A small-scale study of predictive validity

Gibson Ferguson and Elizabeth White
University of Edinburgh[1]

Abstract

Because the International English Language Testing System (IELTS) test is used to screen the English language proficiency of students seeking admission to undergraduate and postgraduate University courses, predictive validity is an important consideration. The research project reported here is one of a number of small-scale pieces of research designed to investigate IELTS' predictive validity.

Twenty-eight Masters degree students studying science subjects at Edinburgh University were followed over an academic year. Their initial IELTS scores were collected, they were re-tested in June and both sets of scores were correlated with their academic results. Students and supervisors were also asked to complete several questionnaires. The aim here was to document students' problems and to identify factors which in individual cases may have contributed to poor academic performance.

Initial IELTS scores correlated 0.39 with academic outcome and 0.46 on re-testing in June. The figures are similar to other predictive validity studies in indicating a weakly positive relationship between language proficiency and academic outcome, and in that sense they are satisfactory for IELTS predictive validity. We find that a variety of factors are implicated in academic failure: academic ability, language proficiency, personal circumstance and traits; but these are inter-related in complex ways. The main problems students report relate mainly to the intensive character of a one year Masters course. Academic writing, particularly in exam conditions, emerges as a significant concern for many students. We

---

[1] For further information contact Dr. Gibson Ferguson, Department of Theoretical and Applied Linguistics, 14 Buccleugh Place, University of Edinburgh. Email: g.ferguson@holyrood.ac.ed.uk

also discuss issues of adjustment to a different academic culture, and problems concerning mixed groups of home and overseas students.

1. Introduction: Predictive validity

The International English Language Testing System (hereafter IELTS) is one of many tests used to screen applicants for employment or educational opportunities. Effective performance of this function rests on the degree to which the test is capable of differentiating those who will be successful in the criterion situation from those who will not. This is a matter of predictive validity, which we will define here as the degree to which scores on a test predict subsequent performance in the criterion situation. It is usually estimated by correlating test scores with some independent measure of that which the test is intended to predict.

There are, however, a number of well-known difficulties in estimating predictive validity, especially when, as with IELTS, one is interested in the test's utility as an indicator of subsequent academic performance. These may be summarized as follows:

a) the 'pre-selection' problem. This arises from the simultaneous use of the test as both predictor and selection instrument as happens when, as with IELTS, the test is already operational. Put another way, the prior use of the test to exclude some potential university students leaves one with a truncated sample where the range of scores on both the test and the criterion measure is curtailed. The likely effect is a depressed validity coefficient. Short of admitting students whatever their IELTS score and seeing how they perform on the criterion, there seem to be few practical ways of overcoming this difficulty.

b) The relationship between the test and the criterion. The use of a language proficiency test to predict the subsequent academic performance of overseas students is problematic in that the criterion is complex, clearly depending on a number of traits over and above language proficiency – motivation, scholastic aptitude, previous knowledge of the subject, for example. It is only to be expected, then, that language proficiency test scores will account for a limited portion of variance in academic outcome, and this, as various studies have shown (e.g. Criper and Davies 1988, Graham 1987), turns out in fact to be the case. The typical reported correlation between test and

criterion in these studies is of the order of 0.3. The question remains, therefore, of the most suitable criterion measure.

Another aspect of this problem is that the relationship between language proficiency and the academic performance may be non-linear. Thus, once a threshold of language proficiency is attained, no further gains from yet higher proficiency might accrue to academic performance. If true, this would mean that measures assuming a linear relation, such as the Pearson product moment correlation, would no longer be useful.

c) the time lapse problem. What distinguishes predictive validity from its criterion-related cousin, concurrent validity, is the time lapse between the test (the predictor) and the criterion. Of itself, this interval may attenuate the predictive power of the test. For example, over a period of months an individual's language proficiency may improve, or even deteriorate, at differential rates; a factor which may, and has served, as grounds for repeating the predictor closer in time to the criterion. Alternatively, events may intervene between the test and the criterion which fundamentally alter the correlation between the two. For example, the IELTS test, like some other EAP tests, has a diagnostic as well as a screening function and is used to serve notice to some individuals that they need remedial language tuition. To the extent that this is successful it will weaken the predictive validity of the original test. As Pollitt (1988: 62) has remarked of IELTS's predecessor, ELTS: *"Any intelligent use of ELTS will reduce its predictive power"*.

Despite these various obstacles to the estimation of predictive validity - from which, incidentally, the present predictive validity study is not immune - investigation of the test's relation to the criterion is still necessary if the test is to be genuinely useful as a screening device. Moreover, even if the correlation of the test with criterion is low – around 0.3 – it may still, as Anastasi (1988) shows, be useful when combined with other indicators. The argument goes roughly as follows: if we admit 100 students whatever their IELTS score we may end up with a successful academic outcome rate of 70%. On the other hand, if we exclude 20 of these students on language grounds, we may - depending on the ratio of false to valid rejections – end up, say, with a 60/80 or 75% academic success rate. The increment of 5% in successful academic outcome is attributable to the use of the language test to screen applicants – a possibly worthwhile

consequence given the financial and personal cost of failure on an academic course of study.

A final point is that establishing a test's predictive validity does not establish a test's overall validity. For that many other kinds of validation evidence are necessary. What it does show, however, is that the test has some utility as a practical screening device, in which case a more apt term for predictive validity might be predictive utility.

Let us turn now from these general opening remarks on the nature of predictive validity to the specific context of the study reported below.

2.   Context of the research

The major purpose of IELTS, like its predecessor ELTS, is to screen the language proficiency of overseas students seeking admission to academic courses in a number of English-speaking countries. Given this purpose, it is appropriate that the test's predictive validity - its validity as a pragmatic device for predicting academic outcome - is central to establishing overall validity and utility.

The research reported in this paper was commissioned by UCLES (University of Cambridge Local Examinations Syndicate) as one of a number of studies of IELTS' predictive validity and was conducted between October 1992 and September 1993. The study involved tracking a cohort of 28 M.Sc. students at the University of Edinburgh following a one year taught Masters courses in Life Sciences subjects (e.g. veterinary science). The limitations of the small but homogeneous sample are discussed below.

3.   Research objectives

Our brief, and key objective, was to investigate the predictive validity of IELTS, given all the limitations of the sample. This included the identification of cut-off points - test scores below which risk of failure would be great enough to render admission. inadvisable.

The second part of the study exploited some of the advantages of a small sample, which are that it allows a closer study of individuals

over an academic year. We took the opportunity to search out reasons for individual students' academic success or failure and to document the problems encountered by the group. The aim here was to contribute evidence to the wider debate on the role of language proficiency in the academic performance of overseas students. This is an issue of obvious relevance to the predictive validity of tests of academic English.

4.   Research design and procedure

4.1.   Overview

The design of the predictive validity component (hereafter referred to as Part One of the study) was straightforward. We collected initial entry scores on IELTS, retested the sample with IELTS close to their final examination and gathered their academic results. We then analysed the relationship between these measures statistically.

In the second part of the study (hereafter Part Two) the style of the research was more 'qualitative'. We investigated through questionnaires and interviews what students and supervisors believed to have influenced academic outcomes. The intention was to inform the predictive validity findings.

4.2.   The sample

4.2.1. Selection criteria

Students were selected for the sample on the following conditions:

(i)   they should be overseas students who had taken the IELTS test and whose test and sub-test scores were accessible

(ii)   they should be students following a taught postgraduate M.Sc. degree course of one year's duration

(iii)   they should have taken the Life Sciences module of IELTS (Module B) and be studying a Life Sciences subject in the Faculty of Science. This group was chosen because the Life Sciences M.Sc. courses have high populations of overseas students; there was also an advantage in that many of these students attended our pre-sessional EAP courses at the Institute for Applied Language Studies (IALS), and so we

were already in contact with students and with their supervisors

(iv)   within one Department there should be a group of not less than 3 students meeting the above criteria (for logistical reasons it was not worth our following single individuals in different departments)

(v)   they should be willing (along with their supervisors) to participate in the study

(vi)   they should be studying in Edinburgh, at Edinburgh University (again, this condition was imposed by budgetary and practical considerations).

Initially, we found 25 students who satisfied all the conditions. One subsequently withdrew because of illness, and was removed from the sample, leaving us with 24. In view of the resulting small sample size, we decided to relax condition (iii) and include a group of 4 students studying in the Department of Artificial Intelligence, who had originally taken the Sciences Module of IELTS.

4.2.2. Nature of the sample

Our sample, then, consisted of 28 students divided among M.Sc. courses as follows:

M.Sc. Course	Abbreviation	Numbers
Seed Technology	Seeds	6
Resource Management and Forestry	RES	5
Tropical Veterinary Medicine	TVM	6
Tropical Animal Production and Health	TAPH	4
Animal Behaviour and Welfare	ABAW	3
Knowledge-Based Systems	AI	4

Table 1: Distribution of Subjects by Field of Study (Department)

Country of Origin	Numbers	Total
Africa:		
Mozambique	3	
Ethiopia	3	6
East Asia:		
Indonesia	1	
Malaysia	1	2
South Asia:		
Bangladesh	3	
Nepal	3	
Sri Lanka	1	7
Middle East:		
Yemen	1	1
Latin America:		
Mexico	4	
Uruguay	1	
Cost Rica	1	6
Europe:		
Spain	2	
Portugal	1	
Poland	1	
Germany	1	
Denmark	1	6

Table 2: Subjects by Country of Origin

Of the 28 students in the experiment group, 6 were from Europe, 6 from Africa, 6 from Latin America, 7 from South Asia, 2 from East Asia and 1 from the Middle East. This distribution was not equal over the courses: all four AI students were from Latin America, and all three ABAW students were from Europe, while other courses had a more even distribution of country of origin.

The median age of subjects in this study was 30.5 years, somewhat older than the Faculty mean of 26-28 for taught M.Sc. courses. Our impression, which was also shared by the subjects (see below) is

that overseas students on such courses tend to be not only older but also to have more work experience than British students.

The men in the group of subjects outnumbered the women by almost two to one, reflecting the general preponderance of male students, particularly from the developing world, on post-graduate science courses. Within the individual courses, however, there were different distributions. AI was predominantly male, while ABAW had a small majority of women overall. The table below shows the distribution of subjects by sex within the different courses:

Course	M	F
TAPH	4	0
TVM	4	2
SEEDS	5	1
RES	4	1
ABAW	0	3
AI	3	1

Table 3: Distribution of Subjects by Course and by Sex

Table 4 on the following two pages gives a general background to each of the subjects of the study, with details of professional and academic backgrounds, previous experience of study abroad, and previous language training.

The previous academic qualification, though far from informative, was the only data we had for assessing the previous academic performance of all the subjects in the sample. The table shows that all but one of the subjects had attained at least a B.Sc. in their subject before coming to Edinburgh. Two had only a teaching Diploma, and were allowed onto their courses on exceptional grounds. Two had M.Sc. qualifications in other, related subjects.

St. No.	Qualification	Employment	Prior OS experience	English language training
1.	B.Sc.	Scientific Officer	None	School & university 3 months British Council 8 weeks pre-course EAP at IALS
2.	B.Sc. M.Sc.	Research Officer	None	School & university 16 weeks British Council and 1 month pre-course EAP at IALS
3.	Diploma	Ministry Official	None	School & university 8 weeks pre-course EAP at JALS
4.	B.Sc.	Ministry Official	Short courses in Europe	School & university 100 hours British Council
5.	M.Sc.	University Researcher	None	School & university British Council EAP & BCE
6.	B.Sc.	Assistant Technologist	None	School & university 100 hours British Council
7.	B.Sc. Diploma	Ministry Official	Short courses in East Asia and Europe	School & university 8 months part-time TOEFL & EAP in Indonesia
8.	B.Sc.	Development Work	3 years work in Tanzania	School & university 3 years with English as working language

...continued

St. No.	Qualification	Employment	Prior OS experience	English language training
9.	B.Sc.	Assistant on NGO	None	School & university Married to an English person
10.	B.Sc.	Assistant Lecturer	None	School & university British Council EAP 3 months pre-course EAP at IALS
11.	B.Sc.	Assistant Lecturer	Short courses in Africa & Norway	School & university Private language classes overseas
12.	DVM	Ministry Officer	None	School & university 3 months pre-course EAP at IALS
13.	DVM M.Sc.	Veterinary Surgeon	None	School & university 7 weeks pre-course EAP at IALS
14.	DVM	Project Co-ordinator	6 years USSR	English-medium school
15.	DVM	Provincial Officer	None	School & university 150 hours study skills overseas 12 weeks BAY at IALS
16.	DVM	Regional Co-ordinator	5 years Ukraine	English-medium school
17.	B.Sc.	Regional Field Officer	5 years Saudi Arabia	School & university 9 weeks (1991) and 13 weeks (1992) pre-course EAP at IALS

...continued

St. No.	Qualification	Employment	Prior OS experience	English language training
18.	Diploma	Teacher	None	School & private classes overseas
19.	B.Sc.	Ministry official	None	School & university 4 months British Council EAP
20.	B.Sc.	Business	6 months work in Japan	School & university BC Business English 1 month pre-course EAP
21.	DVM	Veterinary Surgeon	3 years work in Britain	School & university
22.	B.Sc.	Research associate	1 year work in Britain	School & university
23.	B.Sc.	Teaching	None	School & I year part-time private classes overseas
24.	B.Sc.	None	I year EFL in Edinburgh	School & university 1 year EFL in Edinburgh
25.	B.Sc.	Researcher	Short course in USA	School & university
26.	B.Sc.	Researcher	None	School, university and work
27.	B.Sc.	Programmer	None	School & university 3 months British Council
28	B.Sc.	Computer consultant	7 months in Japan	School & university

Table 4. Subjects by background

The amount of pre-course English tuition varies considerably, and there is no consistent pattern to the English language tuition taken by subjects. Where the pre-course IELTS test result requires a student to take further language classes before beginning his course, these classes were provided either by the British Council in the home country, or by the Institute of Applied Language Studies in Edinburgh, or both.

Most of the subjects had held posts of seniority in their own countries before coming on the course; this implies both that they have a solid professional background to work from, and that they had spent some time away from academic study.

13 of the 28 students had previously worked or studied abroad. This did not always involve studying in English (two of the group had spent long periods in the Soviet Union, during which time they had not necessarily kept up the English they learned at school), but it does mean that they were familiar with the experience of life outside their own country. Section 6.2.3 discusses the problems facing those who had not had this experience, and the implications of these problems for their academic performance.

4.2.3. Strengths and limitations of the sample

The sample size is small (at the lower limit of what is acceptable for most kinds of statistical manipulation) and this makes generalisation difficult. It also precludes more advanced statistical analysis (multiple regression, factor analysis). We hope however that it will be possible to amplify the restricted picture we paint with data from similar predictive validity studies of IELTS being undertaken elsewhere.

A small sample, as already noted, is however not disadvantageous in all respects. Previous validity studies (e.g. Criper & Davies 1985) have occasionally been criticised for an overly heterogeneous sample of subjects from many different institutions and departments. Such heterogeneity has been assumed to complicate the definition of the academic outcome criterion. Our sample, on the other hand, is relatively homogenous in that it consists of students who have all a) studied a science subject at Edinburgh (most of them a life science) and b) followed a postgraduate taught Masters course of one year's

duration. We can therefore be slightly more confident that an academic pass has roughly the same meaning across our sample.

There are other advantages. Larger studies (e.g. Criper and Davies 1988, Lynch 1994) have been obliged to adopt a somewhat coarse outcome criterion: pass or fail (or sometimes a three level criterion). This masks much variation in performance: a M.Sc. pass does not necessarily mean good performance. The size of our sample allowed us to refine our outcome criterion, as explained below. We were also able to follow our subjects much more closely throughout the year than would have been possible with a larger, more heterogeneous sample. We were able to monitor the subjects' academic performance over the year and the evolution of their opinions about their course, their level of English, their academic and other problems. In short, we were able to adopt a qualitative approach which helped to illuminate our quantitative results.

As in other predictive validity studies, our sample is biased in 2 ways (both in the predictor and the criterion):

(i)   students with high IELTS scores are over-represented in our sample. 68% of our IELTS sample achieved band 6 or over. The bias is, however, inescapable since few students with low IELTS scores are admitted to British postgraduate courses.

(ii)   the sample is not equally distributed between academic successes and failures. In our sample of 28, there are only six academic failures, though our sample is, relatively speaking, failure prone. 25% in the sample failed, which is high in comparison with the ELTS study's 19% (2 level outcome). This is unfortunate for the individuals but fortuitous in the context of our study, since it better allows us to examine the correlates of failure.

4.3.   Administering the predictor

Subjects sat for the IELTS (the predictor) on two separate occasions.

The first occasion (Time 1) was prior to the students' arrival in the UK. In most cases the IELTS test was taken in 1992, shortly before departure, but in a small number of cases it had been taken in 1991.

This was because some students who had expected to study in 1991 were in fact delayed.

The second occasion (Time 2) was later in the year close to the time of their June Diploma result. Of the 28 students in our sample, we were able to re-test 26 in June, including 5 of the 6 who did not obtain the 60% pass level. Our decision to retest is explained below.

Predictions of academic outcome on the basis of language test scores at the beginning of the academic year are probably weakened by the long time interval between initial test and academic result (in our case 9-12 months). First, during this time there are differential rates of learning English among the student population. In particular, students with little previous exposure to English (i.e. from an 'EFL' country) may make rapid progress when they are in an environment of constant exposure. Equally, other students may be already near 'their best proficiency' and make little progress. Second, the IELTS test, sometimes together with other locally-designed tests, is used to make recommendations about pre-sessional and in-sessional language tuition. If that tuition is effective, students' proficiency will improve, so weakening the predictive power of the initial test.

4.4. Choosing the criterion

4.4.1. Definition of academic outcome

Defining the academic outcome criterion has been an issue of concern in predictive validity studies. In our study we define the criterion in a variety of ways as follows:

(i)   the final academic percentage attained by the student at the end of his/her Diploma course in June. This percentage combines both coursework and final examination marks. This is our basic and most important criterion measure.

(ii) pass or failure at the end of the Diploma course in June (a 2-level measure).

(iii) academic outcome by 4 bands (a 4 level measure) divided as follows:

50-59% = 1
60-64% = 2
65-69% = 3
70%   = 4

(iv)   final examination scores attained by the student in June examinations at the end of his/her Diploma course.

(v)   final coursework scores attained by the student by the end of the Diploma course in June.

There are two further issues that require comment: the definition of failure and the comparability of academic marks/scores across Departments.

4.4.2. *Definition of failure*

For the purposes of this study we define failure as failure to proceed from the Diploma stage to the Dissertation stage of the M.Sc. programme. At Edinburgh University, M.Sc. level courses typically divide into a diploma stage (October - June) and a dissertation stage (July - September). We are aware that obtaining a diploma is not in formal terms an academic failure. Nevertheless, it is commonly regarded as such by students who enrol on, and are sponsored for, a Masters level course. This is evidenced by their obvious chagrin when they are refused permission to proceed to the Dissertation, and by the number of appeals against such refusals. (One student did in fact win an academic appeal allowing her to proceed to a Dissertation. But she is nevertheless counted as a failure.)

4.4.3. *Comparability of academic outcome*

The issue of comparability of academic outcome merits more extended comment. At the University of Edinburgh academic results on Masters courses are to some extent comparable across Departments because a University-wide scale is in operation as follows:

Below 50%	Failure
50 - 59%	Diploma pass only
60% or above	Proceed to Dissertation for M.Sc.
Pass in Dissertation	M.Sc. pass

In addition, some departments recognise distinction at the diploma stage - for marks in excess of 70%. In this study we identify 60% as the dividing line between success and failure; that is, those below 60% are defined as failures.

A common scale does not of course imply common marking standards. This is a problem besetting many predictive validity studies (c.f.: Criper & Davies 1988: 60, Lynch 1993). It seems likely in our case that there is some variation in marking standards within and across Departments, but we are unable to determine its extent. What we can say, however, is that because we are looking at one institution, and largely at students studying in Life Science subject departments, there is greater consistency in marking and in the standard of an academic pass. Our observations of the students and the Departments across the year also lend weight to our impression of relative consistency of academic measurement.

4.5.   Questionnaires and interviews

For Part One of our study, which was concerned with the predictive validity of the IELTS test, the only essential data were the subjects' IELTS scores and the measure of eventual academic outcome as described above. For Part Two which was concerned with exploring other factors such as motivation, homesickness etc. which had a bearing on academic outcome we gathered more qualitative data derived principally from questionnaires and interviews administered both to students and their supervisors over the course of the academic year (from 10/92 to 7/93). The table below gives the dates and contents of the questionnaires, along with the questionnaire numbers which are used for reference throughout this paper.

No.	Addressed to:	Date:	Contents:
1	Supervisors	Sept. 92	Details of admissions policy Marking schemes and assessment methods Views of the role of English language proficiency
2	Students	Oct. 92	Biodata Self-assessment of English reading/writing/listening/speaking/overall Self-assessment of academic ability Comments on the IELTS test
3	Supervisors	Jan. 93 After first term exams	Assessment of each student in terms of overall ability Motivation/ability in English/improvement in English Ranking of students' expected final performance
4	Students	Jan. 93 After first term exams	Assessment of progress on the course Ranking of problems encountered Assessment of enjoyment of the course Areas of difficulty
5	Students and supervisors	Feb. 93	Factors determining success on M.Sc. course Definitions of success
6	Supervisors	Apr. 93	Comments on teaching non-native speakers of English
7	Students	Apr. 93	Assessment of progress on the course Assessment of enjoyment Breakdown of language difficulties
8	Supervisors	June 93 After final course exams	Assessment of each student: ability/motivation/English ability/progress in English/effect of English ability on performance/success according to expectations/factors affecting performance Comments on the IELTS test

...continued

No.	Addressed to:	Date:	Contents:
9	Students	June 93 After final course exams	Self-assessment of English: reading/writing/ listening/speaking/overall Assessment of improvement in English over the year Comments on the role of English in their study Major problems over the year Advice to future students Comments on the IELTS test

Table 5: Research schedule - details of interviews and questionnaire administration

Some of the questions were repeated over the year; in particular, the students were asked at several points during the year to assess their English language ability and their progress on the course, so that a longitudinal comparison could be made.

The procedures for data gathering varied. The general pattern was that questionnaires were distributed to all subjects. When these had been completed, we would hold short semi-structured interviews with subjects and supervisors. The interview topics derived largely from the results of the questionnaires; where one or more students had brought up points which were worth discussing more widely, we included this in the interviews. And in turn, the discussions sometimes raised matters about which we wanted to have the opinion of all the subjects, thus providing material for the next questionnaire.

The interviews were with individual supervisors or with the group of students from each course. Both interviewers took notes throughout the interviews; these were collated to give the interview data used in the later part of this study.

5.   Findings (Part One)

5.1.   The relationship of IELTS (Time One) to academic outcome

Simple Pearson correlations were carried out for the whole sample between the overall IELTS band and the various versions of the academic outcome criterion - as well as between each IELTS sub-test band and academic outcome. The results are presented below.

r	Criterion
.39	Criterion 1 (Final percentage)
.46	Criterion 2 (pass/fail) Point-Biserial
.32	Criterion 3 (4 - level)
.31	Criterion 4 (exams)
.35	Criterion 5 (coursework)

Table 6: Correlations between IELTS (Time One) and academic
outcome (n = 28)

The correlation between IELTS (Time One) overall band score and the most important academic outcome criterion (Criterion 1) is .39. This is a lowish figure (about 16% of variance) - indicating that language level (as measured by IELTS) is not a powerful predictor of final academic performance. On the other hand, this figure is consistent with that of previous studies which cite a typical coefficient of around 0.3 (Criper & Davies 1988, Pollitt 1988) and may, as these authors have suggested, be as a high as one could expect given the many other factors contributing to academic success.

The table below shows the correlations between academic outcome (Criterion 1) and each of the four IELTS sub-tests. We have provided both Pearson and Spearman (in parenthesis) correlation indices.

Sub-test (n= 28)	Academic outcome (Criterion 1 = final %)
Reading	.28 (.28)
Writing	.26 (.16)
Listening	.40 (.29)
Speaking	.29 (.17)

Table 7: Correlations between IELTS (Time One) sub-tests and academic outcome

As can be seen, the correlations are again rather low but comparable to what one might expect. Listening correlates highest with academic outcome at .40, followed by speaking at .29. Writing correlates lowest at .26. (There is moreover no evidence of a significant difference between the general sub-tests [listening and speaking] and specialist sub-tests [reading and writing] in terms of the strength of the relationship with the academic criterion.)

5.2.   Cut-off points: Risk of failure & IELTS (Time One) score

Studies of predictive validity often identify some level of language proficiency below which the risk of academic failure increases sharply; the purpose is to establish a threshold level of language proficiency which candidates should attain to gain admittance to academic courses.

Table 8 below, therefore, shows the distribution of academic failures by IELTS overall band.

Band 6 seems to represent some sort of cross-over line. Of the nine students who scored below band 6 four failed; a failure rate of 44%. By contrast, of the nineteen students who scored band 6 or higher two failed, a failure rate of 10.5%. At or below band 5 there are actually more students who fail than who pass, so this may be a second lower threshold below which risk of failure increases sharply. A practical implication is that it is very risky to admit students to Life Sciences courses if they score at or below band 5 in IELTS.

IELTS band score	Frequency (cases)	Number of academic failures	Cumulative failures	Failure frequency % at each band score	Failure ratio *
4	1	1	1	100%	1/1 (100%)
4.5	1	0	1	0	1/2 (50%)
5.0	3	2	3	66%	3/5 (60%)
5.5	4	1	4	25%	4/9 (44%)
6.0	5	1	5	20%	5/14 (36%)
Subtotal 14		5	5	36%	36%
6.5	7	1	6	14.3%	6/21 (28.6%)
7.0	4	0	6	0%	6/25 (24%)
7.5	3	0	6	0%	6/28 (21%)
Subtotal 14		1	6	7%	7%
TOTAL 28		6	6		21%

* = *cumulative number of failures at each level/ cumulative total of students at each score level (%)*

Table 8: Distribution of academic failures by IELTS (Time One) overall band score

As usual, the small sample calls for cautious interpretation. Even one more failure at any of the IELTS band levels would considerably alter the percentages above. Nonetheless, our findings are consistent both with those of the larger ELTS validation study (Criper & Davies 1988: 79) which reported band 6 as "*some kind of changeover score*" and with current admissions practice at Edinburgh which typically takes band 6 to be an admissible level of proficiency. They also do not in any way contradict the plausible hypothesis that

there is a non-linear relationship between language proficiency and academic outcome; that is, that there is a level of proficiency below which the risk of failure increases sharply, and that language assumes a more important role in academic performance when proficiency is low.

If language proficiency is one factor in success/failure, it is clearly not the only factor. In our data, we find one student with IELTS 6.5 who failed academically, and there are other cases of students who had high IELTS scores who, although they passed, did not perform well academically. These cases and the factors involved in success and failure are discussed under Part Two (Section 6) below.

5.3.   The relationship of IELTS (Time 2) to academic outcome

Twenty-six of the students in the original sample of twenty-eight were retested in June 1993 and simple Pearson correlations were again calculated to establish the relationship between overall IELTS band score obtained at the second (June) sitting of the test and the various versions of the academic outcome criterion - as well as between each IELTS sub-test band and academic outcome. The results, showing correlations between first and second IELTS sittings and academic outcomes, are set out below:

IELTS (1992) Time One (n=28)	IELTS (1993) Time Two (n=26)	Outcome Criteria
.39	.46	Criterion 1 (Final)
.46	.41	Criterion 2 (Pass/Fail)
.32	.42	Criterion 3 (4 level)
.31	.36	Criterion 4 (exams)
.35 (n = 23)	.44 (n = 21)	Criterion 5 (coursework)

Table 9: Correlations of IELTS overall band (Time One & Time Two)
with academic outcome

The correlations between IELTS and academic outcome on the most important academic outcome criterion (Criterion 1) show a small gain (from .39 to .46) in predictive power from first to second IELTS testing. The improvement, however, is modest - from 16% of the variance in academic outcome to 21%. This finding is again similar to that of the larger IELTS validation study (Criper & Davies 1988).

The correlations between the first academic outcome criterion and each of the IELTS sub-tests are shown below. As can be seen, there is an improvement in predictive power across all 4 sub-tests. Interestingly, this is more marked in the case of the specialist sub-test (reading & writing) than in the non-specialist sub-tests (listening and speaking), where the correlations are only slightly higher.

	IELTS (Time One) (n=28)	IELTS (Time Two) (n=26)
Overall	.39	.46
Reading	.26	.40
Writing	.24	.43
Listening	.40	.49
Speaking	.29	.34

Table 10: Correlation IELTS (Time One and Time Two) Sub-Tests with Academic Outcome (Criterion 1)

Listening is, as before, the sub-test which correlates best with academic outcome (.49), and the correlations with writing (.43) and reading (.40) are stronger than previously. It is not clear why the writing and reading sub-tests have a higher correlation with outcome on second testing; unless it is because these are the skills which undergo greatest change as a result of intensive use during the year.

5.4.   Cut off points: Risk of failure and IELTS (Time Two) scores

Table 11 (below) displays the relationship between IELTS band score and academic failure over the complete range of scores. On second IELTS testing there were 14 students at or below band 6 (the same number as on first testing), and, of these, 5 failed, a failure rate of 36%. By contrast, 12 students obtained band 6.5 or higher, and, of these none failed, a failure rate of 0%.

The figures are fairly similar to those obtained on first testing. Table 12 summarises the position.

IELTS (Time Two) Band Scores	Frequency (cases)	Number of academic failures	Cumulative failures	Failure frequency % at each band score	Failure ratio *
4.5	1	1	1	100%	1/1 (100%)
5.0	2	0	1	0%	1/3 (33%)
5.5	7	2	3	28.5%	3/9 (33%)
6.0	4	2	5	50%	5/14 (36%)
Subtotal	14	5	5	36%	
6.5	0	0	5	0	5/14 (36%)
7	5	0	5	0	5/19 (26%)
7.5	6	0	5	0	5/25 (20%)
8	1	0	5	0	5/26 (19%)
Subtotal	12	0	5		
TOTAL	26	5	5		

* = cumulative number of failures at each band level/ cumulative total of students at each score level (%)

Table 11: Distribution of Academic Failures by IELTS (Time Two) Overall Band Score

	IELTS (Time One)		IELTS (Time Two)	
	Above Band 6	At or below Band 6	Above Band 6	At or below Band 6
Number of students	14 (50%)	14 (50%)	12 (46%)	14 (54%)
Number of failures	1	5	0	5
Failure ratio %	1:14 (7%)	5:14 (36%)	0:12 (0%)	5:14 (36%)

Table 12: Band 6 cut-off point; number of students (and failures) above and below Band 6

Allowing for the small sample size, which qualifies the findings, we can conclude that:

(i)   Band 6 or 6.5 again seems to correspond to some sort of crossover point between lower and higher risk of academic failure.

(ii)   Band 4.5 or 5 again seems to represent a second threshold below which risk of failure is very high.

(iii) there is evidence of a consistent relationship between low IELTS score and increased risk of academic failure.

These findings are consistent with those of the larger ELTS study (Criper & Davies 1988).

6.   Findings (Part Two)

6.1.   Case studies of academic failure

In this study, as already stated, the term failure is used in the technical sense of not achieving the 60% overall mark in coursework and exams which allows the student to proceed to the final M.Sc. dissertation.

The purpose in looking at the failures, of whom there are 6 in our study - a sizeable proportion, was to see if there were factors in common, or whether any conclusion could be drawn as to the factors negatively influencing a student's performance. We also wanted to test a finding of the ELTS validation report (Criper & Davies 1988: 113) that:

> language plays a role but not a major dominant role in academic success once the minimum threshold of adequate proficiency (4.5) has been reached. Thereafter, it is individual non-linguistic characteristics, both cognitive and affective, that determine success.

The small scale of our study allowed us to focus quite closely on the performance of individual students and supervisor's views of their performance. From our data we were able to extract information pertaining to the six students who failed (that is those who did not attain the 60 % pass level), and this is the basis for the case studies presented below.

6.1.1. Academic performance: Coursework and exams

The table below gives the academic performance marks for the six failures, with the figures (in percentages) for the total coursework marks, the exam marks and, where relevant, the oral exam marks. The weighting of coursework/exam work in the final academic score varies from course to course, though the tendency in the departments we studied was for exams to count more heavily. The IELTS writing score is given for comparison of relative performance.

Subject number	Course work %	Exams	Final academic outcome %	IELTS (Time 2) writing score
1	64.6	51.3	55	4.5
12	69.3	54	58	
15	60	49.6	51	4
18	54	52.7	55	5.5
24	70.66	34.5	50	5.5
27	80.64	47.44	57	6.5

Table 13: Academic performance scores for students failing to proceed to the M.Sc. Dissertation

All six subjects achieved a score over 50% in their overall academic scores (exams + coursework), thus qualifying for a Diploma pass. Overall, the coursework scores are higher than the exam scores; this was the case for all but 3 of the full sample. Supervisors commented that this pattern also applies, though much less markedly, to native speakers of English, and agreed with the subjects' frequently expressed opinion that the exam system does disadvantage the non-native speaker of English. Within this sub-group, however, the difference is very sharp, with two subjects having over 30% points of difference; the average is an exam mark 18% lower than a coursework mark. In one of the extreme cases, the very low exam mark was mainly due to the subject not completing the paper.

Of the three subjects who had an oral component to their final exam, two did poorly, while one achieved a higher mark than he had

achieved either in coursework or exams. The two who did poorly said, in interviews shortly after the exams, that they had been at a loss for words because of nervousness and worry about their exam performance - but so did many others, who did well enough in their oral exams.

These comparatively poor performances in final exams must be a worrying result for student and supervisor alike. Almost all the subjects, throughout the course, mentioned the difficulty they had in writing exam scripts. Training in this is available within the university to students, but few of the full group of subjects took up the opportunity. It might well be that more explicit training, more practice in writing exam answers, and more mock exams would be very helpful to students, and would perhaps redress the balance between older overseas students and students who have recently finished a first degree in Britain and who are therefore familiar with the skills and techniques required.

6.1.2. IELTS scores of those students not proceeding to dissertation

The following table shows the scores of the group of failures (by number) for IELTS (Time One) and IELTS (Time Two). Subject 12 could not be contacted to re-sit the IELTS test.

	IELTS (Time One)	IELTS (Time Two)
1	5	4
12	5.5	not available
15	4	5
18	5	5.5
24	6	6
27	6.5	6

Table 14: IELTS scores for students failing to proceed to the M.Sc. Dissertation

Four of the six have an initial IELTS entry score lower than the recommended entry level (6 or 6.5, according to department), and all but one had attended pre-sessional EAP courses in Edinburgh. Over

the course of the year, or between the two testings, only two improved their IELTS scores; one stayed the same, and two actually did worse in the second test.

6.1.3. Supervisors' comments

In discussions with supervisors over the year, all but one (no. 12) of these six subjects had been mentioned as being among the weakest in their groups. Only one (no. 1) was thought likely to fail. The reasons were various, and are discussed below; none were expected to fail purely for lack of ability in English.

At the end of the course, after exam results were known, supervisors were asked (Questionnaire 8) to judge each student's performance on the course in relation to the supervisor's expectations, and to the supervisor's judgement of their ability. Among the failures, all but one (no. 1) were expected to have done better.

The failure of no. 12 was a surprise to her supervisors, who had predicted at the beginning of the course that she would do well, and had ranked her among the better students in the group.

In the students' own self assessments of the progress they were making on the course, four tended to give non-committal answers throughout in questionnaires, and had not been noticeably more self-deprecatory than others in interviews; two (nos. 15 and 24) were unhappy about their chances.

6.1.4. Discussion of individual cases

The following case studies examine the reasons behind the poor performance of each of these subjects, attempting to identify those reasons which contributed most, and to see how far the IELTS assessment of their language ability has predicted their academic outcome. The six case studies are ordered with the lowest IELTS scores first.

<u>Subject Number 15</u>

	IELTS Scores				
	reading	writing	listening	speaking	overall
(Time One)	4.5	4	2.5	4	4
(Time Two)	5	4	4	7	5

Table 15: IELTS scores for Subject 15

This subject had the lowest IELTS entry score of the whole sample. It is unusual for a candidate with so low a score to be admitted to the university; he was given 150 hours British Council tuition in country, and three months pre-sessional EAP training before beginning the course to improve his level. The tutor's reports on his performance during the EAP course are encouraging, finding that his overall ability in English was adequate for study, his motivation was good and that he was a fluent and interesting speaker, willing to venture his own ideas and arguments. The assessment marks they gave bear this out.

He himself was very much less confident, and even at the end of the three months' EAP was uncertain of his ability to follow the course. This lack of confidence in his English continued to be a difficulty. Early on, supervisors were worried about his slow progress both with the course and with his English, imputing it to low expectations of himself and troubled confidence. He said that he found himself very much disadvantaged by his English, and that he had no free time whatsoever, because the course reading was taking all his time.

He did not do well in the coursework or in the exams, doing particularly badly in the final oral exam. In the last interview, shortly before he knew his result, he was not at all sanguine about his chances of success. When asked what might have made a difference to his success on the course, he stated that his major problem had been his English, and that he had never felt able to study easily because of his poor English. He, naturally enough, was anxious to make it clear that he felt that lack of knowledge or ability in his subject had not been a problem.

His course director thought differently: he appeared to have worked well enough, but that lack of academic ability together with weak English and an associated lack of confidence had held him back.

There were other factors operating against him. He came from Lusophone Africa, and had had little previous exposure to English, and little background of Western academic culture - though these factors alone do not necessarily indicate poor performance; other Mozambicans in our sample did very well.

All in all, this seems to be a case where lack of academic ability, weak English as indicated by the low IELTS test score, and an associated lack of confidence had contributed considerably to his eventual failure. Poor English, whilst undoubtedly a contributory factor, cannot be held to be a dominant cause.

Subject Number 1

From the earliest days the prospects of this student had not been bright. His low score on the first sitting of IELTS (5) was not a good augury indicating that he was at risk linguistically and would need extra English tuition. In the end he did not take up the opportunity of extra tuition attending only one class.

The pre-sessional EAP tutors noted that his English was barely adequate for academic study, and noted also that the problem was perhaps not so much one of language as of academic ability.

	reading	writing	IELTS Scores listening	speaking	overall
(Time One)	4	4	6.5	6	5
(Time Two)	4	4	4.4	4	4

Table 16: IELTS scores for Subject 1

From an early stage of the academic course his supervisors identified him as a weak student, rating him the lowest of the group (six in our sample) on academic ability. This assessment was, of course, borne out by his eventual failure to proceed to dissertation.

Throughout the assessments made of this subject's work over the year, there was doubt as to how far the problem was linguistic or cognitive; his writing was characterised by a lack of coherence and by strange misunderstandings; these also appeared in some of his curiously inappropriate oral questions.

Both supervisors commented on his reticence and unforthcoming nature, observing that this made it difficult to help or to identify instances of misunderstanding. EAP pre-sessional tutors had also puzzled over a curious vacancy in his expression.

The subject himself thought that his difficulties were due to "lack of knowledge of system", and would have advised a new student to "have some knowledge about the courses and system"; his comments on the relevance of the IELTS test show that he thought that language was not his main problem:

> When studying in English and writing in English, overall
> English background is essential but I think, maybe not
> necessary good English language.

Ultimately, the supervisors hazarded the opinion that lack of academic ability was the main cause of this failure, and in this they are probably right. His weak English would certainly not have helped but is possibly part of the same factor of a lack of intellectual acuity rather than something independent. The subject's comment on *"lack of knowledge of system"* probably also reflects the fact that unfamiliarity with an academic culture favouring independent thought would not have been an advantage.

It is true that the subject did gain a diploma and thus experienced benefit from his course. On the other hand, his is a case where the relative lack of success seems at its most predictable.

Subject Number 12

| | IELTS Scores | | | | |
	reading	writing	listening	speaking	overall
(Time One)	5.5	5	6	5	5.5
(Time Two)			*not*	*available*	

Table 17: IELTS scores for Subject 12

The failure of this subject, a Bangladeshi woman, was a surprise and a disappointment to her supervisors who in the early part of the course had been pleased with her work and had predicted a successful outcome. They noted her natural ability and confidence and fully expected her to write a good dissertation.

The tutors on the pre-sessional EAP course too had thought her IELTS 1 score of 5.5 was a conservative estimate and had found her very bright, confident, and able. They did not consider that her English would be a handicap.

The subject herself was critical at first of the course, considering it suitable for those who were returning to teaching posts rather than for those engaged in work in the field but this was not repeated after the first term. She did not identify any problems with her English, though she found the amount of reading difficult to cope with:

> *The main problem is the amount of reading. In addition the problem of adjustment with the culture and the environment and also homesickness was disturbing to me.*

The turning point in her fortunes seems to have come in April when, according to her supervisors, a deterioration set in. She was joined by her family from Bangladesh. Her supervisors wondered whether extra family responsibilities might have been the main factor in causing her to lose her way in the final term, and to underachieve in the final exams. In the event she failed narrowly with 58% and was very distraught on receiving the news. She left soon after and could not be contacted to re-sit the IELTS test or to complete the final interviews.

None of the other subjects were in Edinburgh with their families. It cannot be said that family responsibilities necessarily harm academic performance. Many said that they might have worked better if their families had been with them, saying that homesickness and worry over the absent family had prevented them from studying well; one attributed his success mainly to the fact that he had a good family, who looked after themselves well in his absence. It must be remembered, though, that being a woman from her background would have probably entailed considerable domestic responsibilities.

This case seems a particularly unfortunate one. We have a student who by all accounts had sufficient academic ability to do well and whose English, though modest, was not rated a serious handicap. Her failure seems primarily due to personal circumstances - heavy family responsibilities - combined with possibly faltering motivation.

Subject Number 18

		IELTS Scores		
	reading	writing	listening	speaking	overall
(Time One)	4.5	6	3	7	5
(Time Two)	5.5	5	5	5.5	5.5

Table 18: IELTS scores for Subject 18

Number 18 was an interesting case. He found that he was placed on a different course from that which he had expected. This left him with a sense of grievance which he never shed throughout the year, and which led him sometimes to reject what he was studying as irrelevant to his needs, or as knowledge that was already familiar:

> *I'm doing all this about pigs. Pigs pigs pigs But I know I know I know I know about pigs. Almost I am a pig. I don't want to know more about pigs.*

He did not apply himself to working, participating very little in seminars and keeping somewhat apart from the course as a whole.

Part of this rejection may have been a cover for lack of academic ability. His supervisors were not entirely sure of his level of ability. In December 1992 they commented on a "lack of innate ability" but also noted then and in April 1993 that he was not lacking in confidence and had good computer and logical skills. There was perhaps an imbalance of abilities; he had quantitative abilities but was weak in other areas. By February 1993 his supervisors were worried, noting that he was bottom of the group and perhaps "might not go on to the Dissertation stage".

They also commented repeatedly on a quite specific difficulty with writing. The subject was a perfectionist who wrote very slowly and painstakingly, which meant not only that he missed many lectures to work on assignments but that his work was often handed in late. The same characteristic had earlier been noticed by tutors on the pre-sessional EAP course.

The subject himself identified his two major problems as being a lack of confidence to express his ideas, and a lack of style in writing. He was grateful for assistance with writing from one of his supervisors; his advice to new students was:

> *Try to practice a lot with the assistance of the supervisors both to improve writing quickly and to understand what they want from you.*

By his own report he was a perfectionist who set himself very high standards, which appeared to prevent him from writing both coherently and fast;

> *Sometimes when I have to write something I take everything out of my room except the computer and then I draw the curtains and wait until I write. Once I was two days and I did not write anything.*

His final result was a disappointment to his tutors, who felt that he could have done better.

Part of this subject's problem seems to have been an unwillingness to work determinedly at a course he felt was not relevant to his needs:

> *This course is for me more an English course that I think will be useful later on in my research. Perhaps around 25% of the course will be relevant for my future but the rest of the course is not relevant or I had enough knowledge from my previous background to repeat it again.*

This continued rejection showed a less than flexible approach to the course and to academic study in the UK, and an abdication of personal responsibility in his study. He was among the youngest of the group, and a lack of maturity may have contributed to this attitude.

A lack of flexibility is also suggested in his approach to his written work. He certainly did have a language problem associated with writing, and his initial IELTS score of 5 does indicate potential language difficulties. However, unlike some other students with a low initial level, he does not seem to have been able to make progress and increase his writing speed.

We conclude that his lack of commitment to the course allied to a specific difficulty with writing held him back. Weak academic ability and a lack of maturity were also adverse factors.

Subject Number 24

This subject had studied for a year in Edinburgh before beginning her course, attending English classes; this year did not seem to have made her proficient in written English, though her spoken English was fluent enough to mask the many errors she made in writing. She herself was not confident of her English at the beginning of the course:

> *The course is wonderful but I have to confess that the English makes me struggle especially with the writing. Also the problem I found was the short amount of time. We had a big amount of reading. Statistics is really difficult in my mother language, in English much much worse.*

Her supervisor confirmed that her English was not improving as much as expected perhaps because she hardly spoke English at home with her fiancé. This was causing her problems both in writing and in listening.

	reading	writing	listening	speaking	overall
	IELTS Scores				
(Time One)	5	6	5.5	7	6
(Time Two)	5.5	5	5.5	7	6

Table 19: IELTS scores for Subject 24

From early in the course her supervisor had identified her as among the weaker on the course, but her coursework was satisfactory.

Toward the end of the course the supervisor's worries increased; there were doubts about her pulling through the exams, and comments on disorganisation in her written work and muddled thinking.

At her last interview the subject herself was not confident of success, and put her expected failure down to poor English and great nervousness which had both affected her performance in the exams and had caused her to lose time for study earlier in the year. She assessed her progress in English in 'downbeat' terms:

> *I'm conscious of improvement in my English but I want more. English makes you insecure - the more you know, the more you know you need. It's like climbing mirrors.*

> *I thought I would feel more confident by now, but I'm hanging on, I've never been so pressured.*

> *I've always had bad study habits, but I can't leave everything to the last now. I can not and I must not.*

> *When they read my writing, they always ask, 'But what do you exactly mean?'.*

In the event, her failure in the June exams was largely due to a disastrously weak performance in one paper. Her coursework was satisfactory.

We have here, then, a student who seemed reasonably bright, had spent a year in Edinburgh prior to the course, and had a satisfactory initial IELTS score of 6. Given this, her failure seems somewhat surprising. Probably, her low academic ability relative to her peer group was the main factor. But her lack of progress in English, contributing to nervousness, which again affected her capacity to study, would not have been helpful.

We also believe that the fact that she was on a course with a majority of British (native) speakers might have worked to her disadvantage. First, because it would have affected her confidence in her English, which would have seemed all the weaker in comparison with the others; and second, because her weak exam

performance would also have been cast into a even worse perspective by comparison with the stronger majority.

Essentially, her failure can be attributed to lower academic ability, problems with English and poor study habits.

Subject Number 27

	IELTS Scores				
	reading	writing	listening	speaking	overall
(Time One)	6.5	6	5	7	6.5
(Time Two)	6	5	6	7	6

Table 20: IELTS scores for Subject 27

As an AI student, this subject was admitted to our study rather later; so we do not have a full set of data about her early progress. We should also explain that, although she did not attain the necessary 60% in June, she appealed successfully against the markers' verdict and went on to write a successful M.Sc. dissertation.

Despite a respectably high score in her first IELTS test, this subject felt on interview that English, and in particular listening, was a difficulty to her: asked whether she was happy with her performance on the course, she answered:

> *No. I didn't expect to have so much problems with the language. I expected to improve it much faster and hence a better performance on the course. Understanding was a real handicap.*

Her supervisor tended to agree with her assessment. Both in English and in academic study she was rated as well below average in ability. Her own perception of her relative standing may have exacerbated a nervous and unconfident disposition. Before the exams she had personal and health problems which may have affected her performance; it appears, in fact, these were weighed in her successful appeal.

Whilst her coursework marks held up well, her exam results were poor placing her near the bottom of the group of 40.

Although she herself thought she had been seriously hampered by problems with English, it seems unlikely that this was the principal reason for her relatively poor performance. Her IELTS score of 6.5 should indicate a reasonable level of proficiency, and in interviews and questionnaires her English did not appear particularly weak. It may be that because she was on a course with a substantial majority of native speakers of English, and the other non-native speakers in her group were very fluent; that she perceived her language ability in relation to that of others in her group.

It appears, then, that low academic ability in relation to others of high ability, her health problems, and her nervous disposition will have contributed to her failure to do as well as might have been hoped. Just as with subject 24, it is also possible that the high standard of the group highlighted her relative weakness.

6.2.   A summary of factors contributing to academic failure

It is not an easy matter to identify causes for academic failure or success. Failure is in the majority of cases multifactorial, and, it can be difficult to distinguish one factor from another. Take the crucial distinction between language ability and academic ability; Criper & Davies (1988: 58) point out that we cannot establish the contribution of language to academic failure as distinct from academic ability because we lack an independent measure of that ability, or of subject knowledge. Moreover, language and academic ability probably overlap; one tends to go with the other.

There are other difficulties. In any individual case factors will be causally interrelated in complex ways; stresses from academic difficulty will lead, for example, to anxiety in a student's personal life and these may, in turn, feedback into academic performance. The relative influence of the factors may vary over the course of the year; what is true at the end of the course may not apply over the whole year of the student's experience. There are also methodological difficulties of self-report; students experiencing academic problems may wrongly identify the problem as one of language since this involves less loss of face.

Similar considerations apply to the other factors implicated in success or failure: motivation, diligence, adaptability, familiarity with the academic culture, health, personal stability, relationships with tutors and so on. They can turn out to be difficult to measure, and estimating the contribution of any single one is problematic.

Such problems have, however, rightly not deterred the study of the contribution of the various factors to academic success or failure (see Kinnell 1990, Blue 1993). Our own case studies contribute further evidence to the discussion. Small as the sample is, the depth of the study, following a small student group closely over a year, allows us to make reasonably plausible inferences as to cause of failure in individual instances.

6.2.1. Academic ability

A core factor in the majority of our cases appears to be a lack of academic ability. Five of the six were identified by their tutors fairly early on as experiencing problems with their study. In fact, by December 1992 - three months into their course - five of those eventually to fail were ranked near the bottom of the course group on academic ability by their respective supervisors.

It would seem that academic ability is a basic requirement for which proficiency in English and a satisfactory personal life will not compensate.

6.2.2. Affective and personal factors

While academic ability is necessary, it may not be sufficient, and the literature is full of additional factors said to be relevant to success or failure. Among them are a set, sometimes labelled affective factors, which include diverse personal traits, behaviours and circumstances such as motivation, adaptability, diligence, maturity, stability, perseverance. The evidence of our case studies tends on the whole to support the claim for their relevance since most of those who failed were either hindered by personal non-academic circumstances or by unhelpful behavioural traits.

The plainest case perhaps is the student whose family arrived in mid-course. Her supervisor felt fairly strongly that the resulting

distraction was a significant factor in her eventual failure. Another student did not apply himself wholeheartedly to the course feeling that it was inappropriate for him; he was also rather immature and his perfectionist and inflexible attitudes caused him to miss lectures in order to work on written assignments which were sometimes handed in late.

A third student did not communicate well with his tutors and this led to a perception of his reticence, even inscrutability. He tended not to admit misunderstanding nor to seek clarification, and when help was available with language tuition, he did not avail himself of the opportunity.

A fourth student, though bright and personable, was disorganised and on her own admission had poor study habits. Perhaps her disastrous performance in one exam paper, which let her down badly, reflects this conceptual and practical disorganisation. A fifth student suffered health difficulties which, allied to a nervous personality, may have adversely affected her performance.

An interesting aspect of these weaknesses is that many of them were noticed early on during the pre-sessional EAP course. Tony Lynch (personal communication) has remarked that individual idiosyncrasies first observed in pre-sessional courses (e.g. disorganisation, inflexibility, lack of punctuality in coming to class or handing assignments, unwillingness to seek clarification or advice) can herald more serious academic problems later in the course.

We would agree, then, with Tonkyn et al (1993: 48) that, apart from its utility for language tuition, the pre-sessional course has considerable predictive power:

> *We have in our hands a much better instrument (for predicting the academic performance of overseas students) of which we must make more refined use: it is called the pre-sessional course*

An implication is that EAP language tutors can help supervisors make early diagnoses of problem students and by so doing improve the chance of remedy.

6.2.3. Adjusting to a new academic culture

Another influence on student's academic performance often mentioned in the literature is adaptation to a different academic culture.

> *There is...a more fundamental problem for overseas students which is both unanticipated by them and seldom recognised by staff. Students are surprised to find that they must make significant shifts in their approach to study; their styles of learning and their ways of presenting ideas. These shifts are intellectual, rather than linguistic and social. They require a change from previous methods....* (Ballard & Clanchy 1984; cited in Jordan 1993: 72)

There is evidence in our study that at least some students did not understand the need for, or were incapable of, making the shift to more independent and critical ways of thinking. Perhaps this was what subject number 1 was hinting at when toward the end of the year he remarked that:

> [you should] *have some knowledge about the courses and system*

Adjustment to a different academic culture is perhaps more significant near the beginning of the academic year. This seems the case to judge from the comments made by subjects about their often disappointing results in the first term's exams, of which the following is one example:

> *In our country, exams focus on what has been discussed in class. Here they require also general basics which are not discussed in class, they need evidence of thinking and not learning. We didn't know this, we didn't have a training.*

It is, of course, a difficult matter to change one's learning habits or to have to discover that the learning habits which have so far stood one in good stead are no longer valued so highly.

A particular reason for making these comments here is that whereas 14 of the 28 in our sample had studied or worked abroad, 5 of the 6 who failed had not. We hypothesise that this may have worked

against them since living or studying abroad first of all gives an independence and a responsibility which may be applied to further experience of life abroad. Second, and more importantly, it may accustom students to a different system of education and different expectations of supervisors.

Several of the subjects who followed the pre-sessional EAP courses reported that they felt that perhaps the principal benefit of the course was to allow them time to find their feet in this country, and that they would otherwise have spent the first term of the M.Sc. course in settling down. We have no evidence, however, of the influence of the EAP course on adjustment to new study and thinking habits.

6.2.4. Course membership

A further factor to bear in mind in deliberating on the causes of individuals' failures, but one less commonly cited in the literature, is the composition of the student group. In three of the courses in our sample overseas students were numerically dominant, in one there was a rough balance, and in two the overseas students were in a minority. Two of the failures came from these latter courses and the other four from those courses with an overseas student majority. However, whereas these four were demonstrably weak in English with initial IELTS scores of 5, 5.5, 4 and 5, the other two failures, judging by our impression and by their IELTS scores (6 and 6.5) had a higher level of English and on casual impression were not markedly less intellectually able.

These facts lead us to hypothesise that being an overseas student on a course with a home student majority may be slightly more of a handicap than being an overseas student on a course with an overseas student majority. Reasons are suggested below. An underlying issue here may be one of comparability of standards across courses with differing student profiles.

One of the supervisors of a course with a home student majority observed that overseas students tended to occupy the lower ranks at final assessment. Asked to account for this phenomenon, he commented on the obvious differences of language and educational background but added that low levels of postgraduate study funding for home students meant that competition for such funding was

intense and only high calibre students, or the very self-motivated, tended to be admitted. At the same time the drive to income generation through overseas student recruitment may mean that some overseas students are admitted whose academic potential is questionable. The outcome can, as Elsey (Kinnell 1990) also notes, be *"a gap between two ability groups, with the more able being home students and the less able overseas students".*

By itself, these considerations do not explain how a home student majority on a course can adversely affect an overseas student. What we hypothesise, therefore, is that even though academic assessment on M.Sc. courses aspires to criterion-referencing and comparability across courses, it may, in the end, turn out to be norm-referenced in the sense that the norms of the whole group influence how those below the norm are perceived. For example, if there is a substantial number of high performing home students on the course, they can set a standard against which the relatively weak performance of an overseas student may be judged more unfavourably. Had the two failures in our sample from the course with a majority of home students been on a course with a majority of overseas students , it would not have been entirely surprising if they had passed.

Less controversially, it can be dispiriting and undermining of confidence to find yourself on a course where the majority, being native speakers, have better English and therefore seem to cope better and to be readier to ask questions. We suspect that a few students in our sample experienced such feelings, though it seems likely that only those already nervous or underconfident were adversely affected.

6.2.5. Language ability

Finally, we come to the key question of language ability and its influence on academic outcome. This is not easy to determine for, as Criper & Davies (1988: 58) point out, *"the investigator has no independent measure....of the subject knowledge/ability of the students being investigated".*

In the present study 4 of the 6 failures (subjects 1, 12, 15, 18) had IELTS scores below 6 indicating some linguistic weakness, and on interview 4 students (subjects 15, 18, 24, 27), including two who had

IELTS scores of 6 or over, cited language as one of the important factors adversely affecting their academic performance. Subject 27, the AI student with an initial IELTS score of 6.5, mentioned problems with listening; subject 24, with an IELTS score of 6, mentioned problems of writing, and reading speed. Subject 18 had persistent, quite serious problems with writing assignments and exams, and subject 15, with a very low initial IELTS score of 4, lacked linguistic confidence and attributed many of his academic difficulties to his weak English.

In June the supervisors were also asked to indicate for each student whether the student's level of English had adversely affected their academic result. The precise question was: *Do you think the student would have performed better if his/her level of English had been better?* Answers were given on a three point scale (0 = don't know; 1 = probably not better; 2 = probably better). The results are presented in the table below.

Altogether, 9 of the 28 students in the sample were identified by the supervisors as having been hindered by their level of English, and of these 5 failed to proceed beyond Diploma level.

	Don't know	Probably not better	Probably better	
Supervisor 1	0	5	1	6
Supervisor 2	0	3	2	5
Supervisor 3	0	4	2	6
Supervisor 4	0	2	2	4
Supervisor 5	1	1	1	3
Supervisor 6	1	2	1	4
TOTALS	2	17	9	28

Table 21: Supervisors' judgements on the influence of English on student academic performance

Interestingly, there is some evidence of divergence of opinion among subject lecturers; a further 2 lecturers were asked the same question about the same students on their particular course and their verdict

was different from the original one in a number of instances. The fact that within the same department lecturers differ over how much a student's level of language proficiency affects his performance is perhaps indicative of the difficulty of establishing the contribution of language to academic performance in general.

Whilst the evidence above suggests superficially that language has a considerable role in failure, if not success, there remain reasons to be cautious of claiming too much significance for its role. Students may impute the difficulties they experience to the most obvious cause, language; and they may have face-saving reasons for doing so. Also, knowing we were from a language institute, they may have given us answers they believed we wanted. Finally, there is always the danger of language specialists exaggerating the importance of language.

With all this in mind, we would conclude that language will have been one factor in the poor academic performance of these 6 subjects but by no means the most important one. Moreover, difficulties with English are linked to other factors in complex ways. For example, with subject 18, who had obvious difficulties with writing, it is likely that the linguistic problem was compounded by personal traits of inflexibility and perfectionism. And subject 24's poor written English will not have been helped by poor study habits and disorganisation. In the case of subject 1 there seems to have been a double problem of interdependent linguistic and conceptual weakness.

7.   Summary and conclusion

The main focus of the research has been the predictive validity of IELTS (Part One). We found that the first round of IELTS results, derived from the 1992 administration, correlated with our most sensitive measure of academic outcome (final result in %) at 0.39 (about 15% of variance). The figure is comparable to the 0.3 correlation cited in other predictive validity studies (e.g. Criper & Davies 1988). When the predictor was repeated in June 1993, much closer to the criterion, the correlation was 0.46 (about 20% of variance). The increase is visible, but relatively small, indicating that the variance in academic outcome accounted for by language proficiency for our group probably lies somewhere in the range 15-20%.

Of the IELTS sub-tests, listening correlated most strongly with academic outcome on both first and second test administrations (0.40 and 0.49 respectively). On second IELTS testing there tended to be, as one might expect, a closer relationship between each of the sub-tests and academic outcome, though the correlation with speaking remained relatively weak at 0.34.

We also considered the question of the most appropriate cut-off scores for students taking masters level courses in the Life Sciences. The evidence points to a cross-over level of about band 6: above this the numbers failing are negligible but below the risk of failure increases steadily. The small numbers clearly make a definitive conclusion impossible. Nevertheless, our findings are broadly in line with those of the much larger ELTS validation study (Criper and Davies 1988: 79) - "6.0 is some kind of changeover score" - and with policies advising a band score of 6.0 for admission. Put another way, the implication of our findings is that if a student is admitted with a score below 6, the institution or department is taking a greater risk of failure, and below 5 a sharply increased risk.

In Part Two of the study we looked more closely at those students who did not perform satisfactorily on the course. Six of the twenty-eight students in our sample failed to reach the 60% mark level qualifying them to proceed to the Masters degree dissertation. This is a relatively high proportion (21%).

As one might expect, a cocktail of factors contribute to failure, and it is not easy to distinguish one from another. It would appear, however, that a relative lack of academic ability is the dominant factor. That said, in the individual case it is not easy to separate academic and linguistic ability. Four of the six failures had low IELTS scores below band 6 and the other two scored exactly 6. Many of these six, and other students, claimed that low language proficiency obscured, and led to misperceptions of, their real academic ability. While this may be true to an extent; there are, of course, face-saving motives for such claims. Therefore, we hold to the view that academic ability is a basic requirement.

In the individual cases it is noticeable that lack of academic ability is often compounded by personal circumstances or behaviours: worries over health, excessive time looking after family dependants, poor study skills, inflexible attitudes to study, an

unwillingness to admit difficulty and seek help, for example, all have had a role in one or other of the individual failures.

Then, there is the matter of language. The quantitative study suggested that about 15% of variance in academic outcome is accounted for by variance in language proficiency, and the Criper & Davies study (1988) concluded that above a certain threshold of proficiency, language is a minor factor in academic success.

Sympathetic as we are with this viewpoint, we would wish to qualify it. Language may be more important in academic failure than success. Certainly, those who failed attributed many of their difficulties to weak language proficiency, and our case studies of failure show that some individuals had quite specific language problems, principally with academic writing. There are also cases where low language proficiency in relation to others undermined confidence and exacerbated feelings of tension and nervousness.

The supervisors too judged that 5 of the 6 who failed were hindered by language problems and might have done better academically had their English been better. We incline, therefore, to the view that in cases of failure at least language difficulties may have a more significant role than is sometimes argued.

To conclude, we have found with our sample that IELTS predicts academic performance satisfactorily, at least as well as similar tests. However, as our sample is small, caution is required, and it would be sensible to regard what we report above as a small contribution to the wider body of evidence needed to establish the validity of IELTS.

8.   References

Anastasi, A. (1988) *Psychological testing* (6th ed.). London: Macmillan.

Blue, G. (Ed.) (1993) Language learning and success: Studying through English. *Review of English Language Teaching 3,1.* London: Modern English Publications in association with the British Council.

Blue, G. (1993) Nothing succeeds like linguistic competence: The role of language in academic success. In G. Blue (Ed.) Language learning and success: Studying through English. *Review of English Language Teaching 3,1*. London: Modern English Publications in association with the British Council.

Criper, C. & Davies, A. (1988) *ELTS Validation Report 1(i)*. Cambridge: The British Council & the University of Cambridge Local Examinations Syndicate.

Elsey, B. (1990) Teaching and learning. In M. Kinnell (Ed.) *The learning experiences of overseas students*. (pp. 46-62). Buckingham: The Open University Press & SRHE.

Graham, J. G. (1987) English language proficiency and the prediction of academic success. *TESOL Quarterly 21, 3*: 505-21.

Hughes, A., Porter, D., & Weir, C. (Eds.) (1988) *The ELTS Validation Project: Research Report 1(ii)*. Cambridge: the British Council & the University of Cambridge Local Examinations Syndicate.

Jordan, R. R. (1993) Study skills: Experience and expectations. In G. Blue (Ed.) Language learning and success: Studying through English. *Review of English Language Teaching 3,1*: 70-79. London: Modern English Publications in association with the British Council.

Kinnell, M. (Ed.) (1990) *The learning experiences of overseas students*. Buckingham: The Open University Press & SRHE.

Light, R., Xu, M. & Mossop, J. (1987) English proficiency and academic performance of international students. *TESOL Quarterly 21, 2*: 251-261.

Lynch, T. (1994) The Test of English at Matriculation: Validation Report. *Edinburgh Working Papers in Applied Linguistics*. Edinburgh: University of Edinburgh.

Pollitt, A. (1988) Predictive validity. In A. Hughes, D. Porter & C. Weir (Eds.) *The ELTS Validation Report 1 (ii)* (pp. 62-65).

Cambridge: The British Council & the University of Cambridge Local Examinations Syndicate.

Tonkyn, A., Locke, C., Robinson, P. & Furneaux C. (1993) The EAP teacher: Prophet of doom or eternal optimist. In G. Blue (Ed.) Language learning and success: Studying through English. *Review of English Language Teaching* 3,1: 37-48. London: Modern English Publications in association with the British Council.