this particular sense, conversation provides a purer test of oracy since its success does not depend on preparation undertaken before the test.

10:7 CHOICE OF SUBJECT MATTER

The candidate is usually allowed a free choice of subject matter in the talk:

It would be distasteful to impose upon a fifteen-year-old the obligation to talk on a subject for which he must simulate an unfelt excitement; what fails to involve ... will be haltingly and limply articulated. (Examinations Bulletin 11, 1966, p.27)

This means that the examiner may have to assess talks on an extraordinarily wide and uneven range of topics, some of which will be inherently interesting to him, some of which will fascinate by their novelty and others which he will regard as basically trivial and insignificant. For instance, the examiner in Examinations Bulletin 11 (1966) singled out two talks which particularly impressed him: one on bream fishing and the other on a visit to an art exhibition. Although he attributed this to the speakers' enthusiasm for their topics, it has to be recognised that these topics appeal to different age and social groups. However, an examiner may also be presented with a talk taken from the insular world of adolescent culture where the language is full of obscure jargon and the subject matter seems essentially trivial to him. This raises the question of how far the content
of the talk should influence the assessment? The talk is primarily a
device for measuring oral skills but its content plays an important part
in its impact. How easy is it to separate out the success or failure of a
talk from the intrinsic worth of its subject matter? Q043 argued that:

With a free choice of talk being allowed it is so easy for the
candidate to impress or not, by choosing subject matter whose
intrinsic interest affects the examiner who should be
concentrating on treatment.

Examinations Bulletin 11 (1966) felt that:

There was little need to include ... the world of teenage culture,
which tends merely to promote the use of a private adolescent
jargon, and is in itself poor stuff for discussion ... when
teenage culture did come up it remained too threadbare to sustain
language. (p.14)

It pointed out the obstacle to comparability which this variability in
topics can raise:

What kind of discredit attaches to a talk of a frivolous and
tiresome nature? ... It is possible to talk wittily about empty
subjects: how can an examiner discriminate between such an
effort and a laborious and elaborately prepared failure? (p.10)

Bulletin 11 (1966) decided that the inherent quality of the subject matter
should be taken into consideration: "The object is not merely to release a
flow of talk: the imaginative energy ought to be directed towards the
animation of worthy material" (p.28). One of the respondents, Q423,
suggested that since freedom of choice was at the root of the problem, this
should be eliminated:

The talk is difficult to assess because of the variety of topics;
these should be laid down by the examination board. If this was
the case, the teacher would be able to compare one child with
others more satisfactorily.

However, it seems unlikely that this would solve the problem. Both
Examinations Bulletin 11 (1966) and the respondents who took part in the
present study warned about the quality of talks on subjects "mugged up"
specially for the occasion. Furthermore, if a board did provide set topics,
in practice these would need to be sufficiently diverse and open-ended to
cater for a wide range of abilities and interests. This means that set
topics are unlikely to provide the level of standardisation which Q423 believed they could achieve. Thus, although the element of choice does raise some difficulties in evaluating talks, it does seem that freedom in this area is necessary.

10:8 CONCLUDING REMARKS

Although the talk has now been abandoned in the Mode I syllabus, this does not appear to reflect its level of popularity amongst teachers. It elicited a greater level of satisfaction than the reading which emerged as by far the least satisfactory item overall. The retention of the reading and removal of the talk from the Mode I syllabus appears all the more remarkable in the light of the evidence of this survey. Moreover, as a testing procedure, the talk has certain strengths. Firstly, unlike a personal interaction test, the examiner takes no part in the talk and the candidate has the opportunity to display his oral skills uninfluenced by the examiner's personality or flair for administering oral tests. Moreover, the examiner, released from the responsibility of conducting the test, is able to concentrate on giving an accurate assessment. There are, of course, difficulties raised by this method of examining - particularly the internal validity of "giving a talk" and the way in which the distinction between conversation and spoken prose is blurred. However, English teachers, perhaps more than teachers of any other subject, are resigned to the fact that examining is an imperfect science. Within these limitations, the talk has some value as a testing procedure.

11 CONVERSATION ON A ONE-TO-ONE BASIS

11:1 INTRODUCTION*

The use of conversation in the early oral tests (i.e. 1950's - 1965)

* This section deals specifically with conversation on a one-to-one basis. Group discussion is treated separately because it raises its own special problems.
was similar to that of the talk in that some examinations included it and others excluded it. Of the twelve spoken English examinations in operation prior to 1965, which the present study has found detailed accounts of, seven included conversation (five included the talk). Con-
versation was included in the examinations of the Associated Board of the Royal Schools of Music, the Guildhall School of Music and Drama, Durham University Schools Examining Board, University of London Schools Examination Board, the Royal Society of Arts, the Brighton Schools Examination and the Manchester Research Project. The five which did not include conversation were the Northern Universities Joint Matriculation Board, the Union of Educational Institutes, the Tasmanian School Leaving Certificate and the two separate examinations conducted by the English Speaking Board (introduced in 1953 and 1963). Thus, conversation featured in these early tests as a possible item and it did not emerge as the indispensable component which reading aloud quickly established itself as.*

After this relatively modest start, conversation was elevated to a position of importance by CSE. CSE was expressly committed to making examinations more appropriate to the demands of life outside school. As

* This section distinguishes between two different types of exchange which can take place on a one-to-one basis: conversation and interview. This distinction is frequently neglected and all exchanges which take place on a one-to-one basis are indiscriminately placed in the blanket category of conversation. However, conversation and interview are fundamentally different types of oral activity. Conversation involves a genuine exchange of views and information in which both participants "give and take" whereas interviews entail a much more rigid form of exchange in which one participant asks questions and the other participant provides answers. There is no obligation on either party to foster, "the natural growth of conversation" (EMREB Mode I syllabus, 1970, p.62). Thus, an interview employs a more formal and limited question/answer format and the failure to discriminate between these two types of activity provides further evidence of the "confusion" (Allen, 1977, p.15) which has characterized oral examining from the outset. Some of the examinations listed above as not using conversation did use the interview and, although a conversation might ensue from such an exchange, it was not required by the syllabus. Thus, conversation, rather than being a test item in its own right, was a by-product which might arise out of the interview.
Holbrook and O'Malley (1964) asserted, "CSE should not be the kind of examination which does no more than test the capacity of pupils to pass an examination" (p.6). Likewise Examinations Bulletin 1 (1963) urged that, "... the aim should be to put the candidate in a normal language situation" (p. 39). It was this which provided conversation's most powerful claim to inclusion in a spoken English test. After all, there is no other oral activity which offers a more "normal language situation". As Abercrombie (1964) pointed out conversation, "... contains the most natural, the most frequent and the most widespread occurrences of spoken language" (p.11).

It was perhaps this more than any other feature which secured it a prominent position in the new examination. Other widely-used testing procedures such as reading aloud, prepared or impromptu talks, dramatic activities and debates are largely academic and have a limited application to the exigences of real-life communication. As Abercrombie (1964) observed, they are "specialized" skills which most people will need on only a limited number of occasions during their lives.

Hitchman's (1968) survey of the techniques employed by the new regional panels revealed that by far the most widely used procedure was conversation which was available in fourteen of the regions. Even reading aloud, which had traditionally enjoyed the widest popularity, was less widely used than this (it was available in thirteen regions). These figures on availability do not reveal the true extent to which conversation had ousted reading aloud as the most popular method of assessment. A more telling indication of their relative popularity was the extent to which they were made compulsory. Ten of the fourteen areas which offered conversation made it compulsory whereas only seven of the thirteen areas which used reading aloud made it compulsory. There were only two other tests which were made compulsory: the talk in five areas and group discussion in two. The importance attached to conversation is highlighted by these figures. In fact, conversation was a compulsory item in more
areas than most other tests were available in with the exception of reading aloud.

The evidence of the questionnaire suggests that the popularity of the conversation has not waned over the years. Question 23 again provided information about the respondents' level of satisfaction with the conversation (see p.262 for full details of this question).

<table>
<thead>
<tr>
<th>TEST ITEM</th>
<th>MOST SATISFACTORY ELEMENT (%)</th>
<th>LEAST SATISFACTORY ELEMENT (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reading</td>
<td>23</td>
<td>39</td>
</tr>
<tr>
<td>Talk</td>
<td>20</td>
<td>28</td>
</tr>
<tr>
<td>Discussion</td>
<td>48</td>
<td>21</td>
</tr>
<tr>
<td>Invalid response</td>
<td>9</td>
<td>12</td>
</tr>
</tbody>
</table>

The scores for the discussion are superior to the scores for either of the other elements in the Mode I test battery. Nearly half of the respondents (48%) felt that the discussion was the most satisfactory element whereas only a fifth (21%) rated it as the least satisfactory item. In fact, one could devise a popularity scale for the three items which comprised the Mode I test battery based on the combined evidence of Question 23(i) and (ii). The reading would appear at the bottom of the scale since it was judged to be the least satisfactory component overall. The talk would appear in the middle emerging as more satisfactory than the reading but the overall "winner" would be the discussion which was rated far in front of both reading aloud and the talk. Thus, the importance attached to conversation at the inception of CSE has been vindicated by experience.

11:2 INTERNAL VALIDITY

The great strength of conversation is that it entails the speech form which is most naturally and frequently used. This has important implications for the internal validity of this test. For instance, as Q609 observed, "... it is in this area that the candidate relaxes and reveals the quality
of his thinking". In contrast, other less natural activities such as reading aloud and giving a talk are more likely to be hindered by nerves and awkwardness:

Nerves affect the reading - this may be better after the discussion when the examiner has had time to relax the pupil. (Q532)

Reading aloud need not be done with a frightened pupil until he or she is sufficiently relaxed. The talk inhibits a relaxed attitude whereas discussion should bring it about. (Q207)

These findings are endorsed by Examinations Bulletin 11 (1966) where nervousness was found to be "... a hefty obstacle" (p.11) in connection with the talk and where the reading also produced "... difficulties with nervousness" (p.8). In contrast, the conversation promoted a relaxed atmosphere: "... any initial nervousness was soon discarded" (p.14).

A test which releases tension and encourages a relaxed approach may provide a truer picture of a candidate's oral powers and the validity of the conversation is, in this sense, greater than that of the reading and the talk.

In connection with this, some respondents pointed out how gratifying candidates found the conversation. Commenting on a Mode 3 syllabus with a substantial conversation component, I07 remarked:

... something like 70% of the candidates will come to me in the days following their oral examinations, and they've not been examined by me, they've been examined by people from other schools, and say, "I really enjoyed that. We talked about this and we talked about that" ... They're really enthusiastic and they find it the most satisfying part by a long way.

I05 agreed:

It's one of the most enjoyable things that I do during the year ... I think that afterwards a lot of candidates would have been sorry not to have taken part in it. A lot of them do gain from it.

Examinations Bulletin 11 (1966) also emphasized the personal satisfaction to be derived from this method: "A large number of candidates assured their own teachers that they had actually enjoyed the examination" (p.13). It is interesting to note that similar claims were made for the talk
(see pp.284-285). It seems that both forms of oral composition place candidates in a position where they have something to communicate which is valued and the opportunity to convey this to an interested listener may be immensely gratifying.

Another strength of the conversation test is that:

... as in a boxing match he has to improvise on the spur of the moment. No preparation of a position is possible - he has to take what comes and make the best of his opportunities.

(Hitchman, 1966, p.66)

This enhances the validity of the conversation test relative to that of the talk and the reading. In these latter two, the candidate's performance is well-rehearsed. Thus, neither test elicits how well a candidate can cope with a spontaneous use of language. In real life, on the other hand, the opportunities one has to rehearse well-prepared statements are few compared with the infinite number of occasions on which one must improvise with language. In this sense, the conversation is superior to the other two because it tests a candidate's ability to, "... take what comes and make the best of his opportunities."

11:3 INTERNAL VALIDITY : THE INTERVIEW MODE OF COMMUNICATION

The aim of a conversation test is to promote an exchange of ideas and information. There is an obligation on both parties to foster, "... the natural growth of conversation" (EMREB Mode I syllabus, 1970, p.62). In other words, both must "give and take" (Hitchman, 1966, p.66). There is, however, a danger that a very different linguistic pattern to the one outlined above will establish itself. This survey suggests that conversa-
tion easily degenerates into an interview in which the questions are posed by the examiner and the candidate's part is limited to brief replies or, worse still, "yes" - "no" answers.

Most CSE candidates, particularly those of below average ability, find discussion very difficult and this section of the oral examination frequently devolves into a question and answer session. (Q367)
... discussion may deteriorate into questions, yes/no answers. (Q528)

A sustained pupil-teacher conversation on a one-to-one basis is hard to achieve. (Q631)

One point which concerns me is the reliance on teacher skill to ensure an effective discussion rather than just "yes" and "no" answers. (Q689)

Teachers in the East Midland Region are not alone in this experience. As Inglis (1966) observed:

The problem of such conversations is always that the child is deferential and rather brief; he waits for conversational leads to be fed to him and then rejoins in a single sentence. A series of questions-and-answers rapidly sets up a pattern of expectations which is hard to break. (p.221)

Similar problems were reported by Examinations Bulletin 11 (1966):

"The most satisfactory results occurred when, after a series of short questions and answers, the candidate spoke freely and at some length on topics of particular interest. Others were contented to give rather incomplete answers and to rely upon the examiner to seek fuller information from them" ... The problem was that 'interviews continued with this "series of short questions and answers."' (pp.12-13)

The question this raises is: why does the conversation degenerate so easily into an interview?

The "backwash effect" of examinations on classroom practice is a familiar phenomenon but it seems likely that what we are confronted with in this instance is the effect of years of conditioning and expectations built up in the classroom. Barnes (1969, 1976) drew attention to the role of classroom language in promoting the hidden curriculum. He showed how seemingly subtle features of spoken language shape and determine the kind of learning which takes place in a classroom: that what pupils learn in a subject is not simply subject-specific skills and information but concerns the behaviour and responses considered appropriate by the teacher. He found that pupils and teachers exhibited stereotyped linguistic behaviour - that linguistic strategies related to being a "teacher" and a "learner" could be identified in most classrooms regardless of the subject being taught or the
... although there is only one adult in the room, she seems to be
talking more than all the children together. She is the centre
of everybody's attention: she asks many questions, and demands
answers as of right ... In spite of this urgency she seems to
know the answers already, for she dismisses several suggestions
until one comes which she greets with, "That's it. Good answer"...

Her young pupils ask hardly any questions ... Only one child
changes the subject of the conversation ... No child asks
questions like those the teacher asks ... Mrs. Jones uses her
voice to control and shape the thoughts and attention of the
class ... (This builds up) complex patterns of expectation about
who says what to whom and when ... I (am) not just presenting
these as different social relationships but as part of what her
pupils were learning. We cannot make a clear distinction between
the content and form of the curriculum ... The two are
inseparable. (1976, pp.11-14)

Although questioning might have been assumed to be a common activity for
pupils in their role as learners, pupils rarely asked questions. On those
occasions when pupils did ask questions, they were of a specific type:
"... the child is not so much wanting information as reassurance" (1969,
p.45) and often related to a task which the child had been asked to perform.
The questions which teachers asked were equally "role-specific":

... the teacher covertly signals to the pupils what their role
as learners is to be ... Since most questions in the sample were
closed-ended, pupils were seldom invited to think aloud, to
generate new sequences of thought, to explore implications.
(1969, p.27)

The control which teachers exercised over speech in their classrooms had the
effect of limiting the pupils' role as learners. Asking questions is an
effective way of learning but teachers often encouraged their pupils to
become passive recipients of knowledge rather than engaging in active
enquiry. Another lesson which the hidden curriculum quickly transmits to
pupils is that their own personal experiences from outside school are
irrelevant. Barnes (1969) found that, "... those children who come up from
primary schools ready to explore personal experience aloud and to offer
anecdotal contributions to discussion cease to do so within a few weeks of
arrival." (p.25)

Barnes's findings have been described in detail because they may
provide the key to what frequently occurs in oral examinations. Conditioning and expectations built up throughout his school life encourage the pupil to adopt an "answering" role. He does not perceive questioning, initiating talk, extending or directing dialogue with teachers as part of his role. He limits his part to providing answers and often, as Q005 observed, his main concern is to find the "right" answers i.e. what he thinks the examiner wants to hear:

Often the candidates are not sufficiently confident to talk with ease ... and the discussion seems to prompt responses candidates think they are expected to give and not what they truly feel.

Viewed in this light, the fact that candidates sometimes seemed incapable of participating in a genuine exchange, that they were most comfortable when they were providing answers to questions and that they were content, for the most part, to keep these brief is much more understandable.

Conventional classroom roles may play an equally powerful part in determining the teacher's behaviour during the examination. If the teacher is used to playing a dominant role in the classroom, is used to doing most of the speaking, asking the questions and directing the talk, it is likely that when he assumes the role of examiner, he will carry his normal classroom behaviour over into the examination room.

It has been my experience from both internal and external moderation of oral examinations that some teachers are incapable of conducting the discussion in a fair or relaxed way. They have not the proper personality traits to make suitable interviewers or sympathetic listeners. Most simply talk too much. (Q310)

I have witnessed, as a moderator, too many discussions where the teacher/examiner, unwittingly or not, talks too much and asks questions which deny the candidate a full opportunity to express and explain opinions. (Q209)

Discussion is invariably teacher-directed whether in class or examination room. (Q577)

One of things we're bad at as teachers is actually involving children adequately in discussion. It was perfectly clear to me when I was doing the East Midland oral examination that for many children the only experience they had of oral work was the oral exam ... The extent to which a teacher dominates what is said in
a classroom and to which he talks without anyone else talking is, generally speaking, far too high. (101)

Thus, candidate and examiner alike may adopt stereotyped roles which will have a mutually reinforcing effect on each other's linguistic behaviour. An indication of how deeply engrained this pattern of behaviour can be was provided by the reading aloud, one-to-one conversation and group discussion tests used in Examinations Bulletin 11 (1966). Ostensibly these were very different procedures but, "In the event all three methods relied upon question and answer to a larger extent than had been planned or than might have been expected from the description of the methods" (p.21).

The form which the conventional test takes may actively reinforce this speech pattern. The method of initiating the discussion is crucial for establishing the mode of speech for the rest of the test and invariably it starts with questions from the examiner. It may be questions about the reading or about the candidate's talk. Either way, the discussion is launched by a series of questions. Of course, the aim is to, "... foster the natural growth of conversation" (EMREB Mode I Syllabus, 1970, p.62), but the danger is that the interview mode of communication, once established, may be difficult to break:

At first, he will no doubt make a number of fairly trivial remarks by way of setting the candidate at ease. Some of these remarks will necessarily be questions about the candidate, and it would seem likely in the circumstances that the answers will at first be very short. The danger is that this first exchange of question and answer can set up a pattern of expectation in the candidate which it is hard for the examiner to alter. The interview becomes a matter of exchanging comparatively trivial pieces of information, and no communication ever takes place.

(Examinations Bulletin 11,1966,p.24)

To put it bluntly, it seems that the conversation test often starts off "on the wrong foot" and that it would be advisable to find an alternative method of initiating it.

The test devised by Examinations Bulletin 11 (1966) illustrates this tendency. The researchers wished to avoid the, "... arbitrary list of questions about a candidate's accomplishments and biography" (p.12) and so
they decided to use a visual stimulus, a large-scale diagram of a town, which, it was believed, "... could lead easily to conversation on a very wide range of interests and ideas" (p.13). In the event, this technique did not lead as readily to conversation as its advocates had imagined it would. The interview format was still disappointingly marked. However, when this test is viewed in the light of Barnes's findings, it is hardly surprising that this pattern of speech was so pronounced. From the candidates' point of view, this must have been the kind of activity which they had participated in on many occasions in the classroom. They were presented with material selected by the examiner and with which the examiner was already acquainted. They were given time to study the diagram and then the examiner asked questions to which he already knew the answers or at least had a shrewd idea of the kind of replies which would be appropriate. The problem with this approach is that it places candidates in a situation which is familiar to them from the classroom thereby implicitly inviting a certain type of behaviour. Given the opportunity, some more enterprising students may take a more active part in the duologue but it is unreasonable to expect all candidates to do so simply because they have been provided with the opportunity. For many, it is bound to be a false opportunity which contradicts expectations accumulated over many years. As Examinations Bulletin 11 (1966) pointed out, one of the defects of this question/answer procedure as a way of initiating conversation is that it:

... has certain inhibiting effects which may seriously limit an individual's readiness to talk. In an examination, questions must be asked, but it is probably dangerous to isolate this form of utterance ... Communication is not only prompted by interrogation, and when it is, it may be damaged. (p.12)

This difficulty is exacerbated by the necessary brevity of oral examinations. Conventional tests last between ten and fifteen minutes and usually comprise several items. This means that the time available for the conversation is limited. It is not easy to plunge oneself into a sustained and serious discussion, maybe with a total stranger, in such a short period
of time and for no better reason than that the examination demands it. As Examinations Bulletin 11 (1966) noted: "... it is not easy to conduct serious discussion within the terms of oral examining ... each interview cannot last much more than ten minutes, and most worthwhile topics cannot be satisfactorily explored in this time" (p.26).

If conversation is to function as a genuine exchange, an alternative approach might produce better results.

11.4 CONVERSATION IN SIMULATED SITUATIONS

An alternative method of initiating conversation is the simulated situation. An early innovation in this field occurred in the syllabuses of the Associated Lancashire School Board where conversation was employed: 

"'...in fixed circumstances e.g. forgotten latch-key, sheltering from a storm, reporting a minor road accident etc.'" (quoted by Hitchman, 1968, p.220). The importance of this innovation was that it attempted to provide the conversation with a context and a purpose beyond the examination. In most examinations, conversation was a contrived exchange undertaken for its own sake and this was a subject of concern. For instance, Alchon (1964) argued that it is not, "... normal to find oneself engaged in a conversation purely as a technical exercise removed from any social or business occasion" (p.17). The simulated situation, recognising that conversation does not take place in a vacuum, attempted to provide it with the kind of context and purpose which gives the impetus to conversation in real life.

Conversation has been similarly used in other examinations. For instance, the Royal Society of Arts used a simulated telephone call. The situation behind the call was explained to the candidate before the test and a screen and dummy telephones were used to make the situation more authentic. The Manchester Research Project (early 1960's) devised a test in which the candidate was required to make a complaint or an enquiry:

"Complaint - 'You have bought an article from a shop and find it to be
faulty in some way. Return it to the shop and make a formal complaint asking for the money to be refunded or a replacement to be made.'

Enquiry - 'To a new acquaintance on your arrival in a new town, asking what facilities there are for leisure activities'" (quoted by Examinations Bulletin 11, 1966, p.33). Clearly, the range of situations which can be devised to test a candidate's conversational skills without resorting to a rigid interview format is extremely wide. More importantly, the situation can be devised in such a way as to place the onus on the candidate to do much of the talking rather than relying on cues from the examiner. This type of role-play provides conversation with a purpose beyond the vacuum of the examination room and tests specific language skills within a well-defined context.

There was support for a more functional and practically-oriented use of language amongst the respondents:

Not enough use is made of language for specific purposes, such as giving and receiving messages and instructions. Oral examinations ... could be a little more exact than they are now ...
Some oral examining could well be done with a bias toward practical uses of language. (Q121)

Linguistic competence is ... the ability to call to use the language appropriate to the task in its social context ... Any real test of a candidate's linguistic competence requires a range of tasks, each representing a different intended function of language, in a range of contexts, real or simulated. This presents enormous problems in organising assessment, with special preparation of the assessors, not only so that they understand the meaning of "linguistic competence" but also in such techniques as role-play. It is difficult but should perhaps be faced if oral assessment is to make more than very limited sense and tend to exercise a bias towards those best able to wax articulate in a teacher-controlled school context. (Q252)

Moreover, many teachers denounced the existing methods as vague and imprecise:

I have always felt that orals are very vague as far as assessment procedures are concerned. (I02)

Oral examining has always been be-devilled by vagueness and lack of any objective criteria which can be applied ... with consistency. (Q068)
... very little thought is put into what qualities one is assessing when one is assessing students' talk ... I'd like the criteria to be made much more explicit. (Il3)

... it's how to assess oral tests which is the problem ... I'd like a more systematic method of assessment. (Il4)

Likewise, when asked what he regarded as the main difficulties with oral examining, I07 replied: "One of the main ones is defining the criteria for assessment. This is particularly so in the conversation ... it is difficult to compare all sorts of things which might come up in a conversation." This lack of precision has been a subject of concern from the outset. For instance, in 1966 Inglis criticised the "rush" (Allen, 1977, p.17) of orals which had accompanied CSE arguing that: "We all need to be much clearer and firmer about what we expect from spoken English in schools" (p.222).

11:5 CONCLUDING REMARKS

Of the three elements in the Mode I test battery, conversation was judged to be the most satisfactory. It was praised in particular for its high level of internal validity. In spite of the conversation's popularity, the conventional approach to initiating it has serious drawbacks. Other methods of initiating conversation, which do not place the examiner in an interviewing role, need to be explored. In particular, the use of simulated situations merits further attention and may provide a valuable means of testing conversational skill.

12 GROUP DISCUSSION

12:1 INTRODUCTION

Group discussion is one of the most recent and less widely used innovations in oral examining. None of the pioneering public examinations used this method. In fact, of the twelve oral examinations in use prior to 1965, included in this survey, only one included group discussion. The
examination in question was the experimental test battery devised by the Manchester Research Project in the 1960's. This was typical of this method's development prior to 1965 in that it was small-scale and experimental. There were other small-scale, local experiments including the syllabus devised by the Stradbroke and Stowmarket Schools (see Holbrook and O'Malley, 1964). This syllabus required candidates, in groups of four, to discuss topical issues. The Dorothy Stringer School in Brighton also experimented with group discussion allowing children to group themselves according to the topics which they wished to discuss. None of the larger, better-established public examinations used this method. Thus, when CSE was introduced, group discussion was something of an unknown quantity. As Wilkinson (1965) observed: "Very little experiment has been done with such tests" (pp.76-77). The first NATE Bulletin (1964) described the procedure used by the Stradbroke and Stowmarket Schools in detail in an attempt to publicise this innovation and, in this way, CSE helped to disseminate this method. Nevertheless, even under CSE, group discussion remained a little-used procedure. In 1968, Hitchman's survey revealed that of the six assessment techniques commonly used, only one (the debate) was less widely used than group discussion. Less than half (i.e. 7) of the Mode I English Panels opted to include group discussion. The true extent of their tentativeness is revealed by the fact that only two of these seven areas made it compulsory.

The limited use of group discussion was a great disappointment to some commentators who regarded it as an opportunity to introduce radical change into this largely conventional area of testing. Thus, Dixon (1965) praised one syllabus which did include group discussion as, "... a considerable effort to break from the traditional sample ... an imaginative development, particularly the inclusion of group work" (p.32).

In the event, group discussion has never become a widely used testing procedure and its fate in the East Midland Region illustrates this. For
many years, it was not available in the Mode I syllabus but in 1977, with the introduction of dual certification, the marks allocated to oral work were increased from 20% to 30% and group discussion was introduced as an optional item. It is interesting to note that it was offered as an alternative to the talk and conversation and that the remaining item, reading aloud, was as impervious to change as it had ever been. However, its inclusion was short-lived because it was abandoned after only three years. The decision to abandon group discussion reflected its lack of popularity in the region.

Since the value of group discussion as a classroom activity is widely recognised, it is important to explore the reasons why it is so unpopular as an examining technique.

12:2 REASONS FOR ITS LACK OF POPULARITY

The extent of its unpopularity was established by Question 19 which asked respondents which examining technique they preferred (see p.256 for full details of this question). The resulting responses are recorded below.

**QUESTION 19: PREFERRED EXAMINING TECHNIQUE**

<table>
<thead>
<tr>
<th>METHOD</th>
<th>RESPONSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>One-to-one examination</td>
<td>(41%)</td>
</tr>
<tr>
<td>Group discussion</td>
<td>(3%)</td>
</tr>
<tr>
<td>One-to-one + group discussion</td>
<td>(19%)</td>
</tr>
<tr>
<td>Individual performance + peer group audience</td>
<td>(5%)</td>
</tr>
<tr>
<td>Suitability of methods depends on individuals' personalities</td>
<td>(27%)</td>
</tr>
<tr>
<td>Invalid response</td>
<td>(5%)</td>
</tr>
</tbody>
</table>

By far the least popular procedure was group discussion which only a handful of respondents (3%) favoured. However, a more substantial proportion of the sample (19%) would have liked to use group discussion in combination with the most popular method: one-to-one examining.

*Examinations Bulletin 11* (1966) came to similar conclusions arguing that
although, "... there should be more" (p.16) of it, "... it is doubtful if the group discussion method could be used without being supplemented by other methods" (p.21). Together these findings suggest that group discussion is most acceptable when it is used as part of a test battery. Nevertheless, there was little enthusiasm for this method of examining.

It seems that the main obstacle deterring a greater use of group discussion was the assessment difficulties which it poses. Teachers felt that oral assessment was extraordinarily complex even on a one-to-one basis and that to attempt to do it on a group basis would simply compound the difficulties:

Assessment of oral ability is without doubt unreliable because of (1) the ephemeral nature of oracy (2) the non-verbal factors involved (3) the impossibility of appeal or recall (4) the very complexity of the factors involved. All of these difficulties are compounded by an attempt to do it in a group situation. (Q001)

When you've got a group rather than an individual, the difficulties you have with an individual are multiplied. (I02)

The group discussion is a very complex method of assessment, almost impossible to assess accurately. (Q364)

If one is to assess oral performance properly, sufficient time must be given to each candidate on a one-to-one basis. It would be difficult to assess group performance and observe the necessary criteria. (Q619)

I am most unhappy about the group discussion method. How is it possible to assess several people at the same time? (Q341)

I02 illustrated the complexity of the problem when he recounted an experience at a NATE conference in which a group of experienced oral examiners tried to mark a video-taped group discussion:

... when we got to the end of the tape we discovered that we couldn't even agree on the number of contributions that each person had made. It varied from eight to fourteen because immediately you were wondering whether they had to speak a sentence for that to be a contribution, whether "Oh!" or "Rubbish" was a contribution ... If a group of experienced people can't even agree on the number of contributions and you've got five or six other criteria to apply to each contribution, how on earth do you start to be reliable? This is how complex that matter is.

Clearly, many of the respondents regarded assessment in a group setting as
12:3 THE ROLE OF THE EXAMINER

One of the drawbacks of the one-to-one conversation is that the examiner is personally involved in that which he assesses making it difficult for him to devote sufficient attention to assessment. In theory this problem should be eased in the group setting because there are many more participants. Once the discussion has gained momentum, an examiner ought to be freed from the need to make frequent contributions and should be able to concentrate on assessment. In practice, however, candidates frequently rely on the examiner to sustain the discussion and treat him as the focus of their talk. Thus, the examiner who took part in Examinations Bulletin 11 (1966) noted that: "Candidates preferred to speak to me or through me" (p.15) and confessed that:

"I was reasonably satisfied with the contribution of the candidates in the early groups until my observer pointed out that the pattern was of a gently probing question and extended reply rather than of a discussion. I attempted to make plain to following groups ... that they would be encouraged to talk to each other without reference to me". (p.15)

In spite of this, the observer reported that:

"... it proved very difficult to get any discussion going and it generally remained an oral examination in which the children answered the examiner's questions". (p.15)

This pattern, of the examiner directing the talk through the use of questions and the candidates relying on him and limiting their part to supplying answers, has a familiar ring to it. Similar behaviour emerged as a characteristic of the one-to-one conversation (see Section 11:3). In both cases, the behaviour seems closely related to the classroom stereotypes identified by Barnes (1969, 1976). It is hardly surprising that this pattern of behaviour should emerge in the group discussion because children are familiar with being in a group situation dominated by a teacher's talk. The fact that children expressly instructed to discuss a topic amongst
themselves should rely on the adult present illustrates how deeply-engrained these stereotyped roles are. Moreover, this behaviour appears to be just as engrained in examiners. Thus, the examiner in *Examinations Bulletin 11* (1966) admitted that he was unaware that he had lapsed into a questioning-controlling role until this was pointed out to him by his observer. It seems likely that the success of these testing procedures is strongly linked to the kind of activities pupils and teachers are used to engaging in in the classroom. Indeed, there was evidence in *Examinations Bulletin 11* (1966) that where schools used these procedures as a normal part of their work, the tests were more successful (see p.15). Thus, if group discussion and conversation are to play a more meaningful role in the testing of spoken English, the starting place for reform must be the classroom.

It is not only the validity of the test which is threatened by the tendency to lapse into the interview mode. An examiner who has become the focal point of the discussion may become so preoccupied with directing it that his ability to make an assessment may be seriously undermined. Admittedly, this is also a difficulty in the private conversation but in a group discussion, where an examiner may need to make five or six times as many separate assessments, the problem is magnified.

Group discussion also illustrates another "fundamental contradiction" (Allen, 1977, p.15) in oral examining. Most English teachers have had the satisfying experience of initiating a discussion in class which gathers momentum and becomes capable of sustaining itself without their support. When this happens, most teachers are glad to relinquish control of the dialogue to their pupils. Likewise, when this technique is carried over into the examination, it is hoped that the discussion will generate its own dynamic and that the examiner will be able to fade into the background. However, this illustrates the confusion which can occur when an activity which is valuable in the classroom is carried over into the examination
There is a fundamental contradiction between the examiner's desire to become as unobtrusive as possible and the need for him to impose his authority on the discussion in order to make it a fair test for all concerned. Whereas the class teacher may be content to allow certain pupils to dominate the discussion knowing that he will be able to provide other opportunities for less-forthcoming students to talk about different issues and in different situations, for the examiner, the group discussion is part of a "one shot sudden death" examination (Rowlands, 1974, p. 87) and it is vital that he should apportion the talk fairly between the candidates. The fact that he stage directs the discussion means that he is imposing his authority on it. No matter how skilfully he redirects the dialogue and no matter how unobtrusively he draws retiring candidates into the conversation, he is still exerting a subtle form of control over it. This is bound to hamper the natural growth of discussion and to encourage deference to him. This contradictory state of affairs is illustrated by Examinations Bulletin 11 (1966) where the hope that the examiner, "... should be able to behave more or less as a spectator" (p.15) was juxtaposed with the admission that it would be, "... his responsibility to entice the timid into the rush of talk" (p.15). Likewise, the examiner instructed the candidates "... 'to talk to each other without reference to me'" (p.15) and yet it was conceded that, "... the examiner was often obliged to control discussion in order to maintain it" (p.16). It seems that there is a discrepancy inherent in group discussion as a testing procedure. The role which an examiner must play in its assessment undermines its growth as a genuine exchange between peers.

Furthermore, rather than releasing the examiner to concentrate on his assessment, it may make even greater demands than the one-to-one conversation. The administrative demands are considerable: not only must the examiner sustain the flow of speech, he must also devise ways of drawing retiring candidates into the conversation and limiting the part played by
outspoken candidates. Some respondents pointed out that the fair administration of a group discussion makes large claims on the examiner's skill:

Great skill is needed on the part of the examiner to give all participants equal opportunity to express their ideas. (Q156) Examinations Bulletin 11 (1966) also pinpointed this as a source of great difficulties:

The examiner's report admirably presents the very delicate problems of adjustment and distribution of attention needed, problems which make assessment extremely difficult when there is so much else to be borne in mind. (pp.15-16)

Clearly, the greater the administrative burden, the more difficult assessment becomes. Since group discussion is particularly difficult to administer, it follows that its assessment is also peculiarly difficult. As Inglis (1966) observed:

... all the onus falls on the teacher-examiner ... he must mobilise the conversation, encourage the timid, control the exuberant and the whole time be as unobtrusive as possible. This, of course, is to require no more than that he behave as a good teacher always should, but if he is also examining then the responsibilities become more peculiar. He is then obliged to provide everyone with a more or less equal share in the conversation without allowing it to revolve around himself. The difficulties of assessment are acute ... (p.222)

The interplay of personalities is another example of the way in which problems faced in the one-to-one test are compounded in a group test. This was found to be, "... a major source of error" (Hitchman, 1966, p.41) in one-to-one conversation tests but it is bound to be more problematical when there are five or more people involved in the discussion. Clearly, many of the problems raised by group discussion are not different in kind to those created by other personal interaction tests but the scale on which they occur magnifies their complexity.

12:4 THE POSITIVE FEATURES OF GROUP DISCUSSION

One advantage of group discussion is that because it occurs naturally in everyday life, it has a good level of internal validity. Moreover, the
prospect of performing alone before an examiner may be more daunting than being allowed to contribute from within the supportive framework of the peer group. This was perceived as one of the strengths of group discussion:

One valuable addition was the opportunity for ... group discussion ... instead of merely the one-to-one relationship of teacher-examiner and pupil which, to my mind, is a bit unnatural. (I10)

I12 was of a similar opinion. He favoured the group setting because he felt that in the conventional, one-to-one test, "... nerves and the strangeness of the situation mean that they do not, in fact, perform as well as they might." Whereas less natural tests such as reading aloud and giving a talk were noted for awkwardness and stilted speech, group discussion was found to:

... combat what may be thought of as the distasteful "artificiality" of certain other procedures ... Group discussion commends itself as encouraging a "natural" argument between contemporaries ... (it) can encourage a desirable energy, freedom and warmth ... self-consciousness is more totally lost in this than in any other method. (Examinations Bulletin 11, 1966, pp.14 and 16)

Another positive feature of group discussion which helps to relax the atmosphere is the length of the test. Because a group of candidates is assessed simultaneously, the length of the test can be increased. As I12 pointed out: "... because you were dealing with five people at once, you could allow three quarters of an hour." This contrasts with the necessary brevity of a one-to-one test which it is impracticable to extend much beyond ten minutes. Thus, an examiner must accomplish a great deal in a short space of time. In order to assess the candidate's linguistic competence, he must establish a rapport with him. This means relaxing nervous candidates and "drawing out" shy ones in addition to initiating a serious conversation. All of this must be achieved in the five or so minutes which are normally available for the conversation and this places an enormous burden on the skill and tact of the examiner.

The fact that the discussion takes place between peers is also
important. A candidate's contemporaries will be quick to point out irrational or inconsistent arguments. Therefore, the need to argue clearly and justify one's standpoint is that much keener in the presence of peers. The "cut and thrust" of a lively discussion amongst contemporaries can provide an acid test of an individual's linguistic competence. As Examinations Bulletin 11 (1966) remarked:

Group discussion commends itself as encouraging a "natural" argument between contemporaries ... In such a situation, as argument gathers momentum, then fluency, vigour and variety should increase with the heat of the discussion. (pp.14-15)

Thus, a successful group discussion should show how well a candidate can explain or justify his standpoint. It should be difficult for him to rely on prepared statements and so his ability to use language to cope with the exigencies of the moment should be clearly revealed. As 101 pointed out, it is also important for pupils to learn the social skills which are tested by a group discussion:

... it is important that pupils ... learn the etiquette of group discussion and the qualities that go with it like tolerance of other people's points of view and the capacity to express their own point of view in as formidable a way as they can.

Although most respondents agreed that group discussion was peculiarly difficult to assess, there were some arguments in favour of it as an assessment procedure. For instance, when asked whether he thought that group discussion was more difficult to assess than one-to-one conversation, 112 replied:

I think they're both difficult to assess but if you've got ... five people involved, it's much easier to put those five in a rank order. The difficulty then is to tie that rank order with all the other little groups that you have assessed but I would have thought that it was no more difficult than in the present method of individual interviews.

Thus, the opportunity to rate a number of candidates simultaneously and reach a "within group" order is an advantage of assessing in a group setting. In contrast, when an examiner assesses each candidate independently, by the time he reaches the sixth candidate, his memory of the first candidate's
performance may have become blurred.

Despite its extreme unpopularity, group discussion does have some positive attributes. Perhaps the most significant tribute of all came from a group of candidates who had experienced group discussion:

... the majority seemed at ease and answers to our enquiries suggested that many had enjoyed the experience and had not found it inhibiting. (Examinations Bulletin 11, 1966, p.15)

12:5 CONCLUDING REMARKS

Group discussion illustrates a dilemma which is central to so much of English between that which is valuable in the classroom and that which raises problems for assessment. Group discussion undoubtedly raises many of the characteristic problems of personal interaction tests in a particularly acute form in that the problems are similar in kind but greater in magnitude. Nevertheless, it does display some positive features. Despite enumerating its drawbacks, Examinations Bulletin 11 (1966) concluded that it could be both "desirable and profitable" (p.16) and that, "... there should be more of it" (p.16). At the very least, schools which are enthusiastic about the approach should have the opportunity to use it. As I10 argued:

Although it has not been used extensively, as far as I know in the region, I think it is something which ought to be there.

13. THE CONTINUOUS ASSESSMENT OF ORACY

13:1 INTRODUCTION

One of the most significant innovations pioneered by CSE was continuous assessment. Indeed, Dixon (1977) regarded it as: "... the main new contribution from the experiment of the last decade" (p.8). However, in 1965, it was something of an unknown quantity. In comparing it with conventional examinations and objective tests, Eggleston (1967) described it as: "... least implemented and least well-known" (p.60). Its application
to oracy was especially novel since oral examinations were themselves at an experimental stage. Nevertheless, there were some opinion leaders who felt that continuous assessment had an important role to play in spoken English. For instance, the NATE document, *Criteria of Success in English* (1965), argued:

... a satisfactory scheme for the assessment of oral work will almost certainly need to take the form of a cumulative internal assessment of speech in its natural uses throughout the fifth year, complemented by externally moderated tests in the course of the year. (p.32)

Moreover, it was extremely critical of conventional testing techniques describing them as "unnecessary", "dangerous" and "too limited" (p.30). *Examinations Bulletin 1* (1963) was much more guarded in its approval of continuous assessment. It conceded that:

Some people would claim that in many ways course work is the better test of a pupil's ability since continuous assessment is more reliable and valid than a once for all examination. (p.22)

However, its tone was cautious and against these advantages were placed: "... the obvious difficulties of establishing comparability and the optimism of some teachers about their own pupils' standards" (p.22). Nevertheless, it acknowledged that in subjects of a transitory nature, continuous assessment had a part to play (see p.22).

From the outset, continuous assessment played a part in the Mode I oral examination in the East Midland Region. Like reading aloud, it is an element which has survived all radical alterations of the Mode I syllabus and its details have remained virtually unaltered throughout the years. One change which it has experienced is an increase in its mark weighting. In the original syllabus, it accounted for a quarter of the marks allocated to oracy (i.e. 5 marks out of 20). However, in 1977, due to the reorganisation of the English syllabus into separate language and literature syllabuses, the marks awarded for spoken English were increased from 20 to 30. Five of these extra marks were given to the class teacher's assessment which has accounted for one third of the oral assessment since that time.
The application of continuous assessment to spoken English represented a fresh departure in 1965 but teachers in the East Midland Region now have many years' experience of this procedure behind them. Therefore, it seems timely to enquire how this procedure has worked out in practice.

13:2 GENERAL ATTITUDES TO THIS PROCEDURE

Question 22 provided a broad indication of attitudes towards this procedure. It stated:

The class teacher has ten marks out of a total of thirty marks in the Mode I oral assessment allocated for his/her evaluation of a candidate's oral performance throughout the course. This allocation:

(i) represents an adequate proportion of the total mark
(ii) should be increased
(iii) should be decreased

The format of this question was similar to that of Question 21 which dealt with attitudes towards the mark weighting of oracy in general. Since the question formats were similar, it was possible to compare the responses on each question. This demonstrated interesting differences in the sample's attitudes towards the weighting of oral work in general and of continuous assessment in particular. The resulting responses are recorded below.

<table>
<thead>
<tr>
<th></th>
<th>30% FOR ORAL EXAMINATION</th>
<th>10% FOR CONTINUOUS ASSESSMENT</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADEQUATE</td>
<td>59%</td>
<td>55%</td>
</tr>
<tr>
<td>INCREASE</td>
<td>5%</td>
<td>23%</td>
</tr>
<tr>
<td>DECREASE</td>
<td>34%</td>
<td>19%</td>
</tr>
<tr>
<td>(INVALID RESPONSE)</td>
<td>2%</td>
<td>3%</td>
</tr>
</tbody>
</table>

The differences on the first category (adequate mark weighting) were negligible; similar proportions of the sample felt that the mark weightings were adequate for the examination as a whole and for the continuously assessed element. More striking differences emerged over the next two categories. Only a tiny proportion of the respondents (5%) wished to
increase the mark weighting for oracy in general compared with a much larger group (23%) who wished to increase the weighting of the continuously assessed component. Furthermore, a substantial proportion of the respondents (34%) thought that the mark for the oral examination ought to be decreased whereas less than a fifth (19%) wanted the marks allocated to continuous assessment to be reduced. Overall these responses suggested that the sample favoured continuous assessment as an approach to oral examining and its popularity, relative to the conventional examining procedure, was amply demonstrated. These findings are consistent with the decision to allocate extra marks to the teacher's assessment at the time of the reorganisation of the English syllabus (1977). Half of the additional marks allotted to oracy were given to the course assessment which had the effect of upgrading it within the syllabus (i.e. from a mark weighting of 25% to one of 33%).

13:3 EVIDENCE OF DISSATISFACTION

Despite the favourable attitude towards continuous assessment revealed by Question 22, the general tenor of the written comments was critical. Of those respondents who volunteered comments on this component, the vast majority were dissatisfied with it. Initially this discrepancy between the qualitative and quantitative data came as a surprise. However, much of the dissatisfaction derived not from the inherent nature of continuous assessment but from the way in which it had been put into practice in the East Midland Region. This may, in part, explain the discrepancy between the qualitative and the quantitative data. It seems that while many teachers were committed to continuous assessment in principle, some were dissatisfied with the way in which it had been translated into practice in the East Midland Region.

Although the course assessment accounts for a third of the marks allocated to oracy, it is entirely unmoderated. Admittedly, the CSE
regulations allow for up to 10% of the total mark to be unmoderated; nevertheless, this was the principal complaint against this component:

I am unsure about the allocation of 10% to the teacher for oral contribution ... many teachers abuse this facility in my opinion. (Q214)

10% teacher's contribution widely misused. (Q212)

The class teacher's assessment is open to abuse; it should be dispensed with as soon as possible. I suspect that marks awarded by individual teachers (even after internal moderation) are most unreliable. If the marks within a school are unreliable, the marks between schools are very unreliable. (Q341)

I'm not keen on 10% oral assessments that are totally teacher-controlled without moderation. I'm suspicious whether people do get a good spread of marks or, if they do, what motives they use and what criteria they adopt in obtaining these marks and I don't think they're necessarily oral criteria. There are so many other factors which can influence a teacher's judgement on the ten marks. (I14)

I would like to reduce the oral mark to 20 and the bit that I would like to get rid of is the unmoderated teacher's assessment ... I think that's bound to be unreliable. (I02)

The class teacher's 10 out of 30 is by definition (i.e. not moderated or standardised) unreliable. I consider that any public examination should be above all reliable. (Q001)

... one third of the oral mark is not subject to moderation which seems to me too high a fraction and unfair. (Q036)

Clearly, an item which is unmoderated is likely to command less respect amongst users of an examination. Many respondents feared that, without moderation, a procedure is open to abuse and this undermines its standing in the eyes of users. The practice of including unmoderated items attracted criticism from the outset. Thus, although Dixon (1965) strongly recommended the use of continuous assessment (see p.32), he argued that its efficacy depended on an effective system of moderation. Referring to the decision by several boards to operate unmoderated items, he observed:

"... these grades are unlikely to be as trustworthy ... in the final event, moderation must be systematic and statistically sound" (p.32). Likewise, Examinations Bulletin 1 (1963) urged that: "Course work should always be open to moderators' inspection" (p.22). Thus, although the inclusion of a
10% unmoderated element is authorised by CSE, this should not be taken to imply that it has proved acceptable to those who use the examinations.

A second aspect of this component which has caused much concern is its lack of definition. From the outset, the class teacher's assessment has been an exceedingly vague component of the syllabus. The early syllabuses instructed teachers to: "... award marks on the basis of the candidate's skill in spoken English as revealed during the course" (Mode I Syllabus, 1965-1973). However, no attempt was made to specify those skills or the criteria by which they might be judged. The imprecise nature of this assessment was made even more explicit in the 1974 syllabus which informed teachers that: "This mark is not moderated and may be awarded on whatever criteria seem to the class teacher appropriate" (p.11). This licence to award marks, "on whatever criteria seem ... appropriate" was included in the syllabus until 1980 and it clearly lay at the root of the dissatisfaction which many of the respondents expressed:

More guidance on marking is required for teacher's oral assessment. (Q428)

Teacher's oral mark is so obscure as to be useless. It's no more than a pat on the head for the most able students. (Q296)

Oral examining has always been be-devilled by vagueness and lack of any objective criteria which can be applied by teachers with consistency. (Q068)

I find it very difficult to allocate marks out of 10 for a candidate's oral performance throughout the course. (Q004)

I'm torn between saying it should be vastly increased and that it should be scrapped altogether. It all depends on what it's marking. At the present, it seems to me that it's marking something which is extremely vague ... it's just a kind of bonus the teacher can give ... (I12)

I believe that the marks awarded under this section are often only distantly related to the candidate's real performance. (Q150)

Some respondents argued that one way to combat this vagueness would be to stipulate the inclusion of specific tests:

... the oral is a vital and significant element of the course... but I am concerned as to the methods of assessment. I should like to see more specified tests to be completed during the two
years as a method of providing more genuinely and readily assessable course work. (Q129)

There are so many factors influencing a teacher's judgement on the ten marks. This is why I'd like to see a series of little tests during the course. (Il4)

The EMREB system seems inadequate to form an opinion of oral ability: there should be much greater provision for oral assessment during the course. (Q133)

Although CSE was intended to promote freedom for individual teachers, it is clear that many teachers were hostile to what they regarded as vagueness. They preferred their tasks to be carefully defined, detailed and precise. The full significance of this lack of definition only becomes apparent when one realises that no other area of the syllabus was so imprecise. Indeed, the vagueness surrounding this element provides a striking contrast to the other continuously assessed component, the literature folio, where the type of work considered suitable for inclusion was specified as were the criteria by which it should be judged. The absence of moderation would seem to make the provision of guidelines for assessment all the more necessary. This 10% oral component seems even more incongruous in this light.

It is interesting to note that in the more recent syllabus for 1981-82 an attempt has been made to define the assessment criteria. Therefore, of the two sources of dissatisfaction discussed in this section (i.e. lack of moderation and lack of definition), this last is relatively easy to rectify. Providing an effective system of moderation is, however, a more insuperable problem.

13:4 MODERATION : THE LOGISTICAL PROBLEMS

It has already been noted that oracy raises special problems for assessment and nowhere is this better illustrated than in the difficulties created by providing an effective system of moderation for a continuously assessed component. This point may be clarified by comparing continuous
assessment in the oral and the written medium using the literature folio and the class teacher's oral assessment. The literature moderation is facilitated because performance during the course automatically leaves behind it a permanent, written record. Thus, during a single visit, the moderator can look at work which represents performance/assessment at different points during the course. In fact, candidates are required to include sixteen pieces of work with at least one from each term during the course which gives the moderator access to assessments made at many points during the course. Oral performance, on the other hand, is transitory and all existing methods of recording it are open to serious objections. Therefore, to provide a comparable system of moderation for oracy would entail vast numbers of teachers leaving their classes to visit other people's classrooms throughout the course. The logistical problems which such a system would create make it impracticable. Some boards have attempted to overcome this by carrying out a single moderation but such a solution has obvious drawbacks. Firstly, even such limited moderation entails a considerable amount of time and administrative effort and, as many respondents pointed out, in its present form, the oral component is already a drain on resources. Despite the extra effort it would involve, such an exercise could only go a short way to providing an effective moderating system for as Il2 pointed out: "... if you're going to do a proper marking exercise, you'd have to do that more than once ... You ought to get a better idea of the teachers' standards than you would get on just one marking". Some respondents regarded this problem as insuperable:

The constant difficulty with the internally awarded oral mark is that it is incapable of effective moderation. (Q201)

The 10% allocation by class teachers cannot be moderated ... and is therefore unsatisfactory. (Q596)

Moderation is always the greatest problem with any oral assessment. (Q223)

Clearly, a properly moderated continuously assessed oral component would be
a great drain on a resource which is finite: teachers' time. Ultimately the confidence it commands depends upon the amount of time, money and administrative effort an examining board is prepared to invest in it. The crux of the problem in the East Midland Region is that such a demanding procedure as continuous assessment is used as a supplement to a final examination which carries most of the marks. The question this raises is: does the board have sufficient resources at its disposal to do both efficiently?

13:5 PERSONALITY FACTORS

Another feature of the assessment over which respondents expressed concern was the influence of personality factors. It was tempting to assume that the personality problems of oral assessment would be alleviated by continuous assessment where the disconcerting final examination is replaced by assessment in the natural setting of the classroom. Moreover, taking a "movie" rather than a "snapshot" (Rowlands, 1974, p. 44) of performance ought to provide a more rounded and accurate assessment. It was rather surprising, therefore, to find that some respondents regarded "personality" as a critical factor affecting continuous assessment. Q036 noted that oral performance throughout the course:

... can vary greatly and be affected by personal factors. Some candidates will make little oral contribution with a particular teacher when with another member of staff they may have plenty to say. (Q036)

Oral performance during the course should be a part of the final assessment but assessment and internal moderation do present problems. In particular, one wonders how important the teacher and his/her relationship with both the group and the individual is in determining how candidates perform over the two years. (Q143)

The main problem ... is the occasional unease a pupil may feel (whether conscious or subconscious) on talking to a certain teacher ... There is no solution I suppose but I believe it causes some pupils to underachieve in this part of the examination. (Q215)
A teacher assessing his class on oral contribution over the year is assessing what his own personality has largely produced. Some pupils will find him an acceptable audience; others will not. Some will find him intimidating even when talking to their peers in his presence; others will not. This obvious interaction invalidates his assessment. (Q174)

Because it is done by the class teacher, in some respects, it can be an examination of the class teacher. If a teacher... asks the right questions then the candidate will respond but with some teachers the child has less opportunity to show his oral ability (I10)

In fact, it is hardly surprising that similar difficulties arise in continuous assessment as occur in other personal interaction tests since both involve the interplay of personality between assessor and assessed. In both cases, the examiner's personality provides a confounding variable in the measurement of oracy. However, it was not only the effects of the examiner's personality which bothered respondents. Some expressed concern about inhibiting personal qualities which the candidate might possess. A number of them identified a particular type of pupil who caused them anxiety arguing that the retiring but able candidate may be at a serious disadvantage in this part of the assessment:

... one of the biggest problems is that the teacher's mark is awarded for the candidate's "contribution throughout the course." Sometimes a good candidate may be shy and say very little within the course, even though attempts are made to combat this... the teacher may find out from lunch-time conversations with the candidate that the latter has far more oral ability than is revealed in lesson time. (Q341)

The 10% for oral contribution is very difficult to assess as some quiet pupils are cogent and articulate when they do eventually speak whereas others, who frequently contribute to class discussion, rarely say anything worthwhile. (Q033)

The best girl in my CSE set hasn't a hope in the oral examination as she can hardly speak, it seems. Yet I feel she is capable of holding up a good office job that doesn't involve too much oral involvement but her qualifications for this job may be affected by her oral examination performance. (Q396)

I feel that the assessment of oral contribution often turns out to be a reward for extroversion and confidence. (Q447)

Clearly, teachers are sometimes faced with the unwelcome task of marking down candidates of considerable ability on their course contribution despite
their knowledge that these candidates are articulate in one-to-one exchanges. Some teachers expressed surprise at the extent to which such pupils often "opened up" in one-to-one situations. Although this small band of pupils presented their teachers with serious difficulties, on the whole it seems that continuous assessment can provide a more relaxed and natural atmosphere in which to evaluate oracy.

13:6 CONCLUDING REMARKS

Continuous assessment is, in effect, a multiplicity of assessments and this compounds the specialised problems which examining in an oral medium raises. Nevertheless, its right to inclusion in the examination on the grounds of internal validity can hardly be disputed. However, a properly moderated procedure makes considerable demands on an examining board's resources and, therefore, it seems that if continuous assessment is to be included, it should form the principal component of the assessment rather than supplementing a more heavily weighted final examination which takes up most of the limited resources available.
CONCLUSION

This final section draws together what seem to have been the main achievements of CSE English in the East Midland Region and indicates the outstanding problems with which teachers in the survey were concerned.

CSE was the first national examination which succeeded in involving large numbers of practising teachers in decisions about syllabuses and examining procedures. One important consequence of this is that examinations have become more responsive to the changing needs of schools. Whereas previous examinations were sometimes criticised for becoming fossilized, EMREB's Mode I syllabus illustrates the ease with which CSE has evolved in response to changing needs. In little more than a decade, certain procedures were tried and abandoned (e.g. single certification, the thematic approach to literature, group discussion and the talk). Other procedures were expanded (e.g. the literature assessment and the oral assessment) and others were abandoned only to be brought back again (e.g. the Creative Writing Folio). Thus, change and development have been constant features of the examination. Another measure of the success of CSE is the infrequency of complaints about backwash effects. Prior to 1965, perhaps the most common and bitter complaint about examinations was their backwash effect on teaching. Moreover, it appears that, for some teachers, closer involvement in assessment has also had far-reaching consequences for their sense of professional worth.

It must also be noted that EMREB pioneered one of the boldest experiments with continuous assessment to emerge from Mode I: a 100% continuously assessed literature course, an innovation which the
respondents judged to be a great success. Respondents were also enthusiastic about oral assessment, almost all of them regarding it as a vital component of any future 16+ examination.

In spite of what appears to have been its considerable success, some aspects of CSE persist in giving cause for concern. A number of respondents were worried about what they saw as the limited amount of teacher involvement and IO5 went so far as to claim that, even under CSE, an examining élite had emerged in the region. Many teachers felt that it was only reasonable that a greater level of commitment should be expected from them if more account were taken of the demands which involvement in examining made on their limited time. Over 75% of the sample felt that the demands of CSE on their time and energy were either "extremely high" or "higher than the demands made by their other teaching commitments".

Not only did CSE elevate practising teachers to a position of power; it also charged them with the task of making revolutionary changes. In this respect, teacher involvement has sometimes proved to be disappointing. CSE made practising teachers the pivot of the reform while failing to take account of their lack of examining expertise. Many teachers had little or no experience of examining, and those who did mostly derived their experience from GCE and its imitators which made them ill-equipped to make the far-reaching changes expected of them.

The problem of whether English is a unified whole or an amalgamation of separate components persists. Recently, EMREB has abandoned single certification in favour of separate certificates for language and literature, although most respondents claimed to be committed to the unified approach to teaching. However, even when EMREB employed single certification, the traditional divisions into language, literature and oracy were maintained. Respondents questioned the validity of this fragmentation - in particular, the justification for excluding spoken English from the literature assessment and confining it to language. The role which oracy and
literature may play in each other's assessment needs further investigation. If a unified approach to English is preferred, a whole new concept of "examination English" which transcends conventional divisions needs to be developed.

Continuous assessment suffers, in several aspects, from being viewed as a soft option, as more open to abuse and as being less reliable than conventional, final examinations. Its uncertain status may explain why so many continuously assessed courses have such demanding written requirements. The number of syllabuses where course work requirements were felt to be excessive, and where these requirements had been allowed to persist, suggested that the problem went deeper than a mere miscalculation. Indeed, if this alone was the difficulty, it would be relatively easy to correct. Underlying prejudices against this method need to be tackled.

Excessive course work requirements have had adverse consequences in the classroom, which suggests that class-based assessment does not always eliminate backwash effects - it simply alters the form which they take. In an external examination, the teaching syllabus tends to contract to become synonymous with an examining syllabus which is necessarily narrow and selective. In continuous assessment, on the other hand, the teaching syllabus is liable to become overloaded by an over-ambitious examining syllabus. Thus, it has proved to be remarkably easy for the tyranny of external examinations to be replaced by the tyranny of the course work folder. Although it is sometimes argued that there should be a clear distinction between a teaching syllabus and an examining syllabus and that teachers should not allow one to dominate the other, in practice this is difficult to avoid.

Although CSE has gone far towards making English and examinations more compatible, important problems remain to be resolved.
RESEARCH QUESTIONNAIRE ON THE ASSESSMENT OF ENGLISH (LANGUAGE AND LITERATURE) BY THE C.S.E. EXAMINATION

Compiled by Valerie Brooks

LAY-OUT OF THE QUESTIONNAIRE

The questionnaire is divided into four sections. The first section is concerned with brief details about yourself and your school. Section Two deals with assessment techniques and, bearing in mind current uncertainty surrounding the 16+ proposals, the composition of 'C.S.E. type' examinations in English. Section Three focuses on the assessment of oral performance and Section Four concludes by surveying the assessment of the study of Literature.

HOW TO FILL IN THE QUESTIONNAIRE

There is only one type of response required from you in this questionnaire. In each question you will be asked to place a tick (/) in the appropriate box or boxes. There are two types of question employed. The first type takes the form of a list of statements, each of which is to be rated on a 5 point scale. You will be asked to rate each statement on the scale by placing a tick in the appropriate box. The second type of question consists of a list of alternatives with a column of boxes placed alongside it. Here you are asked to tick the appropriate box or boxes.

The only exception to this is that at intervals throughout the questionnaire spaces are left blank for you to amplify your point of view on any of the issues raised in the preceding questions. These open questions, requiring a written response, are not obligatory. However, if any of the preceding questions do touch on experiences that you have had, or opinions that you hold, these will be very welcome. It must be stressed that any information or views that you are able to express in these questions will provide a valuable aid to the investigation.

DEFINITIONS OF SOME TERMINOLOGY EMPLOYED IN THE QUESTIONNAIRE

Whenever 'the C.S.E. ability range' is referred to, it should be taken to mean the 40th - 80th percentile in the overall ability range. (Alternatively, you may be more familiar with this group being described as the 20th - 60th percentile, although both sets of numbers describe the same ability band. In both definitions the top 20% and the bottom 40% of the ability range are excluded.)

Whenever the terms 'most able', 'average' or 'below average' are used they may be defined thus: 'most able' refers to those pupils who can reasonably be expected to achieve a Grade 1 or 2, 'average' describes those candidates who are likely to gain a Grade 3 or 4 and 'below average' refers to those who are expected to obtain a Grade 5 or 'Ungraded'.

The initials E.M.R.E.B. represent the East Midland Regional Examinations Board.

Please ignore the boxes in the right hand margin of each page; they serve to facilitate computation procedures.
SECTION I  PRELIMINARY INFORMATION

In the questions which follow, please place a tick or ticks in the appropriate box(es).

1. For how many years have you been a teacher of English?
   (i) Up to two years
   (ii) 2+ to 5 years
   (iii) 5+ to 10 years
   (iv) 10+ to 15 years
   (v) 15+ to 20 years
   (vi) 20+ to 25 years
   (vii) Over 25 years

2. For how many years in all have you been teaching C.S.E. English?
   (i) Up to two years
   (ii) 2+ to 5 years
   (iii) 5+ to 10 years
   (iv) 10+ to 15 years

3. Which public examination in (i) English Language and (ii) English Literature are taken by fifth formers in your school?
   (i) in English Language
      (a) C.S.E. only
      (b) 'O' Level only
      (c) Both C.S.E. and 'O' Level
   (ii) in English Literature
      (a) C.S.E. only
      (b) 'O' Level only
      (c) Both C.S.E. and 'O' Level

4. Which Mode(s) of C.S.E. English do you personally teach? Please tick the appropriate box (i) for Language and (ii) for Literature.
   FOR LANGUAGE    FOR LITERATURE
   (a) Mode I
   (b) Mode III
   (c) Mode III

5. If you use a Mode III syllabus, is it:
   (i) An individual school Mode III syllabus
   (ii) Group of schools Mode III syllabus
   (iii) Both group and individual school Mode III syllabuses

326
6. What proportions of continuous coursework assessment and final external examination would you prefer to see in the assessment of English Language and English Literature? Below are lists of options for both the Language and the Literature assessment.

(i) Which of these alternatives provides the most satisfactory means of assessing English Language? Please place one tick in the appropriate box.

<table>
<thead>
<tr>
<th>INTERNAL COURSE WORK</th>
<th>EXTERNAL EXAMINATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>100% - 0%</td>
<td>1</td>
</tr>
<tr>
<td>75% - 25%</td>
<td>2</td>
</tr>
<tr>
<td>60% - 40%</td>
<td>3</td>
</tr>
<tr>
<td>50% - 50%</td>
<td>4</td>
</tr>
<tr>
<td>40% - 60%</td>
<td>5</td>
</tr>
<tr>
<td>25% - 75%</td>
<td>6</td>
</tr>
<tr>
<td>0% - 100%</td>
<td>7</td>
</tr>
</tbody>
</table>

(ii) Which of these alternatives provides the most satisfactory means of assessing English Literature? Please place one tick in the appropriate box.

<table>
<thead>
<tr>
<th>INTERNAL COURSE WORK</th>
<th>EXTERNAL EXAMINATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>100% - 0%</td>
<td>1</td>
</tr>
<tr>
<td>75% - 25%</td>
<td>2</td>
</tr>
<tr>
<td>60% - 40%</td>
<td>3</td>
</tr>
<tr>
<td>50% - 50%</td>
<td>4</td>
</tr>
<tr>
<td>40% - 60%</td>
<td>5</td>
</tr>
<tr>
<td>25% - 75%</td>
<td>6</td>
</tr>
<tr>
<td>0% - 100%</td>
<td>7</td>
</tr>
</tbody>
</table>
If the examination system at 16+ is revised, which aspects of the present examination procedures would you like to see retained for the C.S.E. ability range? By the side of the list of items below are boxes for Language and Literature. Please place ticks in the Language boxes of all of those items which you would like to see included in Language examinations at 16+. Put ticks in the Literature boxes of all of those items which you would like to see retained in Literature examinations at 16+.

<table>
<thead>
<tr>
<th>FOR LANGUAGE</th>
<th>FOR LITERATURE</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>(i) The opportunity for teachers to submit their own Mode III syllabuses</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>(ii) Optional external syllabuses devised by the examining board</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>(iii) Democratic teacher control at all levels of the examination procedure</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>(iv) The assessment of oral ability</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>(v) An externally prescribed syllabus, but one which allows wide choice</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>(vi) A narrowly prescribed syllabus</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>(vii) A content-free syllabus involving guidelines as to the aims and objectives of the course</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>(viii) The assessment of aural ability</td>
<td>□</td>
<td>□</td>
</tr>
</tbody>
</table>

When you mark a piece of continuous prose (e.g. an essay or composition) in Language and Literature, which method of assessment do you prefer to use? Please tick one box for Language and one box for Literature.

<table>
<thead>
<tr>
<th>FOR LANGUAGE</th>
<th>FOR LITERATURE</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>(i) An impression mark, bearing in mind certain criteria</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>(ii) An analytical mark where each separate criterion is marked separately and the total mark is arrived at by the addition of these marks</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>(iii) A combination of both of these method</td>
<td>□</td>
<td>□</td>
</tr>
</tbody>
</table>
If you do not teach Mode I, please ignore Question 9 and go on to Question 10.

9. In the Mode I English Language Syllabus for 1981, certain changes have been made which are described in the following statements. Please rate each statement on a 5 point scale ranging from 'Strongly Agree' to 'Strongly Disagree'. For each statement please place one tick in the box which best reflects your own attitude to the proposed changes.

<table>
<thead>
<tr>
<th>Statement</th>
<th>Strongly Agree</th>
<th>Agree</th>
<th>Uncertain</th>
<th>Disagree</th>
<th>Strongly Disagree</th>
</tr>
</thead>
<tbody>
<tr>
<td>(i) The coursework option to one of the final examination papers (Written English or Creative Writing Folio) has been removed. Thus the assessment will be based entirely upon the final examination</td>
<td>27</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(ii) The section of the examination devoted to continuous writing (e.g. essay or composition writing) will account for approximately 16% of the total mark instead of the 30% it accounts for at present</td>
<td>28</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(iii) Letter-writing has been made into a separate, compulsory assessment accounting for approximately 7% of the total mark</td>
<td>29</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

10. If you have any comments to make on any of the issues raised by the preceding questions, please use the space below for this purpose.

________________________________________________________________________
________________________________________________________________________
________________________________________________________________________

11. How great are the demands made by C.S.E. English courses on your time and energy as opposed to the demands made by the rest of your teaching commitments? Are they:

   (i) Extremely high

   (ii) Higher than the demands made by the rest of your teaching commitments, but acceptable

   (iii) No different from the demands made by the rest of your teaching timetable

   (iv) Less demanding than your other teaching commitments

If you have ticked items 11(iii) or 11(iv) i.e. that the demands made by C.S.E. English are not high, please go to Question 14.
12. If the demands made by a C.S.E. English course are high, please place ticks in the boxes of all of those items which demand considerable time and effort.

(i) The planning of appropriate work  
(ii) The marking and ranking of work which contributes to continuously assessed coursework folders  
(iii) The chasing up of work which has not been handed in, but is necessary for the completion of the coursework folder  
(iv) The Agreement Trials, Cross Moderation or Standardization Meetings etc. which are necessary in order to establish a departmental or group standard  
(v) The sheer volume of work required by the syllabus  
(vi) The volume of work necessitated by final internal moderation procedures  
(vii) Please specify any other contributory factors

13. If the demands made by a C.S.E. English course are high, how often do you find it difficult to give as much extra or special preparation to the rest of your teaching commitments as you wish to give? Please tick one box.

(i) Frequently  
(ii) Sometimes  
(iii) Rarely  
(iv) Never

14. If you have any comments to make on any of the issues raised by the preceding questions, please use the space below for this purpose.

__________________________________________________________________________________________
__________________________________________________________________________________________
__________________________________________________________________________________________
__________________________________________________________________________________________
The following statements about continuous assessment (i.e. regular assessment of work completed during the course) are to be rated on a 5 point scale ranging from 'Strongly Agree' to 'Strongly Disagree'. For each of the statements, please place one tick in the box which best reflects your own attitude.

(i) Continuous assessment is capable of motivating pupils to work conscientiously throughout the course of study

(ii) Continuous assessment gives a more rounded and accurate picture of an individual's real ability

(iii) Continuous assessment redresses the balance for those pupils who are prepared to work hard over an extended period of time, as opposed to those pupils who are capable of doing well by working hard just before an examination

(iv) The fact that so many C.S.E. courses involve an element of continuous assessment exerts considerable pressure on some pupils, in that they feel the need to produce a large quantity of work of a consistently high level of achievement

(v) The duties and responsibilities involved in continuous assessment are particularly demanding of a teacher's time and energy

(vi) One of the functions of a teacher is constantly to encourage pupils to give of their best. The need to make realistic assessments of their performance during the course conflicts, at some points, with the attempt to motivate them

(vii) Whenever an individual acts as both teacher and examiner, there is always the possibility that a variety of personal factors may influence the assessment to some extent

(viii) Continuous assessment of the study of Literature avoids certain inherent tendencies of Literature examinations; the learning of notes and regurgitation of secondhand opinions

(ix) Continuous assessment is particularly well-suited to the study of Literature because it evaluates a pupil's response to a text at the time when it is made, rather than asking him to duplicate that response during an examination which may be quite remote from the time when the Literature was studied

(x) Large-scale continuous assessment, operated over an entire region, makes it difficult to be confident that the marks awarded really do reflect accurately the regional standard
The following statements, on general issues, are to be rated on a similar 5 point scale ranging from 'Strongly Agree' to 'Strongly Disagree' by placing ticks in the boxes which best reflect your own attitude. Please give one tick for each statement. (Please note the distinction between the issues of the actual teaching of English raised in items (i) and (ii), as opposed to the certification referred to in items (iii), (iv) and (vii)).

<table>
<thead>
<tr>
<th>(i) English is one complex and indivisible discipline and therefore an integrated approach to the teaching of Language and Literature is most appropriate</th>
<th>STRONGLY AGREE</th>
<th>AGREE</th>
<th>I AM UNDECIDED</th>
<th>DISAGREE</th>
<th>STRONGLY DISAGREE</th>
</tr>
</thead>
<tbody>
<tr>
<td>(ii) English Language and Literature are most effectively taught separately</td>
<td>49</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(iii) The provision of separate certificates for Language and Literature is the best method of certificating English</td>
<td>50</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(iv) The method of single certification which provides one certificate for English (Language, Literature and Oral) is the best method of certificating English</td>
<td>51</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(v) The method of dual certification in Language and Literature, with compulsory entry in both, provides the best means of certificating English</td>
<td>52</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(vi) A C.S.E. English course permits a different kind of teaching with 4th and 5th Year examination classes to 'O' Level English courses</td>
<td>53</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(vii) C.S.E. English courses can accommodate a more flexible approach to English teaching than 'O' level courses can. ('Flexible' is used to define a more adventurous or a more exploratory or a more experimental approach to English teaching)</td>
<td>54</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(viii) The E.M.R.E.B. arrangement, with its Local Subject Panels and Regional English Panel, is an effective means of representing the views of practising English teachers on the syllabus and examination</td>
<td>55</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(ix) C.S.E. English courses provide a satisfactory preparation for 'A' Level English courses</td>
<td>56</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(x) English Language and Literature demand different skills and capacities</td>
<td>57</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(xi) C.S.E. is capable of giving an examination objective to the average and below average pupil and thus motivates him for longer during his school career</td>
<td>58</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(xii) The meetings, panels and committees etc. involved in the running of the C.S.E. examination have brought additional benefits for practising teachers, in that they provide the opportunity for the sharing of ideas and the dissemination of expertise.</td>
<td>59</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

END OF CARD ONE

332
17. The following statements about oral assessment are to be rated on a 5 point scale ranging from 'Strongly Agree' to 'Strongly Disagree'. For each of the statements, please place one tick in the box which best reflects your own attitude.

(i) Your present Oral Examination assesses oral ability as reliably and accurately as possible
(ii) Oral performance is, by its very nature, ephemeral and therefore difficult to assess
(iii) The present system of Visiting Oral Moderators, operated by EMREB, provides a satisfactory means of moderating oral assessments
(iv) The presence of the Visiting Oral Moderator, during an examination, rarely has an adverse effect on a candidate's performance
(v) A formal examination of oral ability makes the situation in which the candidate has to perform unreal and artificial
(vi) Home environment and social background affect oral competence and make it difficult to isolate and assess the candidate's own individual achievement and ability to communicate effectively

In each of the questions which follow, please place ticks in the boxes which best reflect your own attitude. Please give one tick per question.

18. For which candidates is the inclusion of an oral assessment most useful in aiding the assessment of their overall ability?

(i) The most able candidates
(ii) The average candidates
(iii) The below average candidates
(iv) All children benefit equally, regardless of ability
(v) Personal characteristics are a more important factor than ability in determining how much an oral assessment aids the overall assessment of a candidate
19. Which of these methods of examining oral performance do you prefer?

(i) An oral examination on a one-to-one basis between the teacher/examiner and the candidate

(ii) The group discussion method whereby candidates are assessed on their contribution to a discussion

(iii) A combination of the group discussion method and the individual oral examination with the teacher/examiner

(iv) An examination in which the candidate performs individually but an audience is provided from the candidate's peer group

(v) The suitability of each method depends largely on the personalities of individual pupils

20. If you have any comments to make on any of the issues raised by the preceding questions, please use the space below for this purpose.

____________________________________________________________________________________

____________________________________________________________________________________

____________________________________________________________________________________

21. The Mode I Language Examination allocates 30% of the total mark to the oral assessment. This allocation:

(i) represents an adequate proportion of the total mark

(ii) should be increased

(iii) should be decreased

22. The class teacher has 10 marks out of a total of 30 marks in the Mode I oral assessment allocated for his/her evaluation of a candidate's oral performance throughout the course. This allocation:

(i) represents an adequate proportion of the total mark

(ii) should be increased

(iii) should be decreased

If you do not teach Mode I, please leave Question 23 and go on to Question 24
23. The Mode I Oral Examination comprises three elements:— (a) a reading from an extract (b) a talk on a topic chosen by the candidate (c) a discussion with the teacher/examiner.

(i) Which element of the assessment are you most satisfied with?

(a) The reading  
(b) The talk  
(c) The discussion

(ii) Which element of the assessment are you least satisfied with?

(a) The reading  
(b) The talk  
(c) The discussion

24. Please use this space if you wish to amplify your point of view on any of the issues raised in Section III. We would welcome any comments you might make here.
The following statements about the Literature assessment are to be rated on a 5 point scale ranging from 'Strongly Agree' to 'Strongly Disagree'. For each of the statements please place one tick in the box which best reflects your own attitude. (Those statements which refer specifically to Mode I should only be completed by those teachers who use Mode I).

<table>
<thead>
<tr>
<th>Statement</th>
<th>Strongly Agree</th>
<th>Agree</th>
<th>Uncertain</th>
<th>Disagree</th>
<th>Strongly Disagree</th>
</tr>
</thead>
<tbody>
<tr>
<td>(i) A C.S.E. Literature Syllabus should stipulate the inclusion of all the main branches of English Literature (Prose, Poetry and Drama)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(ii) The effort necessary by some candidates to produce the required volume of work in the Mode I Literature Folio (16 separate pieces of work of about 500 words each) is extremely high</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(iii) It is regrettable that only pieces of continuous prose are considered suitable for inclusion in the Mode I Literature Folio, leaving out other types of response (e.g. a creative response or comprehension questions requiring extended answers)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(iv) The EMREE approach to Literature is particularly appropriate to the C.S.E. ability range because it allows the teacher freedom to choose Literature which is appropriate to his pupils' needs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(v) The approach to Literature fostered by C.S.E. is less academic than that generally demanded by 'O' level, in that C.S.E. tends to value the genuine personal response whereas 'O' level requires more critical evaluation and in-depth academic analysis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(vi) C.S.E. candidates who go on to study 'A' level English find it difficult to cope with the academic approach which is required</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(vii) Compact and condensed language is a characteristic feature of most poetry. Thus, poetry is often difficult to teach to the C.S.E. ability range because the language is not easily accessible</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(viii) It is difficult for the majority of pupils in the C.S.E. ability range to derive real enjoyment from most poetry</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(ix) The Mode I requirement that candidates should have studied 7 books or an equivalent amount of material in shorter works makes appropriate demands on pupils in terms of the amount of reading they should accomplish during the course</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(x) Many C.S.E. pupils benefit considerably from the Mode I approach to examining Literature, which allows the work produced during the course to determine the final grade</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
In the questions which follow, please give one tick per question

26. The majority of pupils in the C.S.E. ability range benefit most from studying:
   (i) mostly modern literature written during the twentieth century
   (ii) mostly traditional literature written prior to the twentieth century
   (iii) equal proportions of modern and traditional literature

27. The majority of pupils in the C.S.E. ability range benefit most from studying:
   (i) a small number of texts (e.g. 4 or less) in great detail
   (ii) a larger number of texts (e.g. 5 or more) with less emphasis on detail
   (iii) one or two texts in depth plus a number of other texts read less closely

28. Which of the following alternatives for the content of a Literature Syllabus do you prefer?
   (i) Complete freedom for the teacher to choose the Literature most appropriate to his pupils' needs
   (ii) Freedom for the teacher to choose the Literature most appropriate to his pupils' needs, providing that he must include all the main branches of Literature (Prose, poetry and drama)
   (iii) A number of set texts to be chosen from a syllabus of prescribed texts
   (iv) A syllabus of prescribed authors which provides freedom concerning the actual works studied
   (v) A syllabus which takes a thematic approach, outlining a number of themes and suggesting relevant material

29. Please use the space below if you have any final comments to make

________________________________________________________________________
________________________________________________________________________
________________________________________________________________________
________________________________________________________________________
________________________________________________________________________
________________________________________________________________________

Your assistance is greatly appreciated.
Thank you.
PRELIMINARY INFORMATION FOR QUESTIONNAIRE

1. C.S.E. English teaching is/is not undertaken at __________________________ School (Delete where not applicable)

2. Which Mode is taken for C.S.E. English? Please tick the appropriate box(es)
   Mode 1
   Mode 2
   Mode 3

3. Does your School teach English as part of a Humanities or Integrated Studies course?
   Yes/No

4. Please list the names of all teachers whose responsibilities include C.S.E. English in the space provided below (N.B. The name of any teacher who will no longer be teaching at the School at the beginning of the next academic year i.e. October is not to be included).

   1. __________________________  3. __________________________
   2. __________________________  4. __________________________  5. __________________________
   6. __________________________  7. __________________________  8. __________________________
   9. __________________________ 10. __________________________ 11. __________________________
   12. __________________________ 13. __________________________ 14. __________________________
   15. __________________________

(Your own name where appropriate)
Dear

I am writing to ask for your help in a research project, currently being undertaken at the School of Education, which is designed to evaluate the C.S.E. examination. The project is supported by both the East Midland Regional Examinations Board and the Social Science Research Council and focusses on key subject areas, in this particular instance English. One important aspect of our investigation into C.S.E. English is the collection of data concerning practising teachers' perspectives on the functioning of these examinations in English. In order to achieve this aim, a questionnaire has been devised which deals with central issues in the present examination system. The actual questions have been formulated as a result of interviews with examiners, assessors and panel members throughout the region. Clearly it is essential, in a survey such as this, to gather a wide range of responses from teachers throughout the East Midland Region. Therefore, I wrote to your Head of Department to enquire which of his or her colleagues' responsibilities include C.S.E. English courses.

It is thus that I have written to you, in the hope that you will feel able to help us by completing the enclosed questionnaire and returning it in the stamped addressed envelope. The issues raised in this questionnaire are felt to be particularly important in view of the current uncertainty surrounding proposals for future examinations at 16+. With this in mind, the questionnaire deals not only with the main features of the present system, but also with the way in which teachers would like to see English examinations for the C.S.E. ability range develop in the future. May I stress at this point that all of our work is strictly confidential and that total anonymity is guaranteed to both teachers and schools. If you have any questions or observations that you wish to raise, please do not hesitate to get in touch with me or Val Brooks, the research assistant responsible for this aspect of the study.

There is one feature of the questionnaire's structure which requires explanation. It is appreciated that although Language and Literature are now separately certificated under Mode 1, many Mode 3 syllabuses treat English as one subject. However, for the purposes of this questionnaire it has been most appropriate to deal with Language and Literature as separate items. We apologize if this division appears artificial to you. Nevertheless, the questionnaire does provide you with the opportunity to indicate your views on this matter and these will be taken into account when analysing the rest of the data.

Finally, may I take this opportunity to stress that your assistance in this piece of research will be greatly appreciated.

Yours sincerely,

Gerald Bembaum
Professor of Education and Director
of the School.

Professors of Education: Gerald Bembaum, Brian Simon, Derek Wright
APPENDIX 4

The following list provides a sample of the type of questions used in the interviews.

Can you single out any feature of the English examination which you are/were particularly satisfied with? What are your reasons for this answer?

Can you single out any feature which you are/were particularly dissatisfied with?

Do you prefer single or separate certification? What are your reasons?

Which of the three modes of examining do you prefer? Why?

What proportion of the marks do you think that continuous assessment should account for? Your reasons?

What proportion of the marks do you think should be allocated to spoken English? Your reasons?

What are your views on the Mode I approach to examining literature?

What are your views on the use of final examinations in English?

How time-consuming do you find CSE courses?

(a) very time-consuming
(b) fairly time-consuming
(c) not very time-consuming
(d) not at all time-consuming

Do you regard moderation as a difficult or an easy task? Your reasons?
BIBLIOGRAPHY

GENERAL


Evans, G.D.O. (1968), *The CSE Examination - Three Years On*, *Teacher in Wales* 8(22), July.


Hopkin, K.S. (1967), *English in the New CSE Examination*, *Teacher in Wales* 7(22), July.


MacIntosh, H.G. (1972), *The Teacher as Assessor*, *Secondary Education* 3(1).


Murphy, R.J.L. (1978), *Reliability of Marking in Eight GCE Examinations*, *British Journal of Educational Psychology* 48(2).


**CHAPTERS 1 and 2: HISTORICAL CONTEXT**


Anson-Smith, C. (1960), in Letters to the Editor, *Times Educational Supplement*, 22nd July, p.120.


British Association for the Advancement of Science (1903), *Influence of Examinations - Interim Report of the Committee Appointed to Consider and Report upon the Influence Exercised by Universities and Examining Bodies on Secondary School Curricula; also of the Schools on University Requirements*, London, John Murray.


Correspondent (1960), in *Times Educational Supplement*, 4th November, p.611.


Frenchman, D. (1960), in Letters to the Editor, Times Educational Supplement, 22nd July, p.120.

Harris, R.J. (1958), The Organisation of English in the Secondary Modern School, Use of English 10(2).

Headmistress (1960), in Letters to the Editor, Times Educational Supplement, 5th August, p.178.


Heeley, A.D. (1961), Beloe, But When?, Forum 3(2), Spring.

Holmes, M. (1961), Examinations in the Comprehensive School, Forum 3(2), Spring.


Inglis, F.C. (1963), Reforming the Language Paper, Use of English 15(1).

James, F.R. (1960), in Letters to the Editor, Times Educational Supplement, 21st October, p.532.


Kirkpatrick, W. (1960), in Letters to the Editor, *Times Educational Supplement*, 22nd July, p.120.


CHAPTER 3: CONTINUOUS ASSESSMENT

Banks, B. (1970), Away from Exams, Times Educational Supplement, 3rd April, p.64.


Burchnall, M.L. (1968), The Internal Examining of English Language, Conference 5(1).


Correspondent, (1964), Assessment or Examination? "Justice must be seen to be Done", Times Educational Supplement, 4th September, p.317.


Cowell, B. (1972), in Letters to the Editor, Times Educational Supplement, 16th June, p.16.

Ebel, R.L. (1972), Why is a Longer Test Usually a More Reliable Test?, Educational and Psychological Measurement 32.


Hartley, J. and Braithwaite, A. (1976), All This for Two Per cent: The Contribution of Course-Work Assessment to the Final Grade, Durham Research Review 8(37).


Schools Council (1965), The Certificate of Secondary Education: School-based Examinations, Examinations Bulletin 5, HMSO.

Schools Council (1975), Continuous Assessment in the CSE, Examinations Bulletin 31, Evans/Methuen Educational.

Snowdon, R.L. (1965), Final Examination or Continuous Assessment, Education for Teaching 67.

Spencer, E. (1976), Internal Assessment or External Exam?, Teaching English 9(2).


Whitehead, F. (1958), The Educational Function of Examinations, Education for Teaching 47.

CHAPTER 4 : LITERATURE


Barnes, D. (1968), The Roles of Literature in the Curriculum of the Primary and Secondary School, English in Education 2(1).

Barnes, D. (1969), English with an Examination Fifth Form, Use of English 20(3).


Bazen, D. (1979), Literature Without Examination, English in Education 13(3).


Correspondent (1960), The Muse Examined : Studying Poetry at O-Level, Times Educational Supplement, 18th March, p.553.


Kennedy, R.B. (1957), 'O' Level Literature, Use of English 9(2).


Norris, T.F. (1962), 'O' Level English Literature, Use of English 14(2).


Thompson, D. (1966), What is Literature?, *NATE Bulletin* 3(2).


Tucker, M. (1968), 'Readers' or Literature?, *Use of English* 19(3).


CHAPTER 5: ORACY


Alcock, J. (1972), Students' Questions and Teachers' Questions, *English in Education* 6(3).


Atkinson, D. (1968), A Listening Test for CSE - and After?, *Education for Teaching* 76.

Barnes, D. (1976), From Communication to Curriculum, Penguin Education.


Creber, J.W.P. (1972), Lost for Words : Language and Educational Failure, Penguin.


Cousins, A. (1972), Language to Communicate, English in Education 6(3).

Dobson, M.C. (1972), Learning to Evaluate the Spoken Word, English in Education 6(3).

Foden, F.E. (1972), Examining Talk, Use of English 23(3).


Hadfield, A. (1957), Speeches Lesson, Sir!: Use of English 8(4).


Hitchman, P.J. (1963), Reading and Recitation in Teacher Training in the Nineteenth Century, Researches and Studies 26, University of Leeds Institute of Education.


Hitchman, P.J. (1966), The Validity and Reliability of Tests of Spoken English, British Journal of Educational Psychology 36(1).

Hitchman, P.J. and Daniels, J.C. (1971), *Tape Recording and the Testing of Spoken English in Schools*, University of Nottingham School of Education.

Hughes, F. (1951), Reading Aloud, *Use of English* 2(4).


