Fairness of IELTS Test Scores in University Admission

Motahar Khodashenas Tavakkoly

Islamic Azad University, Science & Research Branch

PO box 1981645561, Saadat-Abad, Tehran, Iran

Tel: 98-912-344-8196   E-mail: Kathy.kathy666@yahoo.com

Abstract

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 students in an EFL context at an Iranian university and the knowledge and beliefs that test users (administrative and academic staff) had about the test. The central issues raised for readers by this paper is the possibility that judgements made on the basis of IELTS may not correlate with the subsequent performance of students, and that flaws or strengths in these performances may be correlated with IELTS scores and the ensuing entry judgements. 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 in universities in Iran or other places around the world.

Keywords: Fairness in tests, High stake proficiency test scores, IELTS interpretation, University staffs as decision makers

1. 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. 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. Chapelle (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). 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 Banjeree 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.

2. 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 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. It is suggested that a generally high level of procedural compliance with university policy and procedures on the part

of staff directly involved in selection is required. 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).

2.1 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 Iranian university. Specifically, the study investigated the following questions:

i. How was the IELTS used to select EFL students and to plan for their future language learning?

ii. What knowledge did administrative and academic staff have and use about the IELTS, English proficiency, and the selection process?

iii. What beliefs did administrative and academic staff have about the IELTS, English proficiency, and the selection process?

3. 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 Iranian university. Although a small case study, it aimed to shed light on the use of the IELTS in Iranian higher education as an educational and administrative social practice, hence its instrumental nature. IELTS score is required to accept students in different fields in undergraduate and postgraduate courses. All the curriculum are taught in English. The study was approved by the university's human research ethics committee prior to its commencement. The research site was a university located in the capital city (Tehran). Twenty staff (11 administrative and nine academic staff) participated 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. Regarding the concept of washback (or backwash), academic staff acknowledged that IELTS is promoting positive backwash as the result of features such as: the standard inclusion of tests of speaking and writing as well as listening and reading; the design of test tasks that simulate academic study tasks; the selection of longer written texts, with some focus on their discourse structure; and the use of a variety of test item types. However, it can not be assumed that these features will have the desired effect of encouraging teachers preparing for the test to develop the range of academic language skills required in graduate and postgraduate studies. For instance, academic staff in this study, found that students who passed IELTS still tended to struggle with longer writing tasks, critical and conceptual analysis, and the amount of academic reading they had to do. 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 semester (from January to June). 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) did 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.

4. Results

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

i. How was the IELTS used to select Iranian 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. The two selection experts indicated that the original setting of IELTS and other English entry levels was based on both a shared sense across Iranian 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 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.

The selection process. The university had clear policy and procedures around the selection of students. In the assessment of applications for both undergraduate and postgraduate programs, the primary emphasis was on applicants' academic qualifications and only secondarily on whether they had met the relevant English language requirements (IELTS scores). 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 IELTS entry scores. 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. 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 fairness and ethical use of the IELTS relates to whether and how it is used to guide students' future learning. The faculty had a provision to admit students who had only achieved an overall band score of 6.0. 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.

ii. 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 1 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 2 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. 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 applicant, then they would have to be much better informed about the selection process in IETLS. The range of criteria taken into account is extremely rich including the applicant's academic background, intellectual capacity, evidence of English language proficiency (IELTS or other recognised measure), 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. Such factors can be either "indicators of success" or "sources of struggle" (Banjeree, 2003, p.236-237). Other factors influencing their decisions include the offer-acceptance ratio, recommendations from other academic colleagues, the reports of agents and scholarship agencies. Successful applicants are described as either a "clear accept", "safe bet" or "risk" and appropriate recommendations are made to students in the second and third categories about undertaking additional English study or seeking additional academic assistance. Banjeree's findings (2003, p.235) 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.

iii. 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 3 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 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. 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 skepticism 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'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.

5. Limitations of the Study

Even the intention to investigate in the future in this setting any correlation between IELTS scores and subsequent performance does not extend to considering correlations between particular aspects of the students' performances and their English proficiency. It is therefore a regrettable omission from the study that the extensive questionnaire to staff did not ask them what aspects of the submitted work of students with low IELTS scores did, or did not, suggest that limitations in their English ability had affected their performance. Another limitation of this study is related to the small sample size.

6. Conclusion

This study has examined the use of IELTS in the selection of EFL students in one faculty of an Iranian university. Clearly, there is a need for other such studies to be conducted so that this issue can be critically examined in the various institutional contexts in which IELTS is employed around the world. Such research is warranted in the interests of identifying valid and ethical selection procedures to inform policy and practice in this area. In methodological terms, this study has demonstrated the importance of careful data collection instruments, such as questionnaire and interview schedules, which have been developed to investigate complex and sensitive issues

relating to the use of the IELTS for high-stakes purposes. It is recommended that future studies incorporate careful trialing of such instruments.

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. 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 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 Iraninan 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. 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 universities around the world 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 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.

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. http://dx.doi.org/10.1207/s15434311laq0201_1

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. http://dx.doi.org/10.1016/S0346-251X(00)00036-1

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*. New York, NY: Routledge. pp. 322-354

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*. Canberra: IDP Education Australia. Vol. 5, pp. 160-207

Davies, A. (2004). Introduction: Language testing and the golden rule. *Language Assessment Quarterly*, 1, 97-107. http://dx.doi.org/10.1207/s15434311laq12&3_2

Graham, J. G. (1987). English language proficiency and the prediction of academic success. *TESOL Quarterly*, 21, 505-521. http://dx.doi.org/10.2307/3586500

Hamp-Lyons, L. (2000). Social, professional and individual responsibility in language testing. *System*, 28, 579-591. http://dx.doi.org/10.1016/S0346-251X(00)00039-7

Hyatt, D., & Brooks, G. (2009). Investigating stakeholders' perceptions of IELTS as an entry requirement for higher education in the UK. In *IELTS research reports*. Manchester, UK: The British Council and Canberra: IDP Education Australia. Vol. 10, pp. 17-68

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. [Online] Available: http://archive.uninews.unimelb.edu.au

International Language Testing Association. (2007). Guidelines for practice. [Online] Available: http://www.iltaonline.com/code.htm

International Testing Commission. (2000). International guidelines for test use. [Online] Available: http://www.intestcom.org/

McNamara, T. (2000). *Language testing*. Oxford, UK: Oxford University Press.

Messick, S. (1996). Validity and washback in language testing. *Language Testing*, 13, 241-256. http://dx.doi.org/10.1177/026553229601300302

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*. Manchester, UK: British Council and Canberra: IDP Education Australia. Vol. 7, pp. 59-136

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. http://dx.doi.org/10.1080/0260293990240406

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*. London: Sage. pp. 236-247

Appendix

Table 1. Staff self-assessment of knowledge: Questionnaire responses

Item	None		limited		Good		Extensive		
	A	B	A	B	A	B	A	B	N
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 2. Staff Knowledge: Questionnaire Responses

Item	Correct Response	Yes A	Yes B	No A	No B	Unsure A	Unsure B	No Response A	No Response B	N
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	0	5	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.

Table 3. Staff beliefs: Questionnaire responses

Item	Yes		No		Unsure		No Response		
	A	B	A	B	A	B	A	B	N
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 a selection interview (face-to-face and/or by phone) would be a useful addition to IELTS scores, if resourced adequately.	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.