IELTS ™
English for International Opportunity

4. An Investigation of the Predictive Validity of IELTS amongst a Group of International Students studying at the University of Tasmania

Fiona Cotton and Frank Conrow
The University of Tasmania

4 An Investigation of the Predictive Validity of IELTS amongst a Group of International Students studying at the University of Tasmania

Fiona Cotton and Frank Conrow
The University of Tasmania

Abstract

The purpose of the present research study was to investigate the relationship between IELTS (International English Language Testing System) and academic outcomes, as well as the extent to which IELTS predicts the kinds of language difficulties international students experience while studying in Australia. Data was collected over one year from questionnaires returned by thirty-three students, from interviews with twenty-three of these students, and from surveys returned by thirty-four academic staff, two international student advisers and two English support tutors.

Correlations were calculated between the IELTS scores of the student group under investigation and three measures of academic achievement: Grade Point Averages, academic staff ratings of student performance and students' self-ratings of performance. No positive correlations were found overall. However, the reading and writing subtest scores correlated at 0.36 and 0.34 with staff ratings of academic achievement, and 0.46 and 0.39 with students' self estimates of academic performance in second semester.

No positive correlations were found between IELTS scores and language difficulties students reported with aspects of their coursework. Qualitative data indicated that language difficulties are one of many variables affecting academic achievement. Several key intervening variables were briefly investigated, namely, the amount of English language tuition received, motivation, cultural adjustment and welfare difficulties experienced by international students. In addition, students and staff reported their views of IELTS, which was generally perceived to be a fair test.

Final discussion focuses on some of the problems inherent in working with a small sample and on how the results may be interpreted. Recommendations are made for further research.

1.0   Introduction

1.1   The Context

The International English Language Testing System (IELTS) is a criterion referenced English language test which is used increasingly widely internationally, to assess the English language proficiency of international students whose native language is not English, and who wish to study at tertiary level in English medium countries. It is used by tertiary institutions particularly in Britain and Australia, as one of several alternative measures to determine the entry level or cut-off point at which the English proficiency of prospective students will hinder their ability to cope with their academic studies. As such, IELTS is an important test, since the careers and tertiary level studies of many depend on their ability to pass the IELTS test. It is evident that the use of such a test in this way rests upon the assumption that proficiency in English is significantly related to academic success. But to what extent is this so? What is the nature of the relationship between proficiency in English and academic success? To what extent does IELTS predict the ability of overseas students to cope with the academic demands of their coursework? The search for answers to these questions is important in order to establish the reliability and validity of IELTS.

1.2   Research Rationale

It has been recommended (Graham, 1987; Burns, 1991) that each institution would do well to conduct its own studies of the connection between language proficiency measures used and academic outcomes. This is important because of the difficulties of generalising findings from previous studies to larger populations. A predictive validity study at the University of Tasmania is appropriate from this perspective alone. In addition, findings from these predictive validity studies in different contexts, as Elder (1992) points out, contribute to the ongoing debate about the nature of language proficiency and the most appropriate ways of testing language, as well as whether performance in one context can be generalised to other contexts.

The accumulation of research findings from such small scale studies may also help to identify the kinds of cut-off points appropriate for different subject areas, which reduces the risk of students being excluded from particular courses for the wrong reasons and decreases the likelihood of students who have a high probability of failure being admitted to courses.

Since the completion of the IELTS predictive validity studies reported in the literature review, there have been several revisions to the format of the IELTS test. One main revision was that the reading and writing subtests are no longer related in terms of content. Candidates are no longer able to refer back to the reading section during the writing tasks. In addition, where there were previously four separate academic modules for the different subject disciplines, there is now only one academic module of the reading and writing subtests for all disciplines. These changes were instituted in April, 1995. A predictive validity study at this time would seem justified on the grounds that to the best of my knowledge, no other studies are yet available which have tested the relationship between performance on the revised version of IELTS and academic outcomes.

1.3 Research Objectives

The present research study was undertaken, with the support of IELTS Australia, as a pilot project to investigate the predictive validity of IELTS using a group of overseas students undertaking their first year of study at the University of Tasmania in 1996.

The objective of the study was to investigate the following research questions:

1. To what extent does the IELTS proficiency rating scale (the independent variable) predict the academic success of non-native speakers of English (the dependent variable) undertaking academic studies within an Australian University context ?

2. What is the strength of the relationship between IELTS bandscores achieved and subsequent academic performance outcomes both in terms of the students' subjective perception of their own performance, and also in terms of more formal objective measures of their achievement?

3. To what extent can IELTS bandscores predict the kinds of language difficulties international students encounter with various aspects of their studies?

4. Which key intervening variables appear to have the most effect on the relationship between IELTS bandscores achieved and subsequent academic outcomes in the population under investigation ?

It is not our intention to investigate the concurrent and construct validity of IELTS. It is beyond the scope of this project to investigate whether IELTS accurately measures language proficiency.

This report describes the present research study and summarises its main findings. Following the introduction, section two reviews the literature relating to research into the relationship between various measures of language proficiency and academic outcomes, and section three outlines the methods and procedures adopted for this study. Section four provides a comprehensive description of the population under investigation. The fifth section reports on the findings in relation to the predictive validity of IELTS, both quantitative and qualitative, and section six reports the findings in relation to IELTS face validity. The final section seven, summarises and discusses the findings and makes recommendations for future research.

2.0 Literature Review

2.1 Predictive Validity Studies

In an article by Graham (1987), a number of predictive validity studies were reviewed which attempted to analyse the relationship between various English proficiency test results and academic outcome. Roughly the same number of studies found no statistical significance (Mulligan, 1966; Sugimoto, 1966; Hwang and Dizney, 1970; Sharon, 1972; Shay, 1975; Wilcox, 1975; Gue and Holdaway, 1973: cited in Graham, 1987) as found statistically significant correlations (Burgess and Greis, 1970; Heil and Aleamoni, 1974; Baldauf and Dawson, 1980; Odunze, 1982; Ho and Spinks, 1985: cited in Graham, 1987). A number of other studies yielded either inconclusive results or gave mixed conclusions (Slark and Bateman, 1982; Bostic, 1981; Mestre, 1981; Light, Xu and Mossop, 1987: cited in Graham,1987). The mixed findings of these studies suggest that the relationship between

proficiency in English and academic outcome is more ambiguous than one might initially suppose.

Graham (op.cit) suggested various reasons as to why the question of the relationship between English proficiency and academic achievement is problematic. First, there is continued debate about the exact nature of language proficiency. The second problem relates to the difficulties of testing language proficiency and how it can be measured with a high degree of reliability and validity. Third, there are a number of moderating variables which affect student performance in the testing situation, and a number of intervening variables which affect students' academic performance. The nature of the relationship between all the variables is complex and therefore not easy to determine. Fourth, the question of what constitutes academic success is open to interpretation and is hard to define. Graham indicated that comparisons between the various studies mentioned, and attempts to generalise the findings to wider populations, are difficult precisely because of these issues.

2.2   English Language Proficiency

First, measures of English proficiency vary from study to study and reflect differing perceptions of what constitutes English proficiency. Thus TOEFL can be seen to reflect a different definition of proficiency to more direct measures of proficiency such as IELTS. The question of the validity and reliability of the various measures of proficiency used needs careful consideration.

2.3   Moderating and Intervening Variables

Predictive validity studies also differ in the degree to which the variables relating to performance in the proficiency test, and also to academic success, are considered and controlled. In his study, Odunze (1982: cited in Graham, 1987) blamed the failure of TOEFL to predict academic outcome on sociolinguistic factors, by suggesting that his Nigerian subjects' strong dislike of multiple choice questions affected their performance in the test. Some studies (Zheng Chen and Henning, 1985; Zeidner, 1987) have found English proficiency tests may exhibit cultural bias, which in turn may affect students' performance in the tests.

Alderson and Urquhart (1985) report research findings showing an interrelationship between background knowledge and linguistic competence which may also affect test results. Not unrelated to this, Light, Xu and Mossop (1987) found that the relationship between TOEFL and academic outcome varied according to students' area of study.

In some studies, aptitude tests and high school subject matter scores (Wilcox, 1975; Graham, 1984: cited in Graham, 1987) appeared to be better predictors of academic success than English proficiency tests. In other studies (Ho and Spinks, 1985: Gue and Holdaway, 1973: cited in Graham, 1987) personality and affective factors such as attitude, motivation and friendships were explored as important influences on academic outcome. More recently, evidence has been found (Jochem, Snippe, Smid and Verweij, 1996) that variables such as the age and gender of overseas students had an effect on academic achievement.

In an interesting study at the Australian National University, Burns (1991) found that the ability of overseas students to cope with the academic requirements of their studies was significantly affected by a number of variables other than proficiency in English. These variables ranged from financial worries, the level of interest and awareness shown by academic staff, the amount of preparation overseas students had for university study, to the

size of the course workload and family pressures on overseas students to perform well. Such factors cannot be ignored in any predictive validity studies.

2.4 Measures of Academic Performance

The criteria used for judging academic success vary in different predictive validity studies. For example, in some studies the criterion for academic success is simply pass or fail. In other studies the amount of work successfully completed is used as the criterion, while in still others, grades or percentages are used. In some studies, where postgraduate students form the sample, the perception of the academic supervisors is used as the main criterion for academic success. This makes any correlations calculated between English proficiency and academic success very difficult to interpret. It was recently pointed out that even GPAs (Grade Point Averages) can be problematic, as these can be calculated over different periods of time and over different numbers of examinations. It is important therefore to use more than one measure of academic achievement in predictive validity studies (Jochem, et al, 1996).

Another reason why some of the studies cited reported no significant correlations between English proficiency test scores and academic outcome may be related to the cut-off points which have been established by the different tertiary institutions. Students with lower scores were not permitted to enrol in courses. It is likely that lower levels of English proficiency are stronger predictors of academic outcomes, and that at higher levels English proficiency ceases to be a significant factor in determining academic success.

2.5 Predictive Validity Studies of IELTS

Most of the studies mentioned so far have not been related directly to IELTS. However, there have been a number of research studies which seek to investigate the predictive validity of IELTS in particular, and also its predecessor ELTS. As with the studies cited earlier, there is continued debate over the strength of the relationship between language proficiency, in this case measured by IELTS, and academic outcome. A number of studies reported a significant correlation (Davies and Criper, 1988; Ferguson and White, 1993; Elder, 1993; Bellingham, 1993) whilst others reported no significant correlation between IELTS and academic outcome (Gibson and Rusek, 1992; Fiocco, 1992).

The best known of these studies and most extensive is the Validation Project for the ELTS conducted by Davies and Criper (1988) which sought to investigate not only predictive validity but also construct, content and concurrent validity of ELTS. A non-representative sample of 720 subjects selected over two years (1984/85) was used, roughly distributed in terms of ELTS module taken, ELTS score, origin, age and gender. In both years, the sample subjects took the ELTS test at the beginning and at the end of the year of study. In addition, subjects were interviewed twice during the year as were their supervisors. Correlational analyses were calculated between the first and the second ELTS test scores and academic outcome, ELTS test scores and the supervisors' and students' self estimates of academic performance, as well as between final academic outcomes and supervisors' and students' self estimates of performance. Davies and Criper concluded that the contribution of language proficiency to academic outcome is about 10%, a correlation of 0.3. This suggests that language does contribute to academic success but does not play a major part.

How accurate a measure this is of the relationship between language proficiency and academic success, is as always clouded by the difficulties inherent in ensuring that all factors involved in such a study have been accounted for. In the proceedings of a conference held to consider the ELTS Validation Project Report, many issues were discussed including the

reliability of the questionnaire used for student self assessment and supervisors' assessment. It was proposed that if data on student self assessment could be collected more reliably, then this might be the most valid form of predictive assessment (Pollitt, 1988). But whatever measures are taken to make the predictive validity research design and implementation more rigorous, Pollitt (1988) stated that 'any intelligent use of ELTS will reduce its predictive power.'

More recently, the predictive validity of IELTS was investigated by Catherine Elder (1992) in a study conducted by the NLLIA Language Testing Research Centre at the University of Melbourne. Data on initial IELTS scores and subsequent academic progress at the end of both first and second semester was obtained from a small sample of overseas students undertaking diploma of education courses at a number of institutions in Melbourne. Whilst the sample was small (n = 32), the research findings indicated that there was a statistically significant correlation between global IELTS scores and first semester course progress ratings of .35 (p=<.05), and a correlation of .40 between the listening subtest and first semester academic ratings. These correlations became much weaker for second semester and were not statistically significant.

In addition, Elder investigated the relationship between IELTS scores and difficulties experienced by NESB (non-English speaking background) students in coping with the English demands of their coursework, but she pointed out that the correlations obtained on the relationship between aspects of coursework and IELTS scores have to be viewed with extreme caution, because of the disappointingly low return rate on the questionnaires (n=17). Even when this is taken into account, there was a correlation between the IELTS writing subtest scores and students' difficulties with writing assignments of .52, suggesting that the IELTS writing subtest may be able to predict subsequent difficulties experienced with written work. Similarly, a correlation of .44 was found for the reading subtest and coursework reading, as well as a correlation of .59 for the listening subtest score and comprehension of lectures and tutorials. With the exception of the speaking test, these results suggest that the IELTS subtest scores can predict subsequent language-related difficulties of coursework. It would be interesting to see if these relationships are corroborated in the present study.

A small scale investigation was conducted by Gibson and Rusek (1992) in South Australia entitled, 'The validity of an overall bandscore of 6.0 on the IELTS test as a predictor of adequate English language level appropriate for successful academic study.' A sample was collected of NESB students (n=63) who sat IELTS between December 1989 and February 1991, before entering one of the South Australian universities. These subjects were asked to complete a questionnaire on the same day they sat the IELTS test. 35 of these same students were subsequently interviewed at the end of the first semester, 1991. Academic supervisors of the 12 postgraduate students in the sample were also interviewed to obtain information on the progress of those students. Gibson and Rusek contributed to the discussion about the difficulties inherent in such a study, such as the degree of bias inherent in self selected subjects who present themselves for interview, and to the discussion about the types of affective variables which impinge on academic outcome. However, in investigating the relationship between IELTS scores and academic outcome at the end of the first semester, the criterion for academic success was very broadly defined. The measure the researchers used for academic success was simply permission to proceed to second semester. Since all the students were allowed to proceed to second semester even where they had failed units or dropped others, they were all deemed to be successful. From this evidence, the researchers cautiously suggested that IELTS scores did not predict subsequent academic success. Had the researchers had access to other more precise measures of academic outcome, they may have been able to make more of their data.

Data from a predictive validity study of the listening subtest, conducted by Denham and Oner (1992) at the University of Canberra was mainly qualitative. From interviews with the 34 subjects in their study, the researchers concluded that there was little relationship between IELTS listening subtest scores and subsequent difficulties with listening comprehension in various contexts. Calculation of the correlation between unit grades and the IELTS listening subtest scores yielded a result of .05 (n=25). Denham and Oner calculated several other correlations between IELTS scores and academic results producing moderately positive correlations for IELTS global scores as well as for the speaking and writing subtest scores. However, these results are problematic statistically. First, they cannot be compared to the correlation for the listening subtest because they were not achieved with the same sample, but with an arbitrary subgroup of subjects in the study whose IELTS scores were 6.5 or above. Second, this subgroup was very small with only 12 subjects, and third, no information was supplied about the characteristics of this subgroup as opposed to the whole sample. There is no way of knowing for example, if the subjects used for these calculations were from one academic discipline or many, whether postgraduate or undergraduate. Without further information it is difficult to assess the merit of these statistical calculations.

In a study of 61 students at Curtin University in Western Australia (Fiocco, 1992), there was a negligible correlation coefficient of .063 between IELTS global scores and the semester-weighted academic results. Nor did Fiocco find any meaningful statistical relationship between IELTS scores and language-related coursework tasks, although her qualitative data appeared to confirm that language proficiency is an important variable influencing academic outcomes. It should be noted that Fiocco included subjects across academic disciplines and from two different populations, NESB permanent residents as well as overseas students, making it difficult to generalise her findings to larger populations.

The predictive validity question was investigated by Ferguson and White (1993) with a group of students taking Life Science Masters courses at the University of Edinburgh. Their findings were somewhat different. Again the study was small with only 28 subjects, but this time the research design was more comprehensive. The subjects took the IELTS test at the beginning and at the end of the year. This allowed correlations to be calculated between scores on both tests and academic outcome, thus testing for both predictive and concurrent validity. In addition, the subjects and their supervisors were interviewed four times each over the period of one year, allowing more data to be obtained on the many and varied intervening variables. In addition, a greater number of statistical measures were used to investigate the research question than in the previous studies. The results indicated that there is a weak relationship between IELTS scores and academic outcome with a positive correlation of approximately 0.3. This is in line with both the results obtained from the much larger validation study by Davies and Criper on the ELTS and the 1992 study of IELTS by Elder.

In a study at UNITEC Institute of Technology, Auckland, Bellingham (1993) conducted a pilot project to investigate the relationship between the IELTS scores of 38 students and first semester academic averages in the National Certificate of Business Studies. The data revealed a moderate correlation between the Global IELTS scores and academic averages of 0.523, a figure which is more positive than other studies. This study appears to be unique for its inclusion of a number of subjects with IELTS scores lower than band 6. In fact, of the 38 subjects involved, 25 had scores lower than 6, and 17 of those had scores below 5. The high correlation would seem to confirm the view proposed by Ferguson and White and others that the lower the bandscore level, the more significant the relationship between language proficiency and academic outcome.

Setting aside the Denham and Oner study (1992), the character of which is not clear, common to all the studies which showed a significant correlation between IELTS scores and academic

outcome, was the use of homogeneous samples of students in terms of academic discipline. Elder's sample involved teacher trainees, Ferguson and White's subjects were all doing Life Science courses, Bellingham's project used students of Business Studies and even the Davies and Criper study grouped students according to academic discipline.

With the exception of the Bellingham study, even where there have been statistically significant correlations between IELTS and academic achievement, these have tended to be weak. As Davies (1988) suggested, it is generally assumed that the length of time which elapses between the measurement of language proficiency and measures of academic achievement allows differential rates of learning and the intervention of a multiplicity of other variables. By repeating the predictor closer to the measurement of academic outcomes, Davies and others (Ferguson and White,1993) hoped to show a much stronger correlation between IELTS scores and academic achievement. Although correlations did improve slightly, they did not do so as much as might be expected. Davies suggested that the effect of the various intervening variables may not be as great as is generally assumed.

However, there are difficulties with repeating the predictor. First, those who agree to sit the IELTS test again, tend to be a self-selecting group. In addition, persuading those whose IELTS scores are lower to take the test a second time is generally a problem. Second, the repeated IELTS test may not be viewed particularly seriously by the participants since nothing is dependent on the outcome for those who are participating. This must have an effect on performance in IELTS and may invalidate data obtained in this way.

In summary, it would seem from the studies cited, that as with other tests of English proficiency, disagreement still exists as to the extent of the relationship between IELTS scores and academic outcome. The area remains open for further investigation. It has been suggested (Graham, 1987; Burns, 1991) that different institutions may need to carry out their own research studies to establish appropriate proficiency level measures across different disciplines within those institutions, if we are to ensure that overseas students are not adversely affected by any lack of linguistic competence and if we are to ensure that students receive the right kind of support once they have embarked on their studies.

3.0 Research Design: Methods and Procedures

3.1 Initial Proposal

It was proposed that the research design follow the steps outlined as follows:

- A sample would be selected with the assistance of the International Students Office, of overseas students who had taken IELTS for admission to the University of Tasmania and who were in the first year of their studies here.

- A questionnaire would be developed to administer to the anticipated sample.

- The questionnaire would be administered to the selected sample.

- IELTS scores would be correlated with first semester outcomes.

- IELTS scores would be correlated with both tutors' estimates and students' self estimates of their academic progress.

- A subgroup of approximately ten of the original sample would be interviewed to investigate in more depth, factors relating to academic outcome and particular difficulties encountered both in relation to IELTS and in relation to their studies. Both quantitative and qualitative data would be collected.

- Academic staff, English language support staff and where applicable, the overseas student advisers, would be interviewed for their perceptions of IELTS as well as perceptions of individual students' performance.

3.2 Sampling Problems

Since information about IELTS is not kept in the central student database, it was necessary to check through all admissions files of the 1996 overseas student intake by hand, in order to identify which students had taken IELTS within the few months prior to entry to the University. The task was made more complicated by the fact that the files had recently been rehoused and had not been completely re-ordered. When all the files had been checked it was discovered that approximately forty-five students at both undergraduate and postgraduate level, had been admitted to the University on the basis of IELTS as a measure of English proficiency. This figure was much lower than had been anticipated, and severely limited our capacity to control for some important factors such as academic discipline, as had been originally planned.

3.3 Research Design Revisions

Because the population was smaller than had been hoped, it was decided to make several revisions to the research design. In addition, a careful study of the literature, consideration of other predictive validity studies, and correspondence with researchers in the field, led us to conclude that a revised longitudinal study might produce more insightful findings than those which could be produced by the original research design. The revisions were as follows:

- All students who had taken IELTS prior to entry to the university to be surveyed, whether undergraduate or postgraduate.

- Greater emphasis to be placed on the interview stage of the project and the qualitative data that it may yield.

- The number of students interviewed to be increased from ten to include all the undergraduates in the population under investigation (n=24).

- The course tutors of the student sample to be surveyed rather than interviewed to allow more feedback on the student sample.

- A small honorarium to be given to those students interviewed.

- Correlations between IELTS scores and both first and second semester results to be calculated.

It had been hoped to compare the sample under investigation with a control group of overseas students who had been admitted to the university on the basis of their TAFE qualifications or on their TCE (Tasmanian Certificate of Education) scores. Considerable time was spent identifying the sample population and sending out the semester one questionnaires. However, the returns on the questionnaire were poor with only eighteen returned (12 TCE and 6 TAFE). It was decided that this mixed group was not a sufficiently representative sample to make comparisons with the IELTS group worthwhile.

3.4   Questionnaire and Survey Design

Serious consideration was given to the design of the semester one questionnaire and the semester two interview schedule, since poorly constructed questionnaires can introduce more problems than they solve. Decisions were made about which of the many and varied intervening variables are likely to have most impact on the relationship between IELTS scores achieved and subsequent academic outcome. For example, the level of motivation possessed by the students, together with their academic aptitude are likely to have differential effects on the rate of student learning. Not all of these constructs could be fully operationalised in the questionnaire, but we attempted to operationalise those we considered were the most important, with the inclusion of carefully constructed questions in the questionnaires. It was felt that the more sensitive questions, such as those which seek to discover whether students are experiencing any personal problems which might be affecting their studies, should be left to the interviews.

An examination of other questionnaires was useful. The research instruments used by Davies, Elder, Fiocco, and especially those of Ferguson and White, have all been influential in shaping the tools used for this project. Consideration was also given to the type of rating scale to use for the questionnaire and the interview schedule. Initially a five point Likert scale was proposed, but after further consideration it was decided to use a seven point scale for the following reasons:

- As subjects generally tend not to use the extreme ends of the scale when giving their responses, this often results in the data being clustered around the middle of the scale (Hatch and Lazaraton: 1991). A seven point scale is thought to allow a greater degree of differentiation, allowing any variations in response to be more clearly perceived.

- As IELTS uses a nine band scale it was thought that a seven point scale might make it easier to calculate correlations between the variables operationalised in the questionnaire and the IELTS scores achieved by the subjects.

- Earlier predictive studies (Davies and Criper: 1988, Ferguson and White: 1993) for the most part used seven point scales. The adoption of a similar scale in our study should make it easier to make comparisons with the findings of these earlier studies.

3.5   The Semester One Questionnaire

The semester one questionnaire (Appendix 4.1) sought information from the subjects under five headings: general information, qualifications, English language ability, course of study, IELTS and English language support. Under the heading of General Information, students were asked to supply information about their age, gender, country of origin, first language, other languages spoken, and date of arrival in Australia. Under the heading, Qualifications, information was sought about the academic qualifications of the students, the amount of English language tuition received prior to the start of their course, English language tests taken and scores obtained, and course units being undertaken.

In the final sections, information was sought about students' perceptions of:

- their English language abilities.
- the extent to which their English has improved since the start of the year.
- their academic performance.
- the relevance of their course.
- the amount of background/subject specific knowledge they bring to their studies.
- the degree of difficulty experienced with their studies.
- the degree to which they are enjoying their course of study.
- the degree of their adjustment to life in Australia.
- the fairness of IELTS.
- the amount of practice undertaken prior to the IELTS test.
- the amount and type of English language assistance the students have received since the start of the year.

3.6   The Semester Two Interview Schedule

The interview was designed in two parts (Appendix 4.2). The first part was similar to the semester one questionnaire in seeking quantitative information under the following headings: English language ability, student welfare, course of study and IELTS. Most of the same questions were asked as in the first questionnaire, but more information was sought about any welfare difficulties students may have experienced and whether these had impaired their ability to do well in their studies. In addition, students were asked to supply more detailed information about any difficulties experienced with the language demands of their course of study. In the second part of the interview, students were asked a series of open-ended questions intended to throw more light on the issues raised in the first part of the interview.

3.7   The Academic Staff Survey

The academic staff survey (Appendix 4.3) was similarly designed in two parts. In the first part, tutors of the student sample were asked to provide information about individual students. They were asked to give estimates of:

- each student's language proficiency.
- the academic performance of the students.
- the degree to which each student's level of English impaired their ability to do well in their studies.
- the level of motivation of each student.
- their degree of cultural adjustment.

The second part of the survey was more general. Tutors were asked for their views on:

- the IELTS test and its fairness.
- the type of assistance given to students experiencing difficulties with the English demands of their studies.
- whether their expectations of overseas students differed from their expectations of the local population.
- measures of a student's level of motivation.
- what constitutes academic success.
- what helps students most to be successful in their academic studies.

3.8   The Surveys for the International Student Advisers and English Support Tutors

Questionnaires similar to those sent to the academic staff (Appendices 4.4) were designed to send to the international student advisers and to English support staff.

3.9   Administering the First Semester Questionnaire

The semester one questionnaire was prepared and trialed by one of the postgraduate support classes and scrutinised by colleagues including staff in the English language centre and the International Student Office as well as several academic staff. The interview schedule was similarly trialed. As a result, several changes were made to the wording of questions and their layout.

The semester one questionnaire was sent to the population under investigation, several weeks before the first semester examinations. A letter was sent to the undergraduate students in the sample. It explained the purpose of the study, requested the students' assistance and assured them of the confidentiality of any information given. It also asked if they would be willing to be interviewed during second semester.

A similar letter was sent to the postgraduates in the sample, but without the request for interviews. At the same time they were asked if they would sign a form giving permission for the researchers to access their examination results and to talk to appropriate staff about their progress. It was decided to delay seeking this permission from those students who were to be interviewed until the time of their interviews in the belief that such a request was more likely

to be granted in a face to face situation. In the event, all students in the sample were happy to cooperate.

After several weeks, a second letter together with a duplicate questionnaire was sent to those students in both groups who had not yet returned the questionnaires. 34 students in the experimental group returned questionnaires, 24 undergraduates and 10 postgraduates. This group formed the basis for our study. Subsequently, one undergraduate moved to the mainland which reduced the population size to 33.

3.10   Conducting Second Semester Interviews

Interviews were conducted with the 23 undergraduate students in late September and early October. The interviews were divided between the two researchers so that the female subjects were interviewed by a female researcher and most of the male subjects by the male researcher. This strategy aimed to minimise any cultural difficulties which might arise by students being asked what could be interpreted as face threatening questions by a member of the opposite sex. The researchers took notes during the interviews, but the interviews were also taped as a means of checking answers against the researchers' notes to ensure an accurate record of students' responses.

3.11   Administering the Staff Surveys

The academic staff, English support staff and two of the international student advisers were contacted in early November several weeks prior to the semester two examinations. 34 academic staff returned questionnaires out of a total of 50. This provided feedback on 30 of the 33 students in the sample. For 19 of these students, we received feedback from more than one course tutor. In addition, English support staff and the international student advisers gave some valuable feedback.

4.0 A Description of the Population

The following tables give a description of the population according to a number of key variables.

Bar:	Element:	Count:	Percent:
1	Malaysia	6	18.18%
2	Japan	1	3.03%
3	Indonesia	3	9.09%
4	Vietnam	9	27.27%
5	Laos	3	9.09%
6	Philippines	1	3.03%
7	Hong Kong	4	12.12%
8	Taiwan	1	3.03%
9	Thailand	5	15.15%

Table 1   Distribution of subjects by country
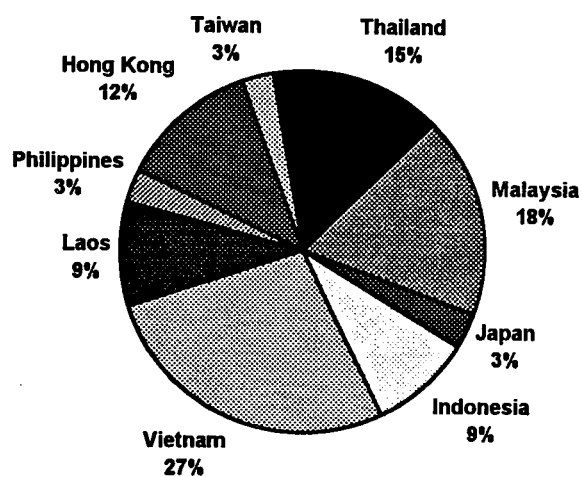
Figure 1   Pie Chart showing distribution of subjects by country

Bar:	Element:	Count:	Percent:
1	Mandarin	4	12.12%
2	Malay	3	9.09%
3	Vietnamese	9	27.27%
4	Japanese	1	3.03%
5	Tagalog	1	3.03%
6	Thai	5	15.15%
7	Javanese	1	3.03%
8	Indonesian	2	6.06%
9	Lao	3	9.09%
10	Cantonese	4	12.12%

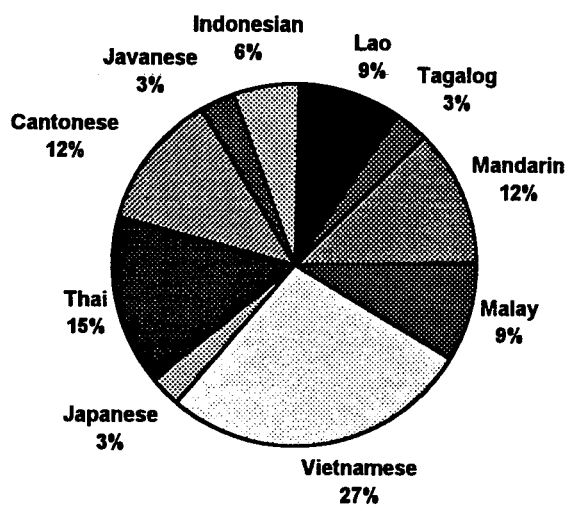*Table 2   Distribution of subjects by first language*

Figure 2   Pie Chart showing distribution of subjects by first language
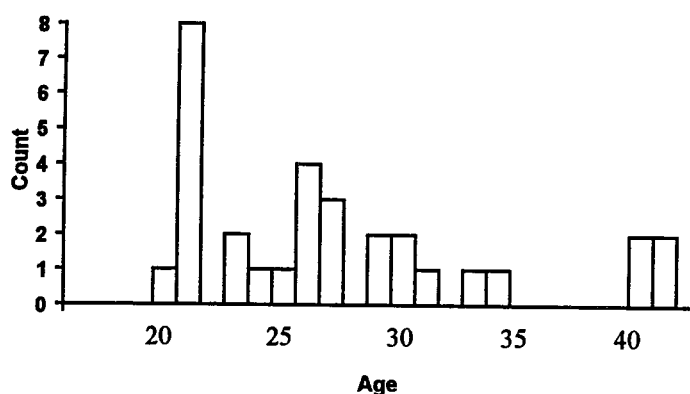
Figure 3   *Histogram showing distribution of subjects by age*

Bar:	Element:	Count:	Percent:
1	Female	15	45%
2	Male	18	55%

Table 3   *Distribution of subjects by gender*

Bar:	Element:	Count:	Percent:
1	Postgraduate	10	30%
2	Undergraduate	23	70%

Table 4   *Undergraduate / postgraduate distribution of subjects*

Bar:	Element:	Count:	Percent:
1	Female	5	50%
2	Male	5	50%

Table 5   *Distribution of postgraduate subjects by gender*

Bar:	Element:	Count:	Percent:
1	Female	10	43%
2	Male	13	57%

Table 6   *Distribution of undergraduate subjects by gender*

Bar:	Element:	Count:	Percent:
1	Architect/Engineer	6	18.18%
2	Commerce/Law	7	21.21%
3	Health Science	6	18.18%
4	Humanities/SocSci	3	9.09%
5	Science/Technology	9	27.27%
6	Visual/Performing	2	6.06%

Table 7   *Distribution of subjects by school*

Female/School

Bar:	Element:	Count:	Percent:
1	Architect/Engineer	1	6.67%
2	Commerce/Law	2	13.33%
3	Health Science	5	33.33%
4	Humanities/SocSci	1	6.67%
5	Science/Technology	4	27.67%
6	Visual/Performing	2	13.33%

Male/School

Bar:	Element:	Count:	Percent:
1	Architect/Engineer	5	27.78%
2	Commerce/Law	5	27.78%
3	Health Science	1	5.56%
4	Humanities/SocSci	2	11.11%
5	Science/Technology	5	27.78%
6	Visual/Performing	0	0%

Table 8   *Distribution of sample by gender and school*

ID	Age	Gender	School	IELTS Global	Reading	Writing	Listening	Speaking
1	23	Female	Science/Technology	7	6	7	7	7
2	20	Female	Science/Technology	7.5	7.5	6	9	6
3	33	Female	Architect/Engineer	6	6	6	6	6
4	39	Female	Humanities/SocSci	7	6.5	7	7.5	7
5	28	Female	Health Science	5.5	5	6	6	5
6	26	Male	Commerce/Law	6.5	6	6	7	7
7	19	Male	Science/Technology	6.5	6	5	6.5	8
8	28	Female	Science/Technology	6	6	6	6	5
9	32	Female	Science/Technology	6.5	6.5	6	6	7
10	24	Female	Visual/Performing	6.5	6	6	7	6
11	21	Female	Visual/Performing	5.5	5	5	6	6
12	20	Male	Architect/Engineer	6.5	7	7	7.5	5
13	20	Female	Commerce/Law	7	6.5	6	7.5	7
14	20	Male	Architect/Engineer	7	8	6	7	6
15	30	Male	Science/Technology	6	6	6	6	6
16	40	Male	Health Science	6	5	7	5.5	6
17	29	Male	Humanities/SocSci	7	6.5	7	7	7
18	22	Male	Science/Technology	6.5	6.5	6	7.5	5
19	20	Female	Commerce/Law	6	6.5	6	6.5	5
20	20	Male	Commerce/Law	6	7.5	5	7	5
21	25	Female	Health Science	6.5	6.5	6	7	7
22	29	Male	Architect/Engineer	5.5	5.5	6	5	6
23	25	Male	Science/Technology	6	7	5	7	4
24	21	Male	Commerce/Law	6	7	5	6.5	6
25	39	Male	Architect/Engineer	6	5.5	6	5.5	6
26	20	Male	Architect/Engineer	6.5	7.5	6	7	5
27	20	Male	Commerce/Law	6.5	7.5	7	6	5
28	21	Male	Commerce/Law	6.5	5.5	6	6.5	7
29	25	Female	Health Science	6	5.5	6	6.5	6
31	25	Female	Health Science	6.5	6.5	5	7.5	6
32	26	Female	Health Science	6	6.5	5	7	6
33	40	Male	Humanities/SocSci	6	6	6	6	6
34	22	Male	Science/Technology	6	6	6	6	6

Table 9   Subjects IELTS scores

Mean:	Std. Dev.:	Std. Error:	Variance:	Coef. Var.:	Count:
6.32	.48	.08	.23	7.62	33

Minimum:	Maximum:	Range:	Sum:	Sum of Sqr.:	# Missing:
5.5	7.5	2	208.5	1324.75	0

Table 10   Mean and standard deviation of subjects' IELTS global scores

X_1: Female - IELTS Global

Mean:	Std. Dev.:	Std. Error:	Variance:	Coef. Var.:	Count:
6.37	.58	.15	.34	9.13	15

Minimum:	Maximum:	Range:	Sum:	Sum of Sqr.:	# Missing:
5.5	7.5	2	95.5	612.75	3

X_1: Male - IELTS Global

Mean:	Std. Dev.:	Std. Error:	Variance:	Coef. Var.:	Count:
6.28	.39	.09	.15	6.24	18

Minimum:	Maximum:	Range:	Sum:	Sum of Sqr.:	# Missing:
5.5	7	1.5	113	712	0

Table 11   Mean and standard deviation of subjects' IELTS global scores by gender

IELTS Listening

Mean:	Std. Dev.:	Std. Error:	Variance:	Coef. Var.:	Count:
6.64	.78	.14	.61	11.8	33

Minimum:	Maximum:	Range:	Sum:	Sum of Sqr.:	# Missing:
5	9	4	219	1473	0

IELTS Reading

Mean:	Std. Dev.:	Std. Error:	Variance:	Coef. Var.:	Count:
6.32	.77	.13	.59	12.17	33

Minimum:	Maximum:	Range:	Sum:	Sum of Sqr.:	# Missing:
5	8	3	208.5	1336.25	0

IELTS Writing

Mean:	Std. Dev.:	Std. Error:	Variance:	Coef. Var.:	Count:
5.97	.64	.11	.41	10.66	33

Minimum:	Maximum:	Range:	Sum:	Sum of Sqr.:	# Missing:
5	7	2	197	1189	0

IELTS Speaking

Mean:	Std. Dev.:	Std. Error:	Variance:	Coef. Var.:	Count:
6	.87	.15	.75	14.43	33

Minimum:	Maximum:	Range:	Sum:	Sum of Sqr.:	# Missing:
4	8	4	198	1212	0

Table 12   Mean and standard deviation of subjects' IELTS subtest scores

5.0 Data Analysis and Interpretation

5.1 Predictive Validity: The Relationship between IELTS and Academic Performance

The main objective of this study is to investigate the relationship between IELTS and academic performance. Three measures of academic performance were used. The first of these was the Grade Point Averages of those coursework students who were assessed in examinations. Grade Point Averages were calculated (according to the University of Tasmania's rules) for the year and for first and second semester. For all students an average result was calculated weighted according to the subject loadings. It should be noted that while most of the sample had a full unit load, seven students took a less than 100% load (typically between 75% to 90%).

The second measure of academic performance was the academic staff ratings of student performance of each student in the sample, collected just before second semester examinations. 34 questionnaires were returned from academic staff. Where more than one questionnaire was returned relating to one student, ratings were averaged. On the whole, staff ratings of student performance were similar and only in a few cases were staff ratings widely different, thus averaging the ratings was thought to give the most reliable picture of student performance overall.

The third measure of academic performance used for this study was the students' own assessments of their academic performance. This was collected from the whole group just before the first semester examinations, and collected a second time from those students who were interviewed in October.

5.1.1 IELTS and Academic Performance: Grade Point Averages.

Using Pearson's Product Moment correlation coefficient, IELTS scores of the sample population were correlated with global academic results, expressed as Grade Point Averages (See Appendix 4.5). These results are set out in Table 13. Of the total sample of 33 subjects, Grade Point Averages were available for 26 coursework students assessed in examinations. (The other seven were assessed by their supervisors.)

As can be seen from the table, only the IELTS reading subtest has a moderate positive correlation with academic results and in the case of the speaking subtest there was in fact a negative correlation.

The first and second semester Grade Point Averages were calculated for those students (n=17) who had taken single semester units as opposed to full year units, and then the correlation coefficients were calculated for the semester averages and IELTS scores. The results are set out in Tables 14 and 15.

IELTS SCORES	ACADEMIC RESULTS
Global	-0.24
Reading	0.42
Writing	0.11
Listening	-0.19
Speaking	-0.55

Table 13   *Correlation coefficients of IELTS global and subtest scores and total 1996 academic results (n = 26)**

IELTS SCORES	SEMESTER 1 RESULTS
Global	-0.62
Reading	0.09
Writing	-0.03
Listening	-0.58
Speaking	-0.41

Table 14   *Correlation coefficients of IELTS global and subtest scores and first semester results (n = 17)*

IELTS SCORES	SEMESTER 2 RESULTS
Global	-0.47
Reading	0.17
Writing	0.05
Listening	-0.56
Speaking	-0.32

Table 15   *Correlation coefficients of IELTS global and subtest scores and semester two results (n =17)*

* 'p' values are not given as the whole population was surveyed and not a random sample

An explanation for the negative correlations may be evident if we look at the frequency distribution of IELTS bandscores and academic results as set out in Table 16.

IELTS	ACADEMIC RESULTS (GPA)*						TOTAL
	5	5-5.5	5.5-6.5	6.5-7.5	7.5-9	9+	
5.5			1	1			2
6			4	1	2		7
6.5	2		3	3	2	1	11
7	2		1			1	4
7.5	1				1		2
TOTAL							N=26

Table 16   The frequency distribution of IELTS bandscores and academic results

*	The figures given for the GPAs are based on the University's system of calculating student averages.

As can be seen from the table, three (12%) students who achieved IELTS scores of 7+ did very poorly in their examinations, whilst two (7%) students who achieved scores of 5.5, obtained good Grade Point Averages. With a small sample, such a distribution is likely to affect the correlations between IELTS scores and academic performance, and indicates the existence of other factors influencing academic outcomes.

The fact that three (12%) of the five students who failed the year had achieved very high scores in the speaking and listening subtests, might go some way to explaining the negative correlations between IELTS speaking and listening subtest scores. These three students had taken single semester units and are included in the sample for the purpose of calculating the correlations between semester Grade Point Averages and IELTS.

It must also be recalled that academic results were weighted according to the University's calculation of relative unit weights. In several cases, the grades for individual units, which were in a number of cases quite high, appear to be lower when calculated in this way. In one case, a subject who achieved a high IELTS score but whose grade point average was recorded as less than 5, did, in fact, pass all but one of the units he undertook.

5.1.2   IELTS and Academic Performance: Staff Ratings

Using the second measure of academic performance, correlations were then calculated between IELTS global and subtest scores and academic staff ratings of academic performance. The sample size is slightly larger, as it includes not only those IELTS students who were assessed in examinations, but also postgraduate students who were assessed by their supervisors. The results are set out in Table 17.

Although there was little correlation between IELTS global scores and academic outcome, using the staff ratings as the criterion for academic performance, there appears to be a weak positive correlation between the reading and writing subtests and academic performance. Staff ratings were then correlated with Grade Point Averages and showed a reasonably strong correlation: $r = 0.73$.

IELTS SCORES	STAFF RATINGS
Global	0.15
Reading	0.36
Writing	0.34
Listening	0.07
Speaking	-0.33

Table 17 *Correlation coefficients of IELTS global and subtest scores and academic staff ratings of academic performance (n = 30)*

Any interpretation of the correlation between staff ratings of academic performance and IELTS scores should also take into account the different measures various staff may have used in assessing student performance. Where staff have assessed written work, their ratings may have been more accurate than when staff have only used performance in tutorials as a measure of academic success. In such instances, if students do not participate, then it is unlikely that staff will give them high ratings. This issue may have some bearing on the negative correlation between the speaking subtest and staff ratings.

Comments made by staff about students' participation would seem to corroborate the perception of low levels of student participation in tutorials.

> *When X does speak, often displays problems in expressing hard concepts.*
> *Very shy.*
> *Great difficulty keeping up in tutorial and making any contribution.*
> *They NEVER volunteered any comments.*
> *Speaks rarely in tutorials.*
> *Limited class participation.*

Another possible explanation for the different correlations using the two different measures of academic performance, may be found if we look at the frequency distribution of IELTS global scores and academic staff ratings of academic performance (Table 18). It should be noted that in order to ensure there was some consistency between estimates of academic performance given by different staff, they were provided with a seven point scale, with 4 indicating average or pass grade performance (See Academic Staff Survey, Appendix 4.3).

No staff feedback was received for two of the three students whose original IELTS scores were 7 or higher, but who nevertheless failed the year, even though questionnaires and reminder messages were sent to all tutors of courses in which the sample subjects were enrolled. Thus these particular students could not be included in the sample used for calculating correlations between staff ratings and IELTS scores.

IELTS	STAFF RATINGS OF ACADEMIC PERFORMANCE						TOTAL
	<3	3	4	5	6	7	
5.5		1	1			1	3
6		2	4	2	2	2	12
6.5		1	2	6	2		11
7				1	3		4
7.5							0
TOTAL							30

Table 18   Distribution of IELTS global scores and staff ratings of academic performance (n=30)

This may be an indication that international students who do not receive sufficient support from academic staff are more at risk of failure, than those who receive higher levels of support. Further investigation is needed to clarify the extent to which varying levels of support affect final outcomes.

In addition to providing a rating of students' academic performance, staff were asked to estimate how much they thought the students' levels of English impaired their ability to do well in their studies.

Not much at all Rankings 1 - 2		A moderate amount Rankings 3-5			A great deal Rankings 6-7	
6	6	8	3	3	3	1
12 (40%)		14 (47%)			4 (13%)	

Table 19   Staff ratings of the degree to which English has impaired individual student's ability to do well in their studies (n =30)

For the sample under investigation, whilst 12 staff (40%) estimated that English levels did not adversely affect academic performance, 14 (47%) estimated that the level of English did impair students' ability to do well in their studies at least moderately, and four (13%) thought that it did so a great deal.

If it could be shown that staff ratings of students' English proficiency were similar to students' IELTS scores, then it might be that staff ratings of the degree to which students' English levels impair performance in academic tasks are a reasonably reliable indicator of the ability of IELTS to predict academic performance, at least to some extent. However, in our study, staff were surveyed almost 10 months after students sat IELTS and sometimes longer, making it unlikely that any relationship can be established. A correlation of staff ratings of student language proficiency and IELTS global scores was calculated, but the correlation was fairly weak: $(r = .29)$.

5.1.4 IELTS and Academic Performance: Students' Ratings

Using the third criterion of academic performance, correlation coefficients were calculated between IELTS scores and students' ratings of their academic performance in both first and second semester.

IELTS SCORES	Student ratings of academic performance (1st semester n = 32)	Student ratings of academic performance (2nd semester n = 22)
IELTS global	-0.28	0.12
IELTS Reading	-0.25	0.46
IELTS Writing	0.28	0.39
IELTS Listening	-0.31	0.16
IELTS Speaking	-0.16	-0.57

Table 20   Correlation coefficients of IELTS global and subtest scores and students' self-ratings of academic performance (semester 1 and 2)

There are a number of factors which need to be considered in any interpretation of correlations based on students' self ratings of academic performance. First, cultural issues need to be considered. To give oneself a high rating may be culturally inappropriate for some cultural groups in the sample, so that the ratings may be depressed towards the lower end of the scale.

Students' self-ratings may also be affected by their perceptions of what constitutes success. High expectations may result in low ratings as students feel that they are not living up to their own expectations.

The negative correlations for first semester may also reflect a measure of culture shock, when students are still struggling to adjust to the new educational context and different academic expectations. They may have given themselves low ratings on academic performance, whilst having scored reasonably well on IELTS.

But perhaps most important of all, first semester student self ratings were provided prior to the first semester examinations so that students had not yet received examination results against which to gauge their own progress. By semester two, they would have had more feedback on their progress, thus making any self ratings of performance more realistic.

For semester two, there appears to be a link between IELTS reading and writing subtest scores and students' self ratings of academic performance, as there was with the academic staff ratings of student performance and IELTS. Given all the reservations which must be borne in mind in any interpretation of the results, it would appear that of all the subtests, the reading subtest has the greatest ability to predict future academic performance.

Discussion

While there appear to be weak correlations for the reading and writing subtests, using two of the three measures of academic performance, there are very low or negative correlations between academic performance and the other subtest scores, as well as between IELTS global scores and academic performance overall. However, there are a number of factors in this study which might explain the absence of a positive link between IELTS and academic outcomes.

First, the group under investigation included subjects from across a number of academic disciplines and lacked the homogeneity of those studies which have shown a significant correlation between IELTS and academic outcome. (Elder 1992; Bellingham 1993; Ferguson and White 1993). In these studies there would have been a measure of agreement about the criterion used for measuring academic outcome within disciplines, whereas in a study across disciplines there is no way of ensuring that academic outcomes have been measured in precisely the same way and are comparable.

Second, it is generally agreed that different academic disciplines require differing levels of English proficiency, even though the degree to which this is so is still not clear. In this study, 16 students (45%) were enrolled in subjects traditionally requiring less English, with six (18%) taking engineering, and nine (27%) doing science and technology subjects. This would surely weaken any link between language proficiency measures and academic outcomes. In fact, a number of those taking engineering, whose scores in IELTS were only moderately good, subsequently did extremely well, being awarded distinctions and high distinctions.

Furthermore, only students who had been admitted to the university were included in the study, thus excluding those with lower IELTS bandscores. The population is therefore skewed and has few subjects with IELTS global scores below band 6. In fact, only two students in the group analysed have scores lower than 6, at band 5.5. Previous studies have shown that the predictive validity of IELTS is greater when IELTS bandscores are lower. (Ferguson and White 1993, Bellingham 1993). Where students' scores fall below 6, the chances of academic failure increase and the correlation between IELTS scores and academic outcomes is stronger. In our study, the use of the IELTS cut off point of 6 for admission to the university would appear to have excluded from our study those students in greatest danger of failure due to inadequate English proficiency.

5.2   Predictive validity: IELTS and Language-Related Difficulties of Coursework.

A further question in relation to IELTS is whether or not it can predict the kind and level of difficulty students experience with the different language demands of their courses. In the semester two interviews, students were asked to rate how difficult they found the English language demands of a variety of coursework activities. The results are presented in Table 21.

	not a problem 1-2		a problem but not major rankings 3-5			a major problem rankings 6-7	
Reading academic texts	2	1	3	9	6	2	0
	3 (13%)		18 (78%)			2 (9%)	
Writing assignments	1	0	5	6	8	3	0
	1 (4%)		19 (83%)			3 (13%)	
Understanding lectures	0	3	6	7	6	1	0
	3 (13%)		19 (83%)			1 (4%)	
Taking notes in lectures	0	4	7	5	5	1	1
	4 (17%)		17 (74%)			2 (9%)	
Taking notes from books	2	7	9	4	1	0	0
	9 (39%)		14 (61%)			0	
Participating in tutorials	0	4	6	4	8	1	0
	4 (17%)		18 (78%)			1 (4%)	
Giving presentations*	2	0	4	2	6	6	1
	2 (9%)		12 (52%)			7 (30%)	

Table 21   *Course work language difficulties reported by students interviewed (n = 23)*

*Two students reported that they did not have to give presentations so gave no ranking.

Giving presentations was perceived to be the most problematic followed by writing assignments and reading academic texts. Taking notes from books was rated as least problematic.

Of the 14 students who made comments, three of them mentioned problems with speaking. As one of them stated:

> *Not always participating in tutorials and seminars (is a problem). The reason is I worry that the answer I would like to voice out is wrong and incorrect. The local students seem very smart and maybe I felt shame to do that. So even I know the answer, sometimes I just keep quiet, and it also happen to other international students.*

Three students mentioned that lectures were no problem because staff provided lecture notes. Several students mentioned the amount of reading as being problematic, and two others mentioned their difficulties with writing assignments.

Correlation coefficients were then calculated for IELTS scores and students' ratings of level of difficulty experienced with language-related coursework tasks. However, no positive correlations were found, and it was not possible to corroborate Elder's perception (1992) of the probable link between various subtest scores and difficulties with aspects of coursework.

Investigating the relationship between IELTS and language-related coursework tasks from another perspective, students interviewed in semester two were asked the following question:

> *Do you think the IELTS test predicted the areas that you were going to have difficulty with? For example, if your lowest score was in listening, is this the skill you're having problems with now?*

13 students (57%) stated that IELTS did predict the areas they were having difficulty with, seven students (30%) stated that it did not accurately predict their problem areas, and three students (13%) stated that they were not aware of the individual subscores and so could not answer the question.

Of those who stated that their weakest subscore predicted their present language difficulties, six students (26%) referred to their low scores for writing and their subsequent difficulties with written assignments. Of the four students whose lowest score was for reading, two reported difficulties with the reading necessary for the course work, but the other two reported that reading was only a problem where there were time restrictions, as in the IELTS test.

It would seem therefore, that there is some relationship between IELTS scores and language-related difficulties of coursework, but the precise nature of the relationship needs further investigation.

5.3   Predictive Validity: Intervening Variables

The relationship between IELTS and academic performance is complicated by a multiplicity of intervening variables which affect differential rates of learning in the period between taking IELTS and achieving academic results. It had been the intention of this study to conduct regression analyses to investigate the effect of a number of these factors, but the small number of cases make it unlikely that such statistical treatment will clarify the issues. However, consideration has been given to a few of the more obvious of these variables, namely; the amount of English language assistance received, motivation, cultural adjustment, and welfare difficulties experienced by the students. No attempt has been made to operationalise the construct of scholastic aptitude. Although this is a major factor affecting academic performance, it was considered to be too complex an issue upon which to obtain valid and reliable data within the constraints of this project.

5.3.1   English Language Assistance

In discussing the predecessor of IELTS, Pollitt (1988) stated that any intelligent use of ELTS will reduce its predictive power. Similarly, this holds true for IELTS. If students seek assistance with, or are provided with tuition for their weakest macroskills, as indicated by IELTS subscores, it would seem likely that their language skills will improve over the course of the year to the extent that their initial IELTS scores are no longer a true indication of their language proficiency.

If in addition, the power of IELTS to predict academic outcomes is greater the lower the bandscore level (Davies and Criper 1988; Ferguson and White1993; Elder 1993) then it would follow that consistent levels of English language assistance are likely to weaken the predictive validity of IELTS, and should be considered in any study of this area.

To investigate this further, information was sought from the students surveyed in semester one, and the students interviewed in semester two about the amount and type of English language assistance they had received from various sources. In addition, feedback was sought

from academic staff and English support staff about the level and kind of assistance given to the international students in our study.

Findings

Widely varied amounts of English language assistance were received by the international students in this study. At the end of semester one, 26 students said they had attended English support classes, but the amount of time spent at these classes varied considerably. To the question of whether they had received any other assistance, 20 students replied affirmatively. The rate of individual assistance from various personnel is shown in Table 22.

Predictably the English support tutors and the academic staff were those who gave the greatest amount of help with English. Of the 23 students interviewed in semester two, only two students (9%) said they had received very little assistance, nine students (39%) said they had had a moderate amount of assistance and 12 students (52%) indicated that they had received a great deal of assistance.

ASSISTANCE FROM*:	Little or no assistance Rankings 1 - 2		A moderate amount of assistance Rankings 3-4			A good deal of assistance Rankings 5-7	
Academic lecturer/tutor	0	3	1	3	8	2	3
	3 (15%)		12 (60%)			5 (25%)	
An Australian friend	6	3	1	2	3	2	1
	9 (45%)		6 (30%)			3 (15%)	
Another international student	2	2	3	4	1	3	2
	4 (20%)		8 (40%)			5 (25%)	
ELSIS support tutor	3	1	0	2	5	3	5
	4 (20%)		7 (35%)			8 (40%)	

Table 22   Amount of English language assistance received (n=20)

* Not all students responded to every category.

In the open-ended questions, particular mention was made of the special English support classes for overseas students enrolled in nursing. These appear to have been much appreciated by the students concerned, all of whom did well in their examinations with credit grades in most of their courses. Other English support classes received generally favourable comments, although six students commented that they were too busy with course work and exam preparation to attend as frequently as they would like. One student who subsequently failed the examinations stated that she was not aware of the availability of English assistance.

5.3.2   Motivation

Whilst motivation has been discussed at some length as an important factor affecting performance, there has been little agreement about what is meant by motivation and how best to operationalise this multi-dimensional construct. For the purposes of this study and for purposes of comparison, students were asked in the initial questionnaire at the end of semester one and again in the second semester interviews, the following questions which replicate those used in the Ferguson and White study (1993):

How much are you enjoying your course?

Do you think your course is relevant to your future career?

It is assumed in this study that the level of enjoyment of a chosen course, and its relevance to future career prospects are two facets of motivation and that these questions attempt to operationalise these two dimensions of motivation.

There was 70% commonality between the responses to the two quasi-variables of motivation, levels of enjoyment and relevance of course to future career. An attempt was made to conduct a regression analysis. These variables appear to have a slight effect on the relationship between IELTS scores and academic performance.

5.3.3 Cultural Adjustment

As with the concept of motivation, clearly agreed definitions of what is meant by cultural adjustment have yet to be determined, but there seems general consensus, especially from student advisers, that an ability to adapt to the social and cultural environment of the university is crucial for the successful academic performance of international students. As one member of staff from the International Student Office commented, when asked her views on IELTS:

The IELTS test provides some indication of people's language ability. However, it does not provide information about how well students will adjust to a new learning style or a new environment. Adaptability to both of these is critical in terms of how a student progresses with their study.

Student advisers and ELSIS support staff were asked to comment on the extent to which difficulties with English affected individual student's ability to do well in their studies. Cultural adjustment difficulties were mentioned as of more significance than language difficulties in comments on eight out of 16 subjects. As one staff member, writing about a representative student, suggested:

Changing context - educational system, cultural environment, lack of family support more significant - if you can separate out from language.

Therein lies the problem when investigating the relationship between language proficiency, in this case measured by IELTS, and academic outcomes. The difficulties in separating out the variables are extensive. Nevertheless, in both first and second semester, the students were asked the question:

How well do you think you have adjusted to life in Australia?

The distribution of their responses was as follows:

Little or no adjustment Rankings 1-2		Moderate level of adjustment Rankings 3-5		High level of adjustment Rankings 6-7		TOTAL
0	1	10	11	8	2	33
1 (3%)		22 (67%)		10 (30%)		

Table 23. Distribution of responses to students' level of cultural adjustment on a 7 point scale: semester 1 (n=33)

Little or no adjustment Rankings 1-2		Moderate level of adjustment Rankings 3-5			High level of adjustment Rankings 6-7		TOTAL
0	1	0	6	12	3	1	23
1 (4%)		18 (78%)			4 (17%)		

Table 24 Distribution of responses to students' level of cultural adjustment: semester 2 (n=23)

A regression analysis of IELTS global results on Grade Point Averages controlling for cultural adjustment was calculated. Much more work is needed to operationalise the construct of 'cultural adjustment' before meaningful results can be produced.

5.3.4 Difficulties Experienced by Overseas Students

The students interviewed in semester two (n=23) were given a list of problems which overseas students often experience and were asked to rate on a seven point scale how much of a problem each one of the difficulties were for them (Interview schedule, question 6). The results are set out in Table 25.

Reported difficulties were mainly language and coursework related. Language problems were seen to be a moderate difficulty by 17 (74%) of the sample and a major difficulty by three students (13%). Too much reading and problems understanding coursework also rate as moderate to major difficulties by the students interviewed. Study skills difficulties were rated by 14 students (61%) as moderate difficulty.

In the open-ended questions (17 and 18) students discussed further their problems and difficulties, and the extent to which these had affected their ability to do well in their studies. Opinion was fairly evenly divided with 12 students (52%) stating that their difficulties had not significantly affected their studies, three (13%) students stating their studies were affected 'a little bit' and eight students (35%) clearly affirming that their difficulties had impaired their ability to do well in their studies.

	Little or no problem Rankings 1-2		A moderate problem Rankings 3 - 5			A major problem Rankings 6 - 7	
Unsat accommodation	8	7	3	2	1	1	1
	15 (65%)		6 (26%)			2 (9%)	
Health problems	10	8	2	1	1	1	0
	18 (78%)		4 (17%)			1 (4%)	
Financial difficulties	9	6	3	2	1	2	0
	15 (65%)		6 (26%)			2 (9%)	
Family difficulties	13	5	3	0	1	0	0
	18 (78%)		4 (17%)			1 (4%)	
Making friends	5	7	5	1	1	4	0
	12 (52%)		7 (30%)			4 (17%)	
Problems understanding coursework	0	4	6	8	3	1	1
	4 (17%)		17 (74%)			2 (9%)	
Too much reading	2	4	4	5	6	1	1
	6 (26%)		15 (65%)			2 (9%)	
Poor teaching	4	4	5	8	1	1	0
	8 (34%)		14(61%)			1 (4%)	
Relat with academic staff	9	7	4	2	1	0	0
	16 (70%)		7(30%)			0	
Language problems	1	2	5	11	1	2	1
	3 (13%)		17 (74%)			3 (13%)	
Inefficient study methods	3	5	4	9	1	0	0
	8 (34%)		14 (61%)			0	
Homesickness	8	9	2	1	2	0	1
	17 (74%)		5 (22%)			1 (4%)	
Loneliness	9	4	6	2	1	1	0
	13 (57%)		9 (39%)			1 (4%)	

Table 25   *Cultural and welfare difficulties students have experienced during their studies (n=23)*

6.0   IELTS: Face Validity

6.1   Student Perceptions

Although the major objective of this study was to investigate the predictive validity of IELTS, students were asked in both the initial survey and in the interviews to give their views on the IELTS test and to give a rating of the fairness of IELTS, thus providing some feedback on the test, its face validity and to a certain extent its content validity. It was thought that this might provide useful information on students' perceptions of the test since its revision and the introduction of a single academic module in the reading and writing subtests. The students' ratings from the semester 1 survey are shown in Table 26.

IELTS SUBTESTS	Not fair at all Rankings 1-2		Moderately fair Rankings 3-5			Very fair indeed Rankings 6-7	
Reading		3	5	6	8	7	4
	3 (9%)		19 (58%)			11 (33%)	
Writing		1	4	5	10	10	3
	1 (3%)		19 (58%)			13 (39%)	
Listening		1	2	1	12	11	6
	1 (3%)		15 (45%)			17 (51%)	
Speaking	1		4	3	11	10	4
	1 (3%)		18 (54%)			14 (42%)	

Table 26   Student ratings of the fairness of IELTS: semester 1 (n=33)

From this table it would seem that the test was perceived to be a fair test overall, with the listening test viewed as the fairest of the subtests. This is interesting to note, given that the listening subtest is a general test of listening proficiency and does not attempt to measure specialised academic listening skills. The reading was viewed as the least fair subtest, with three students indicating that they thought it was not a fair test at all. Several students commented that it was difficult in the time available to read unfamiliar texts of which the examinee has no background knowledge.

Four students commented on the subjectivity of the speaking subtests:

I think the examiners have their own judgements, especially writing and speaking. Therefore different examiners give different score.

Mainly it is a fair test. Nevertheless - especially speaking test is a subjective test.

Writing and speaking test were a bit subjective. They are marked based on the view of the person who are testing.

I don't know how the examiner gives marks on the 'speaking test'. If you can speak, it doesn't mean you can get high marks. Why?

Question 15 of the semester one questionnaire asked students whether IELTS tested their readiness for the English language demands of their particular courses. Eight of 17 students

who wrote a response to this question stated clearly that IELTS did not really test their readiness for the English language demands of their courses as the following comments show:

> *IELTS tested me only in general English not in particular course. However, is not a big problem.*

> *It did not expose me to the technical or scientific words I have to encounter here. The older classification of paper A, B, C and D to Business, Art, Science and Engineering is a much better test.*

> *The topics for reading is very large and it does not focus on my field.*

> *IELTS test is mainly about social topics and general daily life English, not Maths or Physics English.*

> *As IELTS tested general area, it doesn't test the English language requirements for my course.*

In semester two, the undergraduate students were again asked about the fairness of IELTS in their interviews. The distribution of the semester two IELTS fairness ratings are given below (n=23)

IELTS SUBTESTS	Not fair at all Rankings 1 - 2		Moderately fair Rankings 3-5			Very fair indeed Rankings 6-7	
Overall		1	0	8	9	4	1
	1 (4%)		17 (74%)			5 (22%)	
Reading				10	7	5	1
	0		17 (74%)			6 (26%)	
Writing		1	3	10	6	1	2
	1 (4%)		19 (83%)			3 (13%)	
Listening				7	7	6	3
	0		14 (61%)			9 (39%)	
Speaking		2	1	8	4	4	4
	2 (8%)		13 (56%)			8 (35%)	

Table 27   Student ratings of the fairness of IELTS: semester 2 (n=23)

Once again, the table indicates that IELTS is viewed as a fair test, with the listening and speaking subtests ranked as the fairest subtests.

6.1.1   The Reading Subtest

The few critical comments for the reading subtest related to insufficient time to complete the paper and to the reading topics. Comments included the following:

> *The reading test was satisfactory. However, a student should be given a paper related to his future course because reading a general article differs to reading an academic text.*

> *I think that depends on the topic. Some topic are more easy to some people.*

But it can still test people's reading level.

It's a bit long. I coped with a lot of new words in reading, maybe because I am unfamiliar with the topic.

6.1.2 The Writing Subtest

For the writing subtest, six students out of 23 mentioned the time restrictions as being problematic, although one student commented,

Writing is good for time management skills. When I first entered uni, I spent quite a lot of time structuring essays and finding words for my express.

Another six students mentioned the nature of the writing topic, as being either too technical, too general, or not relating to the students' area of study.

6.1.3 The Listening Subtest

For the listening subtest, comments were more diverse. Three students mentioned the general nature of the test. Comments included,

Compare to other three subtests, listening test is a bit too general. It'd be more relevant if it was more academic.

Everyday life is not difficult. It's what you hear every day.

Four other students stated that the test was easy, and three students said that the speakers in the listening test spoke slowly and sometimes repeated themselves. Three students mentioned the test conditions, with one student indicating that performance in the test is affected by how close you are to the tape player and the quality of the tape, and two students suggesting that the use of headphones would improve test conditions.

6.1.4 The Speaking Subtest

For the speaking test, three students interviewed mentioned the subjective nature of the assessment and four students mentioned the general nature of the questions asked. Although two students rated the speaking test as not fair at all, six students spoke very positively of this subtest:

It's a very fair test.

It was the best part of the examination.

I like this one.

It's excellent.

This is also a fair test.

This is my best part because the interviewer is my teacher and she asked me about my job and my subject I'll study.

This last comment is noteworthy since teachers are not supposed to interview their own students and it may indicate a need for IELTS administrators to stress this point in future directives to examiners.

6.2   Academic Staff Perceptions of IELTS Face Validity.

Staff were asked to give their views on IELTS, how fair a test they consider it to be and how it is used in their departments. Twenty-seven staff (79%) stated that they knew nothing at all about IELTS. Of these, two staff commented further:

> *Had students who presumably passed (IELTS) with enormous language difficulties.*

> *I do know significant numbers of overseas student not proficient. One of the major reasons for high failure rate.*

Two staff members based their evaluations on their present students:

> *Experience with two students, conclude test good indicator of proficiency for entry to research higher degree.*

> *I have only encountered one student who has passed the IELTS. Judging by his performance, I'd say the screen is not suitable.*

Two other staff members stated that they thought IELTS was a good test:

> *IELTS is a useful benchmark to maintain. The screening of students for language proficiency is necessary.*

> *It is a good test.*

Another tutor made the following point:

> *IELTS doesn't cover any problems in specific subject areas, eg. psychology.*

The final two tutors gave the following comments:

> *Basically interested in the intellect, not the language skills.*

> *The students struggle in the first few weeks.*

It appears therefore, that IELTS is not well understood by the academic staff surveyed in this study. As the staff surveyed came from every school and ranged from postgraduate to professorial level, one might cautiously hypothesise that IELTS is not well known by most academic staff at the university. As many academics have international connections and if they are consulted in the recruitment of international students, it may be that more informed decisions could be made were IELTS better known. It might also be of value for course tutors to know the subtest scores of their international students in order to alert them early in the year to possible language-related difficulties before these become insurmountable. This was the original intention of reporting the subtest scores separately. If course tutors are aware of potential language problems, they can provide appropriate support, in addition to that supplied by the International Student Office.

7.0   Conclusions and Recommendations

The principal purpose of the present study was to investigate the predictive validity of IELTS employing a group of international students in the first year of their studies at the University of Tasmania. Answers to the following research questions were sought:

- To what extent does the IELTS proficiency rating scale predict subsequent academic outcome?
- To what extent can IELTS bandscores predict the kinds of language difficulties international students encounter with various aspects of their studies?

- Which key intervening variables appear to have the most effect on the relationship between IELTS bandscores achieved and subsequent academic outcomes in the population under investigation?

7.1 Conclusions

Working with a sample of 33 international students who had taken IELTS prior to entry to the University, three measures of academic achievement were used. First, Grade Point Averages in examinations (GPAs) were calculated for the year and for both first and second semester. Second, staff estimates of academic performance were collected just prior to second semester examinations; and third, students' self estimates of performance were collected for both first and second semester.

Correlations were then calculated between IELTS global and subtest scores and GPAs, as well as with both academic staff ratings of academic performance and students' self estimates of academic performance. Although no significant correlations were found for IELTS global scores, there did appear to be weak correlations between the reading and writing subtest scores with two of the three measures of academic outcome. The reading subtest scores in particular, were best able to predict subsequent academic performance.

Discussion focused on difficulties in quantifying these relationships, because of issues connected with the reliability of academic outcome measures, and because of the small size of the population under investigation. This study has demonstrated that a small population may include unusual cases, affecting outcomes. This makes it difficult to generalise the findings to larger populations and to the same population in subsequent years.

In addition, the small number of cases made it difficult to control for certain factors such as the academic disciplines in which the subjects were enrolled. However, the fact that in this study, there was no significant correlation between IELTS *global* scores and academic performance overall, supports the perception discussed earlier, that where this factor is controlled a positive correlation is more likely.

On the other hand, if we accept the perception that low IELTS scores correlate significantly with academic failure, (Davies and Criper 1988; Ferguson and White 1993; Bellingham 1993), the lack of a statistical link between IELTS and the criterion for academic success, may be interpreted as evidence of the effectiveness of IELTS in screening out those students who are more likely to fail. This seems a possible interpretation given that, with the exception of two students all the subjects in our study had IELTS scores of 6 and above.

Nevertheless, several students in our study who had scores of 6+ and 7 subsequently failed the year, and the two whose scores fell below the cut-off point, passed the year. This indicates that language proficiency alone, as measured by IELTS, is no guarantee of success. Other variables may be of equal importance. For example, the lack of feedback from the tutors of two of the five students whose Grade Point Averages fell below 5, may indicate a lack of support, another factor likely to have a marked impact on academic outcomes.

Qualitative data was collected from a questionnaire administered to the students at the end of first semester, in interviews with the 23 undergraduates in the group, as well as from 34 academic staff, two language support tutors and two international student advisers surveyed at the end of semester two. This strengthened perceptions of a link between language proficiency and academic performance and between IELTS and language-related difficulties, despite the lack of statistical evidence. But the qualitative data also provided evidence in support of the view that language is only one of many variables affecting academic outcomes.

The amount of English language assistance received, motivation, cultural adjustment and welfare difficulties experienced are among a number of other variables which are also important. Attempts to show how motivation, cultural adjustment and welfare difficulties affect the relationship between IELTS and academic outcomes were inconclusive. This was due partly to the small number of cases and also to difficulties in operationalising the constructs.

Although not part of the original intention, this study provided qualitative data on the face validity of IELTS. On the whole, the students perceived the test to be fair, but the academic staff surveyed had little knowledge of IELTS. It is suggested that information about the test be provided to academic staff, on the grounds that it might assist them in making informed decisions for the admission of overseas students to their departments, and would be beneficial in alerting course tutors to particular linguistic difficulties overseas students may have.

7.2 Recommendations

The findings from this pilot project indicate that the predictive validity of IELTS is a complex area to research. In order to clarify the relationship between IELTS and academic performance, it is proposed that:

- research be conducted with much larger more homogeneous samples.

- research be focused on the intervening variables using regression analysis.

- the effects of levels of support as an intervening variable be investigated.

- the instruments for operationalising such constructs as motivation and cultural adjustment be refined.

- intercampus collaborative research be supported, using the Internet. It would be possible for a number of researchers to work together to refine the research instruments and use them across different campuses.

- language proficiency measures such as IELTS be recorded on the University student database. This, together with academic Grade Point Averages, would enable an accumulation of data over a number of years to be used in future predictive validity studies.

- academic staff be provided with information on IELTS to enable more informed decisions to be made.

It is hoped that this pilot project, despite its limitations, has made a small contribution to the ongoing debate about the relationship between IELTS and academic performance.

Acknowledgements

Bibliography

Alderman, D. (1981). *Language Proficiency as a Moderator Variable in Testing Academic Aptitude*. TOEFL Research Report 10. Princeton, NJ: Educational Testing Service.

Alderson, J. C. & Hughes, A. (eds) (1981). *Issues in Language Testing*. ELT Documents 111. London: The British Council.

Alderson, J.C. (1988). 'New procedures for validating proficiency tests of ESP? Theory and practice.' *Language Testing*, 5, 2 : 220 - 232.

Alderson, J.C. & Urquhart, A.H. (1985). ' The effect of students' academic discipline on their performance on ESP reading tests.' *Language Testing*, 2, 2 : 192 - 204.

Alderson, J.C. & Clapham , C. (1992). *IELTS research report 2. Examining the ELTS test: an account of the first stage of the ELTS revision project.* London: The British Council, Cambridge:UCLES and Canberra : International Development Program.

Alderson, J.C. (1993). 'The relationship between grammar and reading in an English for academic purposes test battery'. In Douglas and Chapelle: 1993.

Bachman, L.F. (1990). *Fundamental considerations in language testing.* Oxford: Oxford University Press.

Bachman, L.F., Davidson, F. & Foulkes, J. 'A comparison of the abilities measured by the Cambridge and Educational Testing Service EFL Test Batteries'. In Douglas and Chapelle: 1993.

Bellingham, L. (1993). 'The relationship of Language Proficiency to Academic Success for International Students.' *New Zealand Journal of Educational Studies* , 30, 2: 229-232.

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 Blue, G. 1993.

Bong, M. (1996). 'Problems of Academic Motivation Research and Advantages and Disadvantages of their Solutions.' *Contemporary Educational Psychology*, 21:149-165.

Brindley, G. (1989). *Assessing Achievement In The Learner-Centred Curriculum.* Sydney: NCELTR.

Burns, R. (1991). Study and Stress among First Year Overseas Students in an Australian University.' *Higher Education Research and Development,* 10, 1 61-77.

Busch, M. (1993). 'Using Likert Scales in L2 Research.' *TESOL Quarterly,* 27, 4.

Canale, M. 'On some theoretical frameworks for language proficiency'. In Rivera: 1984.

Carlson, S.B., Bridgeman, B. Camp, C. & Waanders, J. (1985). *Relationship of Admission Test Scores to Writing Performance of Native and Non-native Speakers of English* ETS Research Reports #19 Princeton NJ: Education Testing Service.

Chapelle, C. & Douglas, D. (1993). 'Foundations and Directions for a New Decade of Language Testing'. In Douglas and Chapelle: 1993.

Clapham, C. (1993). 'Is ESP Testing Justified?' In Douglas and Chapelle: 1993.

Cumming, A. (1990). 'Expertise in evaluating second language compositions'. *Language Testing,* 7, 31-51.

Cummins, J. (1983). 'Language proficiency and academic achievement' In J. Oller, *Issues In Language Testing Research,* pp 108-129 Rowler, MA: Newbury House.

Cummins, J. (1984). 'Wanted: a theoretical framework for relating language proficiency to academic achievement among bilingual students' In Rivera, C (ed).

Davies, A. (1984). 'Validating three tests of English language proficiency'. *Language Testing,* 1,1: 50-69.

Davies, A.& Criper, C. (1988) *ELTS Research Report 1(i) ELTS Validation Project Report.* London: British Council and Cambridge: UCLES.

De Jong, J. H.A.L. (1995). 'The need for standards in language education.' *System,* Vol 23, 4 : 441-444.

Denham, P. A. & Oner, J. A. (1992). *IELTS Research Project. Validation Study of Listening Sub-test.* IDP/IELTS commissioned report/ University of Canberra.

Douglas, D. & Chapelle, C. (eds) (1993). *A New Decade Of Language Testing Research.* Papers selected from the Language Testing Research Colloquium. Virginia: TESOL.

Elder, C. (1993). 'Language proficiency as a predictor of performance in teacher education.' *Melbourne Papers in Language Testing,* 2, 1 : 68-85.

Ferguson, G. & White, E. (1992). *A Predictive Validity Study of IELTS - a report commissioned by UCLES,* Institute of Applied Linguistics, University of Edinburgh.

Fiocco, M. (1992). *English proficiency levels of students from a non-English speaking background: a study of IELTS as an indicator of tertiary success.* Unpublished research report. Perth: Curtin University of Technology.

Genessee, F. 'On Cummins Theoretical Framework'. In Rivera 1984.

Gibson, C. & Rusek, W. (1992). *The validity of an overall band score of 6.0 on the IELTS test as a predictor of adequate English level appropriate for successful academic study.* Unpublished M A thesis, NCELTR.

Goodbody, M. W. 'Letting the students choose: A placement procedure for a pre-sessional course.' In Blue: 1993.

Graham, J. G. (1987). 'English Language Proficiency and the Prediction of Academic Success.' TESOL Quarterly, 23,1: 505-521.

Grundy, P. 'Student and supervisor perceptions of the role of English in academic success'. In Blue: 1993.

Hamilton, J., M. Lopes, T. McNamara & Sheridan, E. (1993). 'Rating Scales and Native Speaker performance on a communicatively oriented EAP test.' *Language Testing,* 10, 3: 337 - 353.

Hamp-Lyons, E. (1989). 'Language Testing and Ethics.' *Prospect* , 5, 1: 7-15.

Hanania, E. & M. Shikhani (1986). 'Interrelationships Among Three Tests of Language Proficiency: Standardized ESL, Cloze, and Writing'. *TESOL Quarterly,* 20, 1 : 97-109.

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

Hughes, A. (1988). (ed) 'Testing English for university study.' *ELT Documents 127.* Oxford: Modern English Press.

Hughes, A. (1989). *Testing for Language Teachers* Cambridge: Cambridge University Press.

Jensen, C. & Hansen, C. (1995). 'The effect of prior knowledge on EAP listening-test performance.' *Language Testing,* 12, 1 : 100-119.

Jochems, W., Snippe, J. Smid H.J. &Verweij, A. (1996). 'The academic progress of foreign students: study achievement and study behaviour.' *Higher Education* , 31: 325-340.

Johnson, P. (1988). 'English language proficiency and academic performance of undergraduate international students'. *TESOL Quarterly,* 22,1: 164-169.

Jordan, R.R. 'Study skills: Experience and expectations.' In Blue: 1993.

Kinnell, M. (1990) (ed) *The Learning Experiences of Overseas Students.* Buckingham: The Open University Press and SRHE.

Loether, H.J. & McTavish, D.G. (1993) *Descriptive and Inferential Statistics for Sociologist: An Introduction* (4th Ed) Boston: Allyn and Bacon.

Luppescu, S. & Day R. (1990). 'Examining attitude in teachers and students: the need to validate questionnaire data'. *Second Language Research*, 6, 2 : 125-134.

Lynch, B., Davidson, F.& Henning, G. (1988) 'Personality dimensionality in language testing validation.' *Language Testing*, 5, 2 : 206-219.

McDowell, C. (1989). 'Power, Responsibility and the Oral Examiner'. *Prospect* , 4,3 : 95-104.

McNamara, T. (1996). *Measuring Second Language Performance.* London: Longman.

McNamara, T. (1996). 'Modelling Performance-Opening Pandora's Box.' *Applied Linguistics*, 16, 2:159-176.

Morrow, K. (1985). 'The Evaluation of Tests of Communicative Performance.' *Prospect*, 1, 2: 32-4.

Oller, J. (1981). (ed) *Current Issues in Language Testing Research.* Rowley, Mass: Newbury House.

Oller, J. (1983). *Issues in Language testing research* . Rowley, Mass: Newbury House.

Pollitt A. (1988). 'Predictive Validity.' in Hughes, Porter and Weir, 1988.

Reid, J. (1991) 'The dirty laundry of ESL survey research'. *TESOL Quarterly*, 24, 2: 323-338.

Rivera, C (ed) ((1984) *Language Proficiency and Academic Achievement.* Cleveland, Avon: Multilingual Matters.

Saville-Troike, M. (1984). 'What Really Matters in Second Language Learning for Academic Achievement?' *TESOL Quarterly*, 18, 2 : 199-219.

Skehan, P. (1984). 'Issues in the testing of English for Specific Purposes.' *Language Testing*, 1,1 : 202-220.

Spolsky, B. (1985). 'What does it mean to know how to use a language? An essay on the theoretical basis of language testing *Language Testing*, 2: 180-189.

Thiel, T. (1995). *An analysis of the evolution of the IELTS Test and an investigation of its face validity.* Unpublished M.Ed dissertation, University of Tasmania.

Troike, R. 'SCALP: Social and cultural aspects of language proficiency.' In Rivera, 1984.

Vinke, A.A. & Jochems, W.M.G. (1993). 'English proficiency and academic success in international postgraduate education.' *Higher Education*, 26: 275-285.

Walde, B. 'A sociolinguistic perspective on Cummins' current framework for relating language proficiency to academic achievement'. In Rivera: 1984.

Weir, C.J. (1988b) The specification, realisation and validation of an English language proficiency test. In Hughes, 1988b.

Zeidner, M. (1986). 'Are English Language aptitude tests biased towards culturally different minority groups? Some Israeli findings.' *Language Testing*, 3, 1 : 80-97.

Zeidner, M. (1987). ' A Comparison of ethnic, sex and age bias in the predictive validity of English language aptitude tests: some Israeli data'. *Language Testing*, 4, 1 : 54-71.

Zheng Chen & Henning, G. (1985). 'Linguistic and cultural bias in language proficiency tests.' *Language Testing*, 2,2: 155-163.