2. Does IELTS Preparation Work?
An Application of the Context-Adaptive Model of Language Program Evaluation

James D. H. Brown
Hawthorn Campus, The University of Melbourne

2   Does IELTS Preparation Work? An Application of the Context-Adaptive Model of Language Program Evaluation

James D. H. Brown
Hawthorn Campus
The University of Melbourne

Abstract

This paper presents an evaluation of an English Language Intensive Course for Overseas Students (ELICOS) language program, conducted with application of a context-adaptive model of language program evaluation.

It reports on a causal-comparative examination of both an English for Academic Purposes (EAP) course, and an International English Language Testing System (IELTS) course, which sought to examine which of the two programs of study was the most effective in the preparation of ELICOS students for the International English Language Testing System, with particular reference to the academic writing component of the test.

This emphasis was selected due to the significance of writing skills in the academic contexts to which both IELTS and EAP students seek access. The context of the study allows for an examination of the effectiveness of the context-adaptive model of language program evaluation.

An analysis of both the quantitative and qualitative data gathered within the research program is presented. The quantitative data analysis suggests that IELTS preparation programs are more successful in preparing students for the IELTS test than EAP courses. The qualitative data informs an analysis of the causes of this.

The maintenance of focused instructional emphasis on International English Language Testing System writing task completion, together with the development of skills in writing, planning, and systematic tuition in strategies for writing under examination conditions, seem to assist students of IELTS preparation programs to achieve an average gain of one IELTS bandscore over a ten-week course of instruction.

1.0   Introduction

The English language preparation of overseas students seeking entry to the Australian tertiary education system is focused on achieving satisfactory performance in either the Test of English as a Foreign Language (TOEFL) (Educational Testing Service, no date), or the International English Language Testing System (IELTS) tests (University of Cambridge Local Examinations Syndicate et al, no date). English for Academic Purpose courses designed to prepare students for the tests are available in most states of Australia, but the effectiveness of such programs is uncertain.

As Weir and Roberts (1994: 87-88) argue:

> "There is a very important place for British Council/UCLES IELTS results which are used to determine the EAP proficiency of overseas students. This is an area in which there is still an urgent need for research."

The subject of this evaluation design is to examine the hypothesis of no difference that a general Academic English course is as effective as a specific IELTS Preparation course in the training of ELICOS students for the IELTS test. The study was conducted at the Hawthorn English Language Centre in Melbourne, a suburban campus of the University of Melbourne.

1.1   Evaluation Context

The Hawthorn English Language Centre was established in 1981 to provide English for Special Purpose programs for students from South-East Asia sponsored by the Australian International Development Assistance Bureau. ELICOS programs have been offered there since 1984, and IELTS preparation courses were developed following the introduction of IELTS to Australia in November 1989.

The growth of the ELICOS sector in Australia occurred simultaneously with the growth in wider knowledge of and respect for the IELTS test, such that twenty per cent of the ELICOS students in the Centre now follow a specific IELTS preparation course. Hawthorn English Language Centre is a designated IELTS test administration centre for Victoria. Whilst the test is administered separately from the workings of the language centre, some ESL teaching staff also work occasionally as IELTS assessors as is the case in most, if not all, Australian ELICOS centres. However, there has not been any previous formal evaluation conducted of the Hawthorn English Language Centre or its programs.

1.2   Audience and Goals

The consideration of audiences and goals for the evaluation is the initial step in the application of the context-adaptive model. This requires the determination of groups or categories of people who are likely to have an interest in the outcomes of an evaluation of the centre. Multiple audiences can be determined in this evaluation context, but in terms of their perceived expectations of the evaluation outcomes they may be categorised as belonging to one of four groups of stakeholders: *academics, managers, staff* and *students*. The groupings, likely membership thereof in the context of this evaluation, and some of their perspectives are represented in the figure following as Table 1. The identification of audiences and their perspectives in this way allows the easier determination of their rankings as primary or secondary stakeholders. Ranking of audiences assists the consideration of principal stakeholder expectations, particularly in contexts with multiple audiences. The process of

ranking is also of assistance to the evaluator in generating a reminder of the origins of the goals for the evaluation project thereby further focusing the enquiry.

Grouping	Members	Perspectives and Expectations	Ranking
Academics	• Examiners • Researchers • Peers	• Sound & valid research • Academic rigour • Academic conventions	Secondary
Managers	• Institute directors • Marketing manager • Head of Centre • Senior coordinators	• Responsibility for reputation • Commercial concerns • Seeking positive outcomes, & 'good press'	Secondary
Staff	• ESL teachers • IELTS assessors • Curriculum writers	• Maintenance of professional pride • Seeking indicators for program development & improvement	Primary
ELICOS Students	• EAP students • IELTS students • General program students • Prospective students	• Sound tertiary preparation programs • IELTS success • Best program of ESL • Effective language learning	Primary

Table 1   HELC Evaluation : audience profile

Whilst the concerns of both primary and secondary ranks must be considered, the evaluation design was principally responsive to the staff and student audiences in this case, because of their participation in the program as primary stakeholders. The primary audiences are differentiated as they are in this context because there was no external request for the conduct of the evaluation. Whilst the students did not request the evaluation, their active participation was sought and gained through their signing of letters of informed consent to participate in the study.

It may be seen in Table 1 above that the expectations of the secondary audience groups of managers and academics were very different from those of the primary stakeholders. It was recognised that their presence was likely to influence the evaluation, both in ensuring that acceptable academic standards were maintained, and through cognisance of commercial concern.

The importance of identifying the differences in expectations and perspectives of the various audiences to the evaluation lies in the realisation that the evaluator is unlikely to succeed in satisfying the expectations of all audience groups. The evaluator must therefore focus on the expectations of the primary audiences in the preparation of the principal evaluation report.

The goal of this evaluation, that is, to examine which program was the more effective test preparation, arose from a participant hypothesis within the primary audience, thereby enhancing the validity of the project through response to genuine need. However, due to the practical constraints of time for the evaluator to commit to the project it was decided to limit the evaluation to one skill only, writing.

1.3 Context Inventory

Determination of the goal for evaluation allows the evaluator to consider the features characterising the program and to prepare a context inventory. Lynch (1990) proposes nine dimensions of an inventory applicable to this ELICOS setting:

1. Comparison group
2. Reliable and valid measures
3. Timing
4. Selection process
5. Program students
6. Staff
7. Size and intensity
8. Instructional materials and resources
9. Purpose of program

Each of these inventory dimensions is discussed below with regard to their application in the evaluative context of this study. Their value lies in requiring the evaluator to consider thoroughly each of the characteristics of an evaluation context, and thereby to determine their role in the evaluation process.

1.3.1 Comparison Group

The Hawthorn English Language Centre is a part of Hawthorn International Education Ltd, a commercial department of the University of Melbourne. The Centre offers specialised ELICOS programs in both IELTS test preparation, as well as EAP courses of study skills for students pursuing tertiary entrance in Australia. Availability of an appropriate comparison group was thus assured through the provision of courses in EAP as well as IELTS Preparation.

Adoption of a General English course as a comparison group to the IELTS Preparation course, as originally proposed by the teaching staff, would not have been acceptable, due to differences in language proficiency levels and sources of student motivation from the IELTS preparation course. Similarly, use of another IELTS preparation course, but without focused instruction in IELTS preparation, as a comparison group was also unacceptable in the ELICOS context, as variation in the program offered to the comparison group would have been an unethical practice upon students contracting for a full course of IELTS preparation.
The most appropriate comparison group in this context was an EAP course, whose students have similar motivation and language proficiency levels to the IELTS preparation program group, as both groups of students maintain intentions to continue their studies in Australia at tertiary level.

Both IELTS and EAP students were aware of the need for developed writing skills, and could be expected to participate in the research program due to expectations of gain from the instructional program.

The least proficient of the three IELTS classes and the two EAP classes, being IELTS class 3 and EAP class 2, were identified by the program staff as comparable in terms of English language proficiency. This followed pre-course placement testing conducted on arrival at the Centre.

As the least proficient groups of both the IELTS and EAP course streams, it was felt that the research results would be least affected by any prior knowledge or experience of the IELTS test. Later student interviews confirmed that none of the students of either course had any previous IELTS experience.

1.3.2 Reliable and Valid Measures

The goal of this evaluation dictated the use of IELTS test material to determine student language proficiency levels at both commencement and completion of their courses of study. Given the large-scale international administration of IELTS, it is expected that the test has undergone and continues to undergo rigorous validation and reliability studies. One study has already appeared in the University of Cambridge Studies in Language Testing series (Clapham, 1996).

The evaluator was aware of the possibility that use of IELTS might favour the IELTS program group over the EAP group, due to greater familiarity with the testing system despite the lack of familiarity with the test amongst both groups at the outset of the study. However, use of IELTS as a reliable and valid measure of student performance was appropriate in the circumstances, due to the common student need to achieve a satisfactory IELTS result for Australian tertiary entrance.

Use of IELTS exemplar material for use in pre- and post-testing of the two student groups was ruled out, as it could not be ensured that either of the groups would not have access to, or experience of the exemplar materials during self-study periods within the program. Similarly, use of live test materials was excluded for reasons of test security and cost.

The availability of recently redundant live test material in the form of phased out versions of the IELTS test was a significant feature of this evaluation design, allowing access to properly validated real test items to be administered under normal conditions of test use.

Permission to use these phased-out versions of IELTS was gained from IELTS Australia for this purpose. Different forms of the redundant version of the test were used in pre- and post-testing procedures to negate the possibility of invalidating results through practice effect.

1.3.3 Timing

Conduct of the evaluation would ideally be carefully timed to match the study cycle of ELICOS students preparing for tertiary entrance. Such students usually elect to commence their programs of IELTS preparation during the second semester of the Australian academic year, in order to achieve their entrance to tertiary studies at the beginning of the next academic year.

Therefore, the evaluation should ideally be undertaken during the second semester of an academic year, and be maintained for the duration of a complete IELTS preparation program over a ten-week period. Both IELTS and EAP courses are offered during the same student intake periods, and second semester courses are characterised by greater student numbers and more buoyant student motivation levels than is found during semester one programs. Unfortunately, the ideal timing described above was not possible for this study due to project

timelines. The research was carried out during the first full ten-week program of the academic year from 19th February to 26th April 1996.

Accordingly, the courses under study were characterised by slow student arrival to course intake, which delayed pre-testing until the second course week . Furthermore, the program students were typically planning for tertiary entrance in 1997, and were therefore characterised by lower motivation levels than students seeking immediate tertiary entrance. This situation subsequently led to the high levels of student withdrawal from the research program described later in this paper.

1.3.4   Selection Process

Randomisation in the process of selecting students for the evaluation program was not possible, due to the Language Centre's practice of inviting students to select their chosen course of study with advice from course teachers. The evaluation program students thus represented intact predetermined groups formed through expression of their wishes for study focus, and in satisfaction of Centre proficiency level requirements.

Thus, both groups in the evaluation program were characterised by similar language proficiency levels at course commencement, and they shared a common motivation towards tertiary entrance preparation. However, it can be assumed that the students in the IELTS preparation group had a higher level of interest in IELTS specifically rather than academic English in general.

The lack of randomisation in the selection of research participants was a constraint to the statistical analysis of quantitative data. This was an inevitable outcome of both Language Centre policy, and ethical considerations of course-based research that denied the possibility of random student selection in this evaluation context.

1.3.5   Program Students

The majority student group from which the program students were drawn were non-English speaking background ELICOS students from the South-East Asian region, aged between 16 and 25 years, who had completed at least four years of secondary education in their country of origin. Motivated by a desire to undertake tertiary or vocational study in Australian institutions following English language training, they are therefore prepared to be local residents for four to six years, and are often from prosperous socio-economic backgrounds capable of supporting such extended overseas study programs.

The student groups were balanced in gender distribution and were typically enthusiastic, but characterised by diverse mother tongue and cultural backgrounds as well as previous educational experience. A careful record of student attendance within the program groups was maintained to ensure that absenteeism did not influence the program results by falling below the required course attendance level of 90% for ELICOS students.

1.3.6   Staff

The primary goal for this evaluation arose from the comments of program teaching staff, and thus their commitment to the evaluation project was assured.
The two programs within the research study were both staffed with three teachers per program week. Each teaching team comprised one male and two female teachers of similar ages, with matching training levels. All program teachers had post-graduate TESOL qualifications at the graduate certificate level or equivalent. They all had similar lengths of

teaching experience, and were long-standing members of the teaching staff familiar with the evaluation environment.

Both teams of teachers had taught both the IELTS preparation and the EAP courses previously. Coincidentally, none of the research program teachers were either qualified or experienced as IELTS assessors. Their knowledge of the IELTS was thus based solely on in-service staff seminars on the subject, and on experience gained from teaching from commercial IELTS preparation materials.

Beyond this similarity amongst the program teaching staff, it was notable that each team was convinced that their program was the most effective preparation for IELTS specifically and tertiary study generally.

1.3.7 Size and Intensity

All English Language Intensive Courses for Overseas Students (ELICOS) in Australia are full-time programs of study requiring student attendance over twenty- five hours per course week. Instruction is offered through twenty hours of classwork per week, and five further hours of independent study within a self access centre.

The courses are offered in ten-week blocks, and students may elect to continue their studies as performance allows across a range of general and specific purpose courses. The program and comparison groups within this evaluation both utilise an integrated skills approach, and aim to achieve homogeneous proficiency levels within groups through preliminary entry level testing.

Initially both groups were comprised of fifteen students and had similar multi-cultural and lingual backgrounds, as well as a balance of gender. The students were characterised by similar durations of Australian residence, and little previous English language testing experience.

Both groups were monitored through interviews with program staff to establish degrees of motivational difference existing between group participants, in order to determine the extent to which motivational differences might affect the validity of the evaluation. The factor of motivation is discussed further in the analysis of qualitative data.

1.3.8 Instructional Materials and Resources

The programs were taught within the facilities of the Hawthorn Institute of Education. Air conditioned classrooms with computer networks, integrated language laboratories, and extensive educational library resources allow Language Centre teachers to offer courses of diverse construction with methodological variation as appropriate to different client groups.

The Language Centre maintains a comprehensive and specialised library of ESL reference and text books as well as audiovisual resources. Course books are made available to all students upon staff selection, and the Centre is administered and supported by a secretariat.

Discussion of the materials used for the teaching of writing within the two courses of interest (IELTS and EAP) will be addressed in the analysis of qualitative data.

1.3.9 Purpose of Program

The IELTS Preparation Course is an intensive ten-week course focusing on the particular language and study skills demanded by the IELTS examination specifically, and by academic settings in general. Students receive tuition in the format of the IELTS test, and frequent practice of the relevant skills to be tested. The course is open to students who have attained an English language competency level assessed by the Language Centre placement test as intermediate to advanced. The course content appears to balance skill development equally across the macro skills.

The English for Academic Purposes Course is also designed for students of intermediate to advanced levels who are intending to continue their studies in Australia at university or college level. Students are trained in the study skills necessary for effective participation in academic programs. The course content emphasises reading and writing skill development over listening and speaking, but includes training in study techniques, note-taking and computing.

2.0 Research Methodology

2.1 Introduction

This section presents information relating to the methods of data collection.
It describes the thematic framework of the evaluation, the data collection design system and the collection of both quantitative and qualitative data.

Determination of the audience and goals for evaluation, and preparation of the context inventory, allows the evaluator to consider the dominant themes of the evaluation for the third stage of the context-adaptive model: The Thematic Framework. This requires the identification of themes and issues affecting the evaluation.

2.2 The Thematic Framework

The thematic framework was developed through a process of negotiation with the evaluation program staff to determine themes for evaluation which were both desirable in terms of the evaluation goals, and which were feasible within the constraints of this evaluation setting. The thematic framework was as follows:

1. Student levels of English proficiency upon entering the program. (Pre-test)

2. Student levels of English proficiency upon departing the program. (Post-test)

3. Sources of student motivation upon entering the program.

4. Concern that a normal program be taught to both the IELTS and EAP groups.

5. Assessment of gain focused on the 'academic' skill of writing.

6. Use of instructional materials by program staff.

7. Feasibility of sustaining classroom-based qualitative research.

8. Instructional focus on types of writing. In this case, exposition, argument and description.

9. Student attendance levels on the course.

2.3   Data Collection Design System

This evaluation project was conducted with a mixed research design, requiring the collection and analysis of both quantitative and qualitative data (Lynch,1993).   A diagrammatic representation of the research design follows as Table 2.

	Pre test	Qualitative	Post test
IELTS Group	Test A	Program	Test B
EAP Group	Test A	Comparison	Test B

Table 2   Data collection design

Quantitative IELTS test score data generated by the writing sub-tests was to be collected through pre- and post-course testing of both the IELTS and EAP program groups. This data was to be prepared for subsequent statistical analysis to determine the existence and possible significance of student performance gain.

Qualitative data was to be gathered through extensive classroom observation procedures in order to identify the program elements attributable to gain in student writing performance. This included document analysis of staff teaching records, and informal interviews with both staff and students to inform the evaluator's understanding of program element effect.

Analysis of qualitative data required the categorisation and reduction of data through the application of display matrices (Lynch,1990) to ease the evaluator's task of searching for patterns and themes in the data.

2.4   Data Collection

Data collection was undertaken during the first semester of the Language Centre's courses in 1996 as previously described in the context inventory.

As stated above IELTS class 3 and EAP class 2 were selected as subjects of the research due to their apparent similarity in language proficiency. Furthermore, as the least able of the students in ESP streams they were considered to have the least experience of specialised instruction and assessment.

At the outset of the course these two classes had more students enrolled than the more advanced programs, which was an important consideration regarding the validity of the planned statistical analyses. Both the IELTS and the EAP classes had thirteen students each.

Prior to the commencement of the research project, all students in the IELTS and EAP courses were given a letter of informed consent, which detailed the expectations and requirements of students participating in the project. Following explanation of the content of the letter, the students were invited to indicate their preparedness to participate in the project by signing their names to copies of the letter of consent.

It was at this point that the most important variation to the proposed project design arose. Whilst all the IELTS program students were pleased to participate in the research study and signed their consent, only eight of the EAP students wished to take part. The researcher later established through student interviews that those EAP students who declined their participation determined that the research program was "irrelevant" to their study needs.

Conversely, the enthusiastic interest of the IELTS group was born in the anticipation that they might derive benefit from the research focus.

This reaction suggested difference in motivation between the groups which may be understood given the research project focus on IELTS, whilst the EAP students maintained a longer term plan for tertiary entrance, and did not expect to undertake IELTS preparation or testing until later in the year.

The fall in research subject numbers was further exacerbated during research week five, when the Language Centre made the decision to introduce a mid-course intake into the classes to lift the low student enrolment early in the year.

At this point several current IELTS and EAP course students left the ELICOS program as they had gained entry to other courses at other institutions, thereby further reducing the numbers of research subjects to only nine IELTS students, and just five EAP students who continued their courses through to post-testing and course completion.

Thus, circumstances beyond the control and anticipation of the researcher required reconsideration of the statistical tests for quantitative data analysis, and threatened the successful continuation of the research project. However, no further losses of students from the project occurred during the second half of the course, and the research project was continued as planned, albeit with fewer subjects than anticipated in the research plan.

The pre-test was conducted at the beginning of course week two, and the test used was IELTS Version 5, Form 1 module C. The post-test was conducted during the final course week ten, and the test used was IELTS Version 6, Form 1 module C. As had been agreed with IELTS Australia prior to the commencement of the research project, the IELTS test papers were destroyed following the conduct of the post-test to ensure maintenance of test security.

Different test forms were used in the pre- and post-testing procedures to eliminate the possibility of test practice effect. Both pre- and post-testing was conducted under normal examination conditions, fully invigilated and with punctual control of test timing to simulate the IELTS testing experience.

Each student test paper was assessed separately by three independent IELTS assessors, who were selected for the task on the basis of their extensive IELTS assessment experience. The three overall scores for each student were then averaged to obtain the most robust overall result for each research subject at both the pre- and post-tests.

Classroom observations were conducted by the researcher throughout the ten-week research period. All classes designated for writing instruction were observed for both the program (IELTS) and control (EAP) groups. Records of the focus of instruction and the materials and methods used in writing classes were kept. These data records required categorisation and reduction through the use of display matrices. (Lynch, 1990) This process of qualitative data reduction was necessary to aid the researcher's task of searching for patterns and themes in the data which might help to determine the significance of performance gain attributable to the methods and techniques of instruction in academic writing.

3.0 Data Analysis

3.1 Introduction

The data analysis is presented in two parts. The descriptive statistics are first analysed for the presence of statistical significance. Then the qualitative data is both reduced and analysed towards the identification of program elements which may have had an effect on the student performance gain revealed in the statistical analysis.

3.2 Descriptive Statistics

Table 3 displays the raw score data for each subject in the study: the IELTS group, and the EAP group.

Group	Assessment 1A			Assessment 1B			Assessment 1C			Average 1	Assessment 2A			Assessment 2B			Assessment 2C			Average 2
	T1	T2	F	T1	T2	F	T1	T2	F		T1	T2	F	T1	T2	F	T1	T2	F	
IELTS	4	4	4	4	4	4	5	5	5	4.3	6	5	5	6	5	5	6	5	5	5
IELTS	4	5	5	5	5	5	5	6	6	5.3	6	5	5	6	6	6	6	6	6	5.6
IELTS	3	5	4	4	6	5	4	4	4	4.3	6	5	5	5	6	6	4	5	5	5.3
IELTS	4	4	4	4	5	5	5	4	4	4.3	5	4	4	5	4	4	6	4	5	4.3
IELTS	4	4	4	4	4	4	4	5	5	4.3	5	5	5	6	6	6	5	7	6	5.6
IELTS	3	3	3	4	4	4	3	3	3	3.3	6	6	6	5	6	6	5	6	6	6
IELTS	4	4	4	4	5	5	4	5	5	4.6	5	4	4	5	5	5	4	5	5	4.6
IELTS	3	4	4	3	3	3	4	5	5	4	5	5	5	6	5	5	5	5	5	5
IELTS	3	4	4	3	4	4	3	4	4	4	6	5	5	6	6	6	4	5	5	5.3
EAP	4	4	4	4	4	4	5	4	4	4	4	4	4	4	1	2	2	3	3	3
EAP	6	6	6	5	6	6	6	6	6	6	5	4	4	6	5	5	6	4	5	4.6
EAP	5	5	5	4	5	5	6	6	6	5.3	5	6	6	5	6	6	5	6	6	6
EAP	5	5	5	4	5	5	5	6	6	5.3	5	4	4	5	5	5	5	5	5	4.6
EAP	6	5	5	5	6	6	7	7	7	6	5	6	6	7	7	7	6	7	7	6.6

Table 3   Pre- and post-test raw score data

The first column from the left in Table 3 distinguishes between the IELTS and EAP group members. Each column thereafter indicates the IELTS score attributed to each student by each of three independent assessors. Scores were given for both writing task 1 and writing task 2. The shaded columns (F) present the overall score for each student from each assessor. The column titled Average 1 details the average scores from the three assessments for each student at the pre-test. The remaining columns present the score data in the same form for each student at the post-test. Therefore, comparison of the difference between scores at average 1, and scores recorded at average 2 will reveal the gain or loss achieved by each student during their course of instruction.

The three overall scores (in shaded columns) derived from the three assessments for each student were averaged in order to obtain the most robust overall result for each research subject at both the pre- and post-tests. Thus, all statistical figures in this analysis refer to

IELTS bandscores, and may be expressed in decimal point form due to the effect of averaging the three assessments for each student.

GROUP	N	Mean	Median Change	St.Dev.	Min. Change	Max. Change
IELTS	9	0.944	1.000	0.834	0.000	2.700
EAP	5	-0.300	-0.600	0.946	-1.300	0.700

Table 4   Descriptive statistics: pre- and post-tests

Table 4 shows that the IELTS group (n=9) has a mean score change of 0.944, that is, an improvement of nearly one band. The EAP group, on the other hand, showed a mean decrease of 0.3 of a band. (The figures here represented an averaging across raters and are hence not full band scores.). On average the median change in group 1 (IELTS) is +1 bandscore, and in group 2 (EAP) the median change is -0.6 bandscore. The *range* of score changes in the IELTS group was from a minimum of no change to a maximum of +2.7 bands. For the EAP group the change range was from a minimum of -1.3 bands (ie. a decrease in score) to a maximum of +0.7 of a band.

Figure 1 shows the individual score *changes* for each group. The IELTS group students being represented by '0', and the EAP group by '+'. The horizontal axis shows the average score for the pretest, and the vertical axis the score for the post-test. Thus, individuals falling on the diagonal line showed no change from pre-test to post-test, individuals falling above the line showed positive change from pre-test to post-test, and individuals below the line showed negative change. Thus, in the IELTS group two students had no change and seven had a positive change, whereas in the EAP group 3 students had a negative change and two had a positive change.
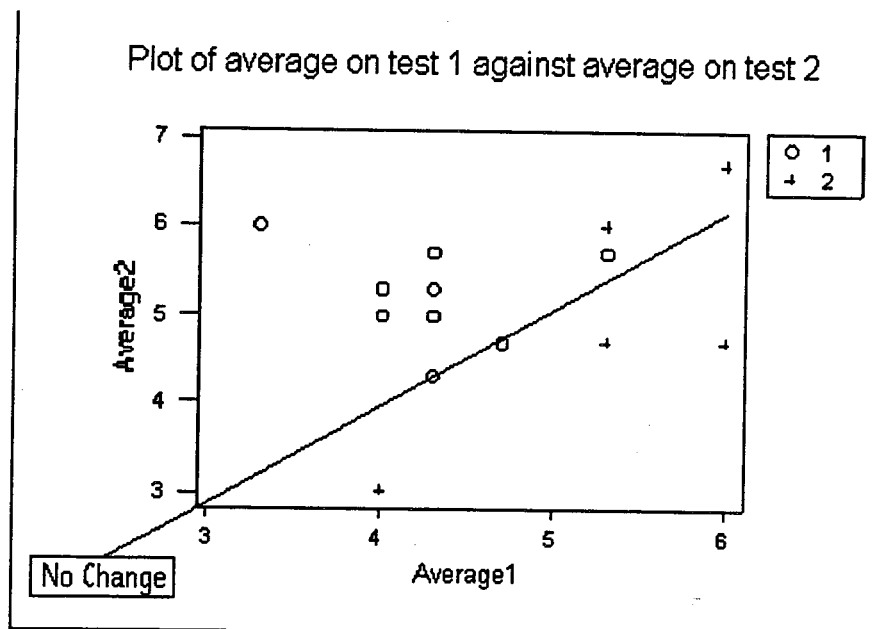

Figure 1   Plot of individual averages. Test 1 vs Test 2

3.3   Tests of Significance

Figure 1 reveals pre-existing differences of ability between the groups from the outset of the research program, with the scores of the IELTS group falling at the lower end of the scale (to the left on the horizontal axis) on the pre-test, and those of the EAP group at the upper end (to the right).   In order to undertake the quantitative analysis of the pre- and post-course IELTS test score data (to determine any existence and significance of gain in performance on the IELTS test following a ten-week program of instruction), the most appropriate statistical test would therefore be ANCOVA, as this test controls for pre-existing differences between groups. However, the lack of randomisation in the formation of the student groups violated the assumptions of this statistical model.

Furthermore, as the low numbers of subjects completing the research program prohibited the use of dependent T statistical analysis by violating the parametric assumptions of the test, another 'equivalent 2' sample test had to be used. The non-parametric equivalent of the T-test is Mann Whitney U, which examines the changes in rank between the pre- and post-test scores. Table 5 details the results of Mann Whitney U:

| IELTS | N = 9 | Median Change = | 1.000 |
| EAP | N = 5 | Median Change = | -0.600 |

Table 5 Mann-Whitney U Test results

Adjusted for ties, the test is significant at $p = 0.0441$. There is therefore evidence against the null hypothesis (of no difference) in the magnitude of the change between the two groups.

However, to return to the issue of pre-existing differences, it could be argued that the degree of change may be related to the differing group abilities at the beginning of the research study, rather than to the program of instruction. This is possible because the IELTS group were the least proficient at the pre-test, and therefore were subject to a greater likelihood of performance gain in the post-test whereas the EAP group were more proficient at the pre-test, and therefore may have achieved little or no gain in the post-test due to 'ceiling effect' (Davies et al, forthcoming). So the performance gain of the IELTS group could be related to the magnitude of the scores of their ability at the outset. Thus, while there is evidence of performance gain in the IELTS group, this cannot be definitively attributed to the program of instruction.

However, intuitively it would seem reasonable to assume some causal relationship between the program of instruction, and the change in scores. This effect is termed 'washback', and while there is a dearth of research into the empirical existence of this phenomenon (Alderson and Wall: 1993), it is nevertheless widely assumed to exist.

We turn now to the qualitative data for evidence to support this belief.

3.4   Qualitative Data

The most significant variation between the two programs of instruction in academic writing may be seen in Table 6 following, which displays the weekly timetables for writing classes of the two groups. The IELTS group received 7 hours of instruction per week, compared to the EAP group's 3 hours per week.

The frequency of writing instruction for the IELTS group was thus more than twice that of the EAP group throughout the period of research. This feature highlights the difference in emphasis placed on the teaching of writing by the staff of the two programs.

Monday	Tuesday	Wednesday	Thursday	Friday
	Writing		Writing	
	Writing		Task 1 Development	
	Task 2 Development			Writing Tasks 1 & 2
Journal Writing		Writing Skills		
			Writing	

Table 6   Weekly timetables

Legend:
IELTS Group (7 hrs instruction per week)
E.A.P. Group (3 hrs instruction per week)

The Program Features Matrix (Table 7) following was prepared from the classroom observation notes to identify program elements which differentiate the conduct of writing instruction on the two programs. The differences identified may help to explain the changes in proficiency revealed by the statistical analysis.

As is suggested by the major difference in the frequency of instruction between the two groups, the objectives of the writing programs differed substantially. The IELTS program sought to develop skills in short essay writing, as is required by the IELTS examination writing tasks. Whereas the EAP program objective was to develop skills in the preparation of an academic assignment of a more substantial length, as is appropriate in an EAP course.

A most substantial difference between the two programs was in the amount of practice of Task 1 and Task 2 writing given to the IELTS students. This group completed both Task 1 and Task 2 practices each course week. Furthermore, the IELTS students were required to complete three timed practices of the writing tasks during the course. They had therefore completed 13 sets of IELTS test writing task practice between the pre- and post-test.

Program Features	IELTS	EAP
Frequency of classes in writing instruction	7 hours per week x 10 weeks = 70 hours per course	3 hours per week x 10 weeks = 30 hours per course
Objective of writing instruction on course	Develop skills in Task 1 and Task 2 short essay writing	Plan, prepare and present one 1,000 word assignment on project topic
Knowledge of IELTS writing assessment criteria	Students given a summary of IELTS writing assessment criteria in course week 2	No reference made to writing assessment
Practice of Task 1 and Task 2 writing on course	All students complete one Task 1, and one Task 2 essay per course week = 10 x Task 1, and 10 x Task 2 practices per course	No task practices of writing during course. Emphasis on student research skills
Timed practice test writing	3 timed practices of both Task 1 and Task 2 during course	No timed practice writing test on course
Writing homework	Students encouraged to complete extra Task 1 & Task 2 essays as weekly homework	No requirement for writing homework - except the completion of course project
Feedback on writing	Correction code distributed to students in course week 2. All student writing subject to teacher application of the code encouraging self-correction by students	Feedback system informal. Teachers correct draft of project writing. Emphasis on content rather than accuracy
Instructional focus on structure of essay writing	All writing instruction oriented to short IELTS writing Task 1 and Task 2 requirements	Writing instruction oriented to long essay planning, research, preparation and organisation
Instructional focus on strategies for timed writing in tests	Emphasis on strategies for writing in examination conditions continuously maintained throughout course	Emphasis on strategies for writing in academic contexts, as in note-taking and summarising
Type of instructional materials	Majority of teaching material drawn from "IELTS Preparation" coursebooks	Majority of teaching material selected from "Academic Writing" coursebooks

Table 7   Program features matrix

At an early stage in the IELTS course the students were made aware of the IELTS writing assessment criteria in the form of an unofficial summary, which had reportedly been distributed at an earlier staff in-service seminar on IELTS preparation. The EAP students did not have reference to any writing assessment criteria.

By contrast, the EAP students had none of this practice, either timed or free. They developed study skills through project work involving literature searches, and instruction in the writing of summaries as well as the preparation of bibliographies.

Motivational differences can be identified between the groups in their approach to homework. The IELTS group was advised by the program staff that any practice of the writing tasks completed as homework would be promptly corrected and returned by their teachers. Some IELTS students took up this opportunity for extra practice and feedback by completing an extra task practice per day towards the end of the course. The EAP students were reluctant to undertake extra writing beyond that required on course, despite the invitation to do so by their teachers.

The presentation of feedback to the student groups was markedly different. Whilst the IELTS group completed many writing task practices, none of these practice papers were formally assessed with an IELTS bandscore. The program teachers were neither trained nor qualified as IELTS assessors. However, all student writing on the IELTS course was comprehensively marked for errors with the use of a correction code distributed early in the course. The pedagogic intention being to encourage student self-correction. Feedback also included comment on the success of the writing as an answer to an IELTS question in terms of its length, structure and accuracy.

The EAP group received feedback on their work in a less formal fashion. The correction system was not codified, and therefore differed from teacher to teacher. Comment was concerned more with the content of the writing in academic terms, rather than its linguistic accuracy.

The orientation of the instructional focus of the two courses was also very different. All of the writing instruction provided to the IELTS group was concerned with the successful completion of the IELTS writing tasks. The writing instruction of the EAP group was oriented differently towards meeting the demands of academic assignment planning, preparation and presentation.

The IELTS group received a continual instructional emphasis on strategies for writing in timed examination conditions, whereas the EAP students were taught strategies for writing in academic contexts, such as the skills of note-taking in lectures and the preparation of summaries.

As could be expected in programs of such different emphases as the IELTS and EAP courses, the teaching materials differed markedly in origin. IELTS program staff utilised a broad range of course material from several of the commercially available IELTS preparation coursebooks. The EAP program staff selected material from an academic writing textbook which had been designated as the coursebook. (Oshima, A. & A. Hogue. 1988).

The features of the two programs identified above represent such differences between the pedagogic experiences of the participants, that the performance gain of the IELTS group suggested by the statistical analysis is almost to be expected.

The intensity of writing task practice undertaken by the IELTS program students led to their completion of the writing tasks at the end of the post-test, precisely on time. Whereas the EAP program students were still writing furiously as the post-test came to an end, having had no training or practice of writing to the clock.

4.0 Summary of Conclusions and Recommendations for Further Research

Whilst the results of this study suggest that IELTS preparation programs may be successful in the preparation of students for the IELTS test, these results should be viewed with some caution due to the low sample size of the study.

In conclusion, the results of this research study suggest that IELTS preparation programs may be successful in the preparation of students for the IELTS test.

Continued instructional emphasis on IELTS writing task completion, together with the development of skills in writing and planning and systematic tuition in strategies for writing under examination conditions can assist IELTS preparation students to achieve an average gain of one IELTS bandscore over a ten-week course of instruction.

However, these results must be viewed with some caution as they were achieved with a small sample size, and therefore may not be generalisable to the wider population. There is a need for further research on this topic to verify these findings through a replication study with a larger population sample.

Furthermore, this research indicates that intense IELTS preparation may benefit IELTS candidates in terms of successful performance in the IELTS test. It does not indicate that such students are necessarily better prepared for the rigours of academic study. Further research is needed to determine whether IELTS or EAP preparation programs are successful in the training of ESL students in the skills required of them for successful participation in Australian tertiary education.

The context-adaptive model of language program evaluation was found to be valuable in proposing both a framework for composing the evaluation of a language education program, as well as ensuring a thorough and systematic examination of the hypothesis through the recursive features embodied within the model.

BIBLIOGRAPHY

Alderson, J.C. and L. Hamp-Lyons. (1993) TOEFL preparation and the communicative classroom. Paper presented at the TESOL Conference, Baltimore, USA.

Alderson, J.C. and D. Wall. (1993) Does washback exist? *Applied Linguistics*, 14, 2. 115-129.
Beretta, A. (1989) Attention to Form or Meaning? Error Treatment in the Bangalore Project. *TESOL Quarterly*, Vol. 23, No. 2, June 1989.

Clapham, C. (1996) The development of IELTS. Studies in Language Testing, No.4. Cambridge: Cambridge University Press.

Davies, A., A. Brown, C. Elder, K. Hill, T. Lumley and T. McNamara (forthcoming) Dictionary of Language Testing. Cambridge: University of Cambridge Local Examinations Syndicate.

Educational Testing Service (no date) Bulletin of Information for TOEFL/TWE and TSE. Princeton, NJ.: Educational Testing Service.

Lynch, B.K. (1990) A Context-Adaptive Model for Program Evaluation.
TESOL Quarterly, 24, 1: 23-42.

Lynch, B.K. (1993) Context-Adaptive Program Evaluation. Paper presented at the University
of Western Australia, Dept. of Education.

Lynch, B.K. (1996) *Language Program Evaluation. Theory and Practice*. Cambridge:
Cambridge University Press.

Nunan, D. (1988) *Syllabus Design*. Oxford: Oxford University Press.

Nunan, D. (1992) *Research Methods in Language Learning*. Cambridge: Cambridge
University Press.

Oshima, A. & A. Hogue. (1988) *Introduction to Academic Writing*. Reading, Ma.: Addison-
Wesley.

Ullman, R. & E. Geva. (1985) Expanding Evaluation Perspective. *The Canadian Modern
Language Review*, 42, 2. 307-323.

University of Cambridge Local Examinations Syndicate, The British Council and IDP
Education Australia (1996) The IELTS Handbook. Cambridge: UCLES.

Weir, C. & J.Roberts.(1994) *Evaluation in ELT*. Oxford: Blackwell.