

1. PS

CRITERIA IN ENGLISH LANGUAGE ASSESSMENT
A SOUTH AFRICAN PERSPECTIVE

ANNE DOMMISSE

A dissertation submitted in fulfilment of the requirements for the degree of Master of Philosophy, specialising in Language Education.

November 1990

The University of Cape Town has been given the right to reproduce this thesis in whole or in part. Copyright is held by the author.

The copyright of this thesis vests in the author. No quotation from it or information derived from it is to be published without full acknowledgement of the source. The thesis is to be used for private study or non-commercial research purposes only.

Published by the University of Cape Town (UCT) in terms of the non-exclusive license granted to UCT by the author.

ABSTRACT

The study recorded in this dissertation was undertaken in the School of Education at the University of Cape Town (UCT) during the period 1986 - 1990. It was motivated by perceived anomalies in the administration of State regulations for endorsement of teachers' diplomas in respect of ability in English (E/e).

The study commences with an analysis of the relevant requirements of the regulations for teacher bilingualism, as set out in sections 10 and 11 of Criteria for the Evaluation of South African Qualifications for Employment in Education, 1988. Theoretical and practical problems of evaluation and endorsement identified at UCT are considered in relation to the concept of test failure, as opposed to testee failure. Responses to a questionnaire sent to other teacher training centres indicate similar concerns elsewhere.

Arising from a review of recent literature on language testing, and against the background of the multilingual target groups tested at UCT, a proposition is put forward for a distinction between communicative competence and language proficiency as criteria in language assessment, depending on whether English is the medium (communicative competence), or the subject (proficiency), of instruction.

Assuming that English will remain a medium of instruction in a changing socio-political dispensation, at least in the short and medium terms, the study then focuses on test design, construction and scoring, where the objective is to test communicative competence in English, rather than proficiency. The role of English in the curriculum in a future South Africa is discussed briefly.

It is concluded that current regulations for language endorsement are in urgent need of review. The following recommendations are made in this regard: that the relevant requirements for teachers in State schools be reformulated to account for one level, rather than two, of endorsement in English as the medium of instruction; that such endorsement be required only in the case of non-English medium graduates, thereby recognising the integrity of the English medium teachers' diploma itself; that procedures for assessment for the purpose of diploma endorsement be standardised; and that the State support further research in this area.

Scrivere è difficilissimo

specialmente per chi

sa scrivere

G. d'Annunzio

CONTENTS

INTRODUCTION	1
CHAPTER ONE	
Regulations for Teacher Bilingualism	6
1.1	Relevant sections
1.2	Analysis of relevant aspects
CHAPTER TWO	
Research Arising from Assessment for E-endorsement 1986 - 1989	12
2.1	Background data
2.1.1	"Failure" - the testee or the test?
2.1.2	An alternative approach to assessment
2.1.3	Proposal for a research programme
2.2	Formulation of a research programme
2.2.1	Preliminary considerations
2.2.2	An approach to the identification of criteria for E-endorsement
2.3	Written tests 1987 - 1988
2.3.1	Analysis of responses
2.3.2	Conclusions
CHAPTER THREE	
The Questionnaire, and Responses to it	41
3.1	The questionnaire
3.2	Responses received and issues raised
CHAPTER FOUR	
Literature Review	53
4.1	Motivation for theoretical analysis
4.2	"(Language) competence" and "proficiency"
4.3	Towards a distinction
4.3.1	A return to a traditional interpretation of "proficiency"?
4.4	From theory to practice

CHAPTER FIVE

Towards a Test Design

76

- 5.1 Preliminary considerations
- 5.2 Teacher needs in the English-medium classroom
 - 5.2.1 Classroom needs - as testees perceive them
 - 5.2.1.1 Task on entrance to HDE course 1990
 - 5.2.1.2 Formulation of test criteria
- 5.3 Basic concepts in testing
 - 5.3.1 Subjectivity versus objectivity
 - 5.3.2 Validity and reliability
 - 5.3.2.1 Reliability
 - 5.3.2.2 Validity
 - 5.3.3 Norm-referenced and criterion-referenced tests

CHAPTER SIX

The Sample Test, and Analysis of Test Items

110

- 6.1 The sample test
 - 6.1.1 Target group
 - 6.1.2 The written test
 - 6.1.3 The oral test
- 6.2 Rationale
 - 6.2.1 Accessibility of text for test items
 - 6.2.2 Selection of tasks and texts for sample test
 - 6.2.2.1 Section A
 - 6.2.2.2 Section B
 - 6.2.2.3 Section C
 - 6.2.2.4 Section D
- 6.3 Focus of sample test items
- 6.4 Inappropriate items - a motivated view
- 6.5 Oral tests: format and rationale
- 6.6 Scoring the sample test

CHAPTER SEVEN

Student Performance in Relation to Test

131

- 7.1 Mark scheme
 - 7.1.1 Theoretical orientation to marking
 - 7.1.2 Rationale for marks per test item
 - 7.1.3 Holistic scoring in the sample test
- 7.2 Analysis of selected responses and scores
 - 7.2.1.1 Section A, part one
 - 7.2.1.2 Section A, part two
 - 7.2.2 Sections B and C
 - 7.2.3 Section D
- 7.3 What do scores mean in pass/fail terms?
- 7.4 Scoring the oral tests

CHAPTER EIGHT
English in the Curriculum - Today and Tomorrow 154

- 8.1 The English class
- 8.2 A historical perspective
- 8.3 English for academic/cultural purposes?
 - 8.3.1 Cultural literacy
- 8.4 Creative teaching
- 8.5 Conclusion

CHAPTER NINE
Conclusion 171

- 1 Review
- 2 Conclusions
- 3 Recommendations
- 4 Directions for Future Research

BIBLIOGRAPHY 181

ACKNOWLEDGEMENTS

My supervisor, Professor Douglas Young, was generous
with time and advice -

the Human Sciences Research Council,
with financial support -

and my husband,
with patience and understanding -

to them all, my grateful thanks.

CRITERIA IN ENGLISH LANGUAGE ASSESSMENT
A SOUTH AFRICAN PERSPECTIVE

INTRODUCTION

This study was undertaken in the Language Education Centre (formerly Language Education Unit) in the Faculty of Education at the University of Cape Town, during the 5-year period 1986 to 1990. It arose in response to problems experienced with the implementation of State regulations for Language Endorsement Requirements (E/e) in respect of the teacher's diploma as required in terms of sections 10 and 11 of CRITERIA FOR THE EVALUATION OF SOUTH AFRICAN QUALIFICATIONS FOR EMPLOYMENT IN EDUCATION, 1988, and specifically with the interpretation of the given criteria for such endorsement, *competence* and *proficiency*.

The immediate aim of the study was to establish a rational basis for English language assessment consistent with current State policy on the one hand, and with practical classroom conditions in South Africa today on the other.

The overall objective of the project is, however, a wider one. Given the rising tide of change in the South African Government's ideology, it would seem a fair assumption that State policy in respect of the bilingual teacher will, perhaps sooner rather than later, change too. In view of the wide currency of English in Southern Africa and in international affairs, it seems likely that English will, however, retain a central role in any new political dispensation, at least for the foreseeable future.

Consistent with this view, this study argues for relevant, practicable criteria for English in teacher education, based on modern linguistic research, but interpreted in accordance with the role of education in general, and of English in particular, in the South African context. The overall aim of this study is to motivate for the standardisation of such criteria.

Chapter One focuses on existing State regulations for teacher bilingualism, and identifies anomalies inherent in them which would seem to be the cause of perceived inconsistencies of interpretation.

In Chapter Two, I focus on problems of assessment experienced in the School of Education of the University of Cape Town during the last four years, with particular attention to the concept of test failure as opposed to testee failure in a linguistically non-homogeneous target group.

Ongoing research aimed at relating the requirements of the regulations (and the given criteria for language ability) to the needs of secondary education in South Africa today in a meaningful way, gave rise, in 1988, to a questionnaire designed to elicit relevant comment from other training centres. The questionnaire, sent out to 42 centres, and a summary of responses to it, constitute Chapter Three.

In Chapter Four I review recent published research in the field of language assessment, and consider in particular the concepts of *(communicative) competence* and *proficiency* as test criteria in English "second language" assessment. I put forward a proposition for a distinction between them to account for different *kinds* (as opposed to *levels*) of ability in English consistent with the notion of official (Afrikaans/English) bilingualism for secondary school teachers.

In accordance with this proposition, and taking into account the changing parameters of education in what is currently perceived as an emergent "new South Africa", Chapter Five addresses wider issues arising from State regulations for E/e endorsement. It considers the requirement for assessment of English in the context of the English-medium Higher Diploma in Education (HDE); and it reflects on test design for a target group which subsumes L1 speakers of all indigenous languages other than English (including Afrikaans). This does not imply acceptance of the principle which relates (Afrikaans/English) bilingualism to teacher career advancement; rather, it is an attempt to rationalise a de facto situation.

A sample of test design, and the rationale which informs it, are put forward in Chapter Six. Test items are analysed against the background of recent studies in language assessment.

The sample test was administered, in September 1990, to a target group consisting of thirty in-service testees. Their performance in relation to it is examined in Chapter Seven.

In Chapter Eight, the focus widens to consider a notion of "cultural literacy", its validity in the South African context today and tomorrow, and what it might imply for teacher education in terms of the proposition put forward in this study.

Chapter Nine concludes the project. It summarises the study, and puts forward recommendations and suggestions for further research.

CHAPTER ONE

REGULATIONS FOR TEACHER BILINGUALISM

1.1 RELEVANT SECTIONS OF THE REGULATIONS

Sections 10 and 11 of CRITERIA FOR THE EVALUATION OF SOUTH AFRICAN QUALIFICATIONS FOR EMPLOYMENT IN EDUCATION, 1988, make provision for teacher language ability as follows:

10.1 The possession of an approved diploma in education *with a language endorsement in respect of Afrikaans and English* shall be a requirement for appointment in a permanent capacity.

11.2.2.1 The ability of a student who wishes to enter the service of an education department to use the two official languages as media of instruction will be determined *in a manner agreed on by the education department and the training institution concerned*, and will be indicated by an endorsement on the diploma by the body issuing it.

11.2.2.3 The *ability to use the two official languages as media of instruction* will be signified by the following symbols:

AE or EA: *Competent to teach both Afrikaans and English-medium classes. The language which the teacher knows best is indicated first.*

Ae *Competent to teach Afrikaans-medium classes, but also able to teach English-medium classes, although with less proficiency.*

Ea *Competent to teach English-medium classes, but also able to teach Afrikaans-medium classes, although with less proficiency.*

A *Competent to teach Afrikaans - medium classes only.*

E *Competent to teach English-medium classes only.*

(My italics)

1.2 ANALYSIS OF RELEVANT ASPECTS OF THESE REGULATIONS

Provision is made for six categories of endorsement - E/A, A/E, E/a, A/e, A, and E - in respect of the teacher's *ability to use the two official languages as media of instruction*. The language the teacher *knows best* is indicated first. In effect however there is also a seventh category, ea (ae?), for students who are unclassifiable in

terms of the given symbol-combinations, as for instance when he or she is unsuccessful in both E and A assessment (as implemented at the various centres), or when neither English nor Afrikaans is the language the teacher "knows best".

Perhaps what is intended is "whichever of these two languages the teacher knows *better*". As formulated however, both the wording and the symbol-combinations imply that the uppercase letters refer to mother-tongue or L1 ability. Statistics included in the 1988 edition of the Annual Report of the Department of Education and Training indicate that 12,077 qualified Black secondary school teachers - 1,837 university graduates and 10,240 with three years post-standard 10 teacher training - in Department of Education and Training schools potentially fall into this "non"-category, with a further 15,812 Black teachers - 2,887 graduates and 12,935 with three years post-standard 10 training - in the National states (excluding Transkei, Bophuthatswana, Venda and Ciskei). The Research Institute for Education Planning (Bloemfontein), in Education and Manpower Development, number 9, 1988, puts the figure at 26,142, including teachers of Transkei and Bophuthatswana (p. 18).

Another area of uncertainty is that of permanent appointment. Subsection 10.1 states that "the possession of an approved diploma in education with a language endorsement

in respect of Afrikaans and English shall be a requirement for appointment *in a permanent capacity*"; and subsection 11 lists three possible categories of such endorsement: AE/EA, Ae, and Ea. In practice however, AE/EA - the so-called "bilingual certificate" - is commonly a condition of permanent appointment. A perceived need for clarity in this regard gave rise, during a meeting of a subcommittee of the Provincial Advisory Committee for teacher training in Provincial Building, Cape Town, on 8th September 1989, to enquiries which elicited the following information: the Cape Department of Education requires

at least Ea for permanent appointment at an English medium school,.....Ae for permanent appointment at an Afrikaans medium school, (and) AE/EA for permanent appointment at a parallel medium school, or for any promotion post. (Minutes of this meeting)

It seems that practical implementation even of this amplified version is rarely consistent with the theoretical import of stated official requirements.

But more directly relevant to this study is the fact that the symbols A/E/a/e on their own are meaningless. They take on significance only in relation to the criteria against which they are "measured" in this context, namely *competence* and *proficiency*. Interpretation of the requirements thus ultimately depends on how the notions of *competence* and *proficiency* are defined; and as in the past, part 11.2.2.1 of CRITERIA FOR THE EVALUATION OF SOUTH AFRICAN QUALIFICATIONS FOR EMPLOYMENT IN EDUCATION, 1988, in effect

delegates this responsibility to the individual teacher training centres concerned.

As one of 42 or more such centres, the School of Education of the University of Cape Town administers the regulations in respect of 275 - 300 students for the Higher Diploma in Education annually, and also offers regular in-service courses for post-qualification certification in English (and/or Afrikaans). It has been our experience that in the absence (to date) of definitive interpretation of "competence" and "proficiency" in the context of language ability in general, and lacking guidelines for the intended relationship between these "criteria" and the given symbols in particular, testers not infrequently resort to endorsement on a more or less "ad hominem" basis, motivated by perceptions of testee ability not necessarily consistent with test results. Given the fact that the target group at this University (and at other centres) is characterised by several South African varieties of English, different expectations from different testees is perhaps, to some extent at least, justifiable on pragmatic grounds. On the other hand, "ad hominem" endorsement, even if only in selected cases, is inconsistent with acknowledged principles of test validity and reliability. It also invites the not unreasonable charge of "double standards" from concerned students.

These and other problems surrounding the issue of assessment for E-endorsement came to my notice when in 1986 I was asked to moderate the marks for the Written Test of English Proficiency administered on course entrance to all HDE students.

CHAPTER TWO

RESEARCH ARISING FROM ASSESSMENT for E-ENDORSEMENT

1986 - 1989

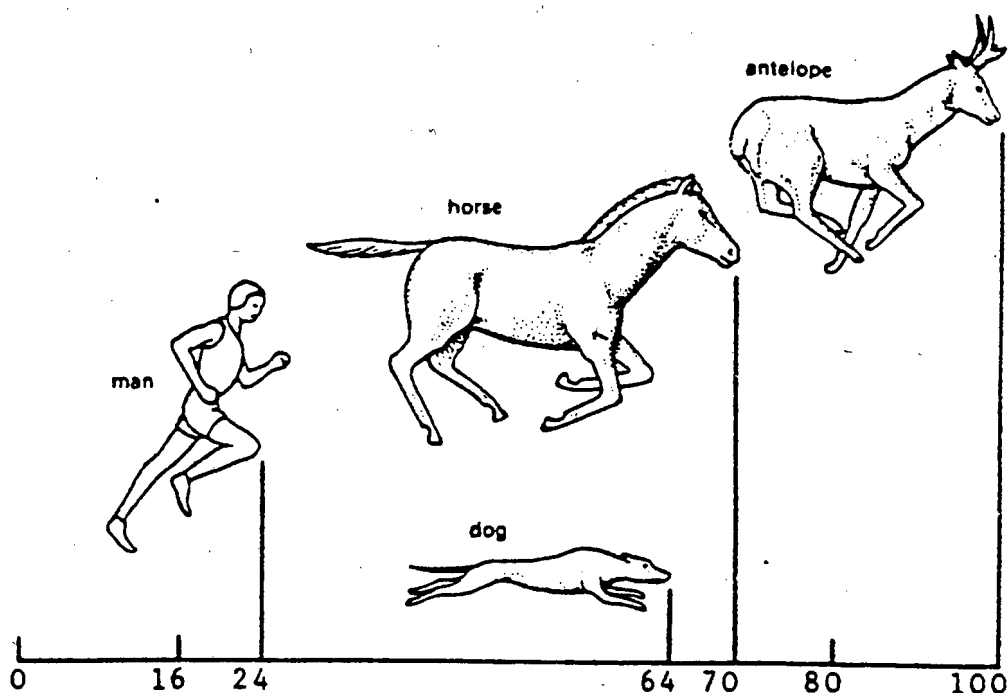
2.1 Background data

In 1986, the Written Test In English Proficiency was administered on course entrance to 245 students registered for the Higher Diploma in Education (Secondary P G). Similar in format and content to tests of previous years, it consisted of four questions, and included a Comprehension Test (15 multiple choice questions each with "the most appropriate answer" and 4 distractors, for 15 marks), a Cloze Test (29 spaces with a redundancy factor of 4, for 25), and an expository exercise on "your subject" (30 marks).

The task required in Question 4 was as follows:

YOU HAVE DRAWN THE DIAGRAM BELOW ON THE BOARD TO EXPLAIN TO YOUR STANDARD SIX CLASS THE RELATIVE SPEED OF MOTION OF MAN RUNNING AT A MAXIMUM SPEED COMPARED WITH ANIMALS DOING THE SAME. WRITE, IN ABOUT 200 WORDS, THE EXPLANATION YOU WOULD GIVE YOUR CLASS OF HOW TO INTERPRET THIS DIAGRAM. (20 MARKS)

(Diagram overleaf)



Kilometres per hour

A mark range from 0 to 90 (instead of percentage scoring) was based on Carroll's 9-band scale (Carroll; 1980). In brief, Carroll's scale consists of a series of numerical ratings from 0 to 9, with corresponding levels, or bands, of "communicative performance" defined by verbal descriptors which relate the bands to (identified) testee needs. In this test of "English Proficiency", a minimum pass score of 60, corresponding with band 6 on Carroll's scale, was intended to denote "competent (communicative) performance" in the test as a whole - which suggested that "proficiency" and "communicative competence" were (in this case) interchangeable concepts. This point is taken up again later.

Scripts were divided for marking, in no particular order, between four higher degree students (about 60 scripts each), and I was asked to co-ordinate the exercise. It is relevant to note also that the test had been designed, and a marker's guide provided, by a sixth person (that is, by someone other than the markers and co-ordinator).

Comments of the markers and informal perusal of the returned scripts suggested inter-marker inconsistency in mark allocation; and a closer scrutiny of the questions, of student responses to them, and of individual marker interpretation of both, was undertaken by the co-ordinator.

2.1.1 "Failure" - the testee or the test?

Preliminary "sampling" in each of the four batches of scripts indicated an inter-marker difference in "failure" rate ranging from 12% to 35% of scripts marked; and interviews with the markers suggested that the reason for this might have been significant differences in the way each had interpreted the tester's guide for mark allocation, particularly for questions 2, 3 and 4.

For the Cloze test, the markers had been told that either the original (deleted) word, or a word close enough to retain the exact meaning of the text, was acceptable. There

was however no prior agreement on specific acceptable substitutes for the deleted word, and marking was uneven. For example, in the passage

The child brings into the.....everything which he has learned in his..... - habits, attitudes, beliefs, etc - and the teacher.....the school will respond to these

school, classroom, and class for the first slot, *family, home, community and environment* for the second, and *and, in and at* for the third were all popular testee choice, all of which were accepted by some markers (sometimes inconsistently from script to script) but not by others. It was clear that on the basis of marker perceptions in this and other sequences, rather than on comparable testee responses, some passed the Cloze test while others who had given similar responses "failed" it, - in some cases also "failing" the test as a whole as a result of poor marks for this question.

In Question Three (the expository exercise) the instruction to "imagine that you are a teacher in a standard six class" was intended to elicit discourse of an "appropriate level" - which does not account for the multilingual matrix of state education in this country. What is "appropriate" language use in one standard six class may not necessarily be so in

another, as suggested by inconsistent inter-marker perceptions of "appropriateness" (possibly related to dissimilar linguistic backgrounds).

Another interesting aspect of this question raised by some of the scripts was whether testees could reasonably be penalised, in a Written Test of English Proficiency, for a level of language use consistent with the traditional notion of proficiency, even if it seemed "inappropriately high" for a Standard Six class in terms, say, of vocabulary used, or complex syntactical structure. Is it valid to test spoken language ("what you will tell your class.....") in the written medium?

In all events, on the evidence of mark allocation for this question, there was little doubt that what one marker considered "appropriate" in terms of register, vocabulary, syntax (and spelling, which was arguably irrelevant - "what you will tell your class...") was unacceptable to another; and inter-marker inconsistency accounted for a significant difference in the average mark for this question too, in each batch of scripts.

Predictably, perhaps, many of the testees found Question Four confusing. Explanations of "how to interpret this diagram" ranged from brief accounts of how a graph "works" (considerably fewer than 200 words), to long and tortuous

explanations of "the theory of relativity". In general, responses were characterised by a tendency to incoherence; and more than 60% of the class as a whole "failed" this question.

On the basis of this "co-ordinator's scrutiny" I recommended that Question Four be scrapped, and that all scripts be re-assessed by one marker on the basis of questions One to Three.

On re-mark of all 245 scripts (undertaken as a possible basis for a research programme), the "failure" rate was 22% (or 58 students, which included at least three registered for English Method). Of particular significance was the fact that although this was similar to the overall "failure" rate in the first marking, there were significant differences in individual testee scores, suggesting that either the test, or inter-marker perceptions of it, were unreliable, or both.

2.1.2 An alternative approach to assessment

No provision is made in state regulations for tuition in language competence/proficiency, and in 1986 none was offered in the HDE course. It was however unacceptable that without E-endorsement a possible 58 graduates would find

themselves qualified (in English) to teach, but unqualified for permanent employment in English medium state schools, and it was only one marker. ?

Underlying this approach were the following assumptions:

* testees would concentrate on content (sociology) rather than on language use per se; language use would thus be more spontaneous ("natural"/"real world") than when responses are elicited under contrived conditions as in the earlier test;

* as the only real issue was whether or not the testee merited E-endorsement - that is, whether he or she "passed" or "failed" the test - discrete point mark allocation was irrelevant; and only one marker. ?

Underlying this approach were the following assumptions:

* testees would concentrate on content (sociology) rather than on language use per se; language use would thus be more spontaneous ("natural"/"real world") than when responses are elicited under contrived conditions as in the earlier test;

* as the only real issue was whether or not the testee merited E-endorsement - that is, whether he or she "passed" or "failed" the test - discrete point mark allocation was irrelevant; and.

* although "objectivity" is at best a relative concept in testing, and particularly so in the Human Sciences, marker subjectivity would be significantly reduced if all scripts

were scrutinised by the same person rather than distributed among a number of markers.

27 of the 58 testees in this latter test received E-endorsement on marking. However, as the balance of the students all passed test content (sociology), and had therefore communicated successfully (if not proficiently) the moderator upgraded the remaining students to E-level status.

2.1.3 Proposal for a Research Programme

There were too many variables in the 1986 tests to draw any significant conclusions from the results other than that, in general, assessment of any given script seemed to depend to a significant degree on different perceptions, on the part of the different markers concerned, of what was required for the purpose of E-endorsement - or, in terms of the test, of English proficiency.

This, together with some student resistance to the requirement itself (on the grounds that the English medium diploma should need no special endorsement for teaching in English), suggested a study to consider all the implications, both theoretical and practical, of State

language regulations and, if appropriate, to put forward recommendations for modification of present requirements.

2.2 Formulation of a Research Programme

2.2.1 Preliminary Considerations

This study was undertaken subject to the following constraints:

- * language endorsement is a professional, not an academic requirement, and is not seen as a function of the full-time pre-service HDE course (although an in-service course is offered)
- * there is at UCT no special assessment for e-level endorsement, and therefore no basis for systematic comparison of e- and E-levels
- * although endorsement is an exit requirement, trainees are assessed in the first semester for administrative purposes
- * some student resistance to the requirement is experienced every year

These constraints are significant for the fact that they are to some extent "UCT specific" and would not necessarily be relevant elsewhere. On the other hand, although there is no explicit requirement in State regulations, as published, for a dedicated test for language endorsement at either level,

some form of assessment in both official languages, at one or both levels, appears to be standard procedure at all training centres, and to this extent at least, this study addresses issues common to all centres.

One such problem is that no distinction is made in the regulations in respect of diploma medium - English-medium graduates are not exempted from either level of English endorsement - and the situation is similar in respect of Afrikaans graduates. The implication seems to be that the kind of language ability required to teach classes in any given medium is somehow different from the kind of ability consistent with the teaching qualification itself.

An obvious way to characterise the two abilities (if there is a difference between them) would seem to be in terms of purpose - English (or Afrikaans) for practical teaching purposes on the one hand, and for academic purposes on the other. The trouble with this argument is that the HDE is not only an academic qualification, it is also a professional one; and the School of Education is a practical training centre within a University Department.

Alternatively then, does endorsement imply different levels of language ability?

It seemed clear that rational interpretation of the regulations would ultimately depend on commitment to one or other of these propositions - that endorsement for competence to teach in the medium of the HDE itself implies either a different *kind*, or a different *level*, of language ability to that consistent with course requirements; and assessment depends on the identification of relevant criteria.

A review of recent developments in language testing early in 1987 suggested a useful perspective on language regulations as set out in the Criteria for the Evaluation of South African Qualifications for Employment in Education, 1988, in that the notion of "communicative competence" seemed in accordance with the spirit, if not with the letter, of the requirement of "competence".

The difficulty with this point of view was that the theory of communicative competence had (and has) not yet been fully developed. As several writers (Widdowson, 1978:162; Romaine, 1984:241; Candlin, 1985:9, among others) have pointed out, the term does not always seem to mean the same thing to different researchers. But then neither have competence or proficiency "found a completely satisfactory expression" (Stern, 1983:356; also Ingram: 1985); and the tendency of some writers to use the terms interchangeably further confuses the issue. Then too, most of the available

literature on language testing, whatever the given criterion, is associated with a specific course or teaching programme on which it is "parasitic" (Davies, 1985). Such tests are "empirical" in the limited sense that criteria reflect course objectives; they do not represent "communicative competence" or "proficiency" per se.

Any attempt to interpret the regulations in the light of recent research must therefore be determined to some extent by which of several theoretical approaches one espouses. The alternative is to regard the notions of competence, and of competence... "with less proficiency", as used in the regulations, as more or less generic terms. They then mean whatever the tester says they mean - which, to judge by responses to our questionnaire as discussed in Chapter Three would seem to be the de facto situation at present.

A more detailed review of current literature is recorded in Chapter Four. In the meantime, as Hulstijn notes in his Comment on Ingram (1985:277),

testers cannot wait for full-fledged theories.....to emerge from language laboratories.....they have to work with taxonomies that seem to make sense.

It was the view of this study that it made sense to distinguish between "proficiency" and "communicative competence" as criteria in language testing (as proposed in

Chapter Four); and in the interests of focused endorsement assessment, and motivated by current trends in purpose-related testing, to work with a taxonomy consistent with the basic principles of communicative competence.

This is not intended to suggest that language proficiency and communicative competence are mutually exclusive concepts, nor that the skills and abilities associated with each can be separated into discrete groupings. On the contrary, studies show that the characteristics associated with each are closely interwoven and that divisions between them are to a great extent arbitrary in nature.

Nevertheless, the fact that since first proposed in the early seventies, Hymes' theory of communicative competence has engaged professional interest and triggered a steadily growing volume of literature in language testing circles all over the world, in itself suggests that the traditional criterion of proficiency, as represented in proficiency tests prior to the seventies, is inadequate to account for dimensions of language defined in recent sociolinguistic research. It is these dimensions that concern us here.

2.2.2 An approach to the identification of Criteria for E-endorsement

From an historical point of view, it is interesting to note that in 1977, Alan Davies affirmed that

proficiency tests.....are concerned with assessing what has been learned of a known or unknown syllabus.....(they take place) at some time after learning has started (and relate to) future non-language performance. ((Allen and Davies, 1977:46-47)

His examples of proficiency tests include

the Cambridge Proficiency Exams, the Michigan Tests, the Test of English as a Foreign Language (TOEFL).....(ibid)

in all of which the "future non-language performance" related to, is a course of academic study.

Eight years later (1985:5) Davies identifies the following continuum categories in language testing procedure:

discrete point.....	integrative
indirect.....	direct
norm-referenced.....	criterion-referenced

and remarks that

although different CL (*communicative language*) tests will take up different places on the continua, in general they will be more integrative, direct and criterion referenced than traditional (*proficiency*) tests. (My italics)

I will argue in Chapter Four that to date, explicit or implicit acknowledgement of the tenets of communicative competence notwithstanding, language tests (whatever they

are called) are for the most part located closer to the left than to the right of Davies' continua. Published tests examined in the last three years characteristically reflect an analytical view of language, with linguistic form and communicative content parallel aspects of testing, rather than a synthetic view with form and content integrated and assessed in terms of successful/unsuccessful communication. The reason for this is probably that the former admits if not of "objective" marking, then at least of more objective assessment than the latter, and is thus perceived to be more "fair". Discrete point, norm-referenced tests are also more readily "rationalised" (on a superficial level), than is integrative, criterion referenced assessment; and Davies underlines this problem with "a sceptical view" of communicative language testing which suggests that the latter may not, in fact, be possible. (in Yeld, 1985:3 - 8)

It seemed to me however that significant problem areas in the 1986 Written Test of English Proficiency as discussed above, were accessible only on the admission of sociolinguistic factors to a theory of language assessment. This, together with the proposition that "language ability" is directly related to the purpose it serves (which is the basis of the needs analysis approach of recent years and perhaps the most significant practical outcome of research in the field of communicative competence to date) suggested

an attempt at moving closer to the right hand side of Davies' continua.

2.3 Written Tests 1987 - 1988

Although, as reflected in Hyltenstam and Pienemann (1985), there is to date relatively little research in the field of language testing other than where tests are "parasitic on" (second language) teaching programmes, all studies in "communicative language" assessment share the basic assumption that "real-life" needs can be captured to a significant extent in the testing situation. This is the point of departure of tests for special, or specific, purposes and the (explicit or implied) needs analyses on which they are based.

A move towards the right of Davies' continua, in the context of endorsement requirements, implies in the first place an analysis of teacher needs. To the best of my knowledge there has not to date been any systematic attempt to identify such needs, or to rationalise endorsement procedures on any other basis, even though at least one, and often both, of the languages required for the so-called bilingual certificate is a second (or third) language for all trainees. Neither is there provision in State

regulations for any specific teaching programme related to endorsement requirements.

But if it was clear that endorsement tests are valid only to the extent that they test what they are intended to test, identification of teacher needs "to teach English-medium classes" along lines as first proposed in the seventies by Munby (1978), for example, and developed by Carroll (1980), did not appear to be a practicable exercise for several reasons.

The immediate reason was the Written Test in English Proficiency (HDE, Secondary) scheduled for February, 1987: there was no time for a meaningful attempt to construct a test accounting for language activities relevant to teacher "occupational events" at secondary school level, based on a taxonomy of 54 communication skills. Then too, it seemed unlikely that any one such instrument comprehensive enough to account for the needs of all method groups would lend itself to meaningful sampling in a single two hour test.

In summary, the problem was not so much to identify professional communicative needs in the English medium classroom, as to contain them in a form which could serve as a basis for the design of valid and reliable tests of communicative competence.

Under the circumstances it seemed that a "trial and error" approach to rationalisation of teacher communicative needs on the basis of an experimental taxonomy would be the most realistic advance on the 1986 situation. To this end, the following assumptions "seemed to make sense":

- * the requirement that teachers be "competent to teach English-medium classes" is less concerned with "competence to teach" (which is the domain of the Higher Diploma in Education) than with effective classroom communication.....
- * effective communication has more to do with meaning ("content") transaction than with the formal characteristics of the communicative medium (in this case, English) through which such transaction takes place; and
- * in the sense that all effective teaching depends on effective classroom communication, there is no significant difference between the communicative needs of EL1 and EL2 teachers in the English-medium classroom.

In this first attempt to move "closer to the right" (of Davies' continua) marking was weighted in favour of integrative assessment (only 15 of a possible 90 were "discrete-point" marks, as opposed to 40 in the 1986 test).

A "direct test" in the form of a letter of application for a teaching post seemed appropriate as a test of performance in a "real-life" situation, in contrast with questions 3 and 4 of the 1986 test which were unlikely simulations; for research purposes, it was however considered consistent with "norm-referencing" on the grounds that in the "real-life" situation the post would be awarded to the "best" applicant. An "open-ended" exercise in the form of an expository paragraph on one of three given topics (25 marks), and a summary-type task (25 marks) were considered more consistent with criterion-referenced marking, the criteria being coherence, cohesion and relevance of content assessed on an integrative rather than on an analytical basis. The potential for subjective mark allocation was thus even higher than in the 1986 test (75 as opposed to 50 possible marks).

In an attempt to avoid this

~~On the other hand,~~ the test was designed and all scripts were marked by one person, and both test and marks were moderated by one person. Although this did not in itself guarantee consistent assessment (single marker inconsistency due to error or fatigue could not be excluded), it was a fair assumption that with two people instead of six involved in marking, potential inconsistency due to different subjective perspectives was considerably reduced. It was also considered relevant (in the interests of consistent assessment) that for this experimental test both the tester

and the moderator were language teachers/testers with experience in applied linguistic research.

2.3.1 Analysis of Responses (1987 Tests)

Of the 238 students who wrote the Written Test in English Proficiency in February 1987, nearly 40% misunderstood Question Two. It was intended to elicit a form of summary consistent with a final paragraph as follows:

THE FOLLOWING TEXT IS REPRODUCED FROM THE INTRODUCTION TO VAN GOGH IN ARLES, AN ILLUSTRATED CALENDAR AND DIARY PUBLISHED BY THE METROPOLITAN MUSEUM OF ART. READ IT CAREFULLY, THEN COMPLETE THE FINAL PARAGRAPH IN ABOUT 50 WELL-CHOSEN WORDS, BEARING IN MIND THE CONTEXT INTO WHICH YOU ARE INSERTING THEM.

Vincent van Gogh (1853-1890) lived in Paris from March 1886 to February 1888. There he discovered the world of the Impressionists and the Post-Impressionists; he was influenced by Japanese prints; he pursued theories of colour and increasingly felt that his artistic language could only attain its fullest development in the south of France, where the light and colour would serve his needs. He chose the provincial town of Arles, situated on the Rhone thirty miles from the Mediterranean.

Van Gogh spent almost fifteen months in Arles. He arrived on February 20, 1888 and left on May 8, 1889. It was a remarkably fertile period. he produced some two hundred paintings and one hundred drawings, and he wrote over two hundred letters - to his brother Theo in Paris, to his sister, Wilhelmina, and to his artist friends.

In this calendar we have reproduced.....

(25 MARKS)

Other questions arising from this Test seemed relevant to the research purpose:

* Question Three required ".....a letter of application to the advertiser, giving the relevant information about yourself, your qualifications, and suitability for the post". Whether or not such information was factually correct was irrelevant (in any case, the tester had no basis on which to assess its accuracy). The intention with this question was to test formal letter composition in a "direct" test approximating a "real-life" exercise. On analysis it became clear however that the "directness" of the test was in fact severely limited by the conditions under which it was written. In a true real-life situation, there is ready access to information on appropriate letter format, as well as to other relevant sources (for checking syntax, spelling, register). And even more pertinent perhaps, the endorsement test is not associated with any specific course or professional requirements other than with competence "to teach English-medium classes".

*...Question Four (a comprehension test) was a brief (about 200 words) but fairly sophisticated text in respect of subject matter (an argument for the legitimacy of Broken English, by H G B Casimir), syntax and vocabulary. There were 5 multiple choice questions, each with the correct answer and 4 distractors, for a total of 5 marks. This question as a whole was of little significance in relation

to the overall mark (90), and was primarily intended to elicit some indication of a general level of comprehension. It was interesting that in follow-up interviews, unsuccessful candidates (about 15% with less than 3 out of 5) were in most cases able to correct errors themselves, on first attempt, after re-reading the given sentences. To what extent did "test conditions" (time limit, nervousness) interfere with comprehension in the first attempt? Relevant in a general sense in any test, the question here is rather more specific: how can appropriacy of comprehension text, questions, and time limit for any given group be rationalised? The point is not whether the comprehension test per se is "too difficult", as ^{was} suggested by a delegate from one of the training centres at the meeting of the subcommittee of the Provincial Advisory Committee on 8 September 1989, referred to in Chapter One. It is rather a matter of how far comprehension texts approximate to actual discourse consistent with English-medium classes at secondary level, and how far "they need to do so to be effective". (Widdowson, 1978) There can be little doubt that under present conditions, comprehension texts are on the whole selected more on tester intuition than on the basis of rational interpretation of its purpose as a test.

* Question Five (for ten marks) required correction of one grammatical error in each of 5 given sentences, and the

rephrasing of 3 ambiguous sentences. The purpose was to try to obtain some indication of the inter-relationship of knowledge and performance (or of usage and use as defined by Widdowson, 1979, Chapter 1), by comparing item responses in Question Five with written discourse elicited in Question One. In fact, there was little evidence for any assumptions in this respect. Overall, testees who scored above 60 out of 90 for the Test, scored 6 and above in Question 5. In the case of scores below 60, there was no discernible consistent relationship between grammatical knowledge as elicited in Question Five and grammatical performance as elicited in Question One. Possibly of more significance was the fact that given the opportunity during personal interviews, testees were frequently able to correct their mistakes, which suggests that they might respond to intensive remedial tutoring focused on identified areas of need, in contrast with the graded objectives approach usually associated with special purpose programmes.

Although not required in terms of State regulations, such tutoring was offered to about thirty trainees who failed to obtain 60 or more in the Test and were thus at risk of graduating without E-status. Eighteen of these were subsequently upgraded on the basis of assignments submitted for credits in different areas of course work - that is, on performance in direct tests in the true "real-life" sense. Ten of the remaining twelve trainees (who did not accept

tutoring) obtained E-status on performance in the Written Test of English proficiency prepared for the In-service Course in September 1987.

This later test consisted of a formal letter of application for a bursary for a possible 25 marks, a traditional comprehension test (not MCQ) for 25 marks, a summary exercise for 25 marks, and a Cloze test (15 marks). (Formal letter writing had been identified as an area of need by in-service students themselves, and had been dealt with during the course, as had comprehension tests, summary technique and the Cloze principle of expectancy grammar). Marks were thus heavily weighted in favour of expository discourse, nearly 75 of a total of 90 marks being allocated to open-ended questions (eight of the ten questions set on the comprehension piece required explanation or discussion not directly retrievable from the text). 14 of the 16 in-service students were upgraded to E-level endorsement on the basis of this written test; and assessment in an individual 5 - 10 minute oral test with two testers corresponded with assessment in the written medium in all cases.

2.3.2 Conclusions

There was little evidence for any significant conclusions about the significance of adopting a criterion of

communicative competence for the testing of English, or about the effectiveness of the given tests as predictors of competence "to teach English-medium classes". On the other hand, in the interests of greater consistency of endorsement assessment, it seemed reasonable to formulate certain tentative assumptions on the grounds of observed testee performance, and to test these assumptions in subsequent HDE groups.

It was assumed that:

- * assessment of all testee responses in the 1987 tests by one tester had ensured a higher level of consistency of marking than had been the case in 1986;
- * the higher success rate in the initial test (10% higher in 1987 than in 1986) was due to the tester move towards Davies' perspective on language ability as communicative competence rather than to any significant "improvement" in testee ability per se;
- * assessment on an integrative, criterion referenced basis in direct tests (course assignments) provided valuable opportunities for correlation of course requirements and tester perspectives in respect of language ability;
- * it "makes sense" to relate the approach to assessment to the purpose for which it is required.

The tentative proposition in 2.2.2^{p.25} above that endorsement assessment be perceived in terms of Davies' continuum categories was now developed as follows:

where teacher needs are perceived primarily in terms of text cohesion, logical progression, register appropriacy, syntax and vocabulary, assessment is consistent with traditional requirements of proficiency as represented by the left-hand side of the continua, and particularly with indirect tests and discrete point ("objective") marking procedures; and

where teacher needs are perceived primarily in terms of "total message transaction", assessment is consistent with the concept of communicative competence as represented by the right-hand side of the continua, and in particular with direct tests and integrative (impressionistic) marking procedures, based on a socio-linguistic view of language as "meaningful communicative behaviour" (Widdowson, 1978;p.3).

As has been noted before, and is central to this study, proficiency and communicative competence as characterised here are not seen as mutually exclusive approaches to language ability. On the contrary, as Widdowson points out, "meaningful communicative behaviour", or "use", of English is closely related to "knowledge of the language system of English", or "usage". The issue is rather "how and where" to locate testing procedures along Davies' continua in order

to yield results consistent with teacher requirements in a multilingual target group.

In-service tutoring in 1988 (four groups ranging from 6 - 8 teachers in each) was based on needs as identified by teachers themselves at the beginning of, and during, the course. Tests were similar to those administered in 1987, with about 75% of marks allocated on the basis of discourse elicited in expository texts, summaries and comprehension tests, and 25% on the basis of closed-ended tests as suggested in Practice Tests for Cambridge Certificate of Proficiency in English, 1986.

Unsuccessful HDE students were again tutored in a series of individual interviews (successful students were not interviewed for two reasons: neither discrete point nor integrative assessment discriminated seriously against them, and they were therefore unlikely to contribute significantly to the study; and having nothing "at stake", they would also have been less motivated to co-operate in the research programme).

And as part of the ongoing collection of data relevant to the concept of *test* failure as opposed to *testee* failure in a linguistically non-homogenous target group, a questionnaire was sent to the 42 training centres affected by endorsement regulations.

CHAPTER THREE

THE QUESTIONNAIRE, AND RESPONSES TO IT

3.1 The questionnaire

In September 1988 the following questionnaire was sent to 42 teacher training centres:

Dear Sir/Madam,

Language requirements for teachers, as formulated in sub-sections 10 and 11 of CRITERIA FOR THE EVALUATION OF SOUTH AFRICAN QUALIFICATIONS FOR EMPLOYMENT IN EDUCATION, in effect leave assessment in terms of these requirements to the training institutes themselves. Recent developments in applied linguistics and our own experiences in the field in question have led us to re-evaluate our interpretation of, and response to, the formulation as it stands.

Research in progress in the Language Education Centre addresses this issue, and gives rise to the enclosed questionnaire. We would be very grateful if you could find the time to complete it and return it to us as soon as possible, or by mid-October at the latest.

We appreciate your co-operation.

QUESTIONNAIRE

Please indicate with a tick the response most nearly consistent with your own.

1. The terms *competent* and *proficiency* as used in the formulation of language requirements set out in the CRITERIA FOR THE EVALUATION OF SOUTH AFRICAN QUALIFICATIONS FOR EMPLOYMENT IN EDUCATION refer to

a) identical abilities/skills and are used merely for stylistic variation.

b) different and discrete abilities/skills.

2. Which of the following do you accept?
- a) language skills = language abilities
 - b) competence refers to skills
competence refers to abilities
 - c) proficiency refers to skills
proficiency refers to abilities
 - d) practical considerations make these theoretical distinctions irrelevant.
3. With reference to the criteria of competence and proficiency, do you
- a) consciously and consistently apply them in your E/e assessments, according to a pre-determined "checklist"?
 - b) assess students on an ad hoc and experimental basis?
 - c) apply them in conjunction with the criterion of *communicative competence* or any other? (Please specify)
4. Do you implement the requirements for the E/e endorsement?
- a) Yes
 - b) No
5. If you implement the requirements for the E/e endorsement, do you do so
- a) by means of special tests at the beginning of the diploma course?
 - b) by means of special tests at the end of the diploma course?
 - c) on the basis of general student performance in the medium of English?
 - d) by some other method (please specify)

6. Would you be interested in the results of our study?

- a) Yes, very interested.
- b) Fairly interested.
- c) No.

Thank you very much for your co-operation. If you have any comments to add to the above, please do so in the space below.

.....

3.2 Responses Received and Issues raised

We received only 16 responses to this questionnaire. They did however constitute a cross-section both of the language groups, and of the types of training centres, to which the regulations apply: 7 from Training Colleges, 4 from Universities in the Republic, (representing English, Afrikaans and dual medium centres), and one each from the Universities of Transkei and Venda. Three completed questionnaires were returned without identification. At about the same time, perhaps coincidentally, we received a request for advice from the Department of General Didactics, University of Durban-Westville, with reference to the regulations in question, as follows:

Do we endorse our Zulu/English students' certificates with a ZE, a Ze, a ZEa, a ZAe or do we ignore the fact that Zulu is "the language which the teacher knows best" and just indicate the official language proficiency?

Analysis of replies to the questionnaire indicated that:
8 respondents believe that as used in the CRITERIA FOR THE EVALUATION OF SOUTH AFRICAN QUALIFICATIONS FOR EMPLOYMENT IN EDUCATION, 1988, *competence* and *proficiency* are stylistic variants of the same concept and refer to identical abilities/skills;

7 respondents believe that they are intended to refer to different and discrete abilities/skills;

7 respondents believe that practical considerations make any such theoretical distinctions irrelevant; and

7 responses suggest that, at the relevant centres, the criteria of competence and/or proficiency are consciously applied, regardless of whether or not any distinction is intended.

Additional remarks offered further highlight ambivalence towards official requirements. At one training centre they were "puzzled by the wording" (of the regulations); at another, regulations were "not really" implemented at all; and a third respondent commented as follows:

I experience the following problems each year:

1. Student unhappiness because E/e and A/a testing procedures, expectations and "standards" seem to differ markedly.
2. Speakers of other mother-tongues are expected to obtain an "E" or "A" (i.e. at L1 level), whereas (for them) these are "second" or "third" languages.
3. We have students who are specializing in "technical" subjects (woodwork, metalwork etc): for them the bilingual competency test and its purpose seem irrelevant or at best - peripheral.
I would like to see this requirement (E/e and a/A) scrapped.

On the limited basis of data collected, there is evidence that theoretical and practical problems surround the issue of the so-called bilingual certificate, and that different interpretations of "competence" and "proficiency" are a major cause of anomalies^a in administration of the language regulations from centre to centre.

The high level of interest elicited by this study (15 respondents would be "very interested", and one "interested" in the results) suggests that testers are not unconcerned about this situation. On the other hand, additional comment (as exemplified above) would seem to indicate that concern is directed less at inter-centre differences per se (which is perhaps a reflection of academic autonomy), than at the fact that language endorsement is an external requirement, standardised in concept but unresolved in terms of reference. To extend Hulstijn's proposition above, it doesn't seem to "make sense" that the consumer body, or State Education as the prospective employer, should delegate assessment for a compulsory professional qualification in terms of regulations which demonstrably rely more on intuition than on rational interpretation for implementation.

This is not intended to deny the role of intuition in the assessment of language ability, as recognized in the

literature by many researchers. As Yeld records, Alderson suggests that rational analysis "is not enough..."

judgements of item difficulty are notoriously inaccurate, and in the problematic area of test content, judgements of what exactly an item measures are often intuitive. (Alderson, in Yeld;1985:38)

In his paper *Do the Tests meet the Criteria*, he suggests that (communicative) language testing be seen in terms of a standardising procedure, with rational analysis of item content the concern of a Preview Phase, followed by a Trial Phase "where initial predictions are tested".

There can be little doubt that, in principle, a standardised procedure of assessment for a standardised language qualification makes sense. Under present circumstances, with each centre responsible for E/e endorsement of its own trainees, and little or no inter-centre co-operation, the question "do the tests meet the criteria" only elicits more questions: whose tests? whose criteria? And underlying both of these, another, central to linguistic debate in South Africa today: whose English? This study is materially located within this debate. As an English-medium "open" university, UCT draws students not only from

the 8% of the SA populace who classify as English mother tongue speakers

some of whom speak Standard South African English (Young; 1990:126 - my italics), but also from all other South African language groups. Numbers in the latter groups have

grown significantly in the School of Education during the period of this project, and continue to do so. The problem, as demonstrated in Chapter Two, was how to interpret the given criteria for endorsement. Should they reflect international norms of educated English, traditionally identified as English proficiency, with regard to syntax, lexis, style and register and, in the spoken medium, pronunciation - thereby discriminating against many teachers with an English-medium teaching qualification who would, on that basis, not receive E-endorsement? Should the "rules" be bent a little, in a more or less patronising attitude towards disadvantaged students - so that E-endorsement would, with intent, mean different things in different cases? Or does the answer lie in accommodating all varieties (sometimes difficult to distinguish from idiolects) - which would arguably entrench a perceived "downward" spiral of what passes as acceptable English, and perhaps ultimately a South African variety less than useful for communication with the growing English-speaking world?

Mawasha (1986), without overtly identifying the variety of English he had in mind, has proposed that

language proficiency in English should be a requirement in teacher recruitment over and above expertise in subject matter content. The maxim that every teacher who teaches through the medium of English is a teacher of English should apply. Secondly, at school level: it should be impressed upon the learner that proficiency in an adopted language calls for practice.

If by "proficiency in English" Mawasha intends a traditional interpretation of proficiency consistent with published tests for entrance to English and American universities, then as shown above, present circumstances in South African education suggest that achievement of his "ideal" would depend more on long-term teacher training in English (Method) than on unrealistic short-term requirements "in teacher recruitment".

Ndebele (1986) takes a contrasting view. He believes that

in promoting English in a multi-lingual society, it may become increasingly and dangerously difficult to make a distinction between English and Education..... English may be spoken universally, but it does not carry the sum total of the world's wisdom.

He makes a valid point. No single language does this; and for this reason, in education everywhere, great value is attached to knowledge of languages other than one's own - German, French, Italian, indeed any language which serves the purpose of academic growth in any field of "wisdom". On the other hand, to see "the promotion of English in a multi-lingual society" as a "dangerous form of encapsulation" (of education) (ibid) is arguably to take a more limited view of English than would seem to be consistent with the not inconsiderable - and growing - repository of "wisdom" it is acknowledged to embrace in global terms today.

Finally, before turning to a review of recent literature in language testing for support for my proposition, it is appropriate to comment here on the research structure of this project.

A re-evaluation of *competence* and *proficiency* as criteria in language tests would lend itself to research along either of the two paradigmatic paths characteristic of social research as identified by Hitchcock and Hughes (Research and the Teacher; 1989). Given the overall objective of test standardisation, the study would, ideally, operate deductively from "hypothesis formulation" to large-scale trial testing and feedback research (that is, in accordance with the positivistic paradigm).

In the absence of facilities for large-scale testing, the present project had little alternative but to operate inductively, moving from the formulation of a proposition to small group testing in conjunction with techniques such as observation and informal interviews characteristic of the qualitative approach associated with the interpretative paradigm. Dislocation of the paradigm here - the literature review follows, rather than precedes, the data collection phase - is due to the particular function of the review in

this case. It is not a consideration of previous research on the topic of endorsement - there appears to have been little interest in the theoretical implications of State language regulations to date. It is rather an attempt to interpret, or re-evaluate, the notions of *competence* and *proficiency* in the South African context, on the basis of recent studies in language testing.

CHAPTER FOUR

LITERATURE REVIEW

4.1 Motivation for Theoretical Analysis of the Criteria

The formulation of State language regulations pre-dates current trends in applied linguistics and language education, and it could therefore be argued that it is unrealistic to attempt to interpret the given criteria on the basis of recent research, which in effect leaves it to the tester to surmise what the given criteria are, or were, intended to mean. On the other hand, there can be little doubt that it is contrary to fundamental principles of education to ignore potentially relevant developments in order to maintain an inefficient status quo.

The question here, however, is not so much whether to interpret State requirements on the basis of current theories in the field of language assessment, as whether there is in fact any real basis for doing so.

As elicited by the questionnaire discussed in Chapter Three, as recorded in the Minutes of the Meeting of the Advisory Committee to the Cape Education Department in September

for language acquisition (1965). Concerned with formalizing the linguistic knowledge of a hypothetical ideal speaker-listener, Chomsky

made the point that the language user himself must possess intuitively and unconsciously this capacity to abstract from the concrete manifestations of language (Stern;1983: 129)

yes - any language

The distinction between "competence" and "performance" on which this proposition is based, with the Chomskyan emphasis on the study of competence as the task of linguistics, has had, and continues to have, far-reaching effects not only in "pure" linguistics, but also in other fields, notably psychology and philosophy. However, as Stern points out,

many linguists.....did not go along with this highly abstract view of linguistic enquiry. (Stern:146)

In deliberate contrast to Chomsky's *linguistic competence*, Hymes takes up the question of "performance" (in Pride and Holmes;1972:269 - 292). Arising from his study of the language problems of disadvantaged children, Hymes sees Chomsky's claim that

linguistic theory is primarily concerned with the ideal speaker-listener.....(who) is unaffected by such grammatically irrelevant conditions as memory limitations, distractions, shifts of attention and interest, and errors (random or characteristic) in applying his knowledge.....in actual performance

as

almost a declaration of irrelevance. All the difficulties that confront the children and ourselves seem swept from view. (op cit:270)

Hymes argues for a sociocultural perspective on language use which transcends this dichotomous view to embrace both the language user's competence and actual performance in a multi-dimensional but integrated theory of *communicative competence*.

Two aspects of this theory are of particular significance here. Firstly, in contrast with the negative connotations of "performance" in Chomskyan theory, in the sociolinguistic view "language performance" is seen in positive terms of "real-life" communication; and secondly, Hymes suggests that judgements about the various sectors underlying language behaviour fall into four distinct categories:

1. whether (and to what degree) something is formally possible;
2. whether (and to what degree) something is feasible ;
3. whether (and to what degree) something is appropriate, and
4. whether (and to what degree) something is done.

For example, he says, "a sentence may be grammatical, awkward, tactful and rare". (op cit:281)

As used in the context of State language regulations, the notion of competence as a "measure" of language performance (rather than as an ideal abstraction of it) is clearly more akin to Hymes' sociolinguistic orientation than to

Chomskyan theory. Consequently this review focuses on literature recording research which seeks to apply insights arising from Hymes' four basic categories of judgements to language testing in general, and to performance criteria in particular.

As Stern affirms,

the social view of language.....has found increasing acceptance since the middle of the sixties.

(op.cit:111)

But while there is no lack of studies addressing issues in test design research, problems surrounding test design are, for the most part, seen in relation to a specific teaching syllabus. They are therefore not necessarily of direct relevance to problems of assessment in the context of State regulations; neither do the regulations make provision for a teaching programme on which testing could be based.

On the other hand, as Davies pointed out at the British Council Course on Communicative Language Testing in 1985, although

CLT makes teaching sense because of washback effect.....it is not necessary (from the point of view of testing theory) for teaching and testing to be isomorphic.

(in Yeld:1985)

In Davies' view,

an ideal CLT should:

- i) - test communicative competence,
 - ii) test target language needs,
 - iii) incorporate a range of situations,
 - iv) have particular objectives, and
 - v) be reliable and feasible;
- (i) and (iii) should be broad-based, (ii) and (v) should be narrow-focused. (op.cit:5)

But although the demands on CLT testing in these terms are heavy, Davies suggests that testing of communicative language may in reality not be very different from traditional language tests, the critical requirement perhaps being merely a *change of tester focus*. The subject of practical test design is explored more fully in Chapter Five. For the present, I consider the theoretical import of the notion of communicative competence vis-a-vis that of proficiency, with the aim of establishing whether there is any basis for a more critical appraisal of these terms in testing methodology than is currently the practice.

As noted earlier, available literature indicates that there is a strong tendency to use "proficiency" and "communicative competence" in apposition, or even interchangeably, to characterise a particular perception of language ability. Stern notes, for example, that in a formulation prepared as a basis for test development (by Canale and Swain, 1980), *proficiency has been interpreted as communicative competence*; while Stern himself suggests an interpretation of *competence (or proficiency) (as) linguistic and*

communicative competence, i.e. a system of formal and social rules (Stern, 1983:345 - 349). Romaine, in a chapter on *The Measurement of Communicative Competence*, offers "guidelines (for) the testing of language proficiency" (1984; p.231), and Ingram motivates for *proficiency as communicative competence*, in terms of the learner's ability to mobilize language in carrying out communication tasks, i.e. with *language proficiency*. (Ingram, 1985:226) (My italics)

The tendency to identify proficiency with communicative competence thus occurs both on an abstract level (as in the first two examples above) and on the concrete level of practical testing (as in the last two examples).

On the other hand, it is significant that Ingram (1985) sees communicative competence in strong contrast to *traditional* notions of language proficiency. In the same study, he quotes Sollenberger as follows:

The person's so-called language proficiency, while it may have been quite accurate in technical skill terms, did not mean effectiveness in *communication*. In some cases, it may have enabled the person to misrepresent or foul up more effectively...I'm sure we all know people who talk nonsense fluently.

On the other hand I know people who butcher the language, whose accents are atrocious and whose vocabularies are limited. For these reasons we give them a low proficiency rating. Yet, for some reason, some of them are effective communicators. *Red herring*

Finally, although we are here more concerned with the impact of theoretical formulation on testing methodology than on

A Glossary of Testing Terms in a recent edition of *Forum* is, in effect, consistent with Yeld' view: a proficiency test, it says,

measures one's knowledge and/or ability in a foreign language without regard to formal study or text used (Vol. XXVI, 1988:6)

a suggestion that makes it difficult to locate the notion of proficiency, not only in syllabus design but also in current language testing research which tends to recognize a fundamental difference between the two (that is, between "knowledge" and "ability").

Clearly, then, proficiency means different things to different people, or as Hyltenstam and Pienemann put it

many researchers have experienced that the term 'language proficiency' is a chameleon... (1985; p.15)

And indications are that the more closely the notion of proficiency is associated with the theoretical construct of communicative competence, the more chameleon-like it becomes; for there can be little doubt that if the goal of the language learner is still frequently (if nominally) proficiency, the "goal-posts" have moved. Where in a "pre-scientific period in language testing" (Spolsky, 1978) they represented formal, "objectively" measurable, linguistic skills, assumed to be the building blocks of what Stern calls "native-like" proficiency, the new, less prescriptive

orientation to language-in-use has shown this to be a false assumption. It is even likely, as suggested by Sollenberger above, that grammatical expertise is not the most important constituent of the ability to communicate - all of which makes assessment of "proficiency as communicative competence" a much more complex proposition than proficiency in traditional terms used to be. Yeld's Report on Proceedings of The British Council Course on Communicative Language Testing (1985) highlights many of the theoretical and practical issues which make it so.

4.3 Towards a Distinction between Proficiency and Communicative Competence in Language Testing

In contrast with the term "communicative competence", which arose, comparatively recently, in the context of linguistic research, "proficiency" is not a specifically linguistic concept and, as shown above, does not lend itself to debate in terms of any readily identifiable linguistic theory. This is arguably the reason for its wide and continuing currency in all sectors of language education, for it is accessible to all in terms of its dictionary definition: "every-one knows" that to be proficient in something means to do something.

correctly and competently through training or practice, skilled. (The Oxford Paperback Dictionary, 1979:508)

Given this definition, it is easy to see how proficiency became

the learning outcome in which teachers, administrators, curriculum developers, test constructors, researchers, parents and, of course, students themselves are all equally interested. (Stern, 1985:341)

In a sense, it is also the goal of the communicative teaching programme that students learn to use the language "correctly and competently"; but whereas the goal of the language learner was traditionally seen in terms of an ideal of *general* language competence/proficiency akin to that of the mother-tongue speaker, the goal of the communicative teaching programme has come to be seen in the more pragmatic terms of "specific purposes".

This study proposes that it is in this fundamental difference between traditional and modern approaches to the purpose of language learning, that the key to problems of assessment criteria today may lie.

The post-Hymes change in the focus of language teaching has triggered many questions - more than have so far been answered, as continuing attempts to transfer the notion of communicative competence from the research environment to the classroom ensure that it remains a focus of linguistic debate, not least in respect of its semantic content.

Although himself an exponent of "teaching language as communication", Widdowson approaches the term with caution:

Communicative competence is now very much in fashion, (but) it is as well to be wary of it: particularly since it does not seem to be used in the same sense by different writers. (1978; p.163).

The literature confirms this. Significantly however, although (as noted above) Stern, Romaine, Ingram and others who have joined in the debate on semantic issues surrounding this concept, do not, in their own work, make any clear distinction between communicative competence and proficiency as criteria of language ability, they do recognize two closely-related but discrete *kinds* of ability: an ability to carry out communicative tasks (which implies *appropriate* language use); and an ability to select the (*correct*) linguistic features - syntax, lexis, register, discourse structure, pronunciation - required to do so.

These two "kinds of ability" - conceptually related to Widdowson's earlier distinction between *communicative abilities*, defined with reference to the manner and mode in which the linguistic system is realized as use, and *linguistic skills*, or usage, defined with reference to language medium (Widdowson 1978:57 -75) - are clearly complementary in nature:

one cannot acquire the former without acquiring the latter. (op. cit:67)

Nevertheless, as Widdowson demonstrates, they do respond to different elicitation techniques (for example, *use inference* and *usage reference* questions; op.cit:100 - 105).

A principle of dual, complementary abilities underlying language use, rather than a single one, suggests that such abilities may, at least to some extent, operate independently of each other.

"Real-world" use confirms this: as Hymes suggests,

a sentence may be grammatical, awkward, tactful and rare (op cit:281,282);

alternatively (as noted by Sollenberger above), an instance of "language-in-use" may be syntactically and phonetically deviant from the norm, while at the same time effective in semantic intent. I suggest that in both cases, the communicative purpose of the hypothetical utterance so characterised would be served, as implied by the "value judgement" *tactful* in the first case, and *effective* in the second. Arguably, within the context in which they occurred, they would be perceived as communicatively *appropriate*, if linguistically *incorrect*.

Arguably, too, the most significant practical outcome of a re-orientation in language education in response to the

questions upon which Hymes bases his model of language use (is it *possible, feasible, appropriate, done?*), has been the "special purpose" language course. Aimed at addressing the language needs of the (L2) learner, SP courses are characteristically organised as a series of graded objectives with successive levels of ability defined in terms of impressionistic verbal descriptions accounting for the learner's *communicative* competence at each level (often called *proficiency* levels, as in Stern 1983:356; and TOEFL from a Communicative Viewpoint on Language Proficiency, Research Report 17, 1985: Appendix E). SP courses characteristically - and inevitably - account also for *linguistic* competence at each level.

The problem, from the point of view of this study, is that such "proficiency" scales commonly serve a dual purpose: they function both descriptively, characterising levels of competence reached by on-course learners; and prescriptively, as rating scales for the purpose of assessment in terms of specified real-world needs. I look more closely at this second function.

4.3.1 A Return to a Traditional Interpretation of "Proficiency"?

To say that Special Purpose (Proficiency) Scales measure (the learner's) communicative competence against defined levels of needs is not to say very much. As the "special purpose" concept itself implies, there are many possible purposes for which people need language competence, and in this sense also many *kinds* of competence to be accounted for on the various levels of the relevant rating scale.

I suggest here that SP Courses, and the rating scales which accompany them, fall into two broad categories of purpose: one, where language ability is a means to an extra-linguistic end (as in the English-*medium* classroom), which implies that the emphasis in teaching is in the first place on *appropriate use*; the other, where ability in the language is itself seen as the end (as in the *English lesson*) and, correspondingly, the emphasis in teaching is on (correct) *grammatical usage*. (I use the adjectives in the sense that Hymes does, e.g. 1972:277, and the notions of *use* and *usage* as Widdowson defines them, 1978:1 - 20.)

This argument is based on the intuition that essentially different kinds (not necessarily levels) of ability in English are required in the case of the L1 speaker of another language seeking professional employment in an

English-speaking country, and by the non-native speaker applying for entrance, say, to Oxford or Cambridge University - as may be confirmed by even a cursory glance at any text book designed for use by such an applicant (characteristically described by the compilers as *guides to proficiency*; for example, Robinson, 1981; Fowler and Pidcock, 1985; King and Stanley, 1987; John Millington Ward;1980).

I would argue that in the case of the former, testee needs may be defined in relation to a specific job description in an identifiable area of the labour market, where vocabulary, syntax, comprehension, and requirements for the written word, commensurate with the responsibilities of the job, may be predicted with reasonable accuracy. This is in accordance with prototype Special Purpose rationale as put forward in Munby's *Communicative Syllabus Design* (1978) and Carroll's *Testing Communicative Performance* (1980).

In the case of the latter "special purpose", on the other hand, prospective students are characteristically required to demonstrate a comprehensive and complex linguistic ability commensurate with tertiary level studies in English, embracing for example

verbal reasoning abilities or higher-order reasoning abilities (*Cracking The System - The Scholastic Aptitude Test*; 1986:31),

recognition of form, register, style, intention, punctuation, literary figurative language, literary style and appreciation (English for Cambridge Proficiency by Archer and Nolan-Woods, 1986: Foreword),

and

flexible use of syntactic structures, morphological control, appropriate and effective organisation of composition material, wide range of vocabulary, spelling and punctuation skills (Michigan English Language Assessment Battery, Information Bulletin 1989 - 1990:13)

It may perhaps be argued that "English for academic purposes" falls into the same basic paradigm as any other special purpose test in that it is also "a means to an end". The point I would make is that it is here not *primarily* a means to an end; alternatively, that "appropriate language use" in this case embraces English *usage* consistent with internationally accepted norms in tertiary education (which is consistent with my argument that E-endorsement of English-medium diplomates is superfluous).

A logical consequence of this argument is that assessment of EL2 teacher ability "to teach English *medium* classes" (where English is in the first place a *means to an end*) falls wholly within the former category of tests. (In the absence of any indication in the regulations of what "less proficiency" *is less than*, a difference between E- and e-level endorsement is meaningless). On the other hand, assessment of teacher ability to *teach English* - where English is not so much the medium as the material of

teaching - would fall into the latter category. The relationship in Widdowson's terms is clear: in the first case "teachers are required to use (their) knowledge of the language to achieve some kind of communicative purpose"; in the second, they are more specifically called upon "to *manifest* (their) knowledge of the language system of English". (1978:3)

Once again, I do not suggest that these two approaches to language testing are mutually exclusive, merely that the tester need not necessarily see them as two halves of an integrated whole. That Widdowson does so, is due to his orientation to communicative competence as *the aim of the learner*. (1978:144) This study is not directly concerned with the *process* of learning (as noted above, State regulations make no provision for language learning), but rather, with the *outcome* of "what has been learned of a known or an unknown syllabus" (in relation to the needs of the EL2 teacher in the English medium classroom).

The distinction is thus not an arbitrary one; it is a reasoned approach to the intuition that all tests of language ability ultimately fall into one of two categories, depending on the "real-world" purpose which motivates it.

It is but a short step to the proposition I would put forward, that

THE CONCEPTS OF COMMUNICATIVE COMPETENCE AND PROFICIENCY ARE SEMANTICALLY DISTINCT FROM ONE ANOTHER. THEY FOCUS ON DIFFERENT AREAS OF (REALIZED) LANGUAGE AND FUNCTION IN DIFFERENT WAYS AS CRITERIA IN LANGUAGE TESTING.

On the basis of this proposition, and the argument which supports it, I suggest that the criterion for assessment for E-endorsement, as required in terms of current State regulations, is properly *communicative competence*. Tests for teachers of English, on the other hand, and for purposes of internationally acceptable academic qualifications, logically require a criterion of *proficiency*.

4.4 From Theory to Practice

The problem is, however, not so much to formulate a theoretical basis for the proposition as put forward here, as to bridge the gap between theory, on the one hand, and the application of theoretical concepts in practical tests, on the other.

If it is accepted that language tests fall into one of two theoretical categories as suggested above, it follows that there must also be two practically identifiable sets of "goal-posts" (to continue the analogy suggested in 4.2 above), or two taxonomies for testing, corresponding to the categories into which tests may fall.

Widdowson's distinction between *linguistic skills* and *communicative abilities* offers a useful perspective on this problem. Noting that "communicative abilities embrace linguistic skills" he demonstrates that (at least) to the extent that *communicative abilities* are realized at all, they involve the operation of the underlying linguistic skills ("speaking, hearing, composing and comprehending") required to do so (1978:67).

But "skills" and "abilities" do not occur in a constant ratio to one another. The "redundancy principle" of language as defined by Ingram makes this clear:

Because language is redundant, any specific item...is not necessarily essential to a speaker's successful functioning in the language since the deficiency may be compensated for by guessing based on the rest of the syntactic, discourse or semantic content... As learners develop proficiency in the language, so they master more of the systems that make up language (i.e. the higher the level of redundancy in the language)...
(Ingram, 1985:241)

From this perspective, Widdowson's concept of "linguistic skills" (defined with reference to medium) can be seen to

accord with my contention that formal skills ("correct" language use, or a higher redundancy level) are intuitively closely associated with an intra-personal language ability traditionally associated with L1 users and characteristically elicited by tests of English "for academic purposes" (and ⁷ for language teachers), and is thus consistent with my (traditional) interpretation of "proficiency".

Conversely, it also accords with my reservation of the term "communicative competence" for Widdowson's *inter*-personal "communicative abilities" (including, and inseparable from, *underlying* linguistic skills); or as Hymes put it -

there is behavior, and underlying it, there are several systems of rules reflected in the judgements and abilities of those whose messages the behavior manifests. (1971:281)

Sollenberger's comment (in 4.2 above) is relevant here too, not so much for the extremity of his views, but for the point he makes for a separation of the concepts of technical accuracy or *proficiency*, and effective communication or *communicative competence*, in the assessment of language ability.

This perspective on effective communication also accords with Canale and Swain's suggestion that besides Widdowson's communicative abilities of mode and manner (together with

the underlying skills on which they depend for realization) there is a third constituent of communicative competence, namely *strategic competence*, or second language user knowledge of

how to conduct himself as someone whose socio-cultural and grammatical competence is limited (Canale and Swain, 1980).

I suggest that for the purpose of practical testing as put forward in this dissertation, the relationship between strategic competence and communicative competence would seem to be analogous to the relationship between redundancy and proficiency, in that a criterion of *communicative competence* admits compensatory strategies (for example, grammatical and lexical paraphrase, sociolinguistically neutral grammatical forms, and first language discourse patterns, as suggested by Canale, in TOEFL Research Report 17; Appendix B:5 - 6), while a criterion of *proficiency* assumes high level redundancy skills characteristic of the L1 user.

Meaningful large-scale testing, with the facilities that such a testing programme implies in respect of data collection and feedback analysis, is needed to determine whether, or to what extent, my proposition is a valid one. But although evidence offered to support it here is based on a relatively limited study, I put it forward as a way of dealing with urgent problems of English language testing as documented in Chapters Two and Three - problems which will

conceivably become ever more urgent as a new South Africa takes shape - even if only as an interim measure, until linguistic theory provides a more fully developed approach.

In Chapters Five, Six and Seven, I attempt to account for my proposition in practical test design on the basis of the evidence elicited by a sample test. In particular, I attempt to devise a taxonomy of teacher needs in the English-medium class, and a methodological approach to the assessment of the language skills/abilities they imply, and to show how a criterion of communicative competence may be interpreted in "real-world" testing.

CHAPTER FIVE

TOWARDS A TEST DESIGN

5.1 Preliminary considerations

The way one tests (L2) language ability is influenced by the theories one holds about language, and about how it is used, taught and learned.

Until the late sixties and early seventies, language teaching was seen largely in terms of prescriptive grammar, and learning was assessed on the basis of discrete item testing and pre-determined "objective" mark schemes, an approach which would seem to assume an underlying theory of language as the sum of its parts, and of language ability as knowledge of the relevant rules for putting the parts together.

To some extent there still is, and probably always will be, a structural element in language teaching and testing, however integrated and disguised it may be in the "communicative approach" which, since the seventies, has been promoted as a more realistic basis for language teaching methodology. Characterised by non-prescriptive

attitudes to language use, and based on a proposed model of "real-life" communicative interaction, a criterion of "communicative competence" in language testing implies a more holistic approach than was previously considered consistent with the notion of (objective) testing, one which substantially views testee performance from the perspective of the "sum" rather than from that of the "parts" which compose it.

The changing focus in applied linguistic research has far-reaching implications not only for language teaching but more specifically, from the point of view of this study, for language testing methodology. The point is that if real-world communication, characterised by open-ended interaction, is the model on which language testing is to be based, then "communicative competence" as a criterion for assessment is subject to interpretation in relation to the real-world "needs" of the testee in the context to which assessment relates in each case.

Responses to our questionnaire (as documented in Chapter Three) suggest however that testers at the various centres have differing perceptions of the "real-world" objectives of assessment in accordance with State regulations; and the de facto situation is that language endorsement for teachers is characterised more by paradoxes and tensions than by consistent standards and rational methodology. The result,

as shown in this study, is that if "competence with proficiency" is assessed according to traditional criteria, it results either in an unacceptable level of non-endorsement (sometimes even in graduates in the same medium), or in inconsistent standards; and if it is assessed (in the spirit if not in the letter of the regulations) as "communicative competence", it would seem to assume a pre-existing available assessment of relevant teacher needs which does not to date exist. The tensions are an inevitable consequence of the paradoxes, and include uncertain testers on the one hand and, on the other, testees who at best perceive assessment in terms of professional status and job security (rather than in the wider terms of education, which is presumably the purpose of endorsement), and at worst as irrelevant bureaucracy and discrimination against the users of varieties of language other than the "standard" (South African) forms in English and Afrikaans.

In this chapter I propose an approach to assessment for E-endorsement in the light of recent research. At the same time, it is relevant to note that as political dispensations change, so do regulations; and if the immediate purpose of the test put forward in Chapter Six was to assess EL2 In-service testees (September, 1990), test design as considered here anticipates also a possible ongoing role for teaching through the medium of English in a post-apartheid South Africa.

In essence, the test represents an attempt to elicit evidence of how testees use English to communicate in the real world. I argue that there is likely to be high correlation between the results of such a test and the functional ability of the testee/teacher in the English-medium classroom, provided that test items appropriately sample linguistic skills and communicative abilities, as these are defined by Widdowson:

those skills which are defined with reference to medium (speaking, hearing, composing and comprehending) are *linguistic skills*.....those.....which are defined with reference to the manner and mode in which the system is realized as use (are) *communicative abilities*.
Communicative abilities embrace linguistic skills but not the reverse. (Widdowson; 1978:66)

The communicative abilities thus include reading and writing ("corresponding"), saying and listening ("talking") - and implicit in all communication, *interpreting* -

the highest level skill:...the ability to process language as communication.....it underlies all language use (Widdowson; ibid)

Appropriate sampling therefore implies prior identification of the functional needs of teachers in the English-medium classroom in respect of these skills/abilities.

5.2 Teacher Needs in The English-Medium Classroom

It is argued here that as training in, and assessment of, the needs of teachers, including the communicative needs, is a basic premise of teacher training - to teach is nothing if not to communicate - extra-curricular endorsement in respect of course medium (whether English or Afrikaans) is superfluous.

Current (external) requirements for internal interpretation and administration of State regulations for language endorsement confuses the issue of responsibility for teacher training, and admits contradictory practice in that students may be found competent in terms of course requirements, but fail assessment for E-endorsement, which in turn potentially devalues the teacher's diploma itself.

On the grounds that this situation is untenable, that the diploma course is in effect based on an analysis of academic and professional needs, and that training centres are ultimately as responsible for the language needs of trainees *in the medium of the diploma* as for any other teaching needs, test design here is specifically concerned with assessment for E-endorsement of EL2 testees in our English-medium In-service programme.

(Assessment for e-endorsement has never been requested during this study, which suggests that it is of little practical significance for career purposes. In terms of this study it is also irrelevant.)

Arising from new perceptions in language teaching and testing, the Special Purpose approach to (teaching and) testing is based on an explicit analysis of learner/testee needs. Our In-service Course is by definition an SP course, and as such is informed by identified teacher needs in the English-medium class. Course content is however pre-determined in broad outline only, taking its final form each year in response to students' own changing perceptions of their needs as elicited by the tutor. Almost inevitably,

although it is not necessary (from the point of view of testing theory) for teaching and testing to be isomorphic (Davies;1985:4)

course exit tests are substantially isomorphic with course content.

Their own perceptions of their needs (on entrance to the In-service Course), as summarised below, was thus the point of departure for nearly 100 practising teachers in the five successive years up to and including 1990. The great majority of these were Afrikaans speaking, only six being first language speakers of a (Black) African language, which in itself constitutes a severe limitation on the applicability of this study to the E-endorsement "target

market" as a whole. However, even if the in-service groups had been more fully representative in composition, the collective sample would in any case still be too small to support a definitive statement on teacher needs and on how to assess them for endorsement purpose. This project is offered merely as a pilot study for further research in the area of relevant criteria for assessment vis-a-vis the classroom needs of EL2 teachers.

5.2.1 Classroom "Needs" - As Testees Perceive Them on Course Entrance

As a first step towards formulating relevant - and practicable - criteria, students were asked to analyse their own needs for E-endorsement as they themselves perceived them on course entrance. It was assumed that such perceptions would be influenced by previous unsuccessful experience of E-assessment.

"Needs" commonly identified, either as written "shopping lists" or elicited in group discussion (groups never exceeding eight in number) indicated that (initially) assessment was perceived to be more concerned with structural than with functional aspects of language.

"Grammar", tense (especially the past tenses), prepositions, and pronouns were consistently put forward as "problem"

areas; and concern with spelling and punctuation, and the techniques of essay and letter writing, also had more to do with structural conventions of written discourse, than with the communicative objectives of spoken and written English which such conventions serve.

Although not directly related to this study, it is of some significance here that on course entrance (before starting the Language Across the Curriculum module) a number of EL1 students in the 1990 HDE group also saw teacher language needs in terms of "proper" language use - not without reason, it would seem!. The task, and the responses to it on which these observations are based, follow.

5.2.1.1 Task on Entrance to HDE Course, 1990

Immediately before their first lecture in curriculum theory, the 1990 HDE class were asked to spend 15 - 20 minutes on a response to the following task:

IT MAY BE SAID THAT "EVERY TEACHER IS A LANGUAGE TEACHER"
In what ways, if any, would you interpret your role as a classroom teacher in terms of this statement? Write ONE paragraph of about 10 - 15 lines on this question, paying special attention to the responsibilities all teachers have to ensure that their pupils use language effectively for learning and communication.

Students were advised that this was not a test, but a request for information for research purposes only. They

were not required to identify themselves, but were asked to indicate their home language(s), and the medium(s) through which they had been educated at primary, secondary and tertiary levels.

210 responses were received from English, Afrikaans and Black language speakers, all but 25 of whom had previous tertiary level education in English only; ^{of these} and ~~all~~ but five ~~of whom~~ indicated that they had had no previous English-medium tertiary experience at all; ^{had} (20 had dual medium tertiary education).

All but one response demonstrated some degree of understanding of the concept. But if, for the most part, no particular effort was required to follow student "interpretations" of the dictum, the incidence of structural and lexical inaccuracies again reflected the experience of previous years. A brief survey of responses highlights some of the more common structural errors encountered during this study as a whole:

- * dangling modifiers (about 50% of responses received)
- * inaccurate use of determiners (about 25%)
- * uncertain concord (about 25%)
- * lexical inaccuracies (about 20%)
- * inappropriately mixed tense sequences (about 15%)
- * incorrect use of cohesive devices (about 15%)

A few extracts, admittedly out of context, from randomly selected responses, serve to illustrate some such errors in sequences received:

"Every teacher should be able to use *his/her* home language to the best of *their* ability. If teachers are forced to have a command over their language, it would encourage the pupils to *do so* as well". (English medium at all levels of education; Method Courses - Geography/Psychology)

(*his/her/their*: inconsistent use of possessive pronouns;
to have command/to do so: uncertain reference)

I would not be one to advocate the 'Queens English' when teaching because I believe that a certain amount of teaching should be *entertainment requiring the use of* humour, comparisons and *the everyday happenings* of the day. (English medium at all levels; Class Music/Music Practical)

(*entertainment requiring... the use of(?)... the everyday happenings...*: lexical and syntactic inaccuracy)

If every teacher is supposedly a language teacher what happens to *the other academies*? Although language is used in other academies e.g. history geography science etc it does not mean that *their major tasks* is defined or exclusively related to language....If one considers that every discipline has a language of its own then this statement *reflectes the general concept of other disciplines*. It *thus* remains an open question of whether every teacher is a language teacher. (English medium at all levels; History/Biblical Studies)

(*teacher=academy?*; *their/is* - uncertain concord;
reflectes...disciplines:lexical inaccuracy;
It thus...: incorrect use of cohesive device)

The use of language in any classroom situation helps the students understand the subject being taught. The way in which the language is used would determine whether the children understand the subject sufficiently. (English medium at all levels; Guidance/English)

(the use...helps...:the correct use?)

understand sufficiently - sufficiently for what? :

- semantic confusion to the point of incoherence)

Teachers using improper language constantly is extremely disruptive to the pupils perception of the subject being taught to him. It could even make the work hard to understand and *could lead to that the pupil* thinks it is or may be his fault that he cannot do the work being taught.

(English medium at all levels; Mathematics/Physical Science)

(could lead to that the pupil...:uncertain cohesion)

(Response commences) Learning must in the end incorporate some means of conveying such information. Different people however may be better at conveying information than others. However this does not mean that they are more intelligent or that they can learn more. In our society however in order for learning to occur and for evaluation of this to be possible, we rely very heavily on the language part of this. (English medium at all levels; Counselling/ Geography)

(conveying such information/evaluation of this/the language part of this: inaccurate use of determiners)

Standard English is the sought after desire that should be on every pupil's mind.....We have to be on our guard against TV (especially US sitcoms) which debases English. We have a duty to speak proper English. (English medium at all levels: English/Geography)

(is...should be...:confused phraseology/tense sequences))

All teachers should be able to communicate properly to carry the lesson across. The language used should be pure and decent at all times... (English medium at all levels; Mathematics/Physical Science)

(to carry the lesson across: lexical inaccuracy/influence of Afrikaans construction?)

If the teacher has *mispelt* a word..... (English medium at all levels; English/History)

(*mispelt*!)

All teachers would have to be capable of teaching fluently in the specific language. (English medium at all levels; History/Guidance)

(teaching fluently?)

the specific language: unclear reference)

As a teacher it is important that proper language be used.

(Dual medium; Technical Subjects/Mathematics)

(As a teacher it is...: dangling modifier)

Language is essential to distinguish between the objective of a given question. English medium at all levels; Mathematics/Physical Science)

(Language is essential to distinguish...: cryptic to the point of being meaningless;

between the objective: objective(s) - a "mistake"?)

The usage of language is not merely for communication or wooing women, but it is also a means of socialisation..... It is with this in mind that I suggest that language should (be) depoliticised in all its contexts. (English medium at all levels; English/Geography)

(the usage of language is not merely for...: tautological)

There was no immediately discernible pattern in EL1 errors in responses to the above task, or others analysed during the course of this project. The high frequency of deviation from standard usage does however suggest that EL1 errors were not merely "slips of the pen" either. Pit Corder's study, *Error Analysis and Interlanguage*, although specifically concerned with the L2 learner, offers a useful perspective also on L1 errors here, in terms of the distinction he makes between systematic *errors* and non-systematic *mistakes*:

It will be useful..... to refer to (non-systematic) errors of performance as *mistakes*, reserving the term *error* to refer to the systematic errors of the learner from which we are able to reconstruct his..... *transitional competence*. (1981:10)

Mistakes, he says,

are of no significance to the process of language learning (ibid)

To what extent (if any) the writers of the above extracts are "learners" in that ("standard") English is not in fact their "home language" (although given as such), is not known. It is however significant that all but one (who was educated in a dual medium environment) were educated at all previous levels (primary, secondary and tertiary) through the medium of English. Is there perhaps another, "intermediate", category to account for errors in a multilingual environment, where cross-linguistic influences cannot be explained in terms of mother-tongue interference? A category, for instance, for *non-systematic errors* in

(South African) English not (negatively) perceived as *mistakes* by EL1 users; errors which are "tolerated" to the extent that meaning remains substantially negotiable?

In contrast with the random selection of scripts above, I include here two selected for the point they make:

No, I don't think it is my role as a Maths and Biology teacher to teach English as I don't expect them to teach my subjects. I think it is important that my English is *good enough* for me to express myself and so that they can understand me well *but that is enough*.

(English medium at all levels;
Mathematics/Biology)

Perhaps it should be compulsory for non-language teachers to do a *communication/language* course. (English medium at all levels; Biology/Geography)

The point is that language is perceived here as a means to an extra-linguistic end - for communication - rather than as an end in itself. And the fact that, implicitly or explicitly, much the same point is made in many of the "non-proficient" random responses, reinforces the view that although errors such as those reflected here do not in themselves appear to be systematic, there may be "systematic (or *symptomatic*?) tolerance" of performance errors in South African English today. Those students who would promote "proper" English (above), though themselves apparently uncertain of the concept, would seem in themselves to testify to current high error tolerance in South African English.

Jubron

A proposition that high tolerance of non-systematic errors in South African English may be indicative of, and conducive to, the emergence of a "new standard" variety is examined in Chapter Eight.

But whatever the reasons for less than what is traditionally perceived as proficient use of English at tertiary level - and given current problems in South African education, there can be little doubt that valid reasons do exist - language proficiency (or "standard" usage) is still widely perceived as an objective of (language) education. Without teachers at primary and secondary level who are themselves proficient (as suggested by student writing at tertiary/postgraduate level), it would seem to be an unrealistic one. It may be that a "new" South Africa sees the (official) status of English maintained, and teacher proficiency in English may in the future become a major area of research in teacher training. This too is discussed in Chapter Eight.

spell it out!

In the meantime, however, the pressing needs of education in general suggest that there is little to be gained by unrealistic language regulations, difficult to interpret by those required to administer them, and discouraging to teachers not proficient in English, perhaps, but demonstrably competent "to teach English-medium classes"

- and who not uncommonly, in my experience, perceive the "bilingual certificate" as an almost insuperable obstacle to the fulfilment of their professional objectives, with accompanying demotivation and, on occasion, as a contributory reason for dropping out of the profession altogether.

5.2.2 Formulation of Test Criteria for In-Service Assessment

That prospective testees - whether HDE students or practising teachers - commonly see their "needs" in terms of "grammar", as noted above, in itself suggests that even after twenty years of "communicative competence" in applied linguistics and the language classroom, (functional) language ability is still more readily seen from the point of view of form, than from that of content.

The In-service course at UCT has, in the last four years, attempted to redress the situation; and while not disregarding the significance of the relationship between syntax (grammar) and meaning, course work has overtly promoted an integrative interpretation of language ability, with the emphasis on "effective" rather than "good" English.

([?] pedantically "correct" ?

Based on the rationale that all students were language users and therefore possessed some (subconscious) knowledge of how language "works", it was assumed that heightened awareness of such pre-existing knowledge would promote more perceptive self-assessment or, in Krashen's terms, more effective "monitoring" :

Successful Monitor users edit their second language output when it does not interfere with communication....(1981:12)

The stimulation of self-assessment, together with the exploitation of pre-existing ability in English (as this is not a threshold course, some basic ability for use could be taken for granted) were thus seen as tutorial priorities; and oral and written tasks were elicited from students for this purpose. On the other hand, where

monitor users show(ed) overt concern with "correct" language, and regard(ed) their unmonitored speech and writing as "careless", (ibid)

"monitor overuse" was actively discouraged in the interests of maintained discourse.

There was thus no formal pre-determined syllabus other than an outline of practical assignments, and tutorial activities centered mostly around themes of general, rather than overtly "educational", interest.

Oral contributions elicited from students included "histories of place-names" ("places" selected by the students themselves and researched at home), book and film

reviews (students were required to read at least two full length novels by recommended authors during the course), motivated itineraries for hypothetical visitors to the city or country (such visitors representing students' own interests), and a sample lesson from students' own classrooms (which recently included a "lesson" in bricklaying!). Group discussion was encouraged after every presentation, and other oral group activities included debates, collective essay composition, dictation by students to students, group evaluation of anonymous written assignments (that is, assignments from parallel groups), Cloze and "Cloze/elide" tests (the latter requiring deletion of extraneous words included in the text), group exploration of the techniques of summary composition, traditional comprehension tests and multiple-choice formats, and correction of selected cross-word puzzles completed at home. Perceived problems of a "grammatical" nature were dealt with within the context of oral and written discourse, and paradigmatically only on request.

Written homework assignments, usually two a week, included essays, summary and comprehension exercises, formal and informal letters, and on occasion, also exercises reinforcing structural conventions which had arisen in tutorial sessions - for example, direct and indirect speech, punctuation, word order, concord. Predictably, communication "breakdown" occurred more readily in written

discourse than in spoken English. The oral medium, characterised by ready verbal and visual feedback, lends itself to "instant" co-operative problem-solving when breakdown threatens, and it was interesting that as students became more aware of their own potential to communicate despite syntactic problems, self-correction in the oral mode (of both random and habitual errors) increased significantly. To some extent, increasing confidence was noted also in the written mode, as reflected both in a voluntary increase in length of written assignments, and in terms of more creative discourse content. *but what about quality of language?*

In-service course work was substantially assessed more in terms of meaning organisation and coherence, than syntactic organisation and cohesion. The latter were not however ignored. On the contrary, students were made aware that without intra-textual order, or *macro*-cohesion, meaning cannot be "organised". On the other hand, where there was no attendant communication "breakdown", *micro*-cohesion as grammatical ^{require} skill for its own sake was not perceived as necessarily relevant in an English-medium (as opposed to English subject) classroom.

These perspectives on "teacher language use" also informed course exit tests, and are further examined in Chapter Seven.

An analysis of "teacher needs" as developed in co-operation with In-service students during successive courses, and on which the sample test below is based, may be summarised as follows:

? For the purpose of general classroom work
 "Teachers need to"

- * *extract relevant material from source texts* (e.g. reference work; summary technique; comprehension exercises; book and film reviews)
- * *select and organise material* (summary technique; comprehension tests; debates; essays; book and film reviews)
- * *explain and question* ("lesson" presentation; critical assessment of peer tasks; post-presentation group discussions)
- * *simplify concepts* ("lesson" presentation)
- * *communicate beyond the classroom* (the tutorial format provided ongoing opportunities for this)

Given that no single test can account for all "teacher needs in English-medium classes", items in the prototype test below were devised to "sample" testee ability (in respect of *reading, writing, speaking, listening and interpreting*, as suggested by Widdowson, 5.1) in the context of a representative selection of "needs" as summarised above. The objective was to elicit evidence on the basis of which testee ability to meet such needs in real life might be predicted.

What claim, if any, can such an approach make with regard to basic concepts in testing?

5.3 Basic Concepts in Language Testing

In a comprehensive study under this title (in *Testing and Experimental Methods*; Allen and Davies:1977), Elisabeth Ingram highlights three pairs of concepts which, historically and experimentally, occupy key positions in language testing research: subjectivity versus objectivity; validity and reliability; and norm-referenced compared with criterion-referenced tests.

5.3.1 Subjectivity versus Objectivity

She notes as follows:

About the turn of the century, a new form of assessment was developed - the objective test. One of the reasons for this was to avoid.....disagreement between examiners. Instead of having to write shorter or longer essays, the learners were asked a series of quick questions, each having only one correct answer, or a set of previously agreed acceptable answers. (1977:11,12)

The implication here is that "objective" testing is defined in terms of closed questions with pre-determined answers - which suggests a need to consider what it is that one is testing. I have suggested that in this study we are concerned with "competence to communicate" (in EL2).

Formulated as a theoretical concept in a research environment, the notion of communicative competence as a criterion in language tests poses a serious challenge to traditional attitudes; for assessment on the basis of real world inter-personal language use, with all its strategies for effective communication, ideally implies a participating tester in a purposeful, open-ended, communicative situation. The more closely the test approximates this ideal (or the more "direct" the test), the more (subjectively) involved the tester is likely to be. How significant is this proposition of the "involved tester" in terms of the concept of measurement that testing implies?

As Allen, among others, has pointed out, there is no absolute dichotomy between objective and subjective assessment, for

although it is possible to have objective scoring, *there is no such thing as an objective test.....*All decisions surrounding the design and administration of a test are subjective.....(1985:32) (My italics)

Less directly perhaps, Ingram's historical perspective on the "objective test" makes the same point: she records that "disagreement between examiners" focused on how different examiners - or even the same examiner at different times - perceived the same essay (and not on the form of the test as such).

"Subjectivity" is therefore integral to testing, in the sense that the *what* (of all the constituents and aspects of a known or an unknown syllabus), and the *how* (of assessment techniques), will always be decisions based on more or less subjective tester perceptions. On the basis of this argument, Allen feels that criticisms of subjective tests are "not.....valid". (ibid)

This study suggests a still sharper focus than Allen's, in the proposition that it is not so much subjectivity per se, but the potentially idiosyncratic element that subjectivity implies, that gives rise to (invalid) criticisms of the "subjective test".

"Valid criticisms" in the form of editorial scrutiny of proposed test items on the one hand, and the moderator function on the other, are recognized counter-balances not, I argue, to subjectivity as such, but rather to idiosyncratic tester perceptions (for based on the principle that "two heads are better than one", collective responsibility for testing increases, rather than decreases, the complexity of the subjective element by adding to its dimensions).

The difference I see between subjectivity and idiosyncrasy was implied, for example, when in a course on the phases of test construction at UCT in 1986, Brendon Carroll (~~whose~~

recommended "de-bugging" of draft items by means of editorial scrutiny by six or more people "with red pencils". "Red pencils" in itself suggests keen active engagement of the (subjective) critical faculties. In this context, multiple "red pencils" represent multiple critical views aimed at rationalising test items, ideally by consensus; the higher the level of concord between "red pencils", the more likely it is that dissent is due to idiosyncratic views.

From this perspective, it is clear that test objectives ultimately depend on testers for interpretation. Where tests (focused on usage/proficiency) aim to elicit "correct" responses, interpretation of objectives is contained by the notion of "correctness" itself. In contrast, where tests aim to assess communicative competence, interpretation of test objectives raises the less clearly defined issue of "appropriateness". Readily influenced by subjective/ idiosyncratic notions, the question of appropriacy in real-world terms is complicated, in the present study, by the nature of the multilingual target group for which the proposed test is intended.

Given present frontiers of research on the nature of language, and about how it functions in society, it would thus seem that commitment to communicative competence as a criterion in language testing implies a concomitant acceptance of the "subjective" tester. At the same time, it

is a basic concept in testing that the subjective/ idiosyncratic element be accounted for, as far as it is possible to do so, at all stages of the language test - in design, in construction and - particularly in the case of a holistic approach as promoted here - in scoring (as discussed in Chapter Seven). Collective responsibility in the context of a standardised testing programme would seem to be a rational approach to problems arising from the conflicting perspectives of "objective"/"subjective" testers - and to the present "mine-field" of assessment for language endorsement (as it was characterised recently in a private communication with a responsible authority of the Cape Education Department).

Oh, not really.

5.3.2 Validity and Reliability

According to Ingram, a test is valid if it tests what it says it tests, and nothing else; and its reliability is high

when variations in the readings taken represent true differences between the individuals who are being tested. Any other variation represents error. (op cit;1977:15)

In the same publication, Davies argues that although validity is far more important than reliability, which is in fact rather an obvious concept, it is essential to establish reliability first. Otherwise there is no point in considering validity. (1977:57)

5.3.2.1 Reliability

Ingram identifies two contrasting sources of error resulting in reduced test reliability: *extrinsic* sources, namely tester variability, and variability of testing conditions; and *intrinsic* sources, which include errors due to lack of test *stability*, and errors due to lack of test *equivalence*.

Error arising from extrinsic sources, she says,
can be greatly reduced by common sense precautions

(1977:15)

And "common sense" initiatives to ensure inter-marker reliability, and consistent test conditions in general, are arguably implicit in the notion of standardisation. Error arising from intrinsic sources is less summarily dealt with.

In brief, *stability reliability*, also referred to as "test-retest" reliability, or repeatability, is estimated by correlating the results of repeated administrations of the same test to the same group of testees; a test is stable, or has high repeatability, if correlation is high; and *equivalence reliability* refers to the correlation between scores on items or tests designed to measure the same "given content" - if items in test A and test B are designed to measure the same testee characteristics, test A and test B should yield

equivalent results. This is the theoretical notion underlying the item bank characteristic of the standardised test programme: test items are pooled according to the characteristics they are intended to measure; items drawn from the same pools for successive tests should yield equivalent results. Experimental methods for obtaining estimates of equivalence reliability include parallel forms, split half, and variance estimates.

Given the limitations of the present study in respect of sample size, and more significantly, in respect of composition, in comparison with the target group as a whole, it was not felt that formal research in the area of test reliability was appropriate, nor would it yield useful results, and none was undertaken. Informal test-retest situations did however indicate that uncertain stability reliability was a contributory cause of perceived inconsistencies in our own tests, as noted in Chapter Two, which in itself would seem to suggest a need for formal research in this area.

On the principle that the larger the research group (and consequently the more dispersed behavioural factors such as fatigue, erratic performance, and poor memory), and the more representative it is of the target group as a whole, the more reliable the results, there would seem to be little

alternative to an appropriate dedicated research programme. Ultimately, meaningful relative ordering and distance between individual scores in the target group as a whole, depends on meaningful research which in turn implies an ongoing programme of standardised tests across administration centres, including handbooks of instructions for (trained) testers and markers to co-ordinate testing procedure in the various centres, and facilities for data collection and processing (for example, conversion of collected raw scores to a standard score scale such as stanines or percentile ranks).

At the same time, it is relevant to note here that reliability in Ingram's term - *reliability is high if differences in results reflect true differences between testees* - although theoretically sound, does not account for the existence of significant "true differences" in South African varieties of English which, although perhaps perceived as "differences" by interlocutors, are not experienced as obstacles in real-world communication. Arguably, then, tests with high reliability in the sense that they are sensitive to "true differences" in basic communicative patterns of indigenous varieties of English, would in practice discriminate against teachers who may be competent (at least in terms of this study) to teach English-medium classes, but whose English does not conform with standard South African English (if such a variety can

indeed be identified). There would seem to be a fundamental contradiction in terms here, suggesting a research initiative also in this area.

In fact, there is a growing body of research closely related to issues raised by the notion of communicative competence in the South African context - *Language Acquisition or Language Socialization? The Case for Ethnographic Input To An ESL Course For Undergraduates* (Chick;1990:302ff) and *ESL Issues Arising from the "Teach-Test-Teach" Programme* (Bradbury, Damerell, Jackson and Searle, in Chick; 1990:325ff) are but two examples of recently published studies. However, in the absence of a coordinated programme directed towards a common goal, potentially valuable insights in the relevant area of applied linguistics remain largely abstract in nature, deliberated at conferences but lost to the tester and the teacher responsible for ongoing education "in the classroom".

5.3.2.2 Validity

As noted in Chapter One, E-endorsement implies competence (identified here as communicative competence) to teach English medium classes. It follows that assessment for endorsement is valid to the extent that this is what it tests. The problem is that there is as yet no definitive

approach to the practical assessment of communicative competence.

According to Hymes, language ability should be viewed holistically:

in speaking of (communicative) competence, it is especially important not to separate cognitive from affective and volitive factors, so far as the impact of theory on educational practice is concerned. (1972; p. 283).

And Ingram (1977) found that in any case

the maxim of one feature at a time (in language testing) is usually frustrated (for) it is virtually impossible to keep (the skills of production and reception) separate in practice.

In apparent contrast with these views, the prototype needs analysis approach to teaching and testing, as put forward by Munby (1978) and developed by Carroll (1980), identifies a potential 70 or more personal, professional and linguistic factors of communicative competence, each to be separately accounted for in respect of individual testees in any given special purpose context.

The principle on which this approach rests - that an analysis of communicator needs within a defined context provides a rational basis for the selection and organisation of "communicative" test material - is central to this study. At the same time, profiling of individual testees as suggested by Munby and Carroll would be too time-consuming in a testee group the (collective) magnitude of the target

group proposed here, and it would in any case render too fragmented an analysis of testee needs to be of practical value for present purposes.

The attempt in this study to reconcile the notion of test validity with the "communicative" test should therefore be seen against the background of the following constraints:

- * a criterion of communicative competence in language assessment implies a holistic/integrative view of, and response to, elicited samples of language-in-use;
- * which in turn implies a "participating tester" in a "real-world" communicative event.
- * Participant profiles (ideally including that of the tester), although of theoretical relevance, are of practical significance only in certain limited test situations; profiling is not a feasible proposition in large-scale testing programmes.

Accordingly, the approach to test design - in respect of *content validity* - I would put forward here, is essentially an eclectic one: conceptually closer to Hymes and Ingram, in the sense that it takes an holistic (and overtly "subjective") view of language-in-use, and it draws on the "special purpose" approach for the concept of the needs analysis, and for the principle of exhaustive research as the basis on which such an analysis is undertaken.

But, as Davies points out, content validity is only one kind of validity associated with language tests. (1977:58). *Face* validity - "does the test look right?" - although not trivial, is arguably more a matter for "public" debate than for formal research, and is not dealt with in this study. *Construct* validity, on the other hand, is clearly relevant here, in that it is concerned with the theory on which test content is based. This study proposes tests based on a theory of language ability as communicative competence; the sample test put forward in Chapter Six thus possesses construct validity to the extent that it can be shown to be so based, as discussed in Chapter Seven. And finally, *concurrent* validity and *predictive* validity are established by statistical procedure (and were not germane to the present study).

In a sense, of course, all tests are intended, at least in part, for predictive purposes. E-endorsement is entirely concerned with predicting teacher performance (in the English-medium classroom); and the validity, in this more generic sense, of the sample test put forward in Chapter Six, could arguably be established in post-assessment classroom research on successful testees. In view of the still partly elusive nature of a criterion of communicative competence, it is however doubtful if predictive validity in the more scientific sense - as a correlation between test scores and a *quantifiable* criterion - would be possible.

Any claim to validity, in statistical terms, of the test put forward below, would therefore necessarily be a tenuous one. This does not per se invalidate the test. For given the variables integral to all enquiry in the Human Sciences, even statistical evidence is relative in nature. I argue, therefore, that as the product of observations and initiatives recorded in this study, the prototype test proposed in Chapter Six is potentially more valid than the tests analysed in Chapter Two, and as such represents a basis for further research.

With reference to the concept of test reliability, I suggest that the only relevant "true difference" to be accounted for in assessment (for E-endorsement) is the difference between competence to communicate in the English medium classroom, and lack of it, an argument which falls substantially within the field defined by the third of the pairs of concepts under discussion, namely norm- and criterion referenced tests.

5.3.3 Norm-referenced and Criterion-referenced Tests

It is doubtful whether norm-referencing is ever absent from the test situation. Consciously or subconsciously, testers probably always compare testee performances and adjust

assessment accordingly; and it is an integral function of the standardising process to establish scoring norms based on feedback research. Equally, test results would be meaningless if unrelated to pre-determined criteria.

But although norm- and criterion-referencing in scoring are not mutually exclusive activities, the relative emphasis accorded each in any given situation accounts for significant differences in the interpretation of test results.

This study argues that in the case of assessment for E-endorsement, the perceived competence, or otherwise, of the individual testee ("to teach English-medium classes") is of greater relevance in the scoring process than is his or her competence in relation to the testee group to which he or she belongs. This in turn suggests a theoretical bias in favour of criterion-referenced tests.

CHAPTER SIX

THE SAMPLE TEST, AND ANALYSIS OF TEST ITEMS

6.1 The Sample Test

6.1.1 Target Group

Designed within the context of a research project focused on a wider target group, the sample test below was constructed to serve the purposes of current regulations for E-endorsement in respect of the twenty-two practising teachers in the 1990 In-service group at UCT, and eight candidates registered for "examination only".

Teaching practice and written assignments, as core components of teacher training during the diploma course, cater for the communicative needs of the teacher as subsumed in the categories above, in actual real-world terms. For practical reasons, assessment for endorsement purposes cannot repeat diploma procedure in a second language. Nor is there any reason for it to do so, for as a professional rather than an academic requirement, E-endorsement implies *de facto* communicative competence in the medium of the diploma to be so endorsed. Test design here acknowledges the integrity of the diploma; it is not intended as a test

of teaching skills as such (as catered for in the diploma course), and neither oral nor written tests attempt to simulate classroom situations. Rather, the test presented here is intended to elicit evidence of testee language ability on the basis of which his/her classroom performance in English may reliably be predicted.

6.1.2 The Written Test

IN-SERVICE TEST OF COMMUNICATIVE COMPETENCE SEPTEMBER 1990

SECTION A

COMPREHENSION TEST (Recommended time allocation: 35 mins)
Read this passage carefully, then answer the questions that follow.

ART IN MUSIC

More than a quarter of a millenium has passed since the death of Antonio Stradivari in 1737.

Stradivari, the man who developed the craft of violin-making into an art form, was born in Cremona, Italy, in 1644. By the age of twenty-two, while still a pupil of master violin-maker, Niccolo Amati, he was already producing instruments in his own name. He continued to produce, either personally or through the craftsmen under his guidance, until his death at the venerable age of 93.

There were two major differences between Stradivari's instruments and those of his contemporaries.

Firstly, he introduced certain innovations in the form of the violin; and the shape of the neck and the f-shaped sound-holes characteristic of his style have yet to be significantly improved upon.

Secondly, there was his method of varnishing. Although perhaps not generally appreciated, the varnishing of an instrument is far more than purely cosmetic; it is the final touch that dictates the quality of the work. No instrument of the violin family reaches the full potential of its sound until the oil base of the varnish has stabilised. This takes about two years.

The structure and form of a Stradivarius violin may be studied and reproduced, but his varnish formula remains a mystery; and the vindication of Stradivari's genius is that a genuine "Strad" can fetch up to two million dollars (more than five million Rand at current exchange rates) on auction!

Brian Lissus, a professional violin-maker based in the small Cape village of Robertson, has been trying to attain a similar degree of perfection in formulating his own varnish. Following the ideas of American aeronautics engineer, W. Fulton, who first suggested that latter-day craftsmen would never succeed with the finishing of a great instrument because modern materials are too refined, Lissus is slowly attempting to formulate a "Strad" finish for his own violins.

Fulton's research established that wood turpentine, the basic ingredient of oil varnish, was transported to Stradivari's workshops by horse-drawn cart over a period of days. The jolting and disturbance of the turpentine led to oxidation, and the end product was a viscous goo that would make any connoisseur of lacquer shudder. Fulton's attempts at simulation led to a liquid that was thick, oily and resinous - ideal for making violin varnish.

So, following the Fulton trick, Brian Lissus took to making his own varnish - horse-drawn carriage excluded - leaving it in the sun for two months. This resulted in a loss of one-third of the volume and a turpentine of wonderfully thick constitution.

Lissus obtained some of the wood he uses for his violins from his old teacher in England. The century-old tree it came from was felled 150 years ago - a tree that had come to life as Stradivari was nearing the end of his. In the hands of Brian Lissus, these perfect timbers are being brought to a rich destiny.

(Adapted from an article by Sebastian Balic, June 1989)

QUESTIONS

PART ONE

In this part of Section A you are required to tick the most appropriate response in each case.

(2 marks for each correct answer;

NOTE: marks will not be deducted for incorrect answers)

1. A quarter of a millenium is
 - a) 400 years
 - b) more than 200 years
 - c) 250 years
 - d) about 150 years
 - e) 25 years

2. The innovations Stradivari introduced in the form of the violin

- a) have since been greatly improved upon
- b) were not of great significance
- c) didn't make much difference to the sound
- d) are still characteristic features of violins today
- e) have disappeared from the violin-makers craft

3. In the context of this article, "latter-day craftsmen" refers to

- a) modern violin-makers
- b) all craftsmen today
- c) all craftsmen since Stradivari's death
- d) violin-makers in Italy today
- e) craftsmen of the future

4. The phrase a "viscous goo" as used here means

- a) a deadly poison
- b) a gluey liquid
- c) hi-tech wood turpentine
- d) varnish used in aeronautic engineering
- e) varnish used for horse-drawn carriages

5. The "perfect timbers" Brian Lisus uses to make his violins came from a tree

- a) planted 150 years ago
- b) planted 100 years ago
- c) which was 150 years old when it was chopped down
- d) which was chopped down 150 years ago
- e) which was planted when Stradivari was born

PART TWO

In this part of section A you are required to answer the questions in your own words.

1. Did Stradivari work alone? Motivate your answer with reference to the text. (3 marks)
2. What is the meaning of "purely cosmetic"? (2 marks)
3. Give three special features which distinguish a Stradivarius violin. (4 marks)
4. Rewrite in your own words: "a genuine 'Strad' can fetch up to \$2 000 000". (3 marks)
5. What light did Fulton's research throw on the "mystery" of Stradivari's varnish formula? (4 marks)
7. Where did Brian Lisus study, what did he study, and what does he do now? (4 marks)

8. What happened when Lisus left his varnish in the sun for two months? (2 marks)

9. Suggest what the writer means when he says that "in the hands of Brian Lisus these perfect timbers are being brought to a rich destiny"? (3 marks)

(Total for Section A: 35 marks)

SECTION B

SUMMARY (Recommended time allocation: 35 minutes)

Read the following passage carefully, and then summarise it in 150-200 words in the space provided below.

SIDEWALK BOOM

Hawkers and pedlars are a common sight these days. You see them on Church Square in Pretoria, outside the Golden Acre and on Greenmarket Square in Cape Town, in the Floreat Mall in Bloemfontein, in Hoek Street and in front of the Stock Exchange in Johannesburg, in Mamelodi, Bekkersdal, Crossroads, Khayelitsha and Ixopo, at stations and bus stops and on street corners.

They peddle fruit, vegetables and flowers, shine shoes, sell soap, toothpaste and cigarettes, wash motor cars, wash, cut and perm hair, make clothes, knit jerseys, take photographs, throw the bones and provide herbs and remedies. They come from all walks of life and represent all groups. Black, White and Brown alike grab every opportunity to keep the wolf from the door or to earn a welcome extra income.

One thing they do not lack is resourcefulness. They sell commodities in streets, from garages, livingrooms, backyards, market stalls, converted motor car wrecks and the back of light trucks. They convey bread, milk, beverages and food on foot and bicycle and by truck and motor car.

Unlike other retailers, they are not bound by time, distance, premises or licences, usually do not pay rent and taxes and are, in most cases, unpopular with the owners of formal retail outlets.

Although this informal sector or so-called "shadow economy" has existed since the Devil talked Eve into taking the apple, it has only been in the past eight years that we in South Africa have really become aware of this phenomenon. Whatever our feelings about these self-made entrepreneurs, they will not disappear, for they have always been and will always be with us.

What we see is, in fact, only the tip of the iceberg. To determine the exact extent of the informal sector in South Africa is impossible. It is therefore with reason that the outstanding characteristic of this part of the economy is that it is "unrecorded".

Figures in estimates and surveys on the informal sector in South Africa vary widely - like elsewhere in the world - putting its contribution to South Africa's gross domestic product anywhere between 5 and 25 per cent.

But although the "shadow economy" is nothing new in South Africa, the country is only now gradually awakening to its significance. Besides being unrecorded and providing self-employment for people who were previously unemployed, the informal sector is also characterised by the resourcefulness of the entrepreneurs and the small scale of the operations. Large financial inputs and skilled labour are not required, activities in the informal sector are labour-intensive, and at least the most pressing daily needs of the people involved are catered for.

According to Dr Ben Vosloo, managing director of the Small Business Development Corporation, the informal sector is not a substitute for the formal sector, but rather a gateway or access route to it, introducing its participants to the world of free enterprise within an economy with the almost impossible task of reconciling first and third world economic structures.

Mr Lawrence Mavundla, founder and president of the African Council for Hawkers and Informal Businesses, believes the informal sector will become the bedrock of the South African economy in five years' time. According to him the informal sector provides people with opportunities to upgrade themselves. "When people create their own employment, they are constructive and not destructive. Self-sufficiency promotes self-respect, and they start to realise the need for education", he says.

(Adapted from an article by Gert Coetzee, January, 1990)

(Total for Section B: 30 marks)

SECTION C

Composition (recommended time allocation: 40 mins)

In the space provided below, write about 200 words on one of the following:

The role of the teacher outside the classroom

OR

To teach is to communicate - verbally, with body language and with visual aids

OR

Advantages and disadvantages of the chalkboard, and of the overhead projector, as aids to classroom teaching

(Total for Section C: 25 marks)

SECTION D

CLOZE TEST (Recommended time allocation: 10 minutes)

Fill each of the 15 gaps in the following passage with ONE suitable word. The size of the gap is not an indication of the length of the word to be filled on.

NOTE: One mark for each suitable word; maximum score: 10

Unfortunately, as the world's population rises, so.....
 pollution. More waste material has to be disposed.....,
 which causes pollution of the land, sea and.....
 Also, the increasing population demands more and.....
 energy supplies and the production of this.....causes
 some of the worst pollution of..... Until earth's
 natural resources run out, one of.....most difficult
 pollutants to control is oil.year, millions of
 tons of oil are.....into the sea; some deliberately as
 tankers wash.....their tanks, but much of it is
 spilled.....as tankers go aground or collide with
ships. Although oil companies are held responsible
the spillage and have to pay compensation, it is
to the local authorities to clean up the.....

6.1.3 The Oral Test

The written test was followed within a day or two by a short oral test in the form of a 5 - 7 minute interview. Candidates were interviewed individually by a senior lecturer in education actively involved in language teaching. The course tutor was present at interviews, and marked scripts were on hand. Interviews were informal in the sense that there was no prepared agenda, and testees were merely expected to take part in a "real-world" discussion led by the tester on subjects of educational, topical or personal interest.

But tests are only part of the assessment process. As significant as test format and content in respect of test results, and in some ways more so, is the tester's approach to scoring. This point is taken up in 6.6 below.

6.2 Rationale for the above approach to testing

In a table on the "components of proficiency" (subsuming communicative competence) Stern (1983:356) identifies several approaches to language assessment implying rationales ranging from the "relatively abstract" (e.g. Oller) to the "relatively concrete" (e.g. Carroll).

The approach taken here is substantially located in the fourth of Stern's five categories, the "fourfold concept" (of proficiency as communicative competence). "Listening, speaking, reading and writing", and "formal mastery, semantic mastery, communicative capacity and creativity", are two examples given in this category of ways to look at the components of proficiency/communicative competence.

But underlying all the approaches summarised in Stern's table, there is the assumption of a pre-existing level of comprehension relevant to the given components of proficiency - "comprehension" here understood to refer not only to the receptive (usage-oriented) skills of reading which Widdowson assigns to comprehension. In this context, "comprehension" embraces also, and more significantly, the (use-oriented) skills Widdowson associates with interpreting in both the written and spoken modes (1978:64/65).

This study argues that "comprehension" in these terms constitutes an essential component of communicative competence for endorsement purposes, and should be explicitly accounted for, not only in test design, but also, and perhaps more significantly, in scoring (as argued 6.6)

In order to focus on comprehension, it is necessary first to characterise the level appropriate to the objectives of the test. As argued elsewhere in this study, it is illogical to

suggest that comprehension tests are in themselves "too difficult", and that *therefore* they should be discontinued. If the difficulty level of a comprehension test is found to be too high, then it is the given text, and/or the questions set on it, that are inappropriate in that context, not the format of the test.

6.2.1 Accessibility of Text for Test Items

The "difficulty" level of discourse, written and oral, is determined largely by lexical content (which is to some extent dependent on semantic content), and syntactic structure.

For practical purposes, and on the assumption that the "English-medium class" in State schools is drawn from a cross-section of the consumer public catered for by the popular English-medium press in daily newspapers, and in factual articles of general interest in community-oriented magazines such as South African Panorama and Flying Springbok, such publications were considered a realistic source of material on which to base assessment of teacher comprehension level in the written mode. The fact that this assumption is patently unrealistic in the case of many "official" English-medium teachers, particularly in Black schools, has more to do with the irrelevance of the language

requirements under the present system of education as a whole, than with the proposed characterisation of a criterion for teacher comprehension of written English at E-level.

A (relatively realistic) working criterion for teacher comprehension skills having thus been identified, the next consideration was text selection. Newspapers and magazines offer a wide variety of subject matter, little of it overtly "educational" in orientation. But assessment here is "educational" only in the sense that it is administered within the educational context; endorsement refers in the first place to language medium, and is not concerned with professional preparation as such. Consequently, texts were selected for topical content and general interest value rather than for their relevance to "education" (in the narrower, school-oriented, sense of the word).

Culturally sensitive issues which might raise testee anxiety levels (probably already raised by the test situation itself) were avoided, as was subject matter which the testee might perceive to be motivated by objectives other than assessment of his or her competence to use English, or which might be culturally remote from one or another sub-group of the target group as a whole.

6.2.2 Selection of Tasks and Texts for the Sample Test

6.2.2.1

Section A The text for the comprehension test, *Art in Music*, although on first glance perhaps culturally remote from some of the target group, is in fact "remote" only in respect of Italian place and personal names. A non-technical, factual account with some historical interest, the "story-line" pre-supposes no particular interest in music or craftsmanship, and is brief enough to allow full engagement of reading skills in the suggested time allocation for this section of the test.

5 multiple choice questions, although of limited value (as discussed in 5.4.2 below), and 10 "traditional" questions, are designed to elicit evidence not only of comprehension of surface meaning of the text, but also of testee ability to extrapolate information from the text and apply it "extra-textually", as in Part One, questions one and three, and Part Two, questions two, four and nine.

6.2.2.2

Section B The text selected here, "Sidewalk Boom", focuses on what has increasingly become a common sight in daily life, and a common economic topic in both public and "official" consumer circles.

Summary technique demands a more sophisticated implementation of testee language skills than required in Section A. For the purposes of assessment here (given the academic level of the testee), it is assumed that the process of information selection and organisation is a familiar one. This test is designed not so much to elicit evidence of comprehension of the theme of the text as a whole, but rather of testee ability to distinguish the main points in the "story-line" from peripheral detail, and to assemble them in coherent written discourse.

6.2.2.3

Section C...If comprehension is the underlying "skill" of written (and oral, and metalinguistic) communication, communicative competence implies also effective skills of production. For this test of written discourse, testees select one of three topics, all more overtly "educational" in nature than the texts selected for Sections A and B.

Offering testees a choice of subject matter, rather than a single set title, is motivated partly by the proposition that individual preference for writing discourse of a descriptive, imaginative or persuasive nature will significantly affect his or her performance in a composition type essay such as this; and partly to attenuate the "prescriptivism" of the test situation by giving some initiative to the testee him- or herself.

The educational bias of the titles was motivated by similar considerations: it is a reasonable assumption that all the topics are of some relevance to all testees, and therefore offer scope for meaningful "self expression" to all in a way which topics of a more general nature might not do; and the titles are also a reminder of the purpose of the test, and of the teacher's "need" to organise discourse and expound ideas in the (English-medium) classroom.

6.2.2.4

Section D...Cloze tests fall into Stern's "single concept" category of proficiency/communicative competence, as exemplified by "expectancy grammar (Oller)". (1983; p.356) A useful classroom exercise for promoting awareness of the interdependence of structural and semantic aspects of discourse, it is a technique which becomes easier with practice; as a test, it may discriminate against testees unfamiliar with the format. For this reason, the text in Section D above was selected not only for topical theme, but also, and more particularly, for structural and lexical accessibility. 7th word deletion, with some tester discrimination, rendered a Cloze test (of "expectancy grammar") with 15 deletions. In the sample test above, it carries a maximum of ten marks.

Inclusion of a Cloze exercise, motivated by Stern's proposition of a "single concept" perspective on proficiency/communicative competence, adds another dimension to assessment of testee understanding of the principles of discourse coherence and underlying textual cohesion in a relatively short space of time. It can also be seen as a test, on two levels, of appropriacy.

In accordance with the "real-life" orientation of the concept of communicative competence, assessment for E-endorsement is designed to elicit evidence of testee ability relevant to real-life needs as defined above, namely selection and organisation of material; skills of explanation and questioning; concept simplification, and communication beyond the classroom. Given present understanding of the complexities of language use, it would be unrealistic to attempt to cater for all the skills and abilities subsumed in these categories in any one test.

6.3 Focus of Sample Test Items

Test design here has been motivated by the fact that responses elicited by the type of questions included in the sample test have, during successive In-service courses, been found to be useful indicators of teacher ability to read and comprehend English at the identified level, and on the basis

of such comprehension to "activate" existing skills relevant to the selection and organisation of source material for classroom purposes (Sections A and B).

Section C tests skills of "production". This does not imply that the "productive" skills/abilities of written discourse are distinct from comprehension skills, but rather that the focus here is on coherence (and on underlying cohesion, to the extent that negotiation of meaning depends on it); in addition, the suggested topics all involve a degree of "explanation" and/or "concept simplification", as relevant to classroom needs.

As already noted, it has been suggested that the Cloze test in Section D offers a "single concept" view of communicative competence in terms of "expectancy grammar", or testee anticipation of elements of discourse on the basis of contextual information. To the extent that testee results in the cloze test accord with results in the other three sections, the cloze may be seen as a kind of confirmatory index. A longitudinal study would provide more reliable evidence of the value of a cloze test for language assessment, as also for the other tests proposed above.

6.4 Inappropriate Items - A Motivated View

Certain traditional test questions have been avoided. The formal letter of application for a teaching post, a bursary or study leave were considered inappropriate, mainly because in a "real-life" situation, the serious writer of such a letter would be well-advised, in his or her own best interests, to use a dictionary and any other source material necessary. There can be no valid reason why formal letter format should be memorised for test purposes, particularly not in a second language - which raises the issue of the "open-book" examination. On the grounds that in real life dictionaries are readily available, and their use is actively encouraged, it would seem acceptable, in principle at least, that they be permitted in the examination room. The reality of time limits in tests on the one hand, and the likelihood of over-use of "legitimate help" at the expense of time which could be better spent on the other, make it of uncertain value in practice. This too has a bearing on scoring, as proposed below.

And finally, role-play is useful as a teaching device; as a test, it denies the reality of the immediate situation. The test is a "real-world" event; to ask candidates to *imagine* "that you are a standard six teacher....." or "that you are organising a school play....." introduces a dimension that may have little direct relevance to the stated objectives of

the test (ability to teach English-medium classes, not necessarily standard six). In testing imaginative (or cognitive, or even communicative) skills which are not necessarily relevant to the target group as a whole, assessment is in conflict with the notion of test validity. To be valid, a test should test what it says it tests and nothing else. That ultimately this may not be possible is not a reason for confusing objectives by introducing irrelevant elements into test tasks.

6.5 Oral tests: Format and Rationale

Ideally, the oral test of competence to teach English medium classes takes place in the English medium classroom. As a core component of the HDE Course at UCT, the two periods of supervised teaching practice (TP) offer the ideal situation, and assessment of oral communicative competence is an integral aspect of TP 1 and 2.

The number and the different geographic locations of teachers enrolling for in-service language assessment (30 in the E-group in 1990) make it an unrealistic proposition to assess in-service testees in this way.

The private interview format used in this study (individual testees interviewed for 5 to 7 minutes by a senior lecturer

in Education, external to the course) has several advantages over the "group interaction" test used elsewhere. With a group of testees, it is a function of the tester to involve all participants in group discussion, which may lessen tester receptivity to individual contributions; the tester in the interview situation gives his or her undivided attention to the testee.

A concentrated one-to-one discussion (on topics relevant to the concept of testee "needs beyond the classroom", as introduced by the tester), also makes more realistic demands on the testee; in order to function effectively in the English medium class, teachers need to be able to maintain coherent one-to-one discourse also with English-speaking parents (and colleagues).

Marked written scripts are at hand during oral interviews and available to the oral tester during the intervals between interviews. This is of particular relevance in borderline cases, where effective performance in the oral may allowed to compensate for a doubtful written test, but not vice versa. Spoken English is the dominant mode in the classroom and as such is considered the dominant mode also in assessment.

6.6 Scoring the Sample Test

"Communicative competence" refers to a multi-dimensional complex of conscious and subconscious abilities and skills. To suggest that it may be measured on a continuous scale of one to 100, would seem to imply an assumption that these abilities and skills may be analysed into accumulative "intervals of competence". Given present understanding of "how language works", the assumption is premature.

A total possible score of 100, and a "pass mark" of 50 as proposed in the sample test are, therefore, more or less arbitrary points on a more or less arbitrary scale, and as such they are subject to interpretation. The difference between interpretation here, and interpretation in the system of endorsement currently in use, is that this study proposes *consistent* interpretation of criteria, based on identified needs relevant to all potential testees in the multilingual and multicultural target group to which endorsement regulations apply.

But a relevant needs analysis alone does not ensure consistent results. Tester/marker training and updating courses, characteristic features of existing standardising programmes, address potential inter-marker differences, and feedback research highlights problem areas. The sample test above was administered, in September, 1990, to the English medium In-service group at UCT. In the absence of a

large scale research programme, scoring of the sample test, as discussed below, can claim to be consistent only in that the written test was designed, and all scripts were marked, by a single tester. For the oral test, testees were interviewed by one of two testers neither of whom had prior knowledge of the results of written tests.

CHAPTER SEVEN

STUDENT PERFORMANCE IN RELATION TO TEST

7.1 Mark Scheme

Included as a guide to testees, the mark scheme as given in the test resembles discrete point, additive scoring, in apparent contrast with the holistic approach proposed above. That final score had less to do with points per se than with how they were earned, is suggested by the note to testees on the information page which preceded the test:

NB PLEASE OBSERVE THE RECOMMENDED TIME ALLOCATION GIVEN FOR EACH SECTION OF THIS TEST.

Even if you do not finish them all, YOU SHOULD ATTEMPT EVERY SECTION

7.1.1 Theoretical Orientation to Marking

Perhaps the most striking feature of published research in language testing to hand, is the heavy "weighting" of the literature in favour of theoretical argumentation about what constitutes communicative competence, and the form and content of "communicative" tests, with relatively far less

apparent concern about how marks should be allocated on test completion. The reason for this is probably the fact that the former issues remain the focus of continuing debate - with the result that attempts to design and administer tests incorporating the theoretical principles of communicative competence are constrained by the largely unresolved question of how communicative competence may be reflected as a "mark".

Significantly, the Report on the British Council Course on Communicative Language Testing in September 1985 (Yeld) contains less than two full pages (out of forty) on the subject of practical mark allocation; and although the several tests reproduced as appendices (in particular, The Associated Examining Board's *Tests of English for Educational Purposes*, and Royal Society of Arts *Examinations in the Communicative Use of English as a Foreign Language*) include instructions to markers, it is not always clear why a given question should carry, say, four marks rather than two or five, for example.

The volume *Modelling and Assessing Second Language Acquisition* (Hyltenstam and Pienemann; 1985), recommended to me in 1989 by the Director of the International Office of Educational Testing Services (Steve Stupak), Princeton, USA, as the most comprehensive collection of research papers on language testing available at that time, is also almost

entirely oriented towards the theoretical underpinning of tests, with relatively little on practical test design - and less on mark allocation.

Returning to the present test, two questions in particular arise. One, assuming that the principles which inform the sample test are in fact relevant to the notion of communicative competence (and "real-world" communicative interaction), how can mark allocation accurately reflect a "cut-off point" between a teacher who is "competent to teach English-medium classes" in real world terms, and one who is not? And two, assuming that

from a small sample of his behaviour, we make inferences about how he would behave in all other situations demanding the same skills

(Ingram, 1977:15)

can teachers in fact be tested in respect of all the relevant language skills/abilities (as defined in Chapter Five, following Widdowson 1978), and should all "samples" carry the same weight in the mark scheme?

To some extent, these questions are addressed in a short paragraph, *Marking of the Test*, in Morrow's Paper delivered at the British Council Course (in Yeld, 1985:3):

There is a clear difference in nature between the "mark schemes" applied to the test in the two different modes (productive and receptive), which reflects the difficulties experienced by the testers in measuring reading and listening, in terms of "tangible instances of performance" of the sort which can be identified in speaking and writing. Another issue with regard to marking is that of compensation - should "above criterion" performance on one of the tasks in the oral

or writing test allow the candidate to make up for "below criterion" performance in the other? No principled and practical solution to this has yet been found - the feeling of the RSA examiners/test developers is that candidates should demonstrate proficiency on all tasks.

The reference to proficiency (in the context of communicative language testing) aside, it would seem that the objectives of the RSA examination (of English as a Foreign Language) are closer to those of the present test than, for instance, the objectives of the TEEP (Test of English for Educational Purposes) which is more concerned with the use of English for *academic* than for *professional* purposes. (From the perspective of this study, *academic* English, or English *proficiency*, is in accord with the stated objectives of the TEEP as presented in Appendix A, iv - vi, in Yeld;1985, and is the concern of the English-medium diploma course itself.)

In the spirit of the argument that some basis for marking is better than no basis, and on the assumption that the objectives of the RSA Examination "in the communicative use of English" in themselves reflect relevant trends in testing research" essentially in accord with our own, mark allocation in respect of the sample test offered in Chapter Six adopted RSA principles for (testing and) marking, as far as these were retrievable from the Overview and Reports contained in Yeld (1985).

The following comments under the heading *Marking and Mark Schemes* are pertinent here:

The procedures for marking.....raise a number of interesting issues for those interested in the principles *as well as the practice of communicative testing*. The tests in the productive area (oral interaction and writing) are both marked by measuring the performance of the candidates in the specified tasks against the degree of skill required to pass at that level *according to the specifications*. Our experience shows that such measurement (*which is essentially of a yes/no nature*) can be carried out with a high degree of reliability by trained assessors *despite its superficial subjectivity*. The receptive skill tests (listening and reading) are marked by awarding marks on a pre-determined basis for obtaining correct answers in a number of discrete questions, which are normally set *in an objectively markable format*.
(in Yeld, 1985:Appendix B:1 :3)
(My italics)

The full specifications are included in Appendix B:11:i-iv; they take the form of graded scales setting out

at each level.....the operations to be tested, the text types on which they must be performed and the degree of skill which must be demonstrated in this performance.....these specifications are the yardstick against which the content and the format of each paper is measured: they are an explicit statement of what it is necessary to be able to do in order to achieve a pass.....
(ibid)

Characterised by verbal descriptors of skills in the categories *accuracy, appropriacy, range and complexity*, the scales are not in themselves definitive, and it is felt that

many of the obvious problems of the RSA derive from its purpose and the fact that it represents neither the end point of a period of learning nor the beginning of an identified (or identifiable) course of action, making rigorous specifications and post-test analysis therefore far more problematic.

(Allen; in Yeld, 1985:33)

In the light of the above, it is clear that this "pilot" study can claim to do little more than identify some of the relevant issues in current research, and attempt to show how they may be applied to the test of English "for teaching purposes" in the present context. Given the linguistic complexity of this context, I argue that a large scale research programme - and marker training within the context of such a programme - would be the only realistic approach to meaningful normalisation of the current confused situation. *But is it so confused?*

It is in this spirit that the rationale for mark allocation in our sample test is offered here.

7.1.2 Rationale For Marks Per Test Item

Pre-determined correct answers in Part One of Section A, and to some extent also in Part D, account for a not insignificant possible 20% of the overall total. Although inconsistent with the concept of holistic scoring as discussed below, MCQ and Cloze techniques were included in the test for the following reasons:

i) they draw on the inter-personal skills/abilities of *reception* (reading, comprehension), in contrast with expository discourse, which draws more specifically on the

intra-personal skills/abilities characteristic of language production (writing, composing); as such,

iii) they increase the experimental value of the test as a whole, in that results achieved in MCQ and Cloze tests can be compared with those scored holistically, constituting important feedback for future test designs;

iii) they require relatively little time to complete.

The five multiple choice questions in Section A carried two marks each for a total of ten (or 10% of a possible overall total score). 4 distractors per question reduce the "monkey" or guess factor to one in five per question, and one in five for Part One as a whole, which seemed a fair weighting for a relatively undemanding demonstration of reading/comprehension skills. On the other hand, and excluding the lucky guess, all correct answers depended on some testee interpretation in that they could not be directly retrieved from the text.

Part Two of Section A consists of 9 comprehension questions for 2, 3, or 4 marks each, for a total of 25 marks, or 25% of the overall test score. Questions 1, 3 and 8 are what Widdowson might call usage reference questions (1978:100), in that responses can be directly retrieved from the text, and the testee is not called upon to interpret the discourse value of the relevant sentences. All the other questions in this section are "use inference" ones, which not only

require greater engagement of testee skills of comprehension and interpretation, but also, and significantly, greater participation of the holistic scorer as the "receiver" of the response.

Sections B and C, for 30 marks and 25 marks respectively, draw heavily on the productive skills (in the written medium). Accounting for 55% of the overall score, these questions were marked holistically, in contrast with Part Two of Section A where the format of the question dictated a more itemised approach. The significant difference between the two approaches lies in the primary concern of the former (holistic marking) with overall text *coherence*, while the latter focuses more readily on *cohesion* (at the level of the sentence).

Examples of testee responses in relation to given scores are analysed in 7.2 below; but first, the rationale for holistic scoring, as it signifies here.

7.1.3 Holistic Scoring in the Sample Test

Widdowson (1978:27) notes that "discourse is not dependent

on overt cohesion....."; and he illustrates the point by comparing the following exchanges:

1. A: What are the police doing?
B: They are arresting the demonstrators.
2. A: What are the police doing?
B: The fascists are arresting the demonstrators.
3. A: What are the police doing?
B: I have just arrived.

In the first exchange, the anaphoric *they* provides a propositional link with the question, allowing ready reference to *the police*. The second exchange is less cohesive, requiring the assumption that B is referring to something that A has said, and that his remark is a response to it; this allows a semantic link between *the police* and *the fascists* - provided we know that for certain groups of people these two terms are often almost synonymous. The third exchange is less cohesive still, with no apparent connection between the two remarks; yet it is not difficult to recognize that B's response could be entirely appropriate in the relevant context. Why is this so?

As Widdowson observes, it is because we focus our attention on the illocutionary acts the propositions perform.....; "we create a situation in our minds", linking the propositions in a way which will account for a purposeful communicative act.

From this illustration, it is clear that although (semantic) coherence and (syntactic) cohesion are closely related in propositional development, coherent discourse does not necessarily depend on overt illocutionary markers. But I would take the argument further, and suggest that neither does it necessarily depend on other forms of syntactic cohesion normally associated with the notion of "good" English - subject/predicate concord for example, or control of co-ordinating/subordinating propositional organisation, or rational cross-textual referencing, or even punctuation. Non-prescriptive attitudes to "creative writing" have long recognised this.

In contrast with the analytical approach of the discrete point marker, I suggest that the holistic marker is thus less concerned with the separate semantic, syntactic (and phonetic) components of testee discourse, than with discourse as a synthesis of ^{Aw} form and meaning in a purposeful communicative event (or series of events) .

The problem is to reconcile a non-prescriptive approach to the coherent, but less cohesive, communicative event in "real-world" terms, with the prescriptivism inherent in the concept of testing itself; and it becomes more complex when we need to account for the fact that perceptions of coherence are markedly influenced by socio-linguistic factors.

This is of direct relevance in the non-homogenous socio-linguistic context of test construction here - what one tester perceives as coherent, or even acceptable English, another might not. For this test there was only one marker for a total of 31 scripts, and it is thus fair to assume relatively consistent scoring. For large scale standardised assessment of EL2 teachers as promoted in this study, the only way to achieve consistency of approach in respect of holistic marking would seem to be through relevant research - to what extent does it "make sense" to tolerate uncertain cohesion in (otherwise) coherent teacher discourse? - and through tester/marker training.

7.2 Analysis of Selected Testee Responses and Scores

7.2.1.1 Section A, Part One

MCQ 1 - 5: Correct responses were not directly retrievable from the text. Questions 1 and 5 required either lexical knowledge (a millenium=1000years; felled=chopped down), or the ability to infer answers from given discourse; in question 2, the correct response, d), depends on testee ability to recognize the conceptual link between "have yet to be significantly improved upon" in the text, and "are still characteristic features.....today" as one of the five

responses offered; and questions 3 and 4 depend on accurate interpretation the "story-line".

Eight testees, with overall scores ranging from 58% - 85%, scored full marks for this part of the test. The most frequent errors were in MCQ 3 (ten incorrect answers) and MCQ 2 (six). All but 3 testees scored 6 or more for this part of the test (compared with 2 unsuccessful testees in the test as a whole)

7.2.1.2 Section A, Part Two

In an attempt to come to terms with *correct* content and *appropriate* form, this section of the comprehension test was marked "semi-holistically". In effect, this meant that where coherence was a function of cohesion, syntax materially influenced score. Given the brevity of responses in this type of exercise, in comparison with the scope for "message" reinforcement in sustained discourse, holistic marking was, to some extent, impracticable here. The following examples represent a cross-section of testee responses to questions 2 and 7, and of how they were scored. In the sense that pre-existing lexical knowledge was not

assumed, Question 2 is a "use inference" question: testees were required to infer meaning from text content.

Q2: What is the meaning of "purely cosmetic"?
(for 2 marks)

Response A - Something which merely looks attractive, but without necessarily functioning well or being of special quality. 1 1/2 marks

(Content coherent, cohesion acceptable; but here implies a clause rather than a phrase, but this does not interfere with overall meaning)

Response B - only for the beauty of 2 marks

(to the point, coherent and acceptably cohesive in a "question and answer" context)

Response C - varnishing is not just using a brush and a tin of paint and that would be purely cosmetic. It must look like being varnished on the outside. 0 marks

(incoherent and incohesive to the point of being meaningless as a response to the question)

Response D - "Purely cosmetic" means that the varnishing of the violin was more than just a camouflage. The varnishing wasn't only done to make the violin look beautiful. 0 marks

(there is evidence here of comprehension of the text; as a response to the question it is both incoherent and incohesive)

Response E - Only for the beauty/visual appeal of it 2 marks

(coherent, and acceptably cohesive in "real world" terms)

Question 7 is to some extent a "usage reference" question in that most of the response can be retrieved directly from the text; to obtain "full marks" for this question, certain inferences were required:

Q7: Where did Brian Lisus study, what did he study, and what does he do now? (for 4 marks)

Response A: - He studied in Robertson. He was a professional violin-maker and has been trying to attain a degree of perfection in formulating his own varnish.
1 mark

(High cohesion and coherence in purely syntactic/semantic terms - that is, as sentences - as a response to the question it is largely incorrect)

Response B - Brian studied in England. He tried to find the formula of Stradivari's oil varnish. Now he is trying to create violins as close as possible to the one's Stradivari made.
2 marks

(coherent, cohesive and partly correct)

Response C - England. The art of violin making. He also makes violins.
3 1/2 marks

(content correct; a coherent response to the first two parts of the question, less so to the third; the form of the response makes cohesion an almost irrelevant consideration here.)

Response D - He is a professional violin-maker and he studied in England to become a craftsman
3 marks

(coherent and cohesive; only partially correct)

Response E - He studied in England to be a violin craftsman. Brian Lisus is a professional violin-maker.

3 1/2 marks

(coherent and cohesive; is a musician a craftsman?)

By the nature of the exercise, scoring of responses in comprehension tests focuses in the first place on "correctness" of content, and to that extent marks must be allocated on a "discrete point" basis. When test items elicit sustained discourse, as in Sections B and C, responses are more amenable to a holistic view (and to scorer perceptions of the objectives of the test).

7.2.2 Sections B and C

A holistic approach to responses to Sections B and C was informed, in the first place, by the "appropriacy" of response to the task in each case. No attempt was made to define "levels of appropriacy", and marks were allocated "on the basis of a general overall impression" of testee ability, as proposed by Hargreaves (1985:22).

"Impression marking" was influenced by evidence of testee ability to select and organise relevant material from a given source (in the summary exercise), and from his or her own experience in relation to a given topic (in the composition question). Overall coherence, and

acceptability, in "real-world" terms, of (underlying) cohesion also influenced mark allocation. Except where they caused breakdown in testee/scorer communication, errors of syntactic linkage, poor spelling and non-standard varieties of English were not negatively assessed, and marks given in these two sections of the test were not, in the first place, a reflection of testee performance measured against "grammatical" criteria.. Scoring here was, rather, a numerical expression of tester response to the communicative, not the formal, value of testee discourse as a whole.

Three sample segments of testee responses to Sections B and C give some indication of scoring at three "levels" of competence, as represented by overall scores of 60%, 50%, and 40% respectively:

Section B - (62%, concluding paragraph):

The informal sector helps people to upgrade themselves, they become self-sufficient and this promotes self-respect and then they realize the need for education.

(50%, concluding paragraph):

Dr Ben Vosloo, managing director of the S.B.D.C. said the informal sector wouldn't take the place of the formal sector but will form a linck between the 2 and reconcile the first and third world economically.

(40%, concluding paragraph):

Mr Lawrence see this as an opportunity for everybody to upgrade themselves by means of leading his own business and powers to rule the others, on his own business. And they also realise that they must go back to school in order to be a director of his own business.

Section C - (63%, opening paragraph)

To teach is to communicate! How true this is, especially as a Junior Primary teacher. The child, coming to school for the first time, cannot read, write or do sums. The only way he can be taught is by using your voice, visual aids and by demonstrating with your body.

(53%, opening paragraph)

Aids to classroom teaching are very important and helpful. Aids have always their advantages and disadvantages. When one wants to find the advantages and disadvantages of the chalkboard and the overhead projector one must keep in mind why one wants to use the above mentioned aids.

(29%, opening paragraph)

Certainly the main advantage of a chalkboard is, demonstrations placed correctly, ensure that the whole class can see it. Mistakes made on the chalkboard can quickly be erased. Demonstrations on a chalkboard are fruitful for everyone who sees it. Demonstrations on the chalkboard take the fourth position on the cone of experience.

Weakening linkage between coherence and cohesion in each of the above two series of responses (and in the full responses from which these segments were drawn) reflects the drop in overall score as indicated. The third example in each set of three responses, although arguably consistent with perceptions of coherence in the *spoken* mode in "real world" terms was, in this marker's view, lacking in cohesion to the point of being unacceptable in *written* discourse which, lacking the instant verbal and visual feedback of the oral medium, is essentially a "one off" event (and there is thus no possibility of self-correction, as there is in spoken interaction, where communication breakdown threatens or occurs).

7.2.3 - Section D

As noted above, the inclusion of a Cloze test, Section D, was partly experimental in nature, motivated by (among other considerations) the proposition that there might be a high correlation level between scores in a test of "expectancy grammar" and scores on the rest of the test for communicative competence.

In this test, correlation was low. Sixteen (of thirty) testees inserted 10 (or more) correct, or "acceptable alternative", lexical items; ten testees scored 9 or 8, three scored 7. Only one testee scored less than 5 (4).

In contrast, on the test as a whole, six testees scored 80% or higher; eleven scored between 70% and 79%; six, between 60% and 69%; five, between 50% and 59%. Two scored less than 50%.

Of these, one scored 40% on the test as a whole, and 4 on the Cloze test, which correlates well; the other scored 29% on the test as a whole and 8 on the Cloze, which does not..

No meaningful conclusions can be drawn from a test on such a small target group. Neither was the level of the Cloze test necessarily appropriate for the purpose. A longitudinal research programme, with facilities for trial testing and

feedback analysis in a representative sample of the target group, would be a useful way to address these and other questions relating to language qualifications for teachers.

7.3 What Do Scores Mean In Pass/Fail Terms?

A familiar question in applied linguistics is whether or not, in a test of communicative competence, testees should be required to pass all sections of a given test in order to pass the test. In theory, if both the consistency factor and the predictive value of holistically assessed written discourse could be accounted for, communicative competence could be assessed on the basis of one task only, and the question would be irrelevant. In practice, this has not (yet) been achieved, and the question must be dealt with from a more or less informed point of view in each case. In the case of the present test, testees were required to score a total of 50%+ in at least three of the four sections for a clear pass. Below 45% was considered a "non-compensatable" fail. Two testees scored 50% and they were assessed as "borderline" candidates in the oral test. A score of 45% to 50% in the written test was "compensatable" on the grounds that the oral medium is the dominant one in the classroom.

7.4 Scoring The Oral Tests

In the sense that the oral medium lends itself more readily to real world tasks (such as teaching a classroom lesson) than do formal written tests, oral tests are perhaps more "direct" than written ones.

But the "directness" and the "real world" nature of the oral test remain a matter of degree. Even teaching practice during the diploma course is not an entirely realistic context for the assessment of teacher competence: on the one hand, the student-teacher depends on regular teachers for opportunities to teach, and may be over- or under-employed with little control over his or her working day; on the other, he or she takes no real responsibility for pupil performance in the long term. In addition, stress levels are often considerably higher than is characteristic of a regular teaching position, even if only because the student-teacher is conscious of being more or less "on show" at all times (and not only during assessment).

With such considerations in mind, in-service oral tests, like the written component of assessment, constituted an attempt to elicit evidence of predictive value, rather than to formulate "actual tasks" for "direct" assessment.

Criteria for oral tests were specific only to the extent that testees were required to demonstrate ability to recognise and exploit visual and verbal feedback, and to maintain coherent discourse as directed by the tester. At the same time, testee initiative in "turn-taking" was encouraged. Within this framework, testing was essentially impressionistic in nature.

Although norm-referencing is never entirely absent from the multiple testee situation, assessment for E-endorsement is strongly criterion-oriented, and the only consideration of practical significance whether or not the testee merits it. For this reason, scoring in the oral test was seen as complementary, rather than supplementary, to the written test.

Reflecting this, perhaps, scoring in the oral test was to a large extent confirmatory in nature. Testees who had scored 55% and above in the written test were, almost without exception, assessed at a similar level of competence also in the oral medium. The end result being unaffected, there was little point in giving numerical value to oral competence in this category of testees.

Similarly, it was found that testees who had scored below 45% in the written test were less competent to "negotiate meaning" than was consistent with identified standards of

competence for E-endorsement. Two testees fell into this category; in consideration of current regulations, they were assessed at e-level.

Finally, it is argued here that as the oral mode is the dominant one in the classroom, oral competence should be allowed to compensate for "borderline" competence in the written mode. Consequently, a more focused approach to oral competence seemed appropriate in the assessment of testees who score 45% - 55% in the written test; and in such cases the tester was made aware of the written score *before* the oral test. Of the 1990 In-service group, two testees had scored 50% in the written test; in both cases, the tester confirmed E-endorsement.

It was interesting to note that in the case of testees who had followed the In-service course, final scores were consistent with average scores for course work in the second half of the course - which is consistent with course/test isomorphism. It is however of little relevance with regard to the plea made in this study for a standardised testing programme which assumes no prior

period of learning nor the beginning of a.....course of action" (Allen, 1985:33).

On the other hand, assuming that candidates who registered "for examination only" (and who therefore had not followed the course and were not formally "prepared" for the test) were motivated to do so by a measure of self-confidence in

the English-medium classroom, and given that all such candidates were in fact successful, may suggest that the test was at least not inappropriate to the classroom needs of the EL2 teacher.

But the situation remains a complicated one, and is likely to become more so in the near future as changes in the political situation are inevitably accompanied by changes in educational perspectives. If this project does little other than give rise to a meaningful initiative to address problems attaching to criteria in English for teachers in the South African context, it will have achieved its main objective.

At the same time, it would do less than justice to a study on criteria in English language assessment to ignore the fact that the significance of English language tests for teachers goes far beyond administrative requirements for competence in the English-medium classroom. Chapter Eight looks briefly at English "use and usage" in the wider society teachers serve, and at the notion of "teacher English" from this perspective, in terms of the contrast between "the English-medium classroom" on the one hand, and "the English class" on the other.

Chapter Eight

English In The Curriculum

Today And Tomorrow

I have based my argument for a distinction between communicative competence and proficiency (as criteria in language tests) on a proposition that language may be assessed in "bilateral" terms: in terms of *inter*-personal competence on the one hand, and *intra*-personal expertise on the other. At the same time, I have acknowledged the fact that the complex, multi-dimensional nature of human language makes the proposition more readily defensible on pragmatic than on "scientific" grounds. Pragmatism is consistent with the practical objectives which motivated this project in the first place.

But if the distinction is primarily intended to serve the interests of EL2 teachers seeking employment in (English-medium) State education, it has implications also for the *teacher* of English (in contrast with the *English-medium* teacher). The notion of "language across the curriculum" aside (its sociolinguistic implications being substantially in accord with my interpretation of communicative competence), the academic/linguistic requirements of English

"in the curriculum" are addressed in terms of the English Method course. Nevertheless, under current regulations for employment in State schools, English teachers are subject also to E-endorsement. The fact that they are not always successful in such tests as are administered for this purpose, as noted in Chapter Two of this study, raises questions not only about E-endorsement tests, but also about the preparation of teachers to teach English.

For this reason, it seems appropriate to focus here, albeit briefly, on the role of English in the school curriculum, and on the language needs of the teacher to teach it.

I have suggested (in Chapter Four) that the teacher of English as a subject should be assessed against a criterion of proficiency which, in terms of the distinctions I make (based on Widdowson's distinction between usage and use; 1987:1-20) subsumes communicative competence (but not vice versa). Communicative competence, I argue, is not enough.

Enough for what? What is it that the teacher of English needs in respect of "ability in English", beyond competence to communicate in the English-medium itself?

8.1 The English Class

I would argue that the educational/cultural components of the English curriculum (particularly at secondary level) assume certain language skills on the part of the teacher, not least among them the verbal skills for articulate reflection on literary texts and on how "language works" in the creative process. More than this, the teacher needs the linguistic (and communicative) skills for leading others to such articulate reflection. For how can the reader respond to a Shakespearian play or an e.e.cummings poem - even on the emotional level - without the linguistic "equipment" for "making sense" of them beyond the surface organisation of the text? Or "appreciate" figures of speech such as the metaphor, without the lexical and syntactic knowledge for "deciphering" them? Or, for that matter, "break the rules" in creative writing, if he or she has never known them - as Picasso could not have created cubism without first knowing the "rules" of painting?

Teacher "needs" in the "English class" cannot be seen in terms of language ability alone, neither do I suggest that linguistic expertise is the most important attribute of the English teacher; for although curricula differ in content according to the grade on which English is offered, English at secondary level is not intended as an exercise in language teaching. On the contrary, it assumes a pre-

existing basic ability for reading and writing English as a point of departure for the exploration of selected literary texts and the creative role of language in the human experience.

What is of concern here, however, is not so much what "the curriculum" assumes in respect of pre-existing pupil potential for studies in literature and "creative writing", but rather what it anticipates of pupils in terms of curriculum objectives - and, more specifically, of English teachers in catering for pupils' needs in this regard.

Given the orientation of this study in terms of tertiary (post-graduate) teacher education, in a sense I "bridge the gap" between pupil and teacher needs, by focusing here on student/trainee teacher needs in respect of English "for academic purposes" - as consistent with a requirement of English "on the higher grade" for university admission.

The qualifier "on the higher grade" would seem to suggest (at least) the skills of "reception" appropriate for engaged reading and perceptive interpretation of prescribed English-medium texts at tertiary level, and the "productive" skills/abilities relevant to written (and spoken) discourse, as consistent with course work in tertiary education. This, I would argue, implies not only coherent "message" negotiation (*communicative competence*), but also linguistic

skills consistent with syntactic "negotiation" and cohesive discourse (*proficiency*).

Focusing on the second of these propositions (productive skills consistent with English "for academic purposes"), the nearly two thousand HDE curriculum theory assignments (lesson transcripts, self analyses and partner comments) I have assessed on two levels (content and language use) in the last two years (1989 and 1990) suggest that English teachers, for whatever reason, (and there are reasons, discussed below) do not always prepare pupils for tertiary level performance as I have characterised it here. In particular, errors of syntax and confused cross-text reference resulting in uncertain cohesion, similar to errors noted in HDE students' responses in Chapters Two and Five, would seem to indicate that many students complete secondary school and graduate studies in the English medium - some with English as a major - able to produce (relatively) coherent discourse, but without the appropriate grammatical skills for cohesive discourse structure.

It is this apparently irrational but de facto situation, (arising in part from the fact that there is more than one official "standard" national language and several non-official indigenous literate languages) which prompts my proposition for a distinction between *communicative competence* and *proficiency* as criteria for teacher ability

in English in South Africa today, depending on whether English is the *medium*, or the *subject*, of instruction. For on the one hand, it is a self-defeating exercise to set "external" standards of English usage (consistent with a traditional interpretation of proficiency) as a "blanket" requirement for teachers, when a high percentage of HDE's every year - including English Method students - are patently unable to achieve them (although apparently able to satisfy the internal demands of the English-medium HDE itself), as shown in Chapters One to Three of this study; and on the other, it would seem to be equally self-defeating to offer secondary (and tertiary) level courses in English - accommodating the English literary tradition - in an educational context which does not realistically cater for it.

But if it would seem to be unrealistic, at this point in our politico-linguistic history, to require "English proficiency" for permanent employment in State schools, regardless of the teaching post held, how relevant will the notion of proficiency for "English in the curriculum" be in a South African educational dispensation of tomorrow? The answer will no doubt depend to some extent on how the authorities of the future apply the lessons of past experience in respect of "the language issue" in education.

8.2 A Historical Perspective: Past, Present and Future

As Hartshorne observes, nearly a century of government policy

influenced by and often in tension or conflict with.....extra-statal bodies.....exercising pressures on the relative positions and status of.....English, Afrikaans (the two 'official' languages) and the eight African languages recognised for educational purposes
(Hartshorne, 1987)

have sensitised the question of language in education at the level of policy, and troubled it at the level of the classroom.

Significantly, however, following the Education and Training Act of 1979, the use of English in Black Education from Standard Three onwards, and the scrapping of the Afrikaans requirement for university entrance, would seem to indicate growing acceptance, within the present political dispensation, of the role of English (language) in national and international affairs; and as argued elsewhere in this study (and in spite of the fact that English is perceived by many as a language of the oppressor), it is likely that this trend will continue in a post-apartheid South Africa, at least into the foreseeable future, if only for economic and commercial reasons.

To what extent English *literature* will retain its position in a future educational dispensation, remains to be seen. But it would seem to be more predictable that a "new"

South Africa (with English as an/the official language?) would be likely to see strengthening economic and socio-political links with the international English-speaking world, that in itself would seem to carry certain implications in respect of English "in the curriculum".

I am suggesting that continuing - and strengthening - ties with the international English-speaking world, if primarily for pragmatic reasons, will inevitably constitute a matrix for the fostering also of academic and cultural links within the "standard" international English tradition.

8.3 English For Academic/Cultural Purposes?

The extent to which the academic experience may be said to embrace "culture" depends, at least in part, on one's definition of culture.

In dictionary terms, culture means

1. the appreciation and understanding of literature, arts, music etc.
2. the customs and civilization of a particular people or group.....

(The Oxford Paperback Dictionary, 1979:149)

Too succinct to be meaningful from a socio-anthropological point of view, perhaps, this "definition" serves present purposes in that it is consistent with the point I would make: that "culture", essentially something of an umbrella

term for collective human experience, is largely "transmitted" by way of language. From this perspective, it is clear that language is not only a *communicative* code, it functions also as a *cultural* code.

On this basis, I argue that to the extent that *proficiency* in English is consistent with "appreciation and understanding" of English civilization through English literature - in a way that communicative competence as *defined in this study* (for given reasons) is not - it is consistent also with the notion of English for cultural purposes or "cultural literacy".

But to see "academic" cultural literacy as a function of English proficiency is to take too narrow a view of culture in a multi-cultural South Africa. Hirsch (1987), exploring the notion from an American point of view, sees it in a somewhat broader perspective.

8.3.1 Cultural Literacy

According to Hirsch, cultural literacy is

the background information, stored in their minds, that enables competent readers to take up a newspaper and read it with an adequate level of comprehension, getting to the point, grasping the implications, relating what they read to the unstated context which alone gives meaning to what they read. (1987:2)

"Background information" here would necessarily include the linguistic expertise for comprehension beyond the surface meaning of the text; "the unstated context" refers to the shared knowledge of literate Americans, as represented by a list of about three thousand key words, phrases and dates suggested by Hirsch (1987:152 - 215), only a small percentage of them "literary" in nature, or even "cultural" as defined above.

As Professor of English at the University of Virginia, Hirsch is concerned that American literacy rates have not kept pace with rising standards of literacy throughout the developed world; and in a chapter on *Cultural Literacy and the Schools* (1987:110 -133), he blames "the fragmented curriculum":

School is the traditional place for acculturating children into our national life.....Our schools have played this role less well than they should, chiefly because they have followed faulty educational ideas.

Part of the problem, he maintains, lies in the encouragement of multilingualism:

well-meaning bilingualism could unwittingly erect serious barriers to cultural literacy (1987:93)

and,

it is contrary to the purpose and essence of a national language.....that a modern nation should deliberately encourage more than one to flourish within its borders
(Hirsch, 1987:93)

In terms of Hirsch's argument, "cultural literacy" in the South African educational context would be negatively

influenced by bi/multi-lingualism; and the status quo would seem to confirm this. But as in the United States, the reasons for low levels of "national" cultural literacy in South Africa are not purely linguistic in nature; neither does America have the prerogative of "faulty educational ideas".

On the other hand, identifying problems is in itself a positive move; and there are currently initiatives at all levels of South African education for upgrading the qualifications of English teachers in anticipation of the day that all adult readers will

take up a newspaper and read it with an adequate level of comprehension.....(Hirsch, 1987:2)

On the whole, such initiatives tend to focus rather more directly on the *intra*-personal skills of reading and writing (consistent with my interpretation of proficiency) - and on how these may be taught and learned, (in contrast with the emphasis proposed in this study on the *inter*-personal skills of "speaking" and "hearing" where the criterion is communicative competence).

8.4 Creative Teaching

Referring to the work of Professor Jeanne Chall (among others), Hirsch notes that she is

one of several reading specialists who have observed that 'world knowledge' is essential to the development of *reading* and *writing* skills (My italics)

and he goes on:

what she calls world knowledge, I call cultural literacy. (1987:2)

Whether one calls the interface of language and culture *world knowledge*, or *cultural literacy*, or *shared information*, depends on one's own "particular areas of interest" as suggested, somewhat closer to home, also by Bakker (1985). In an eloquent plea for *informed attending* in the teaching of English literature in South African high schools, he quotes Maxine Greene as follows:

Teachers who are concerned about responsibility and freedom will empower the young to gain perspectives on what seduces and conditions them. They will enable them to come in touch with their biographies, to tell their stories, so that they can discover how they have mapped their own landscapes and constituted their worlds. (Greene, 1981)

and Bowers,

the content area and pattern of thinking which underlies the organisation of knowledge influences the student's conceptual map..... (Bowers, 1984)

There can be little doubt that creative teaching starts with teacher concern, as Bakker is concerned to bring the reader to awareness and knowledge of self through exploration and personal experience of the literary text. But if he sees

the study of literature as

a call to dialogue, rather than a demand for scholarship (1985:164)

he also makes the point that

it is not possible to engage in significance without a prior engagement in meaning (ibid).

It is "*engagement with meaning*" that I see as the function of English proficiency, the ability to "comprehend" and interpret discourse beyond the surface level, and to enter into "dialogue" not only with a given text, but with the wider cultural context within which it is located. There can be little doubt that "reading" in this sense depends materially on the *intra*-personal skills of language usage, and to this extent, then, I suggest that proficiency is a function of the English class at secondary level - which assumes that it is also an integral function of the English Method constituent of the HDE course.

Teaching and learning skills for language proficiency does not need to be dry, tedious, uninteresting - all the perceived negative properties which caused "prescriptive grammar" to fall into disrepute. The potential for creativity on all levels and in all spheres of teaching, is a major theme in current directions in educational research; and positive attitudes generated in the pupil-centered classroom, "democratic" teacher/student interaction, and peer-teaching/group-learning techniques, could and, I would

argue should, be exploited for the creative promotion of reading and writing skills from primary level - up to, and including, teacher training.

The HDE programme is, I suggest, the place to start if "grammar in the curriculum" is to be restored to respectability. This is not a gratuitous suggestion. Careful examination, during the last four or five years, of several thousand HDE assignments, as well as interviews with hundreds of students for research purposes, have suggested that in spite of reading for course requirements, and the successful acquisition of course specific vocabulary, students do not automatically acquire the language skills consistent with written (or oral) academic argument. It seems that "reading" for content can be accomplished with reasonable success by scanning for key words and phrases - which is the underlying principle of speed-reading courses; it is not necessarily conducive to understanding of how the written text is "put together", or even to the acquisition of the compositional skills for structural cohesion consistent with international academic standards as suggested by published tests for college and university entrance. Arguably, too, tutor assessment of course work on the basis of such key words and phrases compounds the issue.

I would argue therefore that consideration be given to a proactive initiative for sharpening the reading and writing

skills of prospective English teachers. In the form, perhaps, of an imaginative core module in the English Method course, tutorial support aimed at promoting reading not only for content, but also and particularly for recognition of underlying textual cohesion, might be a useful approach to teacher training in this respect. Reduction exercises in the form of (coherent, cohesive) summaries or abstracts of selected articles, with subsequent directed analysis of (unidentified) peer assignments in tutorial groups is just one such approach which has elicited favourable comment from successive in-service student groups. Small group participation in "collective composition" - of a story for children, for example - is another activity which allows informal exploration of how language is "put together".

And finally, on the level of informal teacher education - and in a sense, at the "other end" of the educational "ladder" - a noteworthy recent addition to creative language teaching in South Africa is put forward by Edmunds in *Searching for Relevance* (Chick;1990:570). Noting the profound negative effects apartheid legislation has had on education in Black schools, she shows that "peripatetic support teaching", a programme of itinerant teachers modelled on similar programmes in Wales, England and Israel, could

play a considerable part in strengthening the effective teaching of.....language (Edmunds;1990:579)

not only in Black schools, but also in a future South Africa with mixed ability, multilingual and multicultural classes comparable with the inner-city primary schools in England where such programmes have been operating for some years.

8.5 Conclusion

With this chapter I have, in a sense, returned to the point of departure of this study, to look again at the notion of "the English-medium class" - which, in terms of current requirements, subsumes the "English-medium *English* class". The fact that no such distinction is made in the regulations is perhaps intended to "simplify" assessment. The evidence, as noted in this study, suggests that it may have the opposite effect, further confusing requirements for assessment based on criteria which are themselves subject to interpretation.

If the concept of English for "academic/cultural purposes" I have put forward here to support my argument for a criterion of proficiency for teachers of English, has been argued on a relatively superficial level, it is merely intended to indicate that the weighting of this study as a whole in favour of "communicative competence" in the English-medium classroom does not suggest "lowering of academic standards". On the contrary, it seeks to redress a situation which takes

little cognisance of the fact that English in South Africa is not "just" South African English, but manifests itself in a number of identifiable - and unidentifiable - varieties.

With this chapter, then, I close my argument for a distinction between communicative competence and proficiency as criteria in the assessment of teacher language ability. Chapter Nine briefly reviews the main points, presents my conclusions, puts forward recommendations for teacher language assessment, and outlines directions for future research.

CHAPTER NINE

CONCLUSION

1. Review

This project was undertaken in the School of Education, at the University of Cape Town. It was motivated by a perceived need for a rational approach to language endorsement for teachers in training, in compliance with State regulations as set out in Criteria for the Evaluation of South African Qualifications for Employment in Education, 1988.

The nature of problems giving rise to the study was, in the first place, theoretical. In particular, the "requirement" at e-level for competence "with less proficiency" is difficult to interpret in a manner consistent with modern linguistic research; and the fact that assessment is, for all practical purposes left to 40 or more training schools working independently of each other, complicates the issue.

A questionnaire to all relevant centres elicited responses which indicated considerable differences between training schools in attitude to, and administration of, the

regulations. There was also a high level of interest in an initiative for normalising the situation.

A literature survey was undertaken to establish recent trends in language testing research. It was found that although language teaching/testing is a primary focus of applied linguistic studies, research in testing *per se* is relatively less well documented. Furthermore, a common tendency, among applied linguistic researchers, to use the terms competence and proficiency more or less interchangeably made it difficult to identify a theoretical basis for the regulations as formulated at present. On the other hand, the relatively recent notion of communicative competence, seemed to offer a basis for an approach to assessment in accordance with the spirit, if not the letter, of current regulations

In this dissertation, I have put forward a proposition in terms of which there is a theoretical difference between the notions of language proficiency and (communicative) competence. This would account for the fact that often teachers who are denied E-endorsement under present regulations are in fact (communicatively) competent to the extent that they are able to satisfy course requirements in respect of ability "to teach English-medium classes", although not necessarily proficient/competent in English in a traditional academic sense.

On the basis of this proposition, I argue

i) that communicative competence in English is an intrinsic function of the English-medium Higher Diploma in Education, and that extra-curricular assessment is therefore superfluous; and

ii) that English proficiency is - or should be - a function of the English Method course.

Further, I argue that

iii) teachers trained through "another" medium (effectively Afrikaans, under present regulations) be tested for ability "to teach English-medium classes" on the basis of identified needs (of EL2 teachers in the English-medium class); and I put forward a prototype core needs analysis for this purpose.

A sample test, constructed on the basis of this core needs analysis, and administered to the thirty teachers registered for the 1990 In-service Test for E-endorsement, is offered in Chapter Six of this dissertation. Results of this test, as assessed by the present writer and two examiners external to the In-service Course, are analysed in Chapter Seven.

Finally, in Chapter Eight, I consider the relevance of the approach to criteria for the assessment of English put forward in this dissertation, to currently changing

political perspectives and, in particular, to the needs of education in an anticipated post-apartheid dispensation, in the light of a theory of "cultural literacy" in a "new" South Africa.

2. Conclusions

On the basis of theoretical and practical research as documented in this study, and argument arising from it, I conclude that State language regulations for teachers, and the fragmented approach to them, with consequent inconsistent attitudes to, and standards for, E-endorsement are theoretically untenable, academically unacceptable and educationally unsound.

It follows that the regulations, and the purpose they are intended to serve, should be re-evaluated in the light of current perspectives in applied linguistics. A theoretical rationale for tests for EL2 teachers only would recognize the integrity of the English-medium diploma and make E-endorsement for EL1 teachers superfluous. An analysis of the language needs of EL2 teachers in the English-medium class would account for the educational objectives of assessment.

Errors characteristic of Afrikaans and "Black" first language speakers, and probably attributable to mother-tongue interference - such as faulty subject-verb concord, and erratic tense sequences (particularly the past tenses), in the case of the former, and uncertain use of determiners and the continuous tenses in the case of the latter - do not account for the high frequency of "non-systematic" EL2- and EL1-speaker errors noted in this study. This high incidence of "non-systematic" errors, particularly in written English, at post-graduate level, as yet "unaccounted for", would be integral to such a research project.

Finally, although the parameters of the problems involved - linguistic, and in respect of collective target group size and composition - may render the present "pilot" study anecdotal, results would at least seem to support a conclusion that "standard" South African English (at tertiary level) is currently fairly widely perceived to be tolerant of non-standard structure (however it may be explained).

This proposition would seem to be of more than passing interest in that it may, in turn, suggest a single South African variety of English more deviant from Standard International English than SSAE is at present, emerging in

response to the several varieties (SA "Black" English, SA "Indian" English, SA "Coloured" English, SA "Afrikaans" English) which currently influence it.

3. Recommendations

Arising from this study as a whole, and from the conclusions presented above, I would make the following recommendations in respect of State regulations for teacher ability "to teach English-medium classes":

1. Existing requirements for English language endorsement for teachers, as set out in sections 10 and 11 of CRITERIA FOR THE EVALUATION OF SOUTH AFRICAN QUALIFICATIONS FOR EMPLOYMENT IN EDUCATION, 1988, should be reformulated to account for one level of endorsement only (E-level).
2. The integrity of the English medium teachers' diploma (HDE) should be recognized, making further endorsement in respect of English superfluous.
3. E-endorsement for EL2/3 trainees/graduates, based on identified classroom needs in English, should be standardised within the context of an ongoing research programme, with the infra-structure that such a programme implies in respect of experimental test design and

construction, marker training, feedback analysis, and washback on language education (at primary, secondary and tertiary levels).

4. As State education is ultimately the responsibility of the State, the standardised programme should be undertaken/ supported by the relevant State Authority, with co-ordinated input as appropriate from language testing experts and academic institutions.

Consideration could be given to voluntary assessment for language endorsement. Incentives of higher status and financial rewards remaining substantially the same as for the present "bilingual" certificate, a voluntary requirement would probably encourage, rather than discourage, applicants in that it would remove current negative perceptions of routine "bureaucratic" endorsement, making it "cosmetically" more desirable than it often is at present.

For administrative purposes, standardised written tests could be offered at convenient intervals, and (by arrangement) at appropriate venues such as teacher training centres. Scoring, by markers trained for the purpose within the context of the standardised programme, could be centralised or decentralised as convenient. Short, informal oral interviews could be arranged at the same venues;

alternatively, an oral interview at the appointing school (as "the consumer") might be a viable alternative.

Finally, proficiency as the criterion for the teacher of English "in the curriculum", as motivated in Chapter Eight, would seem to be correctly the responsibility of teacher training centres, and in particular of the relevant "English Method" departments, as the custodians of English as a "cultural code".

4. Directions For Future Research

The likelihood of the continuing use of English, as a "lingua franca" at least, in an immediate post-apartheid dispensation, suggests that research in the role of English in education, EL1 and particularly EL2/3, is not merely a relevant issue today; it is an urgent one.

This research project suggests that meaningful, co-ordinated theoretical and practical research on the use of English in the emergent integrated South African classroom is hardly possible outside the context of a large-scale research programme.

Rationalisation and standardisation of assessment of teacher ability "to teach English-medium classes", and *in situ*

follow-up research on the nature of communicative competence in the South African classroom, would not only be in the interests of maximum relevance of tests and maximum consistency of scoring under the present bilingual system; it would also constitute a major research initiative for ongoing monitoring of teacher needs in a post-apartheid dispensation.

A symposium "attended by representatives of all examining bodies concerned with bilingualism qualifications.....led by an acknowledged expert in this field", as proposed at a meeting of a subcommittee of the Provincial Advisory Committee for Teacher Training (as recorded in Notice of Meeting, 7 August 1989), and not yet, to my knowledge, convened, would be a possible point of departure for the establishment of a relevant research programme in the field of English in education in South Africa today.

In conclusion, it is perhaps well to contemplate that if the survival of a more or less standard form of English on the southern tip of Africa would seem to be assured in the short (and perhaps medium) term, its long term survival will surely depend on the perceptions - positive or negative - of future EL2 users in respect of its communicative relevance "in the market place" rather than in the schoolroom. To be unrealistically prescriptive about "standard English" for

general teaching purposes would therefore seem to be counter-productive.

At the same time, the rich cultural repository of English, and its significance in international academic terms, suggest that it would "make sense" to nurture it in the English class, ensuring its continuing presence alongside, but distinct from, a post-apartheid South African English

no longer racially self-conscious or ethnically defined.....but one proudly and happily used by all those who have need of it as a means of communication

(Young;1989:140)

BIBLIOGRAPHY

- Alderson, C. 1985. Do the Tests Meet the Criteria?
In: Yeld, N. (ed.) 1985. Report on the Proceedings of the British Council Course on Communicative Language Testing at Lancaster.
Cape Town: University of Cape Town.
- Allen, D. 1985. Testing Written Communication.
In: Yeld, N. (ed.) 1985. Report on the Proceedings of the British Council Course on Communicative Language Testing at Lancaster.
Cape Town: University of Cape Town.
- Allen, J.P.B.; Davies, A. (eds). 1977. Testing and Experimental Methods. Vol 4 of Edinburgh Course in Applied Linguistics. Oxford: Oxford University Press.
- Archer, M.; Nolan-Woods, E. (eds) 1986. Practice Tests for Cambridge Certificate of Proficiency in English. Surrey: Thomas Nelson and Sons Ltd.
- Bakker, N. 1985. Informed Attending: An Alternative Approach to the Teaching of English Literature in South African High Schools. (Dissertation)
University of Cape Town.
- Bowers, C. A. 1984. The Promise of Theory.
New York: Longman.
- Bradbury, J.; Damerell, C.; Jackson, F.; Searle, R. 1990. ESL Issues Arising from the "Teach-Test-Teach" Programme. In: Chick K. (ed.) Searching For Relevance. Durban: SAALA
- Canale, M. 1983. General Outline of Communication Skills
In: TOEFL Research Report 17, 1985.
Princeton, New Jersey: Educational Testing Service.

- Canale, M; Swain, M. 1980. 'Theoretical bases of communicative approaches to second language teaching and testing.' Applied Linguistics 1:1-47. Oxford: Oxford University Press.
- Candlin, C. 1985. Explaining Communicative Competence: The Limits of Testability. In: Yeld N. (ed.) Report on the Proceedings of the British Council Course on Communicative Language Testing at Lancaster. Cape Town: University of Cape Town.
- Carroll, B.J. 1980. Testing Communicative Performance: An Interim Study. Oxford: Pergamon Press
- Chick, K. 1990. Language Acquisition or Language Socialization? In: Chick K. (ed.) Searching For Relevance. Durban: SAALA.
- Chomsky, N. 1965. Aspects of the Theory of Syntax. Cambridge, Mass.: M.I.T. Press
- Criteria For The Evaluation of South African Qualifications for Employment in Education. 1988. Pretoria: Department Of Education And Culture.
- Cummins J.; Swain, M. 1986. Bilingualism In Education. England: Longman.
- Davies, A. 1977. The Construction of Language Tests. In: Allen, J.P.B.; Davies, A. (eds). 1977. Testing and Experimental Methods. Vol 4 of Edinburgh Course in Applied Linguistics. Oxford: Oxford University Press.
- Davies, A. 1985. Communicative Language Testing: A Sceptical View. In: Yeld, N. (ed.) 1985. Report on the Proceedings of the British Council Course on Communicative Language Testing at Lancaster. Cape Town: University of Cape Town.

- Du Plessis, A.; du Pisani, T.; Plekker, S.J. 1988
Education and Manpower Development No.9.
 Bloemfontein: Research Institute for Education
 Planning.
- Edmunds, C. 1990. Spreading Our Resources. In:Chick,K. (ed)
Searching For Relevance. Durban:SAALA
- Forum (Journal). 1988. 'A Glossary of Testing Terms'.
 In: Forum vol. XXVI, no. 1, 1988.
 Washington: English Teaching Forum.
- Fowler, W.S.;Pidcock, J. 1985. New Proficiency
English. Surrey: Thomas Nelson and Sons Ltd.
- Greene, M. 1981a. 'Aesthetic Literacy in General Education.'
 In: Philosophy and Education, Eightieth
 Yearbook of the National Society for the
 Study of Education. Chicago, Illinois:
 University of Chicago Press.
- Hargreaves, P. 1985. The Relationship Between Communicative
Language Testing and Teaching.
 In:Yeld, N. (ed.) 1985. Report on the
Proceedings of the British Council Course on
Communicative Language Testing at Lancaster.
 Cape Town: University of Cape Town.
- Hartshorne, K.B. 1987. Language Policy in African Education
in South Africa 1910 - 1985.
 In: Young,D. (ed.) 1987. Bridging the Gap. Cape
 Town: Maskew Miller Longman.
- Hirsch, E.D. 1987. Cultural Literacy.
 Boston: Houghton Mifflin Company.
- Hitchcock, G. and Hughes, D. 1989 Research and the Teacher
 London: Routledge
- Hulstijn, J.H. 1985. Testing Second Language Proficiency
by Direct Procedures. In: Hyltenstam,K.;
 Pieneman, M. (eds) Modelling and Assessing
Second Language Acquisition. Clevedon, England:
 Multilingual Matters Ltd.

- Hyltenstam, K.; Pienemann, M. 1985. Introduction.
In: Hyltenstam, K.; Pienemann, M. (eds)
Modelling and Assessing Second Language
Acquisition. Clevedon, England:
Multilingual Matters Ltd.
- Hymes, D.H. 1971. On Communicative Competence. In:
Pride, J.B.; Holmes, J. (eds) Sociolinguistics.
Middlesex, England: Penguin Books.
- Ingram, D.E. 1985. Assessing Proficiency: an overview on
some aspects of testing. In: Hyltenstam, K.;
Pienemann, M. (eds) Modelling and Assessing
Second Language Acquisition. Clevedon, England;
Multilingual Matters Ltd.
- Ingram, E. 1977. Basic Concepts in Testing. In: Allen, J.;
Davies, A. (eds). 1977. The Edinburgh Course
in Applied Linguistics. Oxford: Oxford
University Press.
- King, C.; Stanley, N. 1987. Building Skills for the TOEFL.
Surrey: Thomas Nelson and Sons Ltd.
- Krashen, S.D. 1981. Second Language Acquisition and
Second Language Learning. Oxford: Pergamon.
- Mawasha, A.L. 1987. The Problem of English as a Second
Language Medium of Instruction in Black Schools
In Young, D (ed.) 1987. Language: Planning and
Medium in Education. Rondebosch, Cape:
The Language Education Unit and SAALA
- MELAB Information Bulletin 1989 - 1990. English language
Institute: University of Michigan.
- Morrow, K. 1985. What Makes a Test Communicative?
In: Yeld, N. (ed.) 1985. Report on the
Proceedings of the British Council Course on
Communicative Language Testing at Lancaster.
Cape Town: University of Cape Town.
- Munby, J. 1978. Communicative Syllabus Design. Cambridge:
Cambridge University Press.

- Ndebele, N. 1986. 'The English Language and Social Change in South Africa'. English Academy Review, 4. Johannesburg: The English Academy.
- Pit Corder, S. 1981. Error Analysis and Interlanguage Oxford: Oxford University Press.
- Robinson, A. and Katzman, J. 1986. Cracking The System: The SAT (Scholastic Aptitude Test). New York: Villard Books.
- Robinson, C. 1981. Themes for Proficiency. Oxford: Oxford University Press.
- Romaine, S. 1984 The Language of Children and Adolescents - The Acquisition of Communicative Competence Oxford: Blackwell.
- Royal Society of Arts (Report). Examinations in the Communicative Use of English as a Foreign Language, 1984. In: Yeld, N. (ed.) 1985. Report on the Proceedings of the British Council Course on Communicative Language Testing at Lancaster. Cape Town: University of Cape Town.
- Saleemi, A.P. 1988. 'Language Testing: Some Fundamental Aspects.' In: Forum, vol. xxvi, 1988. Washington: English Teaching Forum.
- Sollenberger, H.E. 1978. 'Development and Current Use of the FSI Oral Interview test.' In: Clark, J.L.D. (ed.) 1978. Direct Testing of Speaking Proficiency: Theory and Application. Princeton, N.J.: Educational Testing Service.
- Spolsky, B. 1978. Advances in Language Testing. Arlington, Va: Centre for Applied Linguistics.
- Stern, H. H. 1983. Fundamental Concepts of Language Teaching. Oxford: Oxford University Press.

- Subcommittee of the Provincial Advisory Committee for Teacher Training. 1989. Minutes of Meeting, 8 September.
- The Associated Examining Board. 1984. Test in English for Educational Purposes (T.E.E.P.)
In: Yeld, N. (ed.) 1985. Report on the Proceedings of the British Council Course on Communicative Language Testing at Lancaster.
Cape Town: University of Cape Town.
- The Oxford Paperback Dictionary. 1979.
Oxford: Oxford University Press.
- TOEFL Research Reports, Report 17. 1985. TOEFL from a Communicative Viewpoint on Language Proficiency: A Working Paper. Princeton, New Jersey: Educational Testing Service.
- Ward, J. M. 1980 English Proficiency Through Practice. Surrey: Thomas Nelson and Sons Ltd.
- Widdowson, H.G. 1978. Teaching Language as Communication.
Oxford: Oxford University Press.
- Widdowson, H.G. 1979. Explorations in Applied Linguistics
Oxford: Oxford University Press.
- Yeld, N. 1985a. Proficiency Testing and the Language Needs of First-year Engineering Students.
In: Young, D. (ed.) 1985. U.C.T. Papers in Language Education. Cape Town:
Language Education Unit, UCT.
- Yeld, N. (ed.) 1985. Report on the Proceedings of the British Council Course on Communicative Language Testing at Lancaster.
Cape Town: University of Cape Town.
- Young, D. 1990. 'Whose Language?' In: South African Journal of Higher Education: Special Edition 1990
Pretoria: Committee of University Principals of South Africa.