

This article was downloaded by: [University Of Melbourne]

On: 29 June 2011, At: 21:40

Publisher: Routledge

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



Language Assessment Quarterly

Publication details, including instructions for authors and subscription information:

<http://www.tandfonline.com/loi/hlaq20>

The Interpretation and Use of Proficiency Test Scores in University Selection: How Valid and Ethical Are They?

Kieran O'Loughlin^a

^a The University of Melbourne

Available online: 25 May 2011

To cite this article: Kieran O'Loughlin (2011): The Interpretation and Use of Proficiency Test Scores in University Selection: How Valid and Ethical Are They?, Language Assessment Quarterly, 8:2, 146-160

To link to this article: <http://dx.doi.org/10.1080/15434303.2011.564698>

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: <http://www.tandfonline.com/page/terms-and-conditions>

This article may be used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan, sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

The Interpretation and Use of Proficiency Test Scores in University Selection: How Valid and Ethical Are They?

Kieran O’Loughlin
The University of Melbourne

In recent years there has been growing theoretical interest in exploring the relationship between the interpretation and use of high-stakes proficiency test scores. In these discussions, the role of institutional test users (or test score consumers) has received only limited attention. This may be due, at least in part, to the lack of consensus in the literature about the degree of responsibility test users have for the valid and ethical interpretation and use of test scores. To date, there has also been very little empirical research on the work of these stakeholders. This article reports on a study focusing on how the International English Language Testing System was used in the selection of international English as a Second Language students at an Australian university, and the knowledge and beliefs that test users (administrative and academic staff) had about the test. The results suggested that there were a number of serious flaws in the interpretation and use of test scores at this institution. Recommendations are made for improving the use of English proficiency evidence and the assessment literacy of staff within Australian universities.

INTRODUCTION

The interpretation and use of language tests, in particular the evidence they provide about individuals’ language ability for a specified purpose, is of fundamental concern in any evaluation of test validity. As Messick (1996) pointed out,

validity is not a property of test scores and other modes of assessment as such, but rather of the meaning of the test scores. Hence, what is to be validated is not the test or observation device per se but rather the inferences derived from test scores or other indicators—inferences about score meaning or interpretation and about the implications for action that the interpretation entails. (p. 245)

Test interpretation and its relationship to test use have been closely examined in recent years (Xi, 2008). Bachman (2005) posited an assessment use argument that includes a utilization argument linking interpretation to a decision and a validity argument linking assessment performance to interpretation. In relation to the utilization argument, he suggested that “it is possible for the results of an assessment to be used inappropriately, even though these assessments are valid indicators of the abilities they are intended to measure” (Bachman, 2005, p. 16). He identified four

warrants to justify making decisions on the basis of test interpretation relating to relevance, utility, intended consequences, and sufficiency. Chappelle (2008) discussed the relationship between interpretation and use as the last of six inferential steps (the utilization link) in her validity argument for the Test of English as a Foreign Language (TOEFL). She suggested that the warrant or justification for the utilization inference is that the test scores and other related information provided to users are relevant and useful for making decisions about student selection and appropriate curriculum for English as a Second Language (ESL) learners. In a similar vein, from an explicitly ethical standpoint, Hamp-Lyons (2000) argued that administrators need to be alerted

to the impact of the decisions they make, the uncertainty inherent in the data they base decisions on, the need for them to demand better information from testing agencies, and the need to follow ethical principles of their own professions in using test scores. (p. 580)

But how responsible are test users (also known as test score consumers), such as university staff, for the ways in which they interpret and use test scores in situ? The American Educational Research Association's *Standards for Educational and Psychological Testing* suggests that "the ultimate responsibility for appropriate test use and interpretation lies predominantly with the test user" (American Educational Research Association, American Psychological Association, and National Council on Measurement in Education, 1999, p. 112). The guidelines on test use of the International Test Commission (2000) suggest that competent test users will . . . interpret results appropriately, communicate the results clearly and accurately to relevant others and review the appropriateness of the test and its use. The International Language Testing Association's (2007) Code of Practice stipulates that test users should choose a test that is valid, reliable, and fit for purpose; understand the limitations of the test including its standard error of measurement (an index of score fluctuations due to various imperfect measurement conditions); and justify their decision-making process. Nevertheless, strong debate continues within the field about the relative responsibilities that language testers and test users have for the valid and ethical use of tests (Bachman, 2005; Davies, 2004; Hamp-Lyons, 2000; McNamara, 2000; Shohamy, 2001; Xi, 2008).

Given the "inevitable uncertainty" associated with proficiency test scores in terms of both their validity and reliability (Spolsky, 1995, p. 358), it is incumbent on universities to proceed judiciously with their interpretation and use. Rees (1999) argued that reliance on scores from a single test should be replaced by a "multi-judgment system with reference to international students' language proficiency assessment" (p. 434) involving collection of data including language aptitude, references from their home university or language study institution, an additional discipline related written proficiency assessment, and self-assessment. From an even broader perspective, Chalhoub-Deville and Turner (2000) suggested that admissions decisions should not only take into account measurement error indices that "reflect score fluctuations" but also "consider how language ability, individual factors, and academic requirements fit together to ensure more dependable admission decisions" (pp. 537–538). This is consistent with Graham's (1987) view that "language test scores should not play a disproportionate role in admissions decisions" given that "the relationship between English proficiency and academic success is complex and unclear" (p. 516).

In the university context, test users include the administrative, academic, and marketing staff who interpret and use test scores in a variety of ways. The primary focus of this article is on the administrative staff who assess student applications and, secondarily, on the academic staff who

interpret and use proficiency test scores more indirectly. To date, there has been little empirical research on these test users and, in particular, how they interpret and make use of high-stakes proficiency test scores. One important study is Banerjee's (2003) doctoral dissertation that examined the use of proficiency test scores, including the International English Language Testing System (IELTS) and TOEFL, in the selection of postgraduate degree courses at the University of Lancaster in the United Kingdom. She found that the selection of ESL international students was a complex, holistic decision-making process that was based on the recommendation of an admissions staff member taking into account a wide range of criteria including the applicant's academic background, intellectual capacity, evidence of English language proficiency, work experience, the applicant's own argued case for selection, reports from academic and work referees, personal characteristics (such as motivation, age, and adaptability), and in some instances a follow-up telephone interview. Although complex and time-consuming, the selection process Banerjee described ensures that no single factor such as the individual's academic record or English language proficiency scores dominates the selection process. Her findings indicated that if a risk is taken, it is more likely to be taken with language proficiency scores than with an applicant's academic or professional background.

English Entry Requirements in Australian Universities

Australian universities set their own entry standards for local and international students, although these standards must be consistent with the federal government's regulations for issuing international student visas. To enter a program of study, all applicants must meet two key criteria: the minimum academic and English requirements for that program. These two requirements are treated separately in the selection process, and applicants must demonstrate that they have clearly fulfilled each of them before they are selected. As Banerjee's (2003) study suggests, this straightforward approach is not universal. In the United States, for example, a range of other factors is commonly considered in university admissions such as academic letters of recommendation, a student's written statement of purpose, the results of the Scholastic Aptitude Test, and interviews (Van Nelson, Nelson, & Malone, 2004). In the Australian context, therefore, it is quite possible that previous academic results and a single measure of English proficiency are given too much weight in the selection process.

The English entry requirements of Australian universities can be met in a number of ways. On the basis of the results of a survey of 38 Australian universities, Coley (1999) identified no less than 61 different pieces of evidence that are accepted as fulfilling English language entry requirements for international and local ESL students across all of these institutions. Furthermore, 14 universities stated they accepted other unspecified evidence apart from those included in this total. As well as the IELTS, other widely accepted forms of evidence included the TOEFL, the General Certificate of Education A Level, the International Baccalaureate English, Australian senior secondary-school-level English or ESL, previous study in an English medium higher education institution, university foundation studies programs, and "in-house" English language bridging programs and tests.

Coley (1999) claimed that the IELTS was "a stringent measure of the student's English language ability for university study in Australia" (p. 10). However, she argued that Australian universities were pitching their minimum IELTS entry scores at levels that, in the terms of the

advice given by the IELTS partners, were at best only “probably acceptable.” As in a previous study by McDowell and Merrylees (1998), Coley also found that the IELTS was the only form of evidence accepted by all universities and that it was often termed their “preferred” test. The main reason for this preference has been that, since the early 1990s, the IELTS has been the most favored international English proficiency test by the Australian federal government and Australian universities as a result of their close relationships with the country’s leading international education industry body, IDP Education, which shares the ownership of the IELTS with Cambridge ESOL Examinations and the British Council. IDP Education was established in 1969 and is now jointly owned by 38 Australian universities and SEEK Limited, a major online employment and training company. IDP Education has played a key role in attracting and recruiting full-fee paying international students to Australian universities and colleges, especially over the last two decades. Thus, the relationship between the IELTS and Australian universities is commercially very strong. The IELTS remains the most heavily marketed and accepted test for prospective international students in Australian universities. It is the interpretation and use of the IELTS in the Australian university context that is the main focus of this article.

Interpreting and Using IELTS Scores

To guide institutions in setting minimum entry scores the *IELTS Handbook* (2007) suggests overall band scores, which are acceptable for particular courses. However, it also places the burden of responsibility clearly on the shoulders of test users in stipulating that individual institutions and programs must ultimately decide minimum entry scores “in light of their own courses and their experience of overseas students taking them” (*IELTS Handbook*, 2007, p. 5). As well as the overall band score, receiving institutions are advised to consider specifying scores for the components of the test so that the different language skills (listening, speaking, reading, and writing) can be matched to particular programs. Although not explicitly stated these guidelines imply, as Chalhoub-Deville and Turner (2000) argued, that it is important for test users to monitor and review their minimum entry scores to ensure that the minimum scores on English language tests are appropriate for entry to their academic programs. Finally, the *IELTS Handbook* (2007) suggested that “receiving institutions should also consider a candidate’s IELTS results in the context of a number of factors including age and motivation, educational and cultural background, first language and language learning history” (p. 5). This important recommendation implies that IELTS minimum entry scores should not necessarily be rigidly applied in the decision-making process but, instead, should be interpreted in relation to these kinds of individual factors.

Research into the interpretation and use of test scores in situ needs to include the collection of “subjective” data about the knowledge and beliefs that test users, such as administrative and academic university staff, have acquired in addition to “objective” data such as institutional documents on how the test is employed. These types of data have the potential to provide complementary perspectives on these complex issues. For instance, are detailed selection policies and procedures matched by informed administrative staff who use them carefully to assess student applications and academic staff who use them less directly to gain better understanding of their students’ learning needs? Banerjee (2003) reported that administrative and academic staff were not very knowledgeable about the meaning of proficiency test scores. This finding has since been supported in other studies exploring the knowledge and beliefs of test users about

IELTS test scores such as Coleman, Starfield, and Hagan (2003); Hyatt and Brooks (2009); and Rea-Dickins, Kiely and Yu (2007).

Purpose of the Study

The current study aimed to investigate the extent to which IELTS test scores were interpreted and used in valid and ethical ways for the purposes of student selection in an Australian university. Specifically, the study investigated the following questions:

1. How was the IELTS used to select international ESL students and to plan for their future language learning?
2. What knowledge did administrative and academic staff have about the IELTS, English proficiency, and the selection process?
3. What beliefs did administrative and academic staff have about the IELTS, English proficiency, and the selection process?

METHODOLOGY

The study took the form of an *instrumental case study* (Stake, 1994) that sought to understand the use of the IELTS within a specific faculty of one Australian university. Although a small case study, it aimed to shed light on the use of the IELTS in Australian higher education as an educational and administrative social practice, hence its instrumental nature. The study was approved by the university's human research ethics committee prior to its commencement.

The research site was a university located in a major Australian city. The targeted faculty was chosen because it had the largest number of international students in the university. Twenty staff (11 administrative and nine academic staff) volunteered to participate in the study. All of the administrative staff who participated in the study assessed student applications. Although they did not assess applications, the academic staff also interpreted and used evidence of English proficiency (albeit more indirectly) as the range of English abilities of students who are accepted in their courses informs their teaching and assessment in various ways. The academic staff provided a useful comparison with the administrative staff, who were most directly involved in the selection process.

Data for the study were gathered over the course of an academic year (from March to November). There were four main forms of data collection: (a) search and analysis of relevant university policy and procedures documents related to English entry, (b) statistics supplied by the targeted faculty relating to English entry, (c) a staff questionnaire, and (d) follow-up interviews conducted with selected staff. The questionnaire included both forced-choice and open-ended items. Interviewees were chosen on the basis of the representativeness of their questionnaire responses. The interviews, which were semistructured in format, invited interviewees to clarify and elaborate on their questionnaire responses. The interviews were conducted on either an individual or small-group basis. All 20 of the participants completed the questionnaire, and 10 staff (six administrative and four academic staff) volunteered for the follow-up interviews. Two senior academic staff members of the university, who were experts in the area of student selection, were also interviewed, although they did not complete the questionnaire.

First, the university selection policy and procedure documents were examined to build up a clear understanding of the policy and procedures related to the English language requirements for undergraduate and postgraduate study in the faculty, particularly the IELTS entry scores and the place of English language requirements in the selection process. Second, the questionnaire data were coded and entered into separate databases. Third, following Miles and Huberman (1994), running summaries of themes, issues, and ideas raised in the interviews were made and then checked and cross-referenced with field notes taken during the actual interviews. The principal researcher and research assistant independently checked the interview summaries against the field notes to maximise the reliability of this process. They then met to discuss and resolve any inconsistencies noted by either of them. The main issues raised in each interview were coded and then examined across all of the interviews to establish common themes. Finally, the results gathered from the document analyses, questionnaires, and interviews for each of the three research questions were grouped thematically and then synthesised to ensure that the results reported were as representative as possible of the data collected.

RESULTS

The findings for each of the three research questions are discussed next.

1. How Was the IELTS Used to Select International Students and to Plan for Their Future Language Learning?

This question was examined using relevant institutional documents, interview data, and statistical data supplied by the faculty.

English language requirements. Table 1 shows the minimum IELTS entry scores for the university in general and the specific faculty targeted in the study. In this table, the faculty entry scores are given separately because individual faculties are able to interpret the university guidelines for their own particular needs. Both the overall band score and any individual band score requirements from the four subtests in listening, speaking, reading, and writing are shown.

The two selection experts indicated that the original setting of IELTS and other English entry levels was based on both a shared sense across Australian universities of the required English standards and prevailing market forces rather than determined by any formal standard setting exercise, either then or at a later time. The *IELTS Handbook* (2007, p. 5) advises that an overall score of 6.5 (the required entry score for both undergraduate and postgraduate study in this faculty and many other faculties across the Australian university sector) is less than clearly acceptable for both linguistically demanding and less demanding academic programs. Like other English medium educational institutions around the world, this university seems to have decided to ignore this guideline because of intense competition with other universities for full-fee paying international students.

Given that IELTS is only one of a number of possible ways of satisfying the university's English language requirements, it was also important to establish the percentage of students that used IELTS results to enter the faculty compared to the most other popular forms of evidence

TABLE 1
Minimum International English Language Testing (IELTS) Entry Scores Required
by the University and the Faculty

	<i>University Requirements</i>	<i>Faculty Requirements</i>
Undergraduate	Overall band score of 6.5, with 6.0 in writing. If a faculty has a lower IELTS requirement, students may be required to enrol in, and pass, ESL subjects as part of their degree program.	Overall band score of 6.5, with 6.0 in writing. Students may be admitted with an overall band score of 6.0 provided they enrol in two semester-length ESL subjects in their 1st year of study.
Postgraduate	Vary for individual courses from an overall band score of 6.5 (with 6.0 in writing) to 8.0 (with no individual band less than 7.0). Some faculties may accept students with a slightly lower score. These students are required to undertake additional English as part of their academic programs.	Overall band score of 6.5, with no individual band less than 6.0 for all courses of study (except doctoral programs which required an overall score of 7.0 with no band less than 6.0). No provision for lower entry scores.

Note. ESL = English as a second language.

recognised by the university. Data provided by the faculty indicated that four different types of evidence were used by international students for undergraduate or postgraduate entry: IELTS, TOEFL, local and overseas secondary-school-level English/ESL subjects, and previous study in an English-medium tertiary institution. The results are summarised in Table 2. The table shows the numbers and percentages of undergraduate and postgraduate students submitting each of the five different types of proficiency evidence.

These figures show that IELTS was the most widely used form of evidence of English proficiency in the faculty, although only marginally so at the undergraduate level.

The selection process. The university had clear policy and procedures around the selection of international students. In the assessment of applications for both undergraduate and postgraduate programs, the primary emphasis was on applicants' academic qualifications and

TABLE 2
Type of English Language Evidence Used to Enter the Faculty

	<i>Undergraduate</i>		<i>Postgraduate</i>	
	<i>N</i>	<i>%</i>	<i>N</i>	<i>%</i>
IELTS	48	40	215	52
TOEFL	—	—	30	7
Senior secondary-level English/ESL	47	39	—	—
Previous study in English	18	15	136	33
Unknown	7	6	31	8
Total	120	100	412	100

Note. IELTS = International English Language Testing System; TOEFL = Test of English as a Foreign Language; ESL = English as a second language.

only secondarily on whether they had met the relevant English language requirements (i.e., using IELTS scores or other evidence). Applicants were accepted if they met both the academic and English requirements, made a “conditional offer” if they clearly met all of the academic prerequisites but not the English requirements, or rejected outright if they did not meet the academic requirements. Applicants who received conditional offers were not accepted until a later time when they had clearly met the English requirements. Such a rigid and lockstep approach arguably places too much reliance on the applicant’s academic record and a single piece of English proficiency evidence, which may be more or less accurate than the other acceptable measures of language ability. Contrary to the advice given in the *IELTS Handbook* (2007), the selection policy did not allow for IELTS scores to be considered in relation to other relevant individual factors. The same applied for other forms of proficiency evidence.

Assessment checklists were used to assess student applications, and the minimum IELTS requirements were mostly specified correctly in them. One noted anomaly was that the checklist used for undergraduate applications only specified “an overall band score of 6.5 or more” without any reference to the faculty’s minimum requirement of 6.0 for writing. This omission was of serious concern, as writing is the principal form of assessment in most university courses. Without checking this requirement it would have been quite possible to enter the faculty’s courses with an overall band of 6.5 composed of strong scores in listening, speaking, and reading and a low score for writing. This could have resulted in failure for students with this kind of score profile. This error was brought to the attention of senior selection staff when it was observed.

Monitoring and reviewing the use of the IELTS. One senior administrative staff member indicated that the faculty had not undertaken formal tracking of international students’ academic performance in relation to either IELTS entry scores or the other main types of English evidence shown in Table 2 in recent years. This was also true of other faculties in the university according to the two academic selection experts. Because the study was completed, this issue has been addressed by an English language task force, which examined the academic results of all first-year undergraduate students at the university as a function of the various types of evidence of English proficiency they used in their course applications. Although the results indicated the adequacy of the minimum overall IELTS band score of 6.5, the university accepted the task force’s recommendation that future applicants should have to obtain at least 6.0 in all four subtests of the IELTS rather than just in writing to ensure they have adequate proficiency in all four skill areas.

Linking decisions based on IELTS scores to consequences. A further dimension of the valid and ethical use of the IELTS relates to whether and how it is used to guide students’ future learning. As indicated in Table 1, the faculty had a provision to admit students who had only achieved an overall band score of 6.0 provided that they completed two ESL subjects towards their degree in their 1st year of study. This appeared to be the only instance in the faculty’s policy where an IELTS score was used to guide future English language learning. Most significantly, the *IELTS Handbook* (2007) advice that further English study is needed for students who enter linguistically demanding courses with an overall band score of 6.5 (the normal entry requirement for both undergraduate and postgraduate study in this faculty) had gone unheeded. However, in response to the recommendation of the English task force, the university now mandates that all

undergraduate students in this category must sit for a diagnostic test to determine whether they need additional English study.

2. What Knowledge Did Administrative and Academic Staff Have About the IELTS, English Proficiency, and the Selection Process?

This question was investigated through the questionnaires and interviews. First, staff respondents were asked to rate their knowledge of different aspects of the English language proficiency requirements for international students entering the university and the faculty, focusing on the IELTS, using a 4-point scale ranging *none*, *limited*, *good* and *extensive*. Table 3 provides a summary of staff responses to these items. The results for this section of the questionnaire suggested that administrative staff rated their knowledge of the university's English language proficiency requirements quite strongly. Ninety-one percent (10/11) of them rated their knowledge as good or extensive on Items 6, 7, and 10, and 82% (9/11) on Items 9 and 11. However, only 46% (5/11) rated their knowledge at these levels on Item 8, "How the IELTS overall band score is calculated." Although this result can be explained by the fact that they are not required to know this when assessing applications under the current system, it indicates a low level of understanding of the overall IELTS band score. There was greater variability in the academic responses across the six items, with 78% (7/9) rating their knowledge as good or extensive on Items 6 and 10, 67% (6/9) on Items 7 and 9, 44% (4/9) on Item 8, and only 33% (3/9) on Item 10. This is probably unsurprising given that they do not actually assess student applications.

Table 4 provides a summary of responses to the next section of the staff questionnaire where respondents were asked a series of questions that more directly tested their knowledge of how the IELTS was used in selection. For each statement the correct response is provided followed by the actual responses of staff participants.

TABLE 3
Staff Self-Assessment of Knowledge: Questionnaire Responses

Item		None		Limited		Good		Extensive		N
		A	B	A	B	A	B	A	B	
6	The university's English language proficiency entry requirements	0	0	1	2	4	6	6	1	20
7	The use of IELTS test scores in the selection of international students	0	0	1	3	4	5	6	1	20
8	How the IELTS overall band score is calculated.	3	4	3	1	2	4	3	0	20
9	The IELTS scores that are set for entry into the university	0	1	2	2	6	5	3	1	20
10	The IELTS scores that are set for entry into the programs of the faculty	0	1	1	1	2	6	8	1	20
11	Evidence other than IELTS scores that can be used to satisfy the faculty's English language proficiency entry requirements	0	2	2	4	3	2	6	1	20

Note. N = 20. A = administrative staff (N = 11); B = academic staff (N = 9); IELTS = International English Language Testing System.

TABLE 4
Staff Knowledge: Questionnaire Responses

Item	Correct Response	Yes		No		Unsure		No Response		N	
		A	B	A	B	A	B	A	B		
12	Are the minimum level IELTS overall and/or individual band scores required for entry into the faculty different for undergraduate and postgraduate students?	Yes	5	1	2	5	3	3	1	0	20
13	Are applicants with lower than the minimum required overall IELTS band score sometimes accepted into undergraduate programs in the faculty?	Yes	6	2	1	3	3	4	1	0	20
14	Are applicants with lower than the minimum required overall IELTS band score sometimes accepted into postgraduate programs in the faculty?	No	1	1	6	5	4	3	0	0	20
15	Are applicants with lower than the minimum required individual IELTS band scores sometimes accepted into undergraduate programs in the faculty?	Yes	4	1	3	3	3	4	1	1	20
16	Are applicants with lower than the minimum required individual IELTS band scores sometimes accepted into postgraduate programs in the faculty?	No	1	1	6	5	4	3	0	0	20
17	Are applicants admitted to the faculty with scores lower than the minimum IELTS scores usually required to enrol in additional English language credit subjects?	Yes	6	2	2	3	1	3	2	1	20
18	Is the date of the applicant's last IELTS test taken into account in the selection process?	Yes	11	3	0	0	0	5	0	1	20
19	Is the particular module of the IELTS test (i.e., Academic or General Training), which an applicant has taken, checked in the selection process?	Yes	6	2	1	3	4	4	0	0	20
20	Is the IELTS test more commonly used than other recognised measures of English language proficiency (e.g., TOEFL) by applicants for entry to Faculty courses?	Yes	7	5	1	1	3	3	0	0	20

Note. N = 20. A = administrative staff (N = 11); B = academic staff (N = 9); IELTS = International English Language Testing System; TOEFL = Test of English as a Foreign Language.

Despite the relatively strong self-assessments about English entry requirements in general and the IELTS in particular indicated by the administrative staff responses to Items 6 to 11, the results for Items 12 to 20 revealed a fair degree of uncertainty and/or inaccuracy with respect to the specific aspects of the use of IELTS scores in selection. The only item where all 11 administrative staff gave the correct response was Item 18. For the other items the range of correct responses

was from 64% (7/11) to 36% (4/11). These figures are of great concern given that all questions related directly to their work in selection. The most surprising results were for the relatively straightforward Items 12 and Item 19, with only 46% (5/11) and 55% (6/11) correct responses, respectively. Overall, these results do not inspire confidence that the faculty's English language entry policy and procedures were correctly followed by these administrative staff. However, given that they use comprehensive checklists to assess evidence of English proficiency, they may not retain this information outside of the immediate context of their work. Again, unsurprisingly, the percentages of correct responses across all of these items for the academic staff who do not assess applications are even lower.

Despite the availability of checklists, the levels of ignorance about the use IELTS in selection indicated by the responses to Items 12 to 20 are very troubling, particularly as most of these administrative and academic staff are required to formally and informally advise prospective students about English entry matters. It is very important that these staff can provide accurate information to these applicants about the required entry scores (including the degree of flexibility that is exercised around the overall and/or individual band scores) and what they should do if they have not yet achieved the required scores. They also need to be able to advise newly enrolled students about what their test scores imply for diagnostic testing and possible further language support.

In the interviews, administrative staff often indicated that they operated on a "need to know" basis about the IELTS, that is, they only searched out information about the test as required in their work. As one administrative staff participant suggested in relation to increasing her knowledge about the IELTS, "I think the information is there, and I have actually looked at the [IELTS] website and the reports they send out, but it's not something I have needed to know, so I am not filling my brain with it."

Another administrative staff participant suggested that "people understand what the requirements are, but I don't think they understand what the score actually means in real terms." The problem is that the current selection system does not demand such understanding. If administrative staff were required to make an informed, holistic judgment about the language proficiency of an international applicant, then they would have to be much better informed about IELTS test scores, both how they are derived and what they indicate about the applicant's language ability.

3. What Beliefs Did Staff Have About IELTS, English Proficiency, and the Selection Process?

Staff and student beliefs were examined through both the questionnaires and interviews.

Staff beliefs were examined in the final section of their questionnaire. Table 5 summarises their responses to a range of statements about the IELTS. The fact that two administrative and two academic staff omitted responses to all of these items suggests that they were unwilling or unable to express a view on the statements.

Although there was considerable variability in the responses to most of the statements among both administrative and academic staff, there were a number of trends in the responses. For Item 22, 73% (8/11) administrative staff did not believe that IELTS scores provide accurate evidence about an applicant's proficiency. In the interviews these staff suggested that students' scores

TABLE 5
Staff Beliefs: Questionnaire Responses

Item		Yes		No		Unsure		No Response		N
		A	B	A	B	A	B	A	B	
21	I believe that the current IELTS entry levels are adequate for students entering the faculty's programs.	4	2	3	3	2	2	2	2	20
22	I believe that IELTS scores provide accurate evidence about an applicant's English language proficiency.	1	2	8	2	0	3	2	2	20
23	I believe that some undergraduate courses of study should require higher IELTS entry scores than others.	2	2	4	2	2	3	3	2	20
24	I believe that some postgraduate programs should require higher IELTS entry scores than others.	3	3	2	3	4	1	2	2	20
25	I believe that postgraduate programs should require higher IELTS entry scores than undergraduate courses.	4	3	2	3	2	1	3	2	20
26	I believe that an applicant's IELTS scores should be considered in relation to factors such as age, motivation and language learning history.	0	2	7	2	1	2	3	2	20
27	I believe that a selection interview (face-to-face and/or by phone) would be a useful addition to IELTS scores, if resourced adequately.	6	6	1	0	2	1	2	2	20
28	I believe that IELTS scores are good predictors of academic success.	0	1	7	4	2	2	2	2	20
29	I believe that all staff involved in selection have a good understanding of IELTS test scores.	6	4	2	0	1	3	2	2	20
30	I believe that an applicant's English language proficiency is as important as their academic record in making selection decisions.	6	5	2	0	1	2	2	2	20

Note. N = 20. A = administrative staff (N = 11); B = academic staff (N = 9); IELTS = International English Language Testing System.

were sometimes artificially inflated because of (a) excessive coaching or support leading up to the test, (b) less stringent scoring in some countries, and (c) dishonesty on the part of students and/or examiners. In other words, they had serious concerns about the validity, reliability, and security of the test based on their contact with students when they first arrived in Australia. This, of course, may reflect a lack of understanding of what minimally acceptable IELTS scores imply about English proficiency.

However, in terms of the selection process, the lack of faith the majority of administrative staff had in the accuracy of IELTS test scores may have reflected a judicious scepticism about making entry decisions based on the results of one test. This inference is supported by their responses to Items 26 and 27. For Item 26, 64% (7/11) agreed with the advice in the *IELTS Handbook* (2007) that an applicant's scores should be considered in relation to factors such as age, motivation, and language learning history. For Item 27, 55% (6/11) believed that a selection interview would

be a useful addition to IELTS scores if resourced adequately. These responses indicate that the majority of administrative staff would prefer a more complex decision-making process about English proficiency than is currently the case, notwithstanding their limited knowledge of the faculty's current IELTS requirements reported in the previous section.

There were several administrative staff who dissented from the majority opinion on Items 26 and 27. In the interviews, one of them argued that interpreting IELTS scores in relation to other factors such as age, motivation, and language learning history would "open the floodgates for arguments about student entry . . . and who are the selection officers to make those sort of judgments on somebody's motivation or their language learning ability?" The problem with holding this position is that it leads all too easily to the kind of attitude expressed by an administrative participant who was also speaking for the minority on this issue: "I think that's the nice thing about the IELTS. Whether it's wrong or right, you've got to be . . . black and white, you're in or you're out." This comment starkly foregrounds how arbitrary and potentially unfair the current decision-making process is at this university. While there were no clear patterns in the responses of academic staff to Items 22 and 26, the majority (6/9 or 67%) of them (like most of the administrative staff) supported the inclusion of a selection interview in Item 27.

For Item 28, the majority of administrative staff (64%, or 7/11) correctly believed that IELTS scores are not good predictors of academic success. They were likely to know this was true because another dimension of their work was to gather and report on student progress across the faculty. The academic staff were more divided on this question, with less than half (44% 4/9) adopting the majority administrative staff view. One of the senior academic selection experts who was interviewed referred to "the unthinking reverence with which [IELTS score] are treated. . . . It just continues to amaze me when . . . everyone . . . on [the university's] selection procedures committee keep on being told . . . that it [IELTS] doesn't have this kind of predictive significance that they think it has." If this mistaken belief is shared widely across the university, it may serve to perpetuate the prevailing culture of relying so heavily on a single measure of English proficiency in the selection process.

Both groups of administrative and academic staff were divided on the final statements (Items 29 and 30). The fact that only 54% (6/11) of the administrative staff and 44% (4/9) academic staff believed that selection officers have a good understanding of IELTS scores reinforces the previous finding that there was variable knowledge among these staff about the test and the entry scores accepted by the faculty. Finally, the fact that 55% (6/11) administrative staff and 56% (5/11) academic staff considered English proficiency to be as important as academic record in the selection indicates that these staff took a different view from the one enshrined in the university's selection procedures on this question as discussed in the results for the first research question.

CONCLUSION

There are several major validity concerns arising from the findings of this study. First, it appears that there was no principled basis for originally establishing IELTS minimum entry scores in this context, including use of the test guidelines. There was also no tracking of student success to validate these entry requirements. In other words, the decisions made on the basis of applicants' test scores were poorly informed and were therefore neither valid nor ethical. However, subsequent statistical data collected on the academic performance of 1st-year undergraduates across

the university did indicate that the minimum IELTS overall band score was adequate. Second, applicants' entry scores were not considered in relation to other relevant individual factors as recommended in these guidelines. From this perspective, the interpretation and use of test scores were also invalid. Third, IELTS scores were not used to guide English language learning except for undergraduate students sometimes admitted with an overall band score of less than 6.5, the normally required minimum entry level. This suggested there had been few beneficial educational consequences. This suggested that there had been no beneficial educational consequences flowing from use of the test. However, it was noted that at least undergraduate students entering with an overall band score of 6.5 are now referred to diagnostic testing and possibly additional English support once they are accepted into the university.

At the time this study was conducted, therefore, the selection policy and procedures fell well short of the standards required for the valid and ethical interpretation and use of test scores. The approach to selection discussed here would appear to be the same in most Australian universities. Some improvements across this particular university, however, have occurred more recently in relation to monitoring the appropriateness of minimum test score requirements and linking these with beneficial test consequences. These changes are also beginning to occur in other universities as the English proficiency of international ESL students at both entry and graduation comes under increasing scrutiny by the federal government and the wider public.

The key issue, which has not been addressed to date, is the interpretation of test scores. Given the limitations and uncertainty of test scores, they need to be carefully interpreted in relation to other relevant information about applicants such as their age, first language, and language learning history. Such a shift in approach by Australian universities requires a major change in selection policy and procedures. University policymakers need to be better educated about the valid and ethical interpretation of IELTS test scores and the other accepted type of English proficiency evidence. From a procedural perspective, a key challenge for university policymakers and researchers is to explore how this kind of approach could be adopted in the context of large numbers of applications submitted by international ESL applicants and processed by busy admissions staff who currently lack the necessary training to make such complex decisions. The kind of training they would need to undertake would also need to be carefully examined.

However, it is not only policymakers and administrative staff who need to be better educated about the interpretation and use of test scores and the other accepted measures of proficiency. Commenting on the use of proficiency tests in Australian universities, Ingram (2005) suggested that "all persons involved, from marketers to academics and administrators, need to have a better understanding of what English language proficiency means and of what English language tests (especially the IELTS) do and do not measure" (p. 5). Increased levels of knowledge and awareness among all university stakeholders will enable more informed and principled selection policies and procedures to be developed, implemented, monitored, and regularly reviewed.

In relation to the IELTS specifically, universities must rely on the IELTS partners to inform them about the meaning of IELTS test scores (including their limitations) and advise them about their potential use. Despite the development of the IELTS website (<http://www.ielts.org>) including *the IELTS Guide* (2009), *the IELTS Handbook* (2007), *the IELTS Scores Explained DVD* (2009), and information seminars, there is clearly more work to be done by the IELTS partners to actively build and promote the assessment literacy of all test users at a level appropriate to their various institutional roles. In the final analysis, test agencies and test users must share the responsibility for the appropriate interpretation and use of test scores.

REFERENCES

- American Educational Research Association, American Psychological Association, and National Council on Measurement in Education. (1999). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
- Bachman, L. F. (2005). Building and supporting a case for test use. *Language Assessment Quarterly*, 2, 1–34.
- Banerjee, J. V. (2003). *Interpreting and using proficiency test scores*. Unpublished doctoral thesis, Department of Linguistics and Modern English Language, Lancaster University, Lancaster, United Kingdom.
- Chalhoub-Deville, M., & Turner, C. (2000). What to look for in ESL admission tests: Cambridge certificate exams, IELTS and TOEFL. *System*, 28, 523–539.
- Chapelle, C. A. (2008). The TOEFL validity argument. In C. A. Chapelle, M. Enright, & J. Jamieson (Eds.), *Building a validity argument for the test of English as a foreign language* (pp. 322–354). New York, NY: Routledge.
- Coleman, D., Starfield, S., & Hagan, A. (2003). The attitudes of IELTS stakeholders: student and staff perceptions of IELTS in Australia, UK and Chinese tertiary institutions. In *IELTS research reports* (Vol. 5, pp. 160–207). Canberra: IDP Education Australia.
- Coley, M. (1999). The English language entry requirements of Australian universities for students of non-English speaking background. *Higher Education Research & Development*, 18, 7–17.
- Davies, A. (2004). Introduction: Language testing and the golden rule. *Language Assessment Quarterly*, 1, 97–107.
- Graham, J. G. (1987). English language proficiency and the prediction of academic success. *TESOL Quarterly*, 21, 505–521.
- Hamp-Lyons, L. (2000). Social, professional and individual responsibility in language testing. *System*, 28, 579–591.
- Hyatt, D., & Brooks, G. (2009). Investigating stakeholders' perceptions of IELTS as an entry requirement for higher education in the UK. In *IELTS research reports* (Vol. 10, pp. 17–68). Manchester, UK: The British Council and Canberra: IDP Education Australia.
- IELTS Guide*. (2009). Cambridge, UK: Cambridge ESOL, the British Council and IDP Education Australia.
- IELTS Handbook*. (2007). Cambridge, UK: Cambridge ESOL, the British Council and IDP Education Australia.
- IELTS Scores Explained*. [DVD] (2009). Cambridge, UK: Cambridge ESOL.
- Ingram, D. (2005). English language testing: a pass for proficiency is not necessarily the answer. *UniNews*, 14, 4–5. Retrieved from <http://archive.uninews.unimelb.edu.au>
- International Language Testing Association. (2007). *Guidelines for practice*. Retrieved from <http://www.iltaonline.com/code.htm>
- International Testing Commission. (2000). *International guidelines for test use*. Retrieved from <http://www.intestcom.org/>
- McDowell, C., & MerryLees, B. (1998). Survey of receiving institutions' use and attitude to IELTS. In *IELTS research reports* (Vol. 1, pp. 116–139). Canberra: IDP Education Australia.
- McNamara, T. (2000). *Language testing*. Oxford, UK: Oxford University Press.
- Messick, S. (1996) Validity and washback in language testing. *Language Testing*, 13, 241–256.
- Miles, M. B., & Huberman, M. A. (1994). *Qualitative data analysis: An expanded sourcebook*. Thousand Oaks, CA: Sage.
- Rea-Dickins, P. R. Kiely, R., & Yu, G. (2007). Student identity, learning and progression: The affective and academic impact of IELTS on “successful” candidates. In *IELTS Research Reports* (Vol. 7, pp. 59–136). Manchester, UK: British Council and Canberra: IDP Education Australia.
- Rees, J. (1999). Counting the cost of international assessment: why universities may need a second opinion. *Assessment and Evaluation in Higher Education*, 24, 427–438.
- Shohamy, E. (2001). *The power of tests*. New York, NY: Longman.
- Spolsky, B. (1995). *Measured words: The development of objective language testing*. Oxford, UK: Oxford University Press.
- Stake, R. E. (1994). Case studies. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 236–247). London: Sage.
- Van Nelson, C., Nelson, J. S., & Malone, B. G. (2004). Predicting success of international graduate students in an American university. *College and University Journal*, 80, 19–27.
- Xi, X. (2008). Methods of test validation. In E. Shohamy & N. H. Hornberger (Eds.), *Encyclopaedia of language and education, Volume 7: Language testing and assessment*. (pp. 177–196). New York, NY: Springer Science and Business Media.