6. Investigating the Relationship between Intensive English Language Study and Band Score Gain on IELTS

Catherine Elder
University of Auckland

Kieran O'Loughlin
University of Melbourne

Abstract

This exploratory study examines the relationship between intensive English language study and band score gains on the IELTS (International English Language Testing System). Specifically, it investigates the progress of a sample of 112 students from non English-speaking backgrounds enrolled in intensive English language courses at one of four different language centres in Australia and New Zealand. Progress is gauged in terms of score gains on the academic module of the IELTS which was administered at the beginning and end of a 10–12 week period of intensive English language instruction. Pre- and post-study questionnaires were administered to all participating students and semi-structured interviews were conducted with a subset of 18 students sampled according to their level of gain at the post-test session. Interviews were also conducted with administrators and teachers at each of the participating institutions in order to elicit information about a) the learning environment and b) the factors they saw as critical in determining the English language progress of EAP (English for academic purposes) students in general and of those chosen for in depth interviewing in particular. Data were analysed using both quantitative and qualitative methods. Regression analyses were undertaken to investigate the relationship between a range of personal, instructional and environmental variables extracted from the student questionnaire and score gains on the IELTS over the three-month period. In addition, narrative vignettes were produced for the sub-set of 18 students drawing on their own self-reports, as well as information supplied by their teachers. The results revealed that students made variable progress in English during the three month period with an average gain of about half a band overall. The greatest gains were in listening, whereas reading skills were somewhat less amenable to improvement. A range of factors were found to be linked to improved scores on the test, but these varied considerably from one skill to another. Implications are drawn both for students intending to sit the IELTS and for the institutions where they undertake their English language instruction. Avenues for further research are also identified.

Publishing details

**International English
Language Testing System (IELTS)**

Research Reports 2003

Volume 4

Editor: Robyn Tulloh

1.0 Introduction

This exploratory study examines the relationship between intensive study in English for academic purposes and band score gains on the IELTS (the International English Language Testing System) with the aim of identifying issues for further research. It involves an investigation into the progress of a sample of 112 NESB (non English-speaking background) students enrolled in intensive English language courses at one of four different language centres in Australia and New Zealand. This project is therefore one of many applied linguistics studies which tackle the *"HOW LONG?"* question in relation to achievements in a second language (see Collier 1989, for a comprehensive review of such studies) but differs from many of these in that it is concerned not with long term achievements or ultimate attainment but with the degree of English language improvement learners make over a 10-12 week study period. In this case progress is gauged in terms of score gains on the academic module of the IELTS test, which was administered at the beginning and end of the study period.

The IELTS test is a high stakes selection test used to measure the English language proficiency of NESB students intending to study at a tertiary institution. It is made up of four different subtests: Listening, Reading, Writing and Speaking. On the basis of their test performance candidates receive a test report which includes a band score of between 0 and 9 for each sub test and an overall score which is the average of these results. The minimum score required by different tertiary institution varies with most universities requiring a minimum of 6 or 6.5 for undergraduate study and 7 for postgraduate study.

The popularity of IELTS as an admissions tool has grown dramatically over the past decade, with tertiary institutions in Australia and New Zealand requiring a particular score on IELTS for entry for both international students and, in some cases, for permanent residents who have been in the country for a limited amount of time. IELTS is generally preferred in Australia and New Zealand over other admission tests such as TOEFL (Deakin 1997). IELTS-related teaching activities are now "big business", with large numbers of private and tertiary providers offering "user pays" pre-sessional courses of study geared either directly or indirectly to preparing students for the test (see Read & Hayes [2001] for an account of the teaching operations linked to the use of this test in New Zealand).

Issues of test use, test impact and test fairness are now attracting considerable attention in the field of language testing. There is a growing body of research focusing on the values underlying the tests, including the rationales and intentions for introducing them (e.g. Hawthorne 1997, Spolsky 2001), the respective rights and responsibilities of various stakeholders in the testing enterprise (e.g. Hamp-Lyons 1997), the political, social and psychological messages that tests convey (e.g. Shohamy 2001) and the intended and unintended effects and consequences which tests have on education and society and (eg. Messick 1981; 1989, Wall, 1997).

Given the powerful role of the IELTS as a gatekeeper and its effect on the life chances of those who take it, it is of vital importance that research is conducted into the uses which are made of IELTS test scores and the meanings which are ascribed to them. A number of predictive validation studies have been undertaken in Australia and New Zealand exploring the relationship between IELTS scores and subsequent academic performance (e.g. Bellingham 1993, Dooey & Oliver 2002, Elder 1993, Hill, Storch & Lynch 2000, Kerstjens & Nery 2001). The findings of these studies contribute to debates as to whether prescribed cut-offs in different institutional contexts have been set high enough to protect test-takers from undue risk of failure when they embark on their tertiary studies or whether they are they

unnecessarily stringent, preventing those who might ultimately do well from having a chance to demonstrate their academic ability. There is generally agreed to be a language threshold below which candidates are at risk of failure, but there are further questions worth exploring. Would students with lower scores be likely to succeed if granted provisional entry and given additional time to improve their English? How long will it take for those who score below the recommended threshold for entry to improve their English to a point where they are likely to achieve the requisite score? Are some skills more amenable to improvement than others? What strategies should institutions and individuals adopt to maximize students' chances of success? Should different goals and strategies be formulated depending on students' ability profile? Should some candidates be dissuaded from taking the test on the grounds that they are unlikely to succeed? Accurate advice to test-takers about what level of improvement one can hope for within a limited period of time (which may be all that is available to some learners), how best to achieve their short and long term goals and whether any progress they make is likely to be reflected in higher test scores clearly has important ethical implications. The emotional and financial costs of embarking on an intensive course of study in a new country and possibly failing to achieve one's goals may be considerable. The consequences of learners sitting the test too early should not be underestimated: the test is costly and candidates have to wait a further three months before they are allowed to re-sit.

Although there appear to be good reasons for investigating score gains, to date the number of studies looking at gain scores on the IELTS, or indeed on any other high stakes proficiency test, is fairly sparse. One reason for this could be the controversy within the educational measurement community surrounding the interpretation of the simple score gain as a measure of change. Zumbo (1999) however proposes that many of these difficulties surrounding these analyses can be resolved, provided that the assumptions underlying them are duly tested. Our view is that such a study is worth undertaking, if only to reveal some of the complexities involved. This study therefore sets out to investigate the following research questions:

Research question 1:
How much do students improve on the IELTS test after 10-12 weeks of intensive English study?

Research question 2:
What factors (personal, instructional and environmental) are associated with score gains?

2.0 *Context of the study*

2.1 Participants

130 participants were recruited for this study, all of them enrolled for intensive English language courses of a minimum of 10 weeks' duration in either Australia or New Zealand during the year 2001. Recruitment was restricted to students who intended to take the IELTS test (sooner or later) as a means of securing entry to a tertiary institution. For some students participation in the study gave them an opportunity for IELTS practice, others took the test because they felt they were ready and hoped to achieve their requisite score at one or other test sitting.

Care was taken to ensure that the timing of the pre- or post-test did not jeopardize the student's chance of obtaining their optimum score. For example, if a student was planning to enter course in July, taking the test two months prior to that date might not be in her best

interests because, given the three month minimum requirement, she would not be able to re-sit the test before the due date for admission had passed.

2.2 Characteristics of the sample

While students came from a wide range of countries including China, Japan, Korea, Vietnam, Malaysia, Indonesia Thailand, India, Sri Lanka, Latin America and different parts of Europe, the vast majority of the students were of Asian origin and the largest single group was from mainland China. There were similar numbers of males and females in the study and they ranged in age from 17 to 48 with the average age being 25 years. The average time spent in an English-medium country was just over one year, although the group was highly variable in this respect. The majority of students were on a temporary study visa and many of these had just arrived in the country when the study commenced, but around 25% of the sample were immigrants with permanent resident status who had generally been in the English-medium environment for a longer period. Most of the students were university bound, although not all of them had a clear academic goal in the sense of knowing what they were going to study and when. The educational qualification of the group was consistently high. Just under a quarter of the sample had already undertaken some postgraduate study in their home country. Nearly half had a bachelor degree and the remainder had either completed or almost completed their secondary schooling. The participants were distributed across four different course levels – intermediate, upper intermediate, advanced, upper advanced – with just under 50% in the advanced or upper advanced levels. There were a range of different criteria and procedures used across the four institutions for assigning students to these levels.

In keeping with their tertiary study aspirations, the majority of students were enrolled in English for Academic Purpose (EAP) rather than general English courses. It should be noted however that the definition of what constituted EAP and whether the EAP course was general or specific in its focus (see Table 1 below) varied somewhat from institution to institution. Some of the main points of difference included whether the EAP course included academic content as opposed to topics of general interest and whether it included the study of academic grammar and vocabulary. In two of the centres only the academic language was prescribed in the course syllabus. In one centre the academic content was specified in the syllabus. However, the main focus in this instance was on the study skills required in the academic area rather than the academic content per se. In none of the centres was both the academic and language content prescribed. Notwithstanding this variability amongst the four centres, all participating students undertook intensive English language study over the period they were enrolled in the institution. 'Intensive language study' is defined here as a minimum of 200 hours of teacher-fronted language instruction over 10-12 weeks, irrespective of whether the course of study was of an academic or general orientation.

The English level of the students at the commencement of the study (based on IELTS pre-test results) ranged widely with a handful of students with band levels as low as 1, 2 or 3 on one or other skill, and others with band levels of 7 and 7.5. The mean IELTS Global score of the sample was 5, and while no student had an overall score of less the 3, the majority were more than half a band below the requisite cut-off for university entry either in Australia or New Zealand. The skill profile was fairly even, although Writing and Listening were slightly lower (means of 4.7 and 4.9 respectively) than Reading and Speaking (5.2 in both cases).

2.3 Institutions

The participants were enrolled at four different institutions, two in Australia and two in New Zealand. The institutions differed from one another in terms of location and status (whether or not they are attached to a tertiary institution), staff qualifications, the type of teaching experience required for employment, external accountability requirements, the academic status, structure and typical duration of their courses, the regularity of student intake, the management of curriculum issues and the curriculum focus (i.e., which skills are emphasized and whether students are offered discipline-specific electives), whether or not IELTS preparation was integral to the course or taken as an extra, whether or not self access was compulsory or just an option available to students. It should also be noted that the scale of operations run by each institution differed, with Institution 4 being a fairly small operation in stark contrast to Institution 3, which was more corporate in its structure with branches in many countries throughout the world and affiliations with a number of major institutions. Institutions 1 and 2, in contrast to the other two, are both located on tertiary campuses and their courses provide a potential gateway to tertiary study at their own institution (although not all students take up this option). These differences are summarized in Table 1 below.

Another issue worthy of note is that one of the four institutions was in a state of major flux when the study commenced. The client base had expanded dramatically during 1999 and the EAP program was seen by its host university as a valuable source of income. The program was therefore moved out of its host academic department, and became a self-funding satellite attached to the Faculty of Arts. Accordingly it underwent a major change in management structure, resulting in the establishment of new curricula, new recruitment policies and the appointment of new teachers to meet the growing student demand. The program was also relocated to a building slightly off the central campus. Whether or how these changes impacted on the participant student population during this period remains uncertain, but they are documented here as one of the many possible explanations for the differences in score gains across institutions which are referred to later in this report.

Background information gathered from the participants in this study (see Table 2 below) also indicates that the profile of the student population in each of the 4 institutions is not uniform, with marked differences according to country of origin, age, residential status, educational level, proficiency level at entry and course type (whether EAP or general English). It is therefore inappropriate to compare one institution with another in terms of score gains achieved by their respective student populations, since the results might reveal more about the students than about the nature of the institutional provision. Any such study would need to be undertaken under very controlled conditions (e.g., with students of similar provenance and proficiency). Our interest is in the average score gains that can be expected of students across institutions (see Research Question One). Similarly, the factors contributing to score gains are examined across all institutions and not on an institution-by-institution basis (see Research Question Two).

	1 (NZ)	2 (NZ)	3 (AU)	4 (AU)
Campus location	YES Located close to central university campus.	YES Located on tertiary campus	NO Located approx. 10 kms from main campus of host university	NO No direct university affiliation but located close to a number of tertiary institutions
Staff have ESL teaching qualification	NO Only some	YES	YES At least Cambridge CELTA (as well as a degree)	YES At least Cambridge CELTA (as well as a degree)
All staff have EAP experience	NO Only some	NO Only some	YES	YES
Academic accountability	YES University Advisory Board	YES Dedicated program committee	NO (But external accreditation required)	NO (But external accreditation required)
Language course leads to an accredited qualification	NO	YES Certificate and Diploma in English	NO	NO
10 week model	YES	NO (16 weeks per semester)	YES	NO (20 weeks on average)
Weekly intake	YES	NO Once per semester	NO Every 5 weeks	YES
Centralized EAP curriculum	NO Under development	YES	YES	YES
Academic content part of core syllabus	NO Based on topics of general interest to adult learners. These are the vehicle for developing EAP skills S, L R & W (equally weighted)	YES For advanced students, but focus is on study skills required in particular academic area, rather than content per se	NO Syllabus specifies language content only, with emphasis on reading & writing for advanced students	NO Syllabus specifies language content only, with emphasis on reading & writing for advanced students
Discipline specific modules	NO Topics of general interest only	YES Students can take elective subjects related to their intended field of academic study	NO But students complete projects and tasks related to their proposed field of study at the most advanced level in particular.	NO Topics of general interest only
IELTS prep Integral	NO Optional additional short course	NO Optional additional short course	YES 10 weeks F-T course for advanced level students	NO Optional additional short course
Total hrs of instruction (face-to-face)	230 hrs (23 per week x 10)	216 hrs minimum (18 per week x 12)	200 hrs (20 per week x10)	240 hrs (20 per week x 12)
Compulsory self access	NO Optional extra	NO Optional extra	YES 5 hours per week	YES 5 hours per week

Table 1: Profile of participating institutions

	1 (NZ)	2 (NZ)	3 (AU)	4 (AU)
Number	31	33	37	11
Mainland China	100%	75%	16%	55%
Student visa	100%	82%	100%	100%
New arrivals (< 3 mths)	81%	9%	33.3%	18%
Tertiary qualified	75%	81%	59%	36%
Mean age	24	31	24	20
Mean IELTS global score at entry	5.5	5.5	4.5	5
Enrolled in EAP from the outset	100%	53%	73%	64%

Table 2: Participant profile at each institution

3.0 Methodology

3.1 Instruments

The following instruments were used for this study:

- The IELTS test which was administered free of charge to students at the beginning and end of a 10-12 week study period in the context of a regular administration session. Current rather than institutional versions of the test were used in all cases. The purpose of the pre and post-tests was to gauge score differences over time.

- Pre- and post-study questionnaires were administered to all participants at the beginning and near the end of the same study period. The questionnaires (see Appendix 6.1) contained a combination of forced-choice and open-ended items and were designed to elicit information about the range of variables which, according to our reading of the literature and discussions with EAP practitioners, might be expected to influence performance either positively or negatively.

- Semi-structured interviews with a subset of 18 students sampled according to their level of gain at the post-test session. The sample included a number of students who had improved and a number who had not progressed or whose scores had decreased. The interviews were built around students' questionnaire responses so a separate schedule of questions was not needed. They allowed opportunity for clarification of any ambiguities as well as a more fleshed out description of the variables of concern. Any further issues raised by the interviewees in the context of this discussion were recorded. The purpose of the interview was to probe the nature of students' English study experience and the possible reasons for their progress (or lack thereof) over the three month period. (We were however aware that we could not hope to capture everything which might be influential without more intrusive data gathering techniques and/or a more sustained engagement with the participants concerned). These interviews were conducted by one or both of the researchers, in some instances aided by a research assistant or an ESL teacher.

- Semi-structured interviews with administrators and teachers at each of the participating institutions to elicit information about the learning environment and the factors they saw as critical determinants of their students' English language

performance. Some interviews were conducted on a one-on-one basis and others in the context of a focus group. Teachers were also asked to give feedback on the particular students chosen for in depth interviewing. The interview schedule for administrators and teachers is provided in Appendix 2.

3.2   Administration Procedures

A small-scale pilot was done with a sample of five students prior to beginning the study. The pilot involved a focus group interview with the teachers at one of the four institutions to identify salient categories to be included in the student questionnaire, and trialling of the resultant elicitation instruments on five students at the institution. A more extensive pilot was not possible due to funding and time constraints.

The study proper began in March 2001 and continued through to January 2002. All IELTS testing was done in the context of regular IELTS test administration sessions to maximize the ecological validity of the research. Recruitment of students took place on an ongoing basis with different participants commencing the study at different times to coincide with scheduled test administration dates and in accordance with the constraints operating at each institution. Raters were unaware of which students were taking the test for research purposes.

The first questionnaire was administered as soon as participants had been recruited for the study. The second was administered towards the end of the three month period, in most cases around one week prior to the post-test session. Student interviews, which were all tape-recorded with the subjects' permission, were generally conducted *after* the second test administration but *before* candidates had received their official results. The researchers, however, had access to the students' test results at this point, allowing interview subjects to be chosen on the basis of score gains (negative, zero or positive). For logistical reasons a group of students at one of the four participating institutions were interviewed prior to sitting the IELTS test rather than afterwards. This was to ensure adequate levels of participation. The students has finished their course of study prior to attending the second testing session and it was feared that they would disperse immediately after sitting the test and not be contactable. In this latter case our selection of interviewees was based on teachers' hunches about which students would improve and which would not, rather than on actual improvement at the post-test session. Because we interviewed a larger sample than strictly required we felt assured of ending up with a sufficient range in terms of score gains and this in fact proved to be the case.

Interviews with institutional administrators were conducted fairly early in the study and those with teachers at various stages depending on availability. Field notes were taken during the interviews, which were also tape-recorded. Rough transcriptions were undertaken retrospectively to corroborate and supplement the researchers' field notes.

3.3   Validity constraints

There are a number of constraints stemming from the somewhat complex logistics of this study. These are noted briefly below. First it should be noted that not all of the students would have taken the test, had they not been offered it gratis, so the level of commitment may have been less than under normal testing conditions. Anxiety levels may also have been lower for those students who were using the test as a practice opportunity, compared to those needed to have their Time 2 results for university entry purposes.

It is also important to mention that it was not possible to administer the same test version to all candidates or indeed to synchronize the pre- and post-testing sessions. Given that the versions are statistically equated, this should not (strictly speaking) have a major influence on results. Indeed, one could argue that the use of multiple versions in this study is a strength in that we are deriving information from the tests under normal operational conditions. However, a certain degree of method effect may have resulted from the fact that some but not all of students sat for the Speaking test before and after July 2001 when a new Speaking test format was introduced. The new test consists of three parts: the first, an introduction and interview, the second, an individual long turn and the third, a two-way discussion. Its predecessor comprised five phases: introduction, extended discourse, elicitation, speculation and attitudes and conclusion.

In addition there may have been a practice effect on the Listening test for students at Institution 2 occasioned by the fact that, according to anecdotal feedback from their ESL teachers, these students were given the same version at both the pre- and post-testing sessions. The resultant practice effect, if indeed there was such an effect, appears to have been slight, since it did not result in significantly higher mean listening scores among this group of students compared to those at the other three institutions.

Because this was a longitudinal study, there was inevitably some attrition over the three month period, with a number of the students who presented at the pre-test session failing to show for the second round. These students defected from the study for a range of reasons (eg. family crises, ill-health, achieving their desired IELTS score at the first testing session). Although the initial sample had been boosted slightly in anticipation of this difficulty, time and funding constraints prevented us from achieving our target of 120 students. Pre- and post-test data are available for only 112 of the original 130 participants.

The piloting of the questionnaires (albeit limited in scope) meant that most questions were well understood by the participants in this study. There were nevertheless a few omissions which resulted in the data set being incomplete with respect to a number of the variables of interest. While many of the gaps in the data were able to be filled retrospectively, this was not possible in all cases.

In addition, the interview proved to be somewhat problematic as a means of eliciting rich information about students' language learning experiences. This was particularly true for low proficiency candidates, who sometimes misunderstood our questions or were unable to give an elaborated response. Some candidates also appeared reluctant to be drawn out about their English learning experiences. This may have been partly a function of personality, and partly due to the inhibiting presence of unfamiliar interviewers. It was for this reason that we decided to include students' teachers in some of the interviews conducted at the later stages of this study and this proved to be a successful strategy in so far as it elicited more fulsome responses than had previously been obtained from the subjects. (However there were also drawbacks associated with this latter approach, because the students, in the presence of the teacher, may have been reluctant to proffer information about, for example, the quality of the instruction received.) The timing of the interview (before or after the IELTS test) – see methods section above - may have also produced some variation in responses on attitude-related questions regarding their satisfaction with their course or their sense of their own progress and the reasons for it. (The responses amongst those interviewed after the second test may well have been coloured by the test-taking experience whereas the responses of those interviewed beforehand were more likely to have been based on their experience of language learning in a broader sense.). While the above factors may have had some impact on both the validity and reliability of the interview data, it should be noted that these data

were not used for the quantitative analyses, but simply to give "flesh" to the narrative vignettes.

4.0 Methods of analysis

Gain scores were calculated by subtracting the result of Test 1 from the result of Test 2 for the Global band score as well as for the sub test results for Listening, Reading, Writing and Speaking. Using a formula provided by Zumbo (1999), we then calculated the reliability of both subtest and Global gains. In addition, again following Zumbo, a latent variable modelling analysis was undertaken to determine the extent to which observed results may have been affected by error. Two models were fitted to estimate:

a) the 'true' gain,

and

b) the 'true' gain as a function of the results of the test given before the intensive 10-12 week course.

The standard error of measurement (SEM) used for these analyses was 0.36 (based on data supplied by the Research and Validation group at Cambridge ESOL). Note that multiple versions were used in this study (see above) and that reliabilities vary somewhat according to both version and the sub tests within each version. For practical reasons, the SEM estimate use for our calculations has not been calculated separately for each version or each sub test and is an average estimate based on the *overall* reliability of the academic IELTS test (derived from all versions used in 1999). Although only an approximation, it gives some sense of the impact of error on our findings.

Questionnaire data were coded and entered into a database together with the results of the IELTS test on the two test occasions and the Gain scores (see above). Some of the questionnaire data were categorical (e.g. sex and country of birth) and others were continuous or ordinal (e.g. years of prior English study or level of proficiency on IELTS). The data were carefully scrutinized, cross-tabulated with score gains and then recoded in some cases for the later analyses. For example, different questions on one of the questionnaires were sometimes combined and responses were collapsed into a smaller number of categories where this seemed appropriate. The variable entitled *Interaction*, for example, was the sum of a set of Likert scale responses to three different questions (see Q *7g, h & i* on Questionnaire 2, Appendix 6.1) which individually drew a somewhat narrow range of responses. *Age* was collapsed into four different categories (<20, 20-25, 26-30, 30<) because it did not appear to have a linear relationship with score gains (i.e., both younger and older students performed more poorly than the 20-25 year olds).

Interview data from the subset of 18 candidates were roughly transcribed. Transcripts were coded thematically and cross-referenced to teacher interviews (which were similarly coded). Findings were written up as a series of narrative vignettes (see below) containing the occasional verbatim citation from teacher or student where this was deemed to be more revealing than the summary information compiled by the researcher. The interviews were organized around the factors revealed, on the basis of our quantitative analysis, to be significantly associated with score gains (see below).

The same procedure was followed with teacher and administrator interviews. These were then summarized in the form of an institutional profile which was checked for accuracy by each of the participating institutions.

5.0 Results

5.1 Group findings

Research question 1: *How much do students improve after 10-12 weeks of intensive study?*

The latent structure analysis (referred to under Methods of Analysis above) revealed that the the 'true' score gain estimates (which take reliability into account) did not differ from the observed gain estimates. The only difference between the two modes of estimation was that using the 'true' scores produced slightly different confidence intervals (i.e. these intervals were tighter, as one might expect, for the 'true' score analysis than for the observed score analysis). In addition, when the 'true' gain was estimated as a function of the initial test, although the estimates of the slope were smaller than those derived from the observed score estimation they were nevertheless significant. Since the 'true' score gain analyses yielded similar results to those based on observed gains, we have not included the workings or the results in this report.

Results derived from the observed score gain analysis are presented in Table 3 below.

Columns 1-4 show the maximum and minimum gain (based on observed band scores) together with the median, mean and standard deviation for the 112 students in the final sample. Global band score gains are presented as well as gains on the component sub-tests of Listening, Reading, Writing and Speaking. The two right- hand columns indicate the significance level of the observed score change (calculated by means of the Wilcoxon's sign rank test using Version 10.0 of the SPSS for Windows software) and the reliability of this change (based on Zumbo's formula). These results show that the change in observed scores over the three-month period is statistically highly significant in all cases. The average overall gain is slightly more than half a band. The gain scores are higher for Listening than for the other skill components with a mean of .781. The reliability of the gain score estimate is generally high, although a little lower for the Global score than for the other four skills.

Worthy of note is the fact that some students actually performed worse at Time 1 than at Time 2. The maximum overall gain achieved by any student was 2 bands (2 students) and the minimum was −1.0 (one student). The greatest range of scores was for listening with the maximum gain of 4 (achieved by one student) and a minimum of −1.5 (for 3 of the participants).

Test component	Minimum Gain	Maximum Gain	Mdn	Mean	SD	Z	p	Reliability
Global	-1.0	2.0	.5	.598	.545	7.688	.000	.77
Listening	-1.5	4.0	.5	.781	.972	6.824	.000	.85
Reading	-1.5	2.0	.5	.402	.729	5.100	.000	.85
Writing	-1.0	3.0	.0	.545	.948	5.320	.000	.88
Speaking	-2.0	2.0	.0	.500	.930	4.964	.000	.86

Table 3: Table 3: Overall gains on IELTS (N= 112)

Table 4 shows the number of students at each band level whose score increases at the second testing session compared with those whose score remains the same or goes down. The mean gain for students at each band level is presented in the right hand column.

Level on Test 1	N	Higher score on Test 2	Same score on Test 2	Lower score on Test 2	Mean gain on Test 2
Band 3	1	1	-	-	1
Band 3.5	3	3	-	-	0.5
Band 4	8	7	1	-	0.88
Band 4.5	28	26	2	-	0.71
Band 5	22	15	5	2	0.68
Band 5.5	32	24	6	2	0.56
Band 6	13	6	5	2	0.42
Band 6.5	4	-	4	-	0
Band 7	1	1	-	-	0.5

Table 4: Frequencies of overall score gains across Band Levels

It can be seen from this table that the mean gain score on Test 2 tends to decrease as the students' proficiency increases. (We can perhaps ignore the score gains at the upper and lower end of the continuum since the numbers here are too small to allow any trend to be identified.) The fact that 6 students actually performed worse on Test 2 than at the first testing session could be explained by:

        a) what SLA researchers call the U curve, whereby interlanguage decreases in accuracy because the learner is experimenting with new forms,

and/or

        b) an idiosyncratic performance on the part of the candidate caused by such factors as fatigue or anxiety,

        c) inaccuracies in scoring, either at the first or second testing session,

or

        d) to a difference in difficulty of one version compared to another.

(While versions are statistically equated and therefore supposedly equal in difficulty, it should be noted that difficulty of each version is estimated for the group as a whole. It may be that individual test takers respond differentially to different versions, due to, say, their familiarity with a particular topic.) We can nevertheless take the mean gains (in the right hand column) as a rough indication of the likelihood of improvement at each score level, particularly at the middle band levels (4.5, 5 and 5.5) where the sample is probably large enough to mitigate the effect of any individual aberrations.

Research Question 2: *What factors (personal, instructional and environmental) are associated with score gains?*

Table 5 shows the list of variables (extracted from questionnaire responses which were explored as possible contributors to candidates' improvement on the IELTS test.

Category	Sub Category
Biodata	Age
	Sex
	Country of birth
	Educational level
	Residential status
L2 Proficiency	Prior English study (in home country)
	Recent English study (in Australia/NZ)
	Listening score
	Reading score
	Writing score
	Speaking proficiency
	Overall proficiency
Exposure	Time in English-speaking country
	Accommodation type
	Interactional opportunities
	Media exposure
Instruction	Institution where enrolled
	Course (General/academic)
	Level (Advanced/intermediate)
	Total hours of instruction
	Teacher qualifications
	Change of courses
	Change of teachers
	Attendance
	Studiousness
	Library use
Affect	Specific academic goal/motivation
	Perceived importance of Listening for future academic study
	Perceived importance of Reading for future academic study
	Perceived importance of Writing for future academic study
	Perceived importance of Speaking for future academic study
	Satisfaction with English course
	Perceived progress in Listening
	Perceived progress in Reading
	Perceived progress in Writing
	Perceived progress in Speaking
Test	IELTS practice in class
	Attended IELTS training course
	Awareness of test score required for entry to course
	Importance of test practice for successful performance
	Prior experience of taking IELTS or TOEFL

Table 5: Questionnaire variables identified as potential predictors of performance

Each of these independent variables have been clustered into a number of overarching categories which were identified by our informants as potentially influencing student progress, (e.g. Proficiency and Exposure). The grouping of variables within each category is somewhat arbitrary and we are aware that in some cases a particular variable could occupy more than one category. For example, *Satisfaction with the course* could be classified under Affect in that it was based on individual attitudes or under Instruction in the sense that it is presumably related to the teaching input.

Candidates' responses for each of these variables were regressed one by one against the dependent variable *Global score gain* which was coded on a scale of 6, with –1 indicating that the candidate performed worse at the second round of testing than on the first occasion, 0 indicating that the candidate performed at the same level and 0.5, 1 and 1.5 indicating various degrees of improvement in Global scores. It should be noted that the 1.5 category includes all those candidates who had improved by at least 1.5 band levels. (We noted earlier that some actually improved by 2 bands, but they were two few in number to occupy a separate category.) The multinomial logistic regression statistic (also available within Version 10.0 of the SPSS for Windows software program), which can accommodate both categorical and continuous data, was used for this analysis. Those presented in Table 6 are the ones found to make a significant (p = $\leq$0.05) or near significant difference to candidates' probability of improving their global IELTS band score.

Category	Sub Category	Chi square	DF	p
Biodata	Age	27.45	12	.007
	Country of birth	24.77	12	.016
	Educational level	20.81	8	.008
Proficiency	Reading proficiency	11.52	4	.021
	Writing proficiency	20.12	4	.000
	Global proficiency	15.87	4	.003
Exposure	Accommodation type	22.69	12	.029
Instruction	Institution where enrolled	35.31	12	.000
	Level (Advanced/intermediate)	15.46	8	.051
	Teacher qualifications	25.860	8	.001
	Hours of instruction	17.803	6	.007
Affect	Perceived importance of Writing	8.568	4	.073
	Self-assessed improvement in Reading	10.38	4	.035

Table 6: Factors influencing GLOBAL score gains on IELTS

Three of these variables (*Age, Country of Birth* and *Educational level*) fall into the Biodata category. The optimum age for this group in terms of improving the level of English appears to be 20-25. Students in this category perform significantly better than those who are younger or older. As for *Country of Birth*, it was found that the Chinese (mainly from Mainland China), who were the majority group, performed worse than those in the Japanese, Korean and Other (mainly from Latin America and Europe) categories. *Academic level* was

also influential in the sense that those with higher qualifications at entry to their course made greater progress on IELTS.

Only one variable in the Exposure category was related to overall score gains, namely: *Accommodation*. Interestingly, cross-tabulations reveal that those who made greater gains were students in the 'Other' category who had experienced different types of accommodation, rather than those who had spent the whole three months in a homestay, in a flat with other students or in the family home.

Four of the significant variables have to do with candidates' proficiency on commencement of the study. Cross-tabulations revealed that those who made the greatest gains were students at lower levels of proficiency as measured by their Time 1 IELTS score. The Listening and Speaking scores were the only ones which did not have significant predictive power.

Four of the instruction-related variables were significantly linked to score gains namely: *Institution, Teacher qualifications, Hours of instruction* and *Level of course.* Students at Institution 3 significantly outperformed those enrolled at any of the others as far as improvement from Test 1 to Test 2 was concerned. Students at the institutions with less qualified teachers did less well overall than those enrolled elsewhere, although there were other factors operating at these institutions which might have contributed to this outcome. *Hours of instruction*, interestingly enough, was negatively associated with score gains. *Level of course* mirrors the findings for proficiency in that students at lower levels tended to make greater gains than those who entered the institution with higher scores.

Finally, two affective factors, *Self-assessed improvement in reading* and *Perceived importance of writing* were found to be significant, or in the latter case near significant, predictors of overall score gains. The former indicates students' self-assessment of their own progress whereas the latter indicates the students' opinion regarding the importance of Writing for their future academic career. The difference is in the direction one would predict, with students who were more optimistic about their achievement and attributed greater importance to writing achieving greater score gains.

The above results suggest that score gains can be linked to a range of factors but tell us nothing about the relative contributions of each variable to overall improvement on the IELTS. This was explored by performing a backwards stepwise regression analysis, again using the multinomial logistic regression statistic. In other words, all the key variables[1] (i.e. those which emerged from the previous series of regression analyses as statistically significant or approaching significance) were entered into the model. We then removed, one at a time, those variables which appeared from the results of each successive iteration to make the smallest contribution to the whole. This process continued until we arrived at the combination that met each of the following criteria:

1. the standard errors for each variable at each score level were acceptably low;

2. all variables included in the model made a significant independent contribution to the whole (using $p = 0.05$ as our cut-off);

3. the model had optimum explanatory power i.e. a higher chi square value (relative to the degrees of freedom) than all the other factor combinations.

---

[1] Note however that we excluded global proficiency because it was in fact of a composite of proficiency on the other sub skills, two of which were also significant predictors.

The Chi square value for each variable in column 2 of Tables 7 below indicates the relationship between the final model and a reduced model without this variable present. The overall Chi square statistic reported at the bottom of the respective tables indicates the overall explanatory power of the model, compared with the null hypothesis that all variables entered into the model will have a zero effect. Note however that Chi square tests are asymptotic and thus may give unrealistic estimates of probability with a small sample size such as the one used in this study. The results reported below should therefore be interpreted with caution. They are simply indicative of trends in the data and it is unclear whether similar trends would emerge in a larger data set.

Effect	Chi Square	Df	Sig.
Intercept	.000	0	
Accommodation	28.34	4	.005
Level of course	20.63	12	.008
Educational	19.13	8	.014
Qualification	12.80	8	.012
Time 1 Reading score			

Model fitting information: Chi square 80.41 df 32, p=.000

Table 7: Best predictors of GLOBAL score gain on IELTS

The results in Table 7 above indicate that our best estimate of a candidate's chances of success on IELTS will be achieved by considering the kind of accommodation she has been in over the course of the study period (the more mobile the better), the level of course she is undertaking at the relevant institution (the lower the better) her educational qualifications (the higher the better) and her reading proficiency three months in advance of taking the test (the lower the better).

The same series of analyses were undertaken for the various sub skills on the IELTS in the event that particular variables might be more powerful predictors of some skills rather than others. Score categories on the outcome measure were reduced where necessary in order to enhance model fit.[2]

Those variables yielding significant Chi square values when regressed one by one against the measure of Listening gain are presented in Table 8 below.

---

[2] With a small sample size such as this one, the analysis is destabilized if there are two many levels of the dependent variable. For Listening we used 7 scale points from –1 (a negative gain of 1 band or more) through to +2.5 (a gain of 2.5 or more) whereas for Reading we used only six scale points from –1(a negative gain of 1 or more bands) through to 1.5 (positive gain of 1.5 or more) For Speaking and Writing we used 3 and 4 scale points respectively (-1.00. 0, +1 in the case of Speaking and –2, -1, 0 +1 and +2 in the case of Writing).

Category	Sub Category	Chi square	DF	p
Biodata	Educational level	26.569	14	.022
	Visa status	15.126	8	.034
Proficiency	Listening proficiency	39.549	7	.000
	Reading proficiency	20.020	7	.006
	Writing proficiency	28.798	7	.000
	Overall proficiency	25.830	7	.001
Instruction	Institution where enrolled	67.750	21	.000
	Change of course	24.606	7	.001
	Teacher qualification	42.911	12	.000
	Hours of instruction	42.073	11	.000
Affect	Perceived importance of Writing	15.012	7	.036
	Self-assessed improvement in Speaking	20.860	7	.004

Table 8: Factors influencing LISTENING score gains on IELTS

Educational level remains as a significant predictor with *visa status* also linked to score gains. Those who are permanent residents tended not to improve as much as the more recently arrived candidates with short-term student visas. Note also that *Listening* emerges alongside *Reading* and *Writing* and the *Overall* bandscore on the IELTS test as a significant predictor of Listening improvement over time. Again, as was the case for Global gains, the less skilled in Listening students were at the outset the more dramatic their improvement was.

Candidates' *self-assessed improvement in Speaking* (but surprisingly not in Listening) towards the end of the 12-week period is also a significant indicator of actual progress. Interestingly, the relationship between score gains and the *Institution* students attend is again highly significant (with mean scores again significantly higher at one institution than at the others). As might be predicted, score gains are significantly less likely at institutions with less *qualified teachers*. *Hours of instruction* on the other hand is negatively associated with score gains with those spending less time in class performing better. *Level of course* is not significantly associated with Listening score gains but *Change of course* here takes its place a predictor. Since most of those who change courses move upward (from one level to another) it seems likely that this variable is a surrogate measure of progress as gauged by the students' teachers, who are generally the ones who decide whether or not a student can proceed to another level. As shown in Table 9 below, the variables that, in combination, have the best explanatory power and also meet the model fit criteria referred to above are *Listening* (the lower the initial score the greater the gain) and *Institution* (Institution 1 has a lower proportion of students showing improvement than the other three).

Effect	Chi Square	Df	Sig.
Listening	26.772	6	.000
Institution	51.952	18	.000

Model fitting information: Chi square 91.457 df 24 p=.000

Table 9: Best predictors of LISTENING score gain on IELTS

For Reading the picture is somewhat different as shown in Table 10 below. The only significant predictors of improvement are variables measuring candidates' proficiency level (expressed either as a *Reading, Writing* or *Overall* bandscore on IELTS) at the beginning of the study, together with another Exposure variable, namely: the frequency of visits to the library and a test-related variable, candidates' prior experience of test-taking (which could be either IELTS or another academic proficiency measure such as TOEFL). There is some link between age and reading gains, with those in the 16-19 age group performing at slightly lower levels than the others, but this effect is not significant. When all of the above factors are combined and the model reduced according to the criteria specified above, we find the best predictors of progress to be two variables only (see Table 10 below): first and foremost, the *Reading* level on arrival (the lower the better) and second, reassuringly, the amount of time one spends in the *library* during the 12 week study period (the more the better).

Category	Sub Category	Chi square	DF	*P*
Biodata	Age	24.60	15	.056
Proficiency	Reading proficiency	44.460	5	.000
	Writing proficiency	19.685	5	.001
	Global proficiency	27.30	5	.000
Exposure	Library	10.013	5	.051
Test	Test experience	18.912	10	.041

Table 10: Factors influencing READING score gains on IELTS

Effect	Chi Square	df	Sig.
Intercept	.000	0	
Reading proficiency	47.335	5	.000
Library attendance	30.186	15	.011

Model fitting information: Chi square 72.307 df 20, p=.000

Table 11: Best predictors of READING score gain on IELTS

Moving on to Speaking gains, it can be noted in Table 12 that candidates' *Speaking* level at the outset emerges as a significant predictor of progress on this skill, with *Reading* less strongly associated with score gains, as one might expect. *Time spent in an English-speaking country*, whether New Zealand or Australia, is also influential, whereas this was not the case for Listening. Candidates' *satisfaction with their English course* is linked to speaking test outcomes (although this p value is slightly above the 0.05 level required to claim statistical significance). Interestingly, another affective factor, *Attitude to the IELTS test* as a measure of proficiency, seems to play a role, with negative perceptions associated with lower score gains. These findings are presented in Table 12 below.

Category	Sub Category	Chi square	DF	P
Affect	Satisfaction with the course	7.024	3	.071
Proficiency	Speaking proficiency	33.470	3	.000
	Reading proficiency	7.151	3	.067
Exposure	Time in country	17.385	9	.043
Test	Perception of IELTS	18.912	10	.041

Table 12: Factors influencing SPEAKING score gains in IELTS

Table 13 shows the optimum combination of factors when all of the above variables are put together, namely: a candidate's *Speaking and Reading proficiency* at the beginning of the study (the lower the better) and his/her *attitude to the IELTS test* as a measure of this proficiency (the more positive the better).

Effect	Chi Square	df	Sig.
Intercept	.000	0	
Reading proficiency	13.749	3	.003
Speaking proficiency	38.959	3	.000
Perception of IELTS	10.059	3	.011

Model fitting information: Chi square 56.051 df 9, p=.000

Table 13:   Best predictors of SPEAKING score gain on IELTS

When the same analysis is done for Writing we find that only four of the variables hypothesized to affect performance appear to make a difference to Writing score gains. These results appear in Table 14 below. Of these four, only two emerge as best predictors of Writing gain: candidates' writing score at the outset (the lower the better) and their perceptions regarding the importance of writing for future academic study (see Table 15 below).

Category	Sub Category	Chi square	DF	p
Proficiency	Writing proficiency	36.321	3	.000
	Overall proficiency	8.127	3	.067
Affect/attitude	Perceived importance of writing	11.486	3	.009
	Self-assessed improvement in listening	7.482	3	.059

Table 14: Factors influencing WRITING score gains on IELTS

Effect	Chi Square	df	Sig.
Intercept	.000	0	
Writing proficiency	76.643	3	.000
Estimated importance of writing	60.902	3	.008

Model fitting information: Chi square 39.092, df 6, p=.000

Table 15: Best predictors of WRITING score gain on IELTS

5.1.2 Summary of group findings

It would appear from these analyses that the proficiency one starts with (as measured by the IELTS test) is the most constant indicator of how far one is likely to "travel" over the course of a three-month period of study, with initial *Listening proficiency* being the best predictor of Listening improvement, initial *Reading proficiency* the best indicator of Reading progress and likewise for *Writing* and *Speaking proficiency*. For both the global score and for each of the separate sub skills it is easier to move up from one step to another at the lower end of the IELTS rating scale (i.e., from level 4, and 4.5 than at the upper end of the scale [6 and 6.5]). At the 5.5 level a candidate has a good chance of improving by half a band, but a higher overall score gain is the exception rather than the rule. At the Band 6 level, according to the results reported here, a candidate has less than a 50% chance of increasing her global band score. Although it has been argued in the literature that such a result may be due to regression to the mean, this remains the subject of controversy (see for example Zumbo, 1999: 275). The findings of the latent variable analyses referred to earlier suggest that this is not the full explanation.

Among the Biodata variables, *Educational Qualification* was the most important one (although only for Listening and Overall score gains). It seems reasonable to speculate that one's level of education is a surrogate measure of intelligence and/or aptitude and/or first language literacy and that the more educated one is the better one is likely to cope with academic demands in any language. The *Age* factor was significant on its own as predictor of overall proficiency gains, but not in combination with other factors. While youth seems to be an advantage, in that students in the 20-25 age group outperformed those who were older, very young students in this study (i.e. those below 20) were less likely improve over the three-month period. It may be that a minimum degree of maturity is required to cope with the academic demands of an adult EAP course. As for the older students (i.e. those over 25), it was suggested by some of our teacher informants they tend to be less flexible in their approach to learning, and less willing to adapt to the communicative orientation of the classroom, and also that they are more likely to have family responsibilities which can distract them from their studies. For some of them, moreover, quite a long period has elapsed since they have been students and it therefore may take them time to readapt to a formal learning situation. There appears to be some overlap between *Visa status* and *Age*, with the permanent residents (who are generally older) less likely to improve than those on short term study visas. However, as was the case for *Age*, this latter factor did not appear in any of the combinatory models. *Country of birth* appeared to have some influence on study outcomes with those from mainland China improving less than those from other countries. This may be due to a combination of linguistic and cultural distance as well as to differences in learning style. However it may equally be a function of the high concentrations of Chinese speakers in one of the four institutions resulting in limited opportunities for English interaction. In any case, this factor was not one of the best predictors of score gains.

With regard to the Exposure variables, it was somewhat surprising that self-reported *Media* exposure and opportunities for *Interaction* were not associated with enhanced performance amongst this group of participants. It seems that media exposure on its own is insufficient to bring about measurable improvement within such a limited time frame. It is also likely that quality of interaction is more important than quantity. A more sensitive language contact scale would need to be applied before we can dismiss this factor with any certainty. *Time in Country* and *Accommodation* and were the only Exposure-related variables which had an influence with the former predicting performance on the Speaking component only and the latter predicting IELTS performance gains overall. Recall that it was those who moved around within the time frame of the study, rather than those who remained, say, with their

homestay family, who made significantly greater gains. It is probable that moving from one place to another increases the opportunities for hearing and speaking English with different interlocutors. We might also surmise that the *Change of accommodation* category is a surrogate for initiative, with those who move from one situation to another being the ones prepared to go to some lengths to make the most of their in-country study experience. Welsh (2002) in his recent study of perceptions of the homestay experience amongst EAP students in NZ, found that while students were often dissatisfied with the quantity and quality of communication opportunities available in the context of their homestay, only a few of them took steps to move to another environment. Interestingly he found that Asian students were less likely to leave their homestay than students from non Asian countries.

Speaking and Reading were the only skills where test related factors played a role, with reading improvement more likely amongst those with prior *test experience* and speaking improvement more likely amongst those with a positive *attitude to the IELTS test* as a measure of their English proficiency.

As for the instructional variables, there is no evidence from the questionnaire data that IELTS training courses, or courses which are built around IELTS practice materials, increase the likelihood of improvement overall. This could be seen as an encouraging finding. If IELTS score gains are not unduly amenable to coaching, they are arguably more valid as a measure of proficiency. On the other hand it must be noted that the institution (Institution 3) with the highest rate of improvement was the only one at which a 10-week IELTS practice component was integral to the course of study offered at the advanced level. Before we accept this as the sole explanation for Institution 3's better track record of score gains, we must again caution that score gain differences across institutions may be due to a range of factors (relating to differences in the student profile) and that no firm conclusions can be drawn about the quality of the educational experience that each institution offers from the data reported here. Although *Teacher qualifications* may well have been a factor influencing students' somewhat lower level of progress at Institution 1, it was not one of the best predictors. We should also remember that this was the institution where students had a higher mean proficiency level overall (recall that more proficient students are less likely to progress) and where the population was highly homogeneous (i.e. the vast majority of students studying there were from mainland China). The homogeneity of the student population may in turn have reduced the likelihood of high levels of L2 interaction both inside and outside the classroom (and the teachers at this institution commented on this.) The fact that classrooms may not be a source of rich English language input for learners is also supported by the fact that score gains overall and listening skills in particular were negatively associated with *hours of instruction*. (It should be remembered that fewer classroom hours may result in more outside exposure.) As for the other instructional categories, there is little or nothing to report. *Course type* (academic or general) does not seem to matter although, as noted earlier, the distinction between academic and general proficiency is difficult to draw in teaching contexts such as these where, in all cases, there is a greater focus on literacy or cognitive academic language proficiency rather than on basic interpersonal skills. *Level of course* matters, but probably only because it acts as a surrogate for proficiency (i.e. students at the lower levels improve more.)

Affective factors did not loom large as predictors of score gains. Students own estimates of progress were associated with score gains on IELTS but this self-assessment is more likely to have been as much the <u>result</u> of an increase in proficiency as the <u>cause</u> of the improvement. The perceived importance of writing for future academic study was related to progress in writing (the more important this skill was perceived to be the greater the gain) but this factor was not one of the best predictors of progress overall. The limited role of affective factors may be partly due to lack of variability within the sample. The population overall was very

motivated and highly committed to the task of upgrading their English proficiency (all of them fee-paying students with the ultimate goal of proceeding to higher study). Moreover, we would caution that questionnaire data may not be a good vehicle for eliciting information about student attitudes. Data from the case studies, to which we now turn, gives grounds for believing in the importance of affect, although it clearly interacts with other factors.

5.2 Case studies

Five of the eighteen case studies undertaken for the project, considered to be either illustrative of particular trends in the data, or clear exceptions to the rule, are presented below. The texts are annotated with the factors revealed above to be predictive of progress in italics. The + sign is used to remind the reader of the factors found to be positively associated with score gains and the – sign indicates those which work in the reverse direction. A list of relevant predictive factors is included beneath each profile. Factors revealed to be neither positive nor negative in terms of our previous quantitative analysis are not listed. For example, if a student was found to have neither negative nor positive attitudes to the IELTS test, or to his/her course of instruction, this factor has not been included on the list. Likewise, if a student's proficiency level is around 5.5, this is reckoned to be neither a negative or positive factor and therefore not included, whereas if it is above or below this level it is classed as negative or positive respectively.

5.2.1 Case Study One - Cassio
The first of the case study students has been named Cassio. He was aged 24 (*Age* +) at the commencement of the study and had just arrived from Columbia (*Country of Birth* +). He had already completed his undergraduate studies (*Educational Qualification* +) and travelled to Australia with the intention of completing a further postgraduate diploma in Finance with the clear knowledge that he would need to achieve an overall IELTS entry score for this course of at least 6.5. When he arrived he lived briefly with an Australian family but decided to move into a flat (*Accommodation* +) with his Columbian friends. While this meant that he did not normally speak English at home, he had made some non-Columbian friends at the language centre many of whom he socialised with out of class. He was also beginning to develop Australian friends outside the centre.

In terms of his EAP instruction Cassio changed classes (*Change of course* +) more than once. moving from the lower level EAP 1 up to the higher level IELTS preparation course between July and October and since then up another level into the advanced EAP course.

When interviewed he claimed that English had been a serious problem for him back in Columbia. While he knew a lot of grammar he had very limited opportunities in and out of class to practise using the language and his English level was quite low (*Proficiency* +). He decided to come and study in Australia so that he would improve his mastery of the language. He considered this vital for his future job prospects.

He reported that he was enjoying both the communicative orientation of the classroom (*Satisfied with course* +) as well as the opportunities to speak English with both his classmates and native speakers outside the centre, although with native speakers he still felt quite nervous in interaction. He said that he made friends easily and had established a network of friends from other countries through his contacts at the centre. He regularly went out to bars and clubs in Melbourne with these friends. In general, he indicated that pressure to use English inside and outside the classroom had substantially lifted his motivation to improve.

Cassio clearly had a strong investment in improving his English. At one point he said:

> *I came here because I wanted to learn, I paid to learn. Now I put all of me into learning.*

Cassio's teachers described him as an active, creative learner who had a strong desire to communicate in class. He was also a highly analytic learner who frequently asked for detailed clarification of grammar and vocabulary points. They indicated that he was usually willing to make intelligent guesses in class and that he was not too phased when he made mistakes. The distinguishing features of his classroom learning, however, were his constant willingness to be creative with new language and perhaps, even more importantly, his ability to learn co-operatively. Both his teachers indicated that, like many other students from Latin America, Cassio was extremely popular with his classmates whatever their nationality: they all liked and admired his open, friendly nature and eagerly sought his participation in group and pair work.

Cassio was very pleased with his level of progress (*Perceived improvement* +) over the 3 month period. He estimated that his listening, speaking and writing had improved 'quite a lot' and his reading only 'a bit'. He attributed all of his progress in writing to classroom instruction and his oral skills, especially his listening, to practice opportunities both inside and outside the classroom. On the other hand, he believed that his reading skills were already strong when he arrived (*Reading proficiency* -) and therefore that his improvement was not as marked as in the other areas. He felt overall that speaking was still his weakest area.

Here is a summary profile of Cassio giving positive negative or neutral ratings on the various factors found in our quantitative analysis, to be predictors of improvement.

Age: 24 (+)
Country of birth: Columbia (+)
Educational qualification: Graduate (+)
Visa status: student (+)
Listening, Writing and Speaking Proficiency: Low (+)
Reading Proficiency: High (-)
Changed Accommodation: Yes (+)
Institution: 3 (+)
Change of course: Yes (+)
Satisfaction with course: Yes (+)
Perceived improvement: Yes (+)

It can be seen that Cassio scores positively on nearly all of these factors which helps explain his dramatic improvement as shown below

	Listening	Reading	Writing	Speaking	Global
Test 1	4.0	6.5	5.0	5.0	5.0
Test 2	7.0	7.5	7.0	6.0	7.0

5.2.2 Case Study Two - Kim

Kim is 29 years of age (*Age* -) and worked as a physiotherapist and university lecturer in Korea (*Educational level* +) before emigrating to NZ with her husband, a trained accountant, and their 10-year-old son. She had studied English for many years in Korea and her proficiency level at the beginning of the study was above the mean for the sample as a whole

(Proficiency -). She wished to continue her professional practice as a physiotherapist in New Zealand and therefore needed a locally recognized qualification.

Kim was a highly motivated and dedicated student who spent many hours studying English

> *I get up at 5am and study for 2 hours before I come to class*

In an attempt to improve her level of English Kim usually spent three hours in the language learning centre after she had finished her formal classes. She also reported listening to the radio and using the internet.

Before enrolling at her current institution Kim had employed a tutor to help her with her English. Her teachers described her as highly dedicated. She had nevertheless failed her first writing assignment and had difficulty understanding where she had gone wrong in terms of word order and syntax. Her teachers commented that she seemed unable to self-correct even when given very explicit feedback.

While she was delighted with her course and her teachers *(Satisfaction with course +),* her estimate of her progress was rather conservative, and she seemed least confident about her reading ability *(Perceived improvement -).* She felt she lacked the necessary vocabulary and found the format of the IELTS reading questions perplexing. While she felt her speaking skills had improved she believed that this would not show up in her test result because IELTS was a *"test rather than natural speaking"* *(Attitude to IELTS -).*

Kim lived at home with her family *(Accommodation -)* and talked somewhat apologetically about the fact that she used Korean most of the time in communication with both her son and her husband. Although her husband was proficient in English, she felt inhibited about practising with him because he was better than her. It seemed in fact that there was some role conflict in her domestic situation. Kim said that she liked living in New Zealand because women can be equal but reported that her husband did not share this view. He regretted coming to New Zealand because he had not yet found a job and was ordered by his wife to share with housework and childrearing. While there were limited opportunities for English interaction at home, Kim did report speaking English with her neighbours and with her husband's friends when they visited.

Looking at the summary profile below it would appear that the odds weigh against Kim improving her IELTS score. There are more negative than positive factors. Her high level of education one would expect to work in her favour, but her initial level of proficiency, which is in the middle to high range for all skills except writing, combined with her age and permanent resident status, suggests that dramatic gains are unlikely over the three month period. In fact she makes no progress overall and her reading score is worse at Time 2 than Time 1. The only improvement is in Listening and this is slightly below the mean for the sample as a whole. Given her extraordinary conscientiousness this is likely to be very disappointing and a further blow to her already precarious self-esteem.

Age: 29 (-)
Visa status: Permanent resident (-)
Educational qualification: Postgraduate (+)
Speaking Proficiency: High (-)
Writing Proficiency: Low (+)
Changed Accommodation: No (-)
Satisfaction with course: High (+)
Attitude to IELTS: Negative (-)

Perceived reading improvement: No (-)

	Listening	Reading	Writing	Speaking	Global
Test 1	5.5	5.5	5.0	6.0	5.5
Test 2	6.0	5.0	5.0	6.0	5.5

5.2.3 Case Study Three - Liao

Liao, a 26 year old (*Age* -) male from mainland China (*Country of Birth* -), had been in NZ for 18 months at the start of the study. He had been studying English continuously since his arrival in NZ and had just transferred from another institution. He was enrolled in an intermediate English class when the study commenced and his intention was to proceed to study Business at a university in Auckland. He expected that this would help him get a job back in China but would also help to improve his English in the professional domain. He was aware that entry to this course was conditional on his achieving a score of 6 on IELTS. His stay in NZ was being paid for by his uncle's business and he clearly saw the need to improve his English. However he claimed that he was unable to study hard, and linked the inability to work to stress:

Interviewer:	*Why are you stressed?*
Liao:	*I don't know, I haven't been... feel tired, maybe tired, because I didn't take any holiday and break.*

In fact he defected from his course two weeks early due to burn out "*my head is hurting*" and proceeded to study on his own in preparation for the test. He had planned to enrol for an IELTS practice course but later decided against this.

Liao's use of English in the home was somewhat limited, since he was flatting with other Chinese students. He said that he had moved to a new flat quite recently (*Accommodation* +) and now had a television. He also reported talking to kiwi friends he had met at church. In addition he had been working part-time after class and this had given him opportunities to listen to spoken English. In addition he mentioned listening to the news on the radio, watching movies at the self-access centre and spending time in the library (*Library* +), and expressed satisfaction with the fact that he could now understand a lot of what he read:

During last week I have reading simple.. just the intermediate book.. I understand!

What came across at interview was his high level of anxiety (and this was also mentioned by his teachers). Liao seemed to have low self-esteem and the anxiety he experienced was debilitating rather than facilitative. He reported tackling only one of the two writing tasks at the initial IELTS testing session, because his mind went blank and he couldn't think of anything to say. Although he felt his English was slightly better after the three-month study period, he did not expect a great deal of improvement at the post-test session (*Perceived improvement* -). (Note that he had already sat for the test three months before our study commenced so had a fairly realistic idea of what was involved.) At the second interview he seemed slightly more confident about his writing than before because of rhetorical formulae he had learned in class. (He gave "*not only... but also*" by way of example.) He felt that these pre-rehearsed phrases would help him on the test. In fact, this was the only skill in which he expected to achieve a higher score. As for speaking, our mention of an imminent change in the structure of the IELTS interview did not seem to faze him. He seemed more confident in his speaking than his other skills (although he also reported receiving negative feedback from his teacher).

> *Last time I didn't prepare anything for speaking and my teacher tell me I make a lot of grammar mistakes, so I pay attention.*

Liao made the point that the IELTS test was measuring not just language proficiency but also cognitive ability (*Attitude to the test* -), and this was a further source of stress for him.

Liao does not seem to have enjoyed studying English (*Satisfaction with course* -) and did not like changing teachers because this made him feel insecure. His approach to learning appears to have been conditioned by his experience of English language study in China, where the focus was on acquiring discrete point items of vocabulary and grammar. His comments suggest that he relied a lot on memory and rote learning.

> *And the academic vocabulary is so difficult, I couldn't remember them.*

He felt conflict about the course he was enrolled in, saying that it placed too little emphasis on vocabulary and it was difficult to process the different grammar points taught by the different teachers. (Any focus on form at Institution 1 is incidental, arising out of student need or teachers' perception of this need, rather than in the context of an explicit grammatical syllabus).

> *I would change some of the way to teach. I like traditional sometimes. Communicative approach...good for another person but for me, I want to change the way to improve... teach the more difficult vocabulary for the student.. teacher teach the vocabulary not enough .*

On the other hand, while he acknowledged that he would be more comfortable with a more systematic and traditional method, he realized that this might not be in his best interests and that that the more flexible interactive approach adopted by his teachers was cognitively challenging and provided a good opportunity for practice and for independent thinking.

Below is the summary profile for Liao, who shows a mix of positive and negative ratings on the various predictive variables. It seems that these factors cancel one another out, such that Liao shows no overall improvement at the end of the end of the three-month period. Note however that his performance is somewhat atypical of the sample in that his Reading improvement is substantially higher than the mean (and in this respect his performance corresponds with predictions given that the two variables which emerged as best predictors of reading improvement were low initial reading proficiency and use of the library). The only other skill he improves in his Listening, but here his progress is below the average for the group as a whole.

Age: 26 (-)
Country of birth: China (-)
Visa status: Student (+)
Reading, Writing and Listening Proficiency: Low (+)
Speaking Proficiency: High (-)
Changed Accommodation: Yes (+)
Institution: 1 (-)
Library use: Frequent (+)
Test experience: Yes (+)
Perceived improvement: No (-)

	Listening	Reading	Writing	Speaking	Global
Test 1	4.5	4.5	4.0	6.0	5.0
Test 2	5.0	5.5	4.0	6.0	5.0

5.2.4 Case Study Four - Toshi

The fourth learner has been named Toshi. He was aged 38 (*Age -*) and from Japan. He arrived in Australia in April 2001. He completed a Bachelor degree (*Educational qualification +*) 15 years ago in Japan and came to Australia to learn English after resigning from his job in a quality assurance company. He planned to enter into the field of international training and development. Originally, he intended to complete a relevant postgraduate qualification in this area in Australia, but by the end of the study had decided that he would return to Japan to undertake further studies.

Toshi had lived alone since arriving in Melbourne (*Accommodation -*). He didn't speak English at the centre outside class and mixed exclusively with Japanese people beyond the centre. In general, he claimed that his motivation to learn English had slipped because his Japanese friends in Australia had persuaded him that he didn't really need a very high level of English to move into his new career. Besides, he was more comfortable sticking with his Japanese friends and was reluctant to speak English socially because he found it too stressful. He had tried to talk to his neighbours at home when he first arrived but found that they were not very friendly and that it was easier if he didn't persevere with his attempts to interact with them.

In terms of his EAP instruction he has spent most of the time since arriving at the centre in EAP 1 (an advanced EAP class) but had recently negotiated with his teachers to move down to an Upper Intermediate General English class (*Change of course -*). He did not enjoy his first class (*Satisfied with course -*) and at interview reported finding the lower level class more manageable than the previous one both because of the reduced academic focus and the lower proficiency of most of the students. The only problem was the change of teachers, which he found difficult to adjust to.

In his second questionnaire Toshi indicated that he listened to the radio in English almost every day, always did his homework and spent a lot of extra time in the centre's Individual Learning Centre.

Toshi's teachers indicated that he worked fairly enthusiastically, especially in his current class. They said he was prepared to some risks and make guesses in whole class work but only when specifically addressed. However, they did suggest that he was a somewhat isolated figure in class in so far as he was reluctant to join in group work. When he did so he neither initiated conversation nor responded much at all to others' contributions. This resulted in him being ignored and frequently being left out in the cold by other more sociable members of the class. One of his teachers suggested that students of other nationalities sometimes regarded Japanese students as aloof. The other teacher thought his age (*Age -*) might have been the more significant factor in him being alienated from the rest of the group.

Before sitting for his second IELTS test his self-assessment of his progress over the last 3 months Toshi considered that he had not improved at all in listening, speaking, reading or writing (*Perceived improvement -*). He also believed that his IELTS results would not improve in any of the four sub-tests.

Here is Toshi's summary profile, which shows negative ratings against many of the significant predictors of score gains. Toshi's overall band score was lower at the second testing session than at the first. In his post-test interview he claimed not to be overly surprised by his results although he did feel that the reading and writing tests were more difficult in the second test. (This perception may at least partially explain the decline in his results on these two sub-tests.) On the positive side, it is worth noting that his listening result had improved by 1.5 bands. This indicates that he had benefited from his exposure to spoken English inside and outside class even if he had not actively participated in spoken interaction much himself.

Age: 38 (-)
Educational qualification: Graduate (+)
Change of accommodation: No (-)
Listening, Writing and Speaking Proficiency: Low (+)
Institution: 3 (+)
Level of course: Advanced (-)
Change of course: Moved down *(-)*
Satisfaction with course: No (-)
Perceived improvement: No (-)
Attitude to the test: Negative (-)

	Listening	Reading	Writing	Speaking	Global
Test 1	4.0	5.5	5.0	5.0	5.0
Test 2	5.5	4.0	4.0	5.0	4.5

5.2.5 Case Study 5 - William

William is a mature student (*Age* -) from Ethiopia who came to New Zealand with his wife in 1998. Although he arrived as a refugee he now had permanent resident status (*Visa status* -). Although William had only completed 4 years of secondary schooling he had begun training as a nurse in his home country and his goal was to study Nursing in NZ, Although, with an initial IELTS score of 4.4 (*Overall proficiency* +) he was a long way off the required score of 6.5, he was determined to achieve this goal.

William was described by his teachers as "withdrawn and lacking in confidence" and did not seem optimistic about his progress (*Perceived improvement* -). One teacher commented that he seemed preoccupied and that although his oral proficiency was quite good (*Speaking proficiency* -) he often had difficulty understanding what was required of him in class, as she had found to be the case for many other African students. He also had a slight hearing problem which may have further impeded his participation. His performance on classroom tasks was barely adequate and it was noted that while his vocabulary knowledge was broad, his writing style was "very flowery" and did not conform to the academic model.

He was reported to be lonely. He lived at home with his wife (*Accommodation* -) and had very little contact with other Ethiopians. His interaction outside the classroom tended to be solely with his wife, who was not proficient in English. William commented that other students in his class tended to speak in their own languages so here too there were limited opportunities for conversational practice. It also seemed that he was not exploiting the institutional facilities to the full, preferring to study in a quiet place on his own rather than using the Internet or the language laboratory. However he did practice his writing outside class and read the newspaper regularly.

William chose an IELTS elective in his Diploma course as well as attending a short IELTS training course. Nevertheless he seemed rather unconfident about his prospects for performing well on the IELTS test, and was concerned about the lack of time available to complete the tasks (*Attitude to IELTS -*).

His summary profile below would lead us to predict minimal score gains on the test because there are more factors working against him than in his favour. His level of improvement, which is well above the group average, is therefore surprising, in particular his massive gain of 1.5 bands for Reading and 2 bands for Writing.

Age: 41 (-)
Country of birth: Ethiopian (+)
Visa status: Permanent resident (-)
Educational level: secondary school incomplete (-)
Listening, Reading and Writing Proficiency: Low (+)
Speaking proficiency: High (-)
Changed Accommodation: No (-)
Test experience: None (-)
Attitude to IELTS: Negative (-)
Perceived improvement: No (-)

	Listening	Reading	Writing	Speaking	Overall
Test 1	3.5	3.5	4.0	6.0	4.5
Test 2	4.0	5.0	6.0	6.0	5.5

6.0 Discussion

Most of the profiles can be considered illustrative of the trends in the quantitative data in the sense that the positive predictors we identified are associated with progress as measured by the IELTS test and vice versa with negative predictors. Where there is a mix of positive and negative factors the gains or losses tend to be less dramatic. The case studies also confirmed what was found in the quantitative analysis, namely, that certain qualities such as studiousness and exposure to the media did not, on their own, offer any guarantee of progress in English. Kim, for example, was exceptionally studious and spent lots of time listening to the news, watching television and working on the internet but showed no overall improvement over the three month period. Lest we give too much credence to the findings for the group as a whole, we should also note that there were a number of students like William, whose gains on the IELTS test were unexpected. In contrast to his own predictions William managed to improve by half a band in listening (in spite of his hearing disability) and made dramatic gains in reading and writing (in spite of his concerns about time limitations and his teachers' reports that he was struggling in both of these areas). These unpredictable cases suggest the need for caution in generalizing about what counts towards success in English learning in general and IELTS test performance in particular.

Furthermore, what comes across powerfully when one immerses oneself in this qualitative data is that there are forces at work which have not been captured by our rather crude quantitative analysis which is based on very limited self-report data. Leaving aside language aptitude, which has been found by Skehan (1989) to be the most powerful of all individual influences on language learning, and was not measured in this study, there were also personality traits, such as confidence (revealed by Cassio) and anxiety (displayed by Liao) which appeared to play a part in the language learning process, and indeed may also have

affected their test performance. Anxiety has been shown by a number or researchers to be negatively associated with speaking performance in language classrooms (Horwitz et al. 1986) and to IELTS interview performance in particular (Bond 2000), whereas the opposite has been found to be true for extroversion (Dewaele & Furham 1999, Oya 2002). Differences in culture and learning style also emerged as important in our case study data. A number of Asian students, like Liao, were ill at ease with learner-centred orientation of their classes, and the unsystematic way in which formal features of the language were treated in class. While these issues came up in the context of the interview, they were not captured in questionnaire responses.

Learner strategies of the kind explored in Good Language Learner studies (e.g. Naiman, Fohlich, Stern & Todesco 1987, Beebe 1998) were not directly investigated in this research, although the interviewed teachers did provide some feedback on individual learners' approaches. Summarizing the findings of a number of such studies, Ellis (1994) identifies five characteristics of successful language learners. They:

- attend to form
- attend to meaning
- involve themselves in the language learning task
- show an awareness of the learning process
- assess their needs, evaluate progress and give direction to their learning by making use of metacognitive knowledge.

It is doubtful however whether these characteristics differentiate the successful and unsuccessful learners in our study. Kim, for example, was highly involved in the task of language learning, very articulate about her needs and highly self-directed in her approach to study. She nevertheless failed to show improvement after three months of laborious effort. We would suggest that the above characteristics have limited utility as predictors for a number of reasons. First, they do not take into account the educational level or motivation of the learners concerned (in this case we are dealing with a mature and highly educated group of learners with uniformly high levels of investment in the language learning enterprise). Second, they focus more on the formal and conscious aspects of learning in language classrooms than on the informal learning that takes place in input-rich environments such as the ones we are investigating where English is the official medium of communication. Third, and perhaps most important, they focus too heavily, on supposedly fixed individual attributes of the second language learner at the expense of social influences. Writers such as Lantolf (2000), Donato (1994), Norton (1995, 2000) have examined how the environments in which learners operate actually constrain or facilitate their language development as well as on how successful language learners are able to best exploit their learning and social environments. Recall that neither the most successful (Cassio) nor the least successful learner (Toshi) had much access to native speakers at home: Cassio lived with other Columbian students and Toshi lived alone. However, Cassio was much more successful in developing relationships with students from other nationalities both inside and outside the classroom. This appears to be both a function not only of his more outgoing personality but also his greater acceptability to his classmates (and possibly also to his teachers) in terms of age and cultural background. In particular, the degree to which each of these learners has been accepted by their peers appears to been extremely significant in determining their level of participation in social interaction in English inside the classroom and hence their attitudes to learning. While Toshi appears to be neglected by his peers, Cassio is highly popular and has ample opportunity to engage in classroom interaction.

The situation in class appears to mirror their environments outside: Toshi's initial enthusiasm to engage in social interaction in English appears to have been extinguished by the early

rejection he experienced from his neighbours in particular while Cassio's social life flourished as a result of his widespread popularity. As a result he had increasing opportunities to use English with both native and non-native speakers outside the centre.

These and other participants the study appear to have changed over the three month period, with many of them reformulating their goals for the future. Cassio, while adhering to his plan to embark on post graduate study has decided to study English for a further six months to consolidate his learning achievements. Toshi, contrary to his initial plans has given up on English and returned to Japan relieved to escape the stress of his overseas experience. Kim, is, for better or for worse, in the process of negotiating new identities for herself as wife, mother and student in a country where these roles are played out very differently than they were in her country of origin. Liao has, for reasons which are not altogether clear, changed his mind about returning to China or studying Business in Auckland and has decided to migrate to Canada.

7.0 Conclusion and implications

Attempts to predict success in language learning are inevitably thwarted by the fact that predictions are made a particular point in time and cannot accommodate the dynamic interactions between the individual and his/her social and learning environment, which are in a constant state of flux. It must be stressed that predictions are about groups rather than individuals. It is clear from our case studies, only a few of which are reported here, that individual attributes, such as personality, motivation and confidence are fluid and differentially responsive to the social conditions of language acquisition which are never entirely under the control of the learner. This is particularly true in input-rich environments, such as the ones we have investigated, where the classroom and the institution at which students are based are only one of many sites for language learning. This means that giving firm guarantees about individual progress may be both misleading and unethical. We should also remind the reader that a score gain in itself is not always evidence of a real gain in language proficiency. In fact it must always be borne in mind that with a SEM of .36 (for the overall IELTS band score) there is only a 68% probability that a candidate's score is accurate of within 0.36 of a band.

Nevertheless, there are number of practical implications which can be drawn from this study:

1. The clearest finding emerging from this research is that 10-12 weeks of intensive study in an English-medium environment does make a significant difference to performance, with students on average make moving up half a band during this period, and slightly more than half a band on the Listening component. This is probably the most reliable information that institutions can offer their clients, with the caveat that score gains are not guaranteed, that they may not always reflect real gains in proficiency and that they are less likely at the higher levels of proficiency. Clearly, this information should be made public and if information about individual students' proficiency is available on enrolment, they should be advised about their chances of improving by their institutions. Since many do not arrive with IELTS scores, it is important that the receiving institutions carry out diagnostic tests at entry which are both reliable and broad enough in scope to form the basis for advice about likely rates of progress, or more importantly, about what constitutes an <u>unrealistic</u> expectation. We have some doubts as to whether some of the placement tests currently used are adequate for this purpose, although the findings of this study suggest that, at the institutions we are dealing with here, the level in which students were placed was in the end, a reasonably good predictor of their subsequent progress.

2. Low proficiency students (below 5.5) will be encouraged to learn that an intensive 10-12 week course of instruction in an input-rich environment is likely to bring about measurable improvement in a their IELTS score, even if the degree of improvement will almost certainly fall short of what is required for university entry. The experience of taking IELTS may be valuable for such students, in that it offers them practice opportunities and may serve to encourage them about their progress. Whether these benefits are sufficient to warrant paying the high test-taker fees is however uncertain, given the availability of practice test materials and other sources of feedback about their progress.

3. Test-takers should also be told that their educational level (and to a lesser extent their age) may influence their rate of progress, and opportunities should be provided for those with lower levels of L1 education to familiarize themselves with the requirements and expectations of academia and to develop their world knowledge and cognitive skills to a level that will enable them to function in the higher education environment. Indeed, all test-takers should be advised that the performance on the test, important as it may be for their immediate personal advancement, should not be taken as the sole indicator of readiness for academia (Deakin 1997). A number of teachers and test-takers interviewed for this study expressed doubts about the validity of IELTS as a measure of proficiency, with some feeling they had made progress and could cope with academia in spite of what the test results told them, and others pointing out that high scores on the test were, on their own, an insufficient basis for future study.

4. The finding of higher rates of overall progress amongst those students who moved from one kind of housing is difficult to interpret and may be due to chance factors and/or to the overlap between *Accommodation* and some other variable. In any event it would seem wise for institutions accepting enrolments from overseas and immigrant students to pay attention not only to the provision of high quality teaching but also to aspects of pastoral care, ensuring where possible that students' living circumstances are satisfactory and that they provide rich opportunities for English interaction. Recent studies of students homestays (e.g. Rivers 1994, Welsh 2001) suggest that these are not necessarily as conducive to language learning as one would hope. Students should be encouraged to take active measures to improve their living circumstances if these prove to be unsatisfactory.

5. The fact that self-assessed improvement (or lack of same) seemed to be associated with listening and overall score gains suggests that candidates are more often than not aware of their progress. Self-assessments are both more accessible than objective measures and more indicative of learners' affective state, which may itself contribute to or inhibit progress in language learning. Self-assessments could perhaps be used on a regular basis in intensive English courses or in one-on-one counselling sessions as a means of encouraging learners to take stock of learning progress and to formulate suitable strategies to achieve their short and long term goals.

6. There are no clear indications from this study of what institutions can do to improve their students' chances of success (and indeed, we stated at the outset that it was not our aim to provide this kind of feedback). The profile of students at each institution was too diverse to allow firm conclusions to be drawn on this issue. The fact that highest mean score gains occurred at Institution 3, may be due to the fact that is a well established, has highly qualified teachers, that IELTS preparation is integral to its advanced program (although recall that the advanced students were less likely to improve), that it has a more heterogeneous learner population than the other institutions, that students were in the

optimum 20-25 age group and, above all, began with relatively low levels of proficiency. Few of these factors are easily manipulated, although highly qualified teachers would surely be an advantage and heterogeneous classes worth organizing where feasible in order to boost the chances of English interaction.

As already noted the findings reported in the study need to be treated with caution due to the small sample size. Further, the overlap between some of the variables investigated means that some of the trends reported here may be unstable and should be subjected to further investigation. Some avenues for further research are suggested below.

First, this study demonstrates the value of developing profiles of individual learners both to assist with interpretation of quantitative findings and also to illustrate their limitations. William's case is reassuring since it reminds us that learners do not always conform to the scripts we write for them and that there are many possible pathways to success. The profiling undertaken for this study was somewhat limited, in most cases based on only one face-to-face encounter. Further research involving more frequent conversations, supplemented with journals and retrospective written or oral accounts of language learning histories (preferably recorded in learners' L1) may shed further light on the different configurations of factors which make for successful language learning in an input rich environment.

It is however unclear whether the findings of studies conducted in English-medium environments can be generalized to other contexts. Further studies of IELTS score gains in non English-speaking countries would therefore be worth undertaking. We might expect that Listening skills will be less amenable to improvement in situations where there are fewer opportunities for English language exposure outside the classroom. Indeed, it may be worth conducting an experimental study which includes learners with instruction and no exposure, learners with instruction and exposure, and learners with exposure and no instruction, to determine the relative importance of language contact compared to the language input received in a formal learning environment.

As for the variables explored in this study, the influence of accommodation arrangements and other opportunities for interaction in the wider environment would seem to warrant further investigation. The apparent relationship between ethnicity/country of origin and score gains would also be worth exploring in more depth. In addition, a study comparing the kinds of interaction taking place in ethnically heterogeneous and homogeneous EAP classrooms might indicate whether, as some teachers in this study proposed, and as suggested by the relatively low mean gains for Listening at Institution 1, the composition of the class and the resultant classroom dynamics make a significant difference to opportunities for English practice and exposure.

Finally, any further attempt to investigate score gains using a pre- and post-test design should not overlook the possible effect of the test itself on learning outcomes. It was clear from our interviews that the participants in this study were very concerned about their performance on the IELTS test, even when they were only taking it for practice purposes. The band score that learners received both at the beginning and end of the study was clearly powerful in shaping their self image and predictions about their future progress and may well have influenced their approach and attitude to language learning. It would seem useful to follow up with willing students to determine how or indeed whether the score they received had an influence on their beliefs about language and language learning, the kinds of educational goals they set themselves and their life plans more generally. This is an important aspect of test impact which has thus far been under-researched.

Ackowledgements

Bibliography

Alderson J. C. and Banerjee J (2001) Impact and washback research in language testing. In Elder C., Brown A., Grove E., Hill K., Iwashita N., Lumley T., McNamara T. and K. O'Loughlin (eds) *Experimenting with uncertainty: essays in honour of Alan Davies.* Studies in Language Testing 11. Cambridge: UCLES/CUP.

Beebe, J.D. (1998) Individual differences in context: Japanese secondary students learning to speak English. U.M.I. Dissertation information Service

Bellingham, L. (1993) The relationship of language proficiency to academic success for international students. *New Zealand Journal of Educational Studies* 30, 2: 229-232.

Bond, N. (2001) The relationship between learner anxiety and IELTS test performance. Unpublished MA thesis. Department of Applied Language Studies and Linguistics, University of Auckland.

Collier, V. (1989) How long? A synthesis of research on academic achievement in a second language. *TESOL Quarterly* 23,3: 509-528

Deakin, G. (1997) IELTS in context: issues in EAP for overseas students. *EA Journal* 15, 2: 7-28

Dewaele J.M. and Furham, A. (2000) Personality and speech production: a pilot study of second language learners. *Personality and Individual Differences,* 28: 355-365.

Donato, R. (1994) Collective scaffolding in second language learning in J.P. Lantolf and G. Appel (eds). *Vygotskian approaches to second language research.* Norwood, N.J.: Ablex Publishing Corporation.

Dooey, P. & R. Oliver, (2002) An investigation into the predictive validity of the IELTS test as an indicator of future academic success. *Prospect* 17,1: 36-54.

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

Ellis, R. (1994) *The study of second language acquisition.* Oxford: Oxford University Press

Hamp-Lyons, L. (1997). Washback impact and validity: ethical concerns. *Language Testing* 14, 3: 295-303.

Hawthorne, L. (1997) The politicisation of English: The case of the *Step* test and the Chinese students. In G. Wigglesworth and C. Elder (eds) *The language testing cycle: from inception to washback.* ARAL Series S, 13, ANU: Canberra.

Hill, K., N. Storch and B. Lynch (1999) A comparison of IELTS and TOEFL as predictors of academic success. *IELTS Research Reports* Volume 2: 52-63.

Horwitz, E.K., Horwitz, M.B. & J. Cope (1986) Foreign language classroom anxiety. *The Modern Language Journal* 70:125-132.

Kerstjens, M. and C. Nery (2000) Predictive validity in the IELTS test: a study of the relationship between IELTS scores and students' subsequent academic performance. *IELTS Research Reports* Volume 3: 85-108.

Lantolf, J.P. (2000). Second language learning as a mediated process. *Language Teaching* 33,2, 79-96.

Messick, S. (1981). Evidence and ethics in the evaluation of tests *Educational Reseracher* 10: 8-20

Messick, S. (1989). Validity. In R. L. Linn (ed.) *Educational measurement* (3rd edn) Washington DC: The American Council on Education and the National Council on Measurement in Education

Naiman, N. Frohlich, M. Stern, H.H., and A. Todesco (1978). *The good language learner.* Research in Education Series, 7, Ontario Institute for Studies in Education.

Norton-Peirce,, B. (1995). *Social identity, investment and language learning.* TESOL Quarterly *29, 9-31.*

Norton-Peirce, B. (2000). Identity and language learning: Gender, ethnicity and educational change. *Harlow Longman.*

Oya, T. (2002). *Factors in oral performance of Japanese learners of English in New Zealand.* Unpublished Masters thesis. Department of Applied Language Studies and Linguistics,University of Auckland.

Parry, T.S. and C.W. Stansfield (1990). *Language aptitude revisited.* Englewood Cliffs, NJ: Prentice Hall Regents.

Read, J. & B. Hayes (2001) The impact of IELTS on preparation for academic study in New Zealand. Unpublished report.

Rivers, W. (1998). Is being there enough? The effects of homestay placements on language gain during study abroad. *Foreign Language Annals* 31, 4, 492-500.

Shohamy, E. (2001). *The power of tests: a critical perspective on the uses of language tests.* Longman/Pearson.

Skehan, P. (1989). Individual differences in second language learning. *London: Arnold.*

Spolsky, B. (2001) Cheating language tests can be dangerous. In Elder C., Brown A., Grove E., Hill K., Iwashita N., Lumley T., McNamara T. and K. O'Loughlin (eds) *Experimenting with uncertainty: essays in honour of Alan Davies.* Studies in Language Testing 11. Cambridge: UCLES/CUP.

Welsh, A. (2001). Homestay: the perceptions of international students at a tertiary institution in New Zealand. Unpublished MA thesis. Department of Applied Language Studies and Linguistics, University of Auckland.

Wall, D. (1997). Impact and washback in language testing, In C. Clapham (ed) *The Kluwer Encyclopedia of Language in Education, Vol 7: Testing and Assessment* Dordrecht: Kluwer.

Zumbo, B. (1999). The simple difference score as an inherently poor measure of change: Some reality, some methodology. In B. Thompson (ed.) *Advances in social science methodology, Vol 5.* Greenwich, CT: JAI Press.

Appendix 6.1:

Student questionnaires

Student questionnaire 1

```
┌─────────────────────────────────────────────────┐
│ Office use                                       │
│ Institution:          _____     │
│ Student number:       _____     │
│ Date administered:    _____     │
│ Student's course:     _____     │
│ Level of student's course: _____        │
│ Teacher(s): _____         │
└─────────────────────────────────────────────────┘
```

Please answer the following questions as carefully as possible.

1. Family name _____

2. Given name   _____

3. Sex   _____

4. Date of birth   _____ (day) _____ (month) 19___ (year)

5. Country of birth   _____

6. Nationality   _____

7. First language   _____

8. What other languages do you speak? _____

9. What is your highest level of education? (*Tick one of the following*)

PhD degree	❏
Masters degree	❏
Postgraduate Certificate/Diploma	❏
Bachelor degree	❏
Certificate/ Diploma	❏
Final year of secondary school	❏

10. What English courses did you do *before* coming to Australia (NZ)?

I studied English at secondary/high school in *(name country)* _____ for ___ years ____ months.

I studied English at university in *(name country)* _____ for ___ years ____ months.

I studied English at a private language school in *(name country)* _____ for ___ years ____ months

I had a private tutor in *(name country)* _____ for ___ years ____ months.

11. Why did you learn English *before* you came to Australia (NZ)?
(circle ' yes' or ' no' for each of the following statements).

English was compulsory at school	YES	NO
English was compulsory at university	YES	NO
I needed to know English to travel abroad	YES	NO
I needed to know English to study abroad	YES	NO
I needed to know English for my job	YES	NO

Other reason *(please explain)*

_____

_____

_____

_____

12. How long have you been in Australia (NZ)?
I have been in Australia for _____ years ___ months.

13. Where are you living at the moment? *(Tick one of the following)*

In a flat with other students	❏
In a homestay	❏
With my own family	❏

Other *(please explain)*

_____

_____

_____

7. Have you lived in any other *English speaking* countries?
(circle 'yes' or 'no')

YES   NO

If yes, how long did you live in those countries?

Country 1   ____ years ___ months.
Country 2   ____ years ___ months.
Country 3   ____ years ___ months.

15. What is your current visa status in Australia (NZ) (*tick the appropriate box*)?

Permanent Resident (PR)   ❏
Student   ❏
Tourist/visitor   ❏
Working holiday   ❏
Temporary Resident   ❏
Other (please explain)

_____

_____

_____

16. Have you studied English at other centres in Australia (NZ) before this one? (*please circle*)

YES   NO

If you answered 'yes' to this question, please complete the following:

I studied English for _____ months at (*name centre 1*) _____

Was it (*please tick*)

                    a general course   ❏
OR   an academic course   ❏
OR   an IELTS training course?   ❏

I studied English for _____ months at (*name centre 2*) _____

Was it (*please tick*)

                    a general course   ❏
OR   an academic course   ❏
OR   an IELTS training course?   ❏

17. Why are you learning English now in Australia (NZ)? (*Tick one or more of the following*)

My parents want me to   ❏
For my future job   ❏
To prepare for further studies   ❏
Other reason (*please explain*)

_____

_____

18. If you are planning to do further studies in Australia (NZ)

what course do you want to do?

_____

when do you plan to start? *(give date and year)*

_____

what IELTS score do you need to do this course? *(please circle)*

7   6.5   6   5.5   5   Don't know

do you think you can obtain this score in three months time?
(please circle)

YES   NO

19. Have you taken the IELTS or TOEFL *before* you began your current course?
If so, please list below the dates, which test you took, the places and your overall results:

	Date	Name of test	Place	Result
a				
b				
c				
d				

20. Rate the different language skills below, by circling one of the numbers ranging from '**1. Not important at all**' to '**4, Very important**'
(A skill is important if you often need it for your studies or in your personal life)

Listening	1	2	3	4
Reading	1	2	3	4
Speaking	1	2	3	4
Writing	1	2	3	4
IELTS test practice	1	2	3	4

Student questionnaire 2

Office use
Office use
Date administered:
Centre:
Student's number:
Student's course:
Level of student's course:
Teacher(s):

Please answer the following questions as carefully as possible.

1. Family name _____

2. Given name _____

3. Sex _____

4. Why are you learning English now in Australia (NZ)? *(Tick one or more of the following)*

        My parents want me to   ❑
        For my future job   ❑
        To study at university   ❑
        Other reason *(please explain)*   ❑

        _____

        _____

5. If you are planning to study at a university in Australia

    a)   *what course do you want to do?*

    _____

    b)   when do you plan to start? *(give month and year)*

    _____

    c) what IELTS score do you need to do this course? *(please circle)*

    7   6.5   6   5.5   5   Don't know

    d) do you think you can obtain this score the next time you do the IELTS? *(please circle)*

    YES   NO

6. Where are you living at the moment? *(Tick one of the following)*

In a flat with other students	❏
In a homestay	❏
With my own family	❏
Other *(please explain)*	❏

_____

_____

_____

7. How often did you do each of the following activities in the last three months?   *(Circle your response)*

(a) I attended my English class. Often   Sometimes   Not often   Never

(b) I worked hard during
my English class.   Often   Sometimes   Not often   Never

(c) I did homework.   Often   Sometimes   Not often   Never

(d) I studied alone
outside class.   Often   Sometimes   Not often   Never

(e) I studied in a self access
centre.   Often   Sometimes   Not often   Never

(f) I studied in a library   Often   Sometimes   Not often   Never

(g) I practised English with
my friends inside class.   Often Sometimes   Not often   Never

(h) I practised English with my
friends outside class.   Often   Sometimes   Not often   Never

(i) I practised English with my
homestay family.   Often   Sometimes   Not often   Never

(j) I watched English
television/movies.   Often   Sometimes   Not often   Never

(k) I listened to English
radio programs.   Often Sometimes   Not often   Never

(l) I read English
newspapers/magazines.   Often   Sometimes   Not often   Never

(m) I used the internet   Often   Sometimes   Not often   Never

(n) Other *(please explain)*
_____   Often   Sometimes   Not often   Never.

8. How much do you think your English has improved over the last three months?
(*please circle your response*)

Listening	A lot	Quite a lot	A bit	Not at all
Speaking	A lot	Quite a lot	A bit	Not at all
Reading	A lot	Quite a lot	A bit	Not at all
Writing	A lot	Quite a lot	A bit	Not at all

The next few questions are about your English classes during the last three months.

9. Did you change classes during the last three months? (*please circle*)

YES   NO

If yes, please explain what happened:

_____

_____

10. How many different teachers did you have over the last three months? (*please circle your answer*).

1   2   3   4   more than 4

11. Do you like changing teachers? (*please circle*)

YES   NO

Why?

_____

_____

_____

12. What do you like *most* about your current English course?

_____

_____

_____

13. What do you like *least* about your current English course?

_____

_____

_____

14. How much do you agree or disagree with the following statements about your current English course?

(Note: SA = Strongly agree, A = Agree, D = Disagree, SD = Strongly disagree)

This course has helped me improve my reading.   SA   A   D   SD

This course has helped me improve my writing.   SA   A   D   SD

This course has helped me improve my listening.   SA   A   D   SD

This course has helped me improve my speaking.   SA   A   D   SD

This course has given me confidence in using English outside class.   SA   A   D   SD

I was happy with the teaching on this course.   SA   A   D   SD

Overall, I am very satisfied with my current English course.   SA   A   D   SD

15. Has your current English course included practice on the IELTS? (*please circle*)

    YES   NO

16. Have you completed a special IELTS training course in the last three months? (*please circle*)

    YES   NO

17. How much do you think your next IELTS results will improve since you last did the test? (*Circle your response for each part of the test*)

Listening   A lot   Quite a lot   A bit   Not at all

Speaking   A lot   Quite a lot   A bit   Not at all

Reading   A lot   Quite a lot   A bit   Not at all

Writing   A lot   Quite a lot   A bit   Not at all

18. Do you think the IELTS is a good test of your English language ability? (*please circle*)

    YES   NO
Why?

_____

_____

Appendix 2:

Interview schedules for centre directors and teachers

Interview schedule for directors

1. How many students are currently enrolled in your English program?

2. Is this typical of your usual pattern of enrolment?

3. What categories of students do you cater for?
PR immigrants? international "visa" students? Other?

4. From what countries to these students typically come?

5. For what reasons do students typically enrol in your courses?
Preparation for IELTS test? Preparation for university study?, general English? Preparation for entering NZ workforce?

6. Do you have any formal mechanism for assessing students' needs?

7. What level of student do you cater for?
Give approximate range of abilities in terms of IELTS eg 5- 6.5

8. Do you stream students? On what basis?

9. How long do your courses run*?*
Three months, six months, rolling intake?

10. What is the average period of enrolment for your students?

11. What is the average class size?

12. How many hours per week of instruction are offered?

13. How would you characterize your institution's approach to English language instruction?
(egTheme/content-based? skills based?)

14. Are some skills/areas/topics given greater emphasis than others? And if so why?

15. Is there a central curriculum? If so can we view a copy of this to get a sense of what is covered?

16. Are set texts used? Which ones?

17. What other resources are available?
self access centre? one-to-one tutoring?

18. Are any students concurrently enrolled in other courses?
eg IELTS preparation? School or university subjects?

19. Does your institution organize accommodation for enrolled students?

Eg homestay, apartments

20. How many staff do you currently employ on your EAP program?

21. What qualifications./experience do you require of your teachers?
(native speakers of English? Diploma of Education, RSA Certificate, prior ESL/EFL teaching experience? Other?)

22. What mechanisms do you use to evaluate student progress? What level of progress do you expect after 3 months of instruction?

23. In your view what factors, institutional, social or individual, have the greatest bearing on whether students make progress in their English language learning?

24. What, in your view, are the chief obstacles to English language progress among students enrolled in your course?

Interview schedule for teachers

1. How many students are currently enrolled in your EAP class/es?

2. Can you give us a breakdown of these students according to a) category *(immigrant, international, other)* and b) country or origin?

3. For what reasons are these students studying English?
Preparation for IELTS test? Preparation for university study?, general English? Preparation for entering NZ workforce?

4. What range of abilities are you catering for in your class/es?
Give approximate range of abilities in terms of IELTS eg 5- 6.5

5. How would you characterize your approach to English language instruction?
Eg Do you follow a particular method? Do you emphasize some skills/areas/topics more than others? And if so why?

6. How many of your current students avail themselves of additional resources provided by your institution? *(eg self-access, library)*

7. What other opportunities for English language practice are available to them?

8. How do you evaluate students' progress on this course?

9. In your view what factors (institutional, social or individual) have the greatest bearing on whether students make progress in their English language learning? *(eg attitudes/ motivation? Age? Personality? Personal circumstances? Aptitude? Exposure to English outside the classroom?)*

10. What, in your view, are the chief obstacles to English language progress among your students?

Then for any of the case study students:

1. I'd like to ask you about X and your impressions of his/her approach to English language learning.

2. Was X a regular attender? If not why not?

3. Did X participate actively in class sessions? If not why not?

4. Did X complete required work promptly? Successfully? Please give details.

5. Did you notice anything about his/ her learning style or the strategies adopted to perform various classroom tasks?

6. Did this student avail him/herself of additional opportunities/resources available for learning English?

7. In your view how much progress has X made in terms of Reading, Writing Listening and Speaking?

8. What factors have contributed most powerfully to this progress/lack of progress in your view?

9. Are you aware of any personal circumstances which may have had a bearing on his/her English acquisition? *(eg family troubles, experience of homestay)*

10. Can you offer any other comments about this student that you think are relevant to our understanding of the quality of his/her English language learning experience over the last three months?