

Assessing Entry-Level Academic Literacy with IELTS in the U.A.E.

Christopher Morrow

Abstract While the IELTS exam has earned a reputation for its usefulness in many international educational contexts, it does not appear to fully address the specific needs and issues of students and universities in the United Arab Emirates. IELTS was introduced in 2004 at the main national university, United Arab Emirates University, to complement existing assessment tools in the English Foundation Program. Since then, its influence on instructional and assessment practices has grown at that university in ways that do not seem to be completely congruent with the needs of low-level English learners who are about to begin studies in English-medium universities. Evidence for this comes from the claims IELTS makes about the test, the national results of UAE students, research on English-medium instruction in the UAE, and survey results from English lecturers in the UAE University Foundation Program. This chapter argues that the advanced linguistic demands of the IELTS exam, its equal weighting of scores from the four skill-based sections, and the exam's general communicative orientation are not well-suited for making valid and reliable decisions about the readiness of Arab students to begin college-level studies in English. The disproportionate effect of testing in UAE educational programs makes it imperative that a more appropriate means of assessing students is selected or developed. Alternative assessment options are considered which address these issues more directly and efficiently.

Keywords IELTS exam • Language proficiency testing • Validity research • Washback • English-medium instruction • Admissions testing

C. Morrow (✉)
U.A.E. University, Al Ain, UAE
e-mail: gomorrow@gmail.com

1 Introduction

The rapid introduction of tertiary-level English medium-instruction (EMI) in the United Arab Emirates (UAE) has had major effects on its colleges and universities. The impact this policy shift has had on content learning, Arabic literacy, and English language development has been problematic in several ways, especially since overall levels of English proficiency are quite low (McLaren, 2011). Much of the legitimacy for this policy change is linked to the IELTS testing system as a means of determining students' readiness for EMI in colleges and universities. Unlike most other countries, UAE colleges and universities usually accept a low score of overall Band 5 as evidence of readiness for EMI (Gitsaki, Robby, & Bourini, 2014). It is unclear if this demanding and sophisticated test is appropriate for this purpose considering its high-level texts, its limited ability to address local cultural perspectives, and its effects on classroom instructional practices (viz., washback). It appears that IELTS is being used in ways that are not in total accordance with the realities of undergraduate study in EMI programs in which students have ongoing language development needs. Survey data from English lecturers at United Arab Emirates University (UAEU) is summarized at the end of this chapter to identify some apparent discrepancies between the IELTS exam requirements and students' developmental needs.

This chapter focuses on the use of the IELTS exam at UAEU, the country's main national university (which is referred to as a 'test user' in this chapter). While prospective students normally take at least one course in the UAEU's Foundation Program (FP), a growing number are achieving the minimum requirement (IELTS Band 5) on their own and skipping all or part of the FP English course sequence (Moussly, 2012). This development puts more pressure on institutions to assess students accurately, both by choosing an appropriate test and by setting appropriate minimum requirements in the specific sections of the tests. IELTS has had a disproportionate influence on English teaching throughout the region recently; however, its usefulness is unclear relative to the types of English proficiency needed for success at institutions like UAEU.

2 Entry-Level Testing

Language tests for admissions are crucial for setting the norms of EMI and determining its viability. If education is inherently dependent on language and communication, then achieving adequate levels of English at entry strongly affects the results of EMI (Doiz, Lasagabaster, & Sierra, 2013). Although many UAEU students are still learning intermediate-level academic English, the university typically follows a late immersion approach to education which occurs almost completely in English. At a similar university, Rogier (2012) reported that teachers found the weak writing and listening skills of their students forced them to modify

their instructional and assessment methods which limited their overall effectiveness. As the leading national university, UAEU has programs like linguistics, philosophy, and medicine which are inherently demanding both from cognitive and linguistic perspectives. Several researchers at UAEU have described the obvious difficulties that this causes when students have low levels of English achievement (e.g., McLaren, 2011; McLean, Murdoch-Eaton, & Shaban, 2013). My own teaching in English language education is particularly constrained by students' weak reading skills so I am sympathetic toward colleagues who resort to summarizing as much as possible for students in PowerPoint presentations in order to make sure students comprehend essential course content. Technology offers some new options (e.g., audio and video recordings of lectures), but I consistently have students who lack the necessary literacy skills to manage the course requirements. Ironically, students with the lowest English skills often choose the humanities and education even though these fields have some of the highest linguistic demands.

3 Searching for a Proficiency Test

Prior to 2004, when the IELTS was introduced at UAEU, undergraduate students were nearly all graduates of its own English FP. High school students were placed in one of three levels in the FP by means of a national standardized test called the Common Educational Proficiency Assessment (CEPA). Typically, a very small group of about 5 % got high CEPA scores and were allowed to sit for a TOEFL in order to skip the FP English track completely. The vast majority spent 1 or 2 years in the FP program where they were evaluated using many types of assessment activities which emphasized the achievement of specific English course outcomes. Evidence for students' readiness for EMI came from the FP's conventional tests and assignments, so those who passed the program's highest level of English (Level 3) were deemed eligible to begin their undergraduate coursework. These types of assessments resembled many of the projects and tests they would face later on, so they were authentic in the sense that they were based on classroom activities and academic requirements. This authenticity was valuable, but measurement errors were possible at the two extremes of student performance. The first was that strong students may have lacked appropriate opportunities to show they were qualified to bypass the FP. The second, and more serious one, was that weak students sometimes got through the FP after failing and repeating several times even though their true language proficiency levels were still inadequate. Implementing a reliable external exam seemed like an objective way to address the assessment needs of very weak and very strong students.

4 Introduction of IELTS

TOEFL and IELTS were the leading candidates for becoming university entrance exams at UAEU in 2004. The paper-based TOEFL available at that time was relatively academic in nature and used only multiple-choice items to assess listening, reading, grammar, and structure. In contrast, IELTS incorporated all four skills and adopted a more communicative approach to language proficiency in line with the Common European Framework of Reference. This may have appealed to administrators since it gave students credit for their functional and oral abilities, skills which were not assessed well by the normal FP instruments. Furthermore, IELTS was already widely used throughout the region for university admissions purposes, and several teachers in the FP at UAEU had experience as raters of its speaking and writing sections. Given the intense focus that the government placed on improving English, there was reason to believe in 2004 that standards would eventually rise and IELTS requirements to enter EMI courses would reach levels that were similar to the rest of the world.

EMI had already been commonplace at UAEU for a few years in 2004. The first cohorts to take the IELTS in 2004 only needed to reach a benchmark score of Band 4.5, and 85 % of those who passed Level 3 attained this low requirement. While most other UAE public colleges soon set their minimum score at Band 5, UAEU did not raise its requirements to Band 5 until 2011. By that time, about 60 % of students were normally reaching the overall score of Band 5 (Morrow, 2005).

By setting Band 5 as a minimum score, public colleges and universities in the UAEU are catering to students who are clearly at the low end of the spectrum of IELTS candidates globally. IELTS publishes score results on their website which show how poorly UAE students tend to perform on an international level. Globally, only 10 % of all candidates received scores of Band 5 or lower in 2012 on the Academic version of IELTS (IELTS, 2012). In contrast, 72 % of candidates in the UAE scored in that low range. Similarly, just 12 % of UAE candidates obtained Band 6 or above in 2012, but 76 % of the world-wide cohort achieved that level (IELTS, 2012). These figures cover every candidate who sits for an IELTS in the UAE for any reason; consequently, they are not necessarily an accurate reflection of the nation's overall English proficiency level. Nevertheless, they suggest that UAE students may not be the main target group for the IELTS test. Although IELTS has a well-earned reputation for test quality, the relevance of available validity evidence to low scorers has not been conclusively established. Most testing programs address this issue by producing an array of exams which target successive levels of proficiency. IELTS may function well across several levels but it is very unlikely that it can measure proficiency across all nine levels with equal levels of accuracy.

5 Early Costs of the IELTS

Implementing the IELTS exam on such a large scale required a great deal of support from UAEU staff, lecturers, and administration because of its burdensome financial and logistical issues. At a current cost of \$239 (USD) per candidate, the exam is relatively expensive and time-consuming. Examining hundreds of candidates often takes several days because of the speaking interviews. It is common for the FP to lose 1 week of instructional time at the end of a term to compensate for the fact that IELTS is only available intermittently (typically about three times each month). Preparing students to sit for the test consumed a fair amount of class time in Level 3 since IELTS tasks employ a wide array of question formats. Test preparation became a dominant aspect of many Level 3 classes because students needed to face two major exams at the end of the school term: the standard course exams immediately followed by IELTS. Finally, UAEU was forced to follow the IELTS protocols for handling the results. This included making students wait about 10 days for verified scores to arrive from abroad. These burdensome procedures appeared justified at the time as a way of identifying a small group of low-achieving students who were truly unprepared for EMI despite their success in the FP courses. Interestingly, however, most of those students had actually failed major parts of the FP exams; their weaknesses were identified by existing assessment practices, but they ended up passing the course because of high scores on other assignments.

6 Early Benefits of the IELTS

There were some immediate benefits which seemed to justify IELTS as a benchmark test. For the first time ever, the actual achievements of UAEU students could be compared to others worldwide. Of course, the test challenged even the best students in the FP and this increased their motivation to study much more than the relatively easy FP tests they were used to. By including a speaking section, IELTS helped document a skill area which had been largely ignored previously. FP administrators were able to use IELTS results to check the concurrent validity of their own tests and assessment instruments. Perhaps the greatest benefit of the IELTS for the whole nation was a logistical one: any UAE student could visit their local IELTS testing center and find out if their English was adequate to attend the main national university without attending the English FP. Outsourcing assessment to an international organization brought benefits but unfortunately the costs associated with this practice were considerable.

Of course, the challenges of EMI at UAEU have little to do with the IELTS exam per se; instead, they are readily explained by the fact that low scores are accepted for admission to EMI programs. The current minimum requirement, overall Band 5, is categorized by IELTS as a “Modest” user of the language. This is defined as follows: *The Modest user has partial command of the language, coping*

with overall meaning in most situations, though is likely to make many mistakes. Should be able to handle basic communication in own field (IELTS, 2013). From the perspective of IELTS, no academic courses should be attempted by anyone who does not have at least IELTS Band 6.5. Similarly, IELTS claims that no training courses should be attempted by anyone scoring below IELTS Band 5.5, and even then the courses should be less linguistically demanding. It is possible that these rigorous guidelines may be directed primarily at institutions in English-speaking countries where instructional accommodations for language learners are few. If so, UAE colleges might contend that the linguistic or academic demands of their programs were lower so a Band 5 is adequate. Nevertheless, the exact nature of the linguistic demands in the EMI programs of UAEU are poorly understood (Rogier, 2012). The education students that I teach appear to struggle to read 20 pages a week for a standard course, especially when more technical or abstract language is used (e.g., theoretical discussions of language learning). Belhiah and Elhami aptly described this dilemma in their extensive survey project in which they found that, “The current EMI situation leaves much to be desired with students struggling to learn the subject matter due to their low-proficiency in English” (2014, p. 1).

7 Argument-Based Test Validity

Although the IELTS is widely-recognized as a valid test for the purpose of college admissions, traditional notions of validity as a characteristic of a test itself have been recently revised by many in the field of language testing. Leading figures in this area have proposed that test validity needs to be seen as encompassing all phases of the testing process, including the interpretation of scores and the consequences of subsequent decisions based on those scores (Bachman & Palmer, 2010; Weir, 2005). In the traditional model, reaching a minimum score in IELTS represented a certain level of English proficiency that was sufficient evidence for acceptance and rejection decisions. Argument-based approaches, however, call for additional types of direct and indirect evidence to confirm the soundness of these interpretations and decisions. For example, Weir’s socio-cognitive framework divides such validity investigations into two main areas (Weir, 2005). The first, context-based validity concerns the correspondence of the test items and tasks to the larger domain of target language use (*viz.*, undergraduate EMI). The second, theory-based validity, deals with the soundness of the linguistic and cognitive processing that students engage in. For example, rapid or expeditious reading is a major challenge for low-level candidates on the IELTS exam, but the importance of this kind of reading in current theories of reading and the EMI context of UAEU is unclear.

8 Validity Threats

A good test must have an appropriate mix of tasks and items so that valid decisions can be made based on scores from the tests. Testing experts identify two main threats to test validity: the inclusion of items assessing irrelevant knowledge and skills (construct-irrelevant variance), and the inadequate use of appropriate tasks and items (construct underrepresentation) (Fulcher & Davidson, 2007). A possible example of the first type of threat in the UAEU context is the fact that the weight given to speaking in the IELTS score formula is equal to the other three language skills. Speaking was rarely assessed at all in the FP before IELTS was introduced in 2004, presumably because the other three skills were much more highly valued for academic purposes. Relatively high speaking proficiency scores frequently raise the overall IELTS band scores for some students with low academic literacy. The IELTS website reports that UAE candidates do much better as a national group on the speaking section than the other sections: the speaking mean score for the country is 5.3 but the other skills have means of 4.6, 4.7 and 4.8 for listening, writing, and reading respectively (IELTS, 2012). Orality is highly valued in Arab culture and conversational English predominates in many cities in the UAE (O'Sullivan, 2004; Troudi, 2009). Therefore, many young people have opportunities and incentives to develop their speaking, with the possible exception of females from rural areas. Consequently, students with good speaking skills seem to benefit on IELTS exams in ways that may not be consistent with the FP emphasis on academic literacy. Furthermore, the spontaneous type of speaking assessed in the IELTS interviews seems relatively rare in EMI contexts. Academic discussions and presentations usually occur in contexts where participants have had many chances to organize their speech relative to course readings and topics.

Of course, ignoring speaking does not agree with current conceptions of communicative competence, but there are sound reasons for proposing that the underlying knowledge and skills it requires are quite different in nature than the other three skills. Gu (2014) analyzed TOEFL iBT scores using structural equation modeling and found that two latent components accounted for the results: the ability to speak and the other three skills. Including speaking scores in a language test may interfere with measurement of the second, and arguably more important in the present context, latent variable. Although effective speaking is highly valued in fields like business, giving it equal weight with the other three skills appears inappropriate for academic discourse in the UAEU context. Other UAE public institutions (e.g., Zayed University) have partially addressed this issue by specifying minimum skill area scores for the IELTS and insisting that none of the four Band scores fall below 4.5. This policy is an appropriate way to prevent students who are very strong in one skill (usually speaking) to compensate for severe weaknesses in other areas. However, the expense and inconvenience of testing speaking needs to be reconsidered if speaking scores contribute relatively little to the measurement of academic language proficiency.

A second possible example of construct irrelevance in the IELTS is the inclusion of very difficult reading items that are far beyond the levels of UAEU students. These items are relevant to the broad claims IELTS is able to make about very high levels of proficiency, but they are a source of frustration to UAEU candidates since they encourage students to resort to various guessing techniques. The IELTS website states that a Band 5 in reading on the Academic IELTS is normally achieved by getting at least 15 out of 40 questions correct (IELTS, n.d.). UAEU passers can easily succeed by only answering thirteen items correctly to get Band 4.5 in reading. Of those thirteen, it seems plausible that some guessing techniques can help students succeed on five or six of the items, especially the ones that only have three possible choices (i.e., True/False/No Information). In other words, low-level IELTS candidates may only appropriately demonstrate their actual reading ability by answering about ten out of forty questions correctly. Even though the question formats on IELTS usually minimize the effects of guessing, relying on such a small number of items to assess reading is obviously unwise considering its importance in the EMI context. The skills and strategies needed to improve the accuracy of one's guesses on the difficult IELTS items probably have little to do with normal academic reading processes (Haladyna & Downing, 2004). Nevertheless, successful use of guessing strategies could have a major impact on one's score when the proportion of difficult questions is so large. The widespread popularity of private IELTS language institutes in the UAE indicates that certain kinds of testwiseness are highly valued here; several of my students have told me that guessing was a key to their success. The key issue in this context is not the use of guessing methods per se but the fact that they probably have a disproportionate effect on results given the very small proportion of items that candidates need to answer correctly.

Another possible threat to the context-validity of the reading exam is the type of texts it uses. IELTS reading passages are characterized by a lack of titles, section headings or graphic and typographic aids that are considered essential for top-down processing in reading comprehension. This demanding kind of plain text allows testers to use many item formats which assess students' ability to identify main ideas and text structure. However, such reading passages are the direct opposite of the user-friendly formats which are so common in popular textbooks. Learning to refer to common text features is considered an essential literacy skill for native speakers because they make the reading process more efficient and effective when reading to learn is the focus. Therefore, the ability to read plain text passages may be viewed as a form of construct-irrelevant variance in the UAEU context.

A similar example is the forms of background knowledge needed to succeed on IELTS reading and writing tasks. On one hand, good background knowledge is considered an essential component of literacy so reading comprehension suffers when students lack basic knowledge of major historical, social and scientific trends. On the other hand, certain topics and tasks seem to present unusual challenges for the UAE population because they reflect issues and concerns that are rare in this society. In her study of UAE students sitting for the IELTS reading test, Freimuth (2014) found that the cultural themes related to religion and social values seemed to

interfere with students' reading comprehension. English is increasingly being viewed as an international language in the UAE so the importance of using background knowledge from Western cultures is questionable. The Middle East is an important market for the IELTS program; nevertheless, Arab students have different sorts of background knowledge than students from Asia and elsewhere. Fulcher and Davidson (2007) consider this kind of differential item functioning to be a source of construct-irrelevant variance, and the ability of IELTS to account for this while serving the needs of many different regions of the world is questionable.

9 Construct Underrepresentation

In the case of low-level reading and writing on the IELTS exam, separating the two types of validity threats is not that easy. As mentioned in the previous example, plain text reading passages are clearly very rare in college texts but the processes involved in reading them are somewhat germane to some kinds of academic reading. Similarly, the use of many difficult, high-level reading tasks on the test is not an obvious error, but it is associated with a corresponding reduction in the number of items that are accessible to students with lower comprehension levels. The IELTS reading test would obviously serve the UAE population better if it did not underrepresent lower-level tasks and items. It is very likely that IELTS relies mainly on challenging skills (e.g., reading for inferences) to assess those above Band 7. This seems very appropriate, but it may result in low-level items being underrepresented on the exams. UAE professors like this author recognize the difficulties faced by students of modest ability when performing high-level reading skills. IELTS reading tasks draw on complex combinations of lexical knowledge, syntactic knowledge, discourse knowledge, etc. If they did not, the test would have to be three times longer so that each of these areas of knowledge could be individually assessed at all the relevant levels. This test design, however, leaves low-level items in a minority on the test even though UAEU professors are probably more interested in this aspect of reading skill.

The IELTS reading test is technically impressive because of its ability to discriminate at so many levels of proficiency in just 40 questions. The IELTS writing test has fewer objectionable elements than the reading test, but it also has fewer positive aspects. The main IELTS writing prompt is a simple statement eliciting an essay of 250 words that is worth about two-thirds of the writing score. Opinion essays are commonly used for this section, and students must rapidly produce concepts and language that are relevant to the given topic. The specific wording used in this simple prompt statement is of utmost importance since it sometimes contains a phrase or a concept that is unfamiliar or confusing to many UAEU students. Of course, this undermines the reliability of the test in serious ways that other researchers have recognized (e.g. Gebril, 2009). However, my major concern at this point is the way the two writing tasks underrepresent the other forms of writing that are characteristic of undergraduate EMI programs. Undergraduates are

rarely asked to write using only their personal background knowledge and linguistic knowledge. Moore and Morton (2005) investigated this issue by analyzing 155 writing assignments at Australian universities. They concluded that IELTS tasks have some resemblance to the university essay as a genre, but they are mainly non-academic in nature. The most common academic genre, writing from sources, was used as a task type in older versions of the IELTS but was abandoned in the new versions of the test. Integrated writing tasks which combine reading and writing are clearly more authentic than opinion essays when it comes to undergraduate studies.

Green (2006) compared two types of writing courses to investigate this issue: IELTS preparation courses and English for Academic Purposes courses. He found many similarities between the two, but the former usually avoided writing from sources in favor of genres demanded by the IELTS test: descriptive writing (for Task 1), and evaluation and hortation (for Task 2). He found that IELTS classes emphasized reproducing information from graphs and from memory while EAP classes were more cognitively demanding since they stressed integrating information from sources. Both types of writing can be challenging for low-level students, but the absence of writing from sources is a troubling form of construct underrepresentation in the case of IELTS.

10 Validity Research

To their credit, IELTS has sponsored many studies examining the validity of the exam, most of which lend support to its common usage: selecting high-intermediate level and advanced students for western universities. Nevertheless, the amount of validity evidence that pertains to its specific applications in the UAE is relatively small. It is necessary to collect validity evidence in order to confirm the suitability of a test for a specific purpose, and Davies (2011) emphasizes that this is the responsibility of test users, not researchers or test developers. As informed test users, UAEU administrators realize how weak a Band 5 candidate can be, but they have done relatively little to investigate the effect of poor English skills on the classroom discourse and academic achievement of their EMI students.

This author found only one validity study that specifically addressed the reading processes of lower level candidates. Weir, Hawkey, Green, and Devi (2012) assessed the cognitive reading processes used by 352 subjects by using verbal retrospective techniques. Due to the unusually large number of subjects involved, the researchers were able to validly compare weaker readers (Band 5 and below) with intermediate and stronger ones. The results generally support IELTS validity by noting many more similarities than differences between reading processes and strategies reported by the Band 5 candidates and the other groups of candidates. For example, all groups claimed that their most common strategy was: *quickly match words that appeared in the question with similar or related words in the text*. More evidence like this is needed to ensure that low proficiency candidates in the UAE

are being assessed appropriately. Weir (2005) outlines several useful techniques for collecting such data. They include detailed examinations of the reading and writing tasks required by universities in order to establish their similarities to test tasks. At UAEU, humanities students are often seen carrying around thick anthologies of Western literature which seem far beyond their normal reading levels. The specific reading tasks associated with such texts need to be fully analyzed in order to determine how well they correspond to IELTS reading tasks.

11 Hijacking the Curriculum?

Before the IELTS was introduced in 2004, getting ready for undergraduate EMI was the obvious mandate of the FP at UAEU. Coordinators and instructors consulted occasionally with others (especially colleagues in the English for Academic Purposes program) about the linguistic and academic needs of students, and they also relied on their professional judgment when planning courses and tests. Unfortunately, these local curriculum development processes began declining in importance after 2004 when educational progress was largely defined in terms of IELTS scores. Preparing for the demands of the IELTS tasks became a major focus of the final level of English in the FP, and the normal coursebook was replaced by an IELTS preparation book called *Focus on IELTS Foundations* (O'Connell, 2007). Higher-level secondary students who only need one term in the FP may be ready for this approach, but weaker students probably suffered when test preparation became the focus of classes rather than dealing with the obvious deficits in their global English development. From an educational perspective, it may be plausible that some IELTS preparation activities can be integrated with normal language development activities. Nevertheless, there are far more positive ways to achieve balanced language development using methods that are consistent with best practices in language education, academic literacy training, and formative types of holistic assessment (Pilgreen, 2007). The IELTS exam seems poised to become a key assessment tool for English programs in secondary schools across the UAE but evidence suggests that an emphasis on learning test-taking strategies for the exam may divert students' attention away from balanced language development and preparation for EMI (Gitsaki et al., 2014).

The FP at UAEU attempts to serve underprepared, at-risk students while at the same time challenging those who are ready for some fundamental aspects of academic literacy. Rather than using a test like the IELTS as the arbiter of success, educators need to analyze the exact nature of EMI practices and tasks in various departments and programs of UAEU. In the last few years, I have served on two university-wide committees addressing the challenges of helping undergraduates to continue developing their English in the mandatory English for Academic Purposes courses. Both groups concluded that academic literacy was the major need of most students, and one made a policy shift to replace a module on oral presentations with an increased emphasis on academic writing. Although faculty members from

several departments participated in these committees, some resisted the idea that professors were responsible for helping students improve their English. Only a few from the humanities (e.g., literature) have consistently recognized that ongoing language development is a key aspect of their educational mission.

12 Parallel and Converging Assessment Systems

Since its implementation in 2004, the IELTS has had a major effect on how English is taught and tested in the FP. When it was first adopted, FP exams and IELTS exams seemed to complement each other since the former stressed achievement while the latter assessed proficiency. In recent years, however, FP tests have begun imitating many features of the IELTS test to the extent that the current FP writing exam is a virtual copy of the IELTS writing exam. The FP reading exam still has sections for grammar and vocabulary that represent achievement more than proficiency. Nevertheless, it has become difficult for FP instructors to see their role as extending much beyond that of serving as IELTS test preparation specialists. As educators with graduate degrees, they are qualified to help students develop a wide repertoire of educational skills and strategies. However, their current program narrows that focus almost exclusively to IELTS exam performance. Since instructors in the highest level of the English FP are responsible for both IELTS preparation and course exam preparation, I surveyed them to determine their views of the two types of assessment instruments. Most of the instructors had been hired in the past decade and had only known the dual exam system currently used by the FP. Others, however, recognized that the FP had enjoyed more freedom in the past to teach and test in ways that were more consistent with the needs of future EMI students.

13 Reading Teachers' Views of IELTS

I first surveyed reading teachers to determine their global attitudes toward the IELTS exam as an appropriate assessment tool given the normal demands of the Level 3 curriculum, the instructional needs of the students and the future demands of EMI. These frequencies appear in Table 1. Twelve out of 15 of the reading teachers in Level 3 responded to the survey. Their responses revealed their mixed views of the usefulness and appropriateness of the IELTS reading test and its washback. While 50 % endorsed a statement regarding the alignment of the Level 3 curriculum and the IELTS reading exam, more than half disagreed with a statement about the relevance of the IELTS reading texts and tasks to the future academic work of their students. A total of 42 % of respondents felt that preparing for the IELTS reading test had positive effects on their students but 25 % disagreed with that statement.

Table 1 Teachers’ global opinions about IELTS reading

	Strongly agree (%)	Agree (%)	Neutral (%)	Disagree (%)	Strongly disagree (%)
The IELTS requirements correspond closely to the stated course objectives of Level 3 with regard to reading	8	42	42	0	8
Reading instruction in Level 3 is heavily influenced by the types of tests used to assess reading	25	75	0	0	0
The IELTS reading exam seems relevant to the types of reading texts and tasks my students will face in their future studies	8	33	0	42	17
Overall, activities to prepare for the IELTS reading exam have a positive effect on my students	0	42	33	25	0

In order to investigate their views of specific aspects of the reading exam, I used a seven-point Likert scale in which the extreme points on the scale were labelled either as “Very Appropriate” or “Very Inappropriate” (see Table 2). Respondents were relatively positive about the topics covered on the IELTS reading section, with 40 % indicating that they were appropriate. They were evenly divided, however, concerning the effect of the time constraints on the exam: 40 % felt they were appropriate and 32 % judged them as inappropriate. With regard to the appropriacy of the linguistic levels of the test, the opinions were more strongly critical of the IELTS. Fifty-eight percent judged them to be inappropriate while only 33 % considered them relatively appropriate.

Exam preparation is such a major aspect of the Level 3 course that it is essential that these preparation activities have their own language learning benefits beyond mere testwiseness. To examine teachers’ views about this, I asked respondents to compare the instructional usefulness of IELTS activities to conventional Level 3 test preparation activities (see Table 3). The teachers showed a clear preference for

Table 2 Teachers’ views of specific aspects of the IELTS reading exam

	Very appropriate	⋈	<	Neutral	>	⋈	Very inappropriate	Weighted averages
Weighting	+3	+2	+1	0	-1	-2	-3	
Topics covered	0 %	17 %	33 %	33 %	8 %	8 %	0 %	0.43
Time constraints	0 %	17 %	33 %	8 %	25 %	17 %	0 %	0.08
Linguistic levels	0 %	25 %	8 %	8 %	33 %	25 %	0 %	-0.25

Table 3 Teachers' comparisons of the usefulness of IELTS and Level 3 reading exams

	IELTS is far more useful	≪	<	Equal	>	≫	Level 3 is far more useful	Weighted averages
Weighting	+3	+2	+1	0	-1	-2	-3	
Student engagement/motivation	17 %	8 %	17 %	8 %	17 %	25 %	8 %	-0.07
Vocabulary learning	0 %	17 %	0 %	25 %	25 %	17 %	17 %	-0.76
Discourse awareness	0 %	0 %	25 %	8 %	33 %	17 %	17 %	-0.93
Thinking skills	8 %	0 %	8 %	42 %	0 %	25 %	17 %	-0.69
Preparation for the faculties	0 %	9 %	9 %	27 %	18 %	27 %	9 %	-0.72

the usefulness of Level 3 activities except concerning the area of “Student Engagement/Motivation”. The demands of IELTS readings seem to motivate students in ways that are useful even though their instructional value is unclear. Opinions were evenly divided between those who considered the IELTS more engaging and those who considered the Level 3 assessment tools more engaging.

14 Writing Teachers' Views of IELTS

Similar online surveys were completed by writing teachers in the Level 3 program and they revealed a somewhat more positive attitude toward the IELTS writing exam than the reading teachers had toward the reading exam. Nine out of 20 writing teachers responded to the survey and they saw general alignment between Level 3 and IELTS (see Table 4). Almost all the writing teachers (89 %) recognized a direct correspondence between the demands of the IELTS writing test and the Level 3 writing curriculum. Very few writing teachers had negative attitudes toward either the relevance of IELTS writing to university writing or the effect that IELTS writing had on their students. When asked to judge the specific aspects of the IELTS writing exam, teachers were relatively positive about the Expected Grammar Levels and the Expected Vocabulary Levels (see Table 5). Their opinions about the appropriacy of the Topics and the Time Constraints were very mixed, however. On the whole, these opinions were still slightly positive, but a significant number of teachers thought the writing exam was inappropriate in these two respects.

Although the teachers were slightly positive overall about the IELTS writing exam, they did not compare it favorably to the Level 3 writing assessment tasks (see Table 6). When asked about specific aspects of the instructional usefulness of both exams, a majority of respondents judged them to be equal. However, a few teachers clearly favored the Level 3 writing tests, especially with regard to the development of thinking skills and sentence structure. The only aspect of the IELTS that was

Table 4 Teachers’ global opinions about the IELTS writing exam

	Strongly agree (%)	Agree (%)	Neutral (%)	Disagree (%)	Strongly disagree (%)
The IELTS requirements correspond closely to the stated course objectives of Level 3 with regard to writing	33	56	0	0	11
Writing instruction in Level 3 is heavily influenced by the types of tests used to assess writing	78	22	0	0	0
The IELTS writing exam seems relevant to the types of writing texts and tasks my students will face in their future studies	22	33	22	0	22
Overall, activities to prepare for the IELTS writing exam have a positive effect on my students	25	50	13	13	0

Table 5 Teachers’ views about specific aspects of the IELTS reading exam

	Very appropriate	≪	<	Neutral	>	≫	Very inappropriate	Weighted averages
Weighting	+3	+2	+1	0	-1	-2	-3	
Expected vocabulary levels	11 %	22 %	33 %	22 %	11 %	0 %	0 %	0.99
Expected grammar levels	12 %	25 %	25 %	25 %	13 %	0 %	0 %	0.98
Question formats	11 %	11 %	33 %	33 %	0 %	11 %	0 %	0.66
Topics covered	11 %	11 %	22 %	22 %	33 %	0 %	0 %	0.44
Time constraints	22 %	11 %	11 %	22 %	11 %	22 %	0 %	0.44

clearly more useful for teachers than the Level 3 test was that of “Student Motivation/Engagement”. It appears that the challenging aspects of the IELTS reading and writing tests motivate students in some positive ways.

When asked to comment on the advantages and disadvantages of the writing test, teachers had more negative comments than positive ones. For example, a few teachers questioned how authentic the writing test was in light of the faculty requirements. Several complained that much of their course was spent teaching to the test, and the improvements students made were too superficial. For example, one commented, “It’s more about making a band 4 writer look like they are actually better than genuinely improving their abilities as a writer of English”.

Table 6 Teachers' comparisons of the usefulness of IELTS and Level 3 writing exams

	IELTS is far more useful	≤	<	Equal	>	≥	Level 3 is far more useful	Weighted averages
Weighting	+3	+2	+1	0	-1	-2	-3	
Student engagement/motivation	22 %	11 %	0 %	56 %	11 %	0 %	0 %	0.77
Discourse awareness	0 %	11 %	0 %	67 %	11 %	0 %	11 %	-0.22
Preparation for the faculties	0 %	0 %	0 %	89 %	0 %	0 %	11 %	-0.33
Sentence structure	0 %	0 %	0 %	78 %	11 %	0 %	11 %	-0.44
Thinking skills	0 %	11 %	0 %	56 %	11 %	11 %	11 %	-0.44

In many respects, these findings were similar to those of Lewthwaite's study of IELTS writing washback which was also conducted in the FP at UAEU (Lewthwaite, 2007). He found very positive impressions of IELTS washback from both teachers and students in the English FP. The weakest area of congruence he found was similar to the one identified here: the relevance and usefulness of Task 2 for academic language learning and preparation for the faculties. In his study, many students were neutral on this issue and only 6 out of 16 teachers strongly endorsed his statements to that effect. His data were collected when the IELTS had been in place for just 3 years; since then, based on the present findings, the Level 3 course seems to have yielded to more test-driven pressure.

15 Conclusion

Thus far, there is little conclusive evidence that adopting the IELTS testing system has actually raised levels of achievement in the FP or UAEU in general. The IELTS exam appears to motivate students to study more and try harder in comparison to the traditional FP tests because of its unusual level of difficulty. Nevertheless, it seems quite unsuitable for normal language development activities given the nature of its advanced reading texts and its difficult time constraints. The IELTS writing exam may have fewer negative aspects than the reading exam, but its positive features are not clear in comparison to locally developed tests. The most serious problem with the writing test, in my opinion, is the uneven quality of the prompts for the UAE contexts. There is no doubt in my mind that a locally-produced writing exam could emphasize tasks and questions that are specifically tailored to the needs of UAE students and include relevant rhetorical forms, appropriate types of scaffolding, and suitable background knowledge requirements. Helping students prepare for crucial reading and writing tests should be good opportunities for

ongoing language development and academic study, not just ways to prove what a student has already achieved.

15.1 Pedagogical Implications and Future Trends: An Agenda for Assessment

Even though test-driven schooling is the norm in the Gulf region, there is no reason why good tests cannot accompany high-quality teaching and learning. There are several excellent tests available for students at the Band 5 level that could be used for UAEU students in ways that would probably be more valid and reliable than the IELTS exam. UAEU and UAE secondary school systems should be more realistic about their assessment needs and consider an exam system like the Cambridge First Certificate in English (FCE) if they want an external program. The FCE targets the CEFR B2 Level that corresponds to IELTS Band 5 through Band 6.5. Its linguistic content appears far more accessible than the IELTS, so teachers will be able to combine useful language practice with test preparation in ways that are rarely possible with an IELTS reading exam. Low-level students have a fundamental need for large amounts of relatively accessible input, which Nation (2007) calls “meaning-focused input”. This need is much greater than that of devising coping and guessing strategies for material which is far beyond the testees’ levels. Although the UAE has a large pool of expertise in language testing, they may prefer to coordinate with international testing organizations when creating a national test for university admissions. Alternatively, the UAE can look towards standardized, locally-produced English tests such as are being developed in nations like Italy and Mexico, where international experts are collaborating with local institutions in order to ‘co-certify’ tests and ensure their quality (Newbold, 2012).

Although external exams bring a certain amount of prestige, a more preferable approach on many levels would be for the nation to develop its own entrance exams that reflect the distinctive realities of both secondary schools and tertiary educational processes in this unique bilingual setting. The UAE Ministry of Education has been successfully producing and administering a testing system known as the CEPA exam which is a reliable way to place secondary students in FP English programs. Producing a similar customized test to replace IELTS is not a difficult task considering the fact that all the public universities accept Band 5 as the key qualification for EMI. A test that focuses on one or two key proficiency levels is much easier to produce than a test that covers nine levels. Rather than testing all students with time-consuming speaking and writing tests, the UAE may wish to adopt a more economical approach that divides assessment into two or three stages. For example, results from initial tests of listening and reading could be used to identify three groups: those who are ineligible for higher education, those who need further tests (e.g., writing), and those who could be accepted into higher education directly based solely on their superior levels of listening and reading. If institutions

like UAEU began setting minimum scores in specific skill areas based on analyses of EMI tasks, a test would be needed which allows candidates to retake only those sections in which they were unsuccessful. Ideally, such a test would offer results that would specify the components of students' performance. In the case of reading, it would be useful for students to know how they performed on skill areas like reading for main ideas, reading for details, and reading for inferences. The current IELTS system offers none of this flexibility or diagnostic information and it is not addressed to the local circumstances of the UAE. The great expense and time associated with sitting for an exam like the IELTS needs to be reduced so that many students can monitor their progress through high school, the FPs, and beyond.

Australian universities seem to be leading the way when it comes to recognizing the limitations of entrance test results when addressing the continuous language development needs of non-English speaking students. For example, Dunworth (2010) claims that universities put far too much faith in results like IELTS scores and neglect the realities of academic discourse. She calls for an "institutional process to link the measures that universities accept (on entrance exams) to the lived experience of the tertiary classroom" (Dunworth, 2010, p. 6). To some extent, institutions like the Higher Colleges of Technology in the UAE have done this by integrating language development into normal college courses and using the IELTS at the end (rather than the beginning) of the undergraduate program. Other universities should follow similar approaches since ongoing language development is vital to students' success. Ironically, very few of the graduates of my own program can meet Abu Dhabi's new requirement of IELTS Band 6.5 for government primary teachers because our curriculum does not specifically address English training. Instructional solutions are urgently required in programs like my own, but appropriate and flexible assessment systems will be needed to accompany and support them.

References

- Bachman, L. F., & Palmer, A. S. (2010). *Language assessment in practice*. Oxford, UK: Oxford University Press.
- Belhiah, H., & Elhami, M. (2014). English as a medium of instruction in the Gulf: When students and teachers speak. *Language Policy*, 14(1), 3–23.
- Davies, A. (2011). Kane, validity and soundness. *Language Testing*, 29(1), 37–42.
- Doiz, A., Lasagabaster, D., & Sierra, J. M. (Eds.). (2013). *English-medium instruction at universities: Global challenges*. Clevedon, England: Multilingual Matters.
- Dunworth, K. (2010). Clothing the emperor: Addressing the issue of English language proficiency in Australian universities. *The Australian Universities' Review*, 52(2), 5–10.
- Freimuth, H. (2014). *Cultural bias on the IELTS examination* (Ph.D. Thesis). Rhodes University, Grahamstown, South Africa.
- Fulcher, G., & Davidson, F. (2007). *Language testing and assessment*. London, NY: Routledge.
- Gebriel, A. (2009). Score generalizability of academic writing tasks: Does one test method fit it all? *Language Testing*, 26(4), 507–531.
- Gitsaki, C., Robby, M., & Bourini, A. (2014). Preparing Emirati students to meet the English language requirements for higher education: A pilot study. *Education, Business and Society. Contemporary Middle Eastern Issues*, 7(2/3), 167–184.

- Green, A. (2006). Watching for washback: Observing the influence of the international English language testing system academic writing test in the classroom. *Language Assessment Quarterly*, 3, 333–368.
- Gu, L. (2014). At the interface between language testing and second language acquisition: Language ability and context of learning. *Language Testing*, 31(1), 111–133.
- Haladyna, T. M., & Downing, S. M. (2004). Construct-irrelevant variance in high-stakes testing. *Educational Measurement: Issues and Practice*, 23(1), 17–27.
- IELTS. (n.d.). *How is IELTS marked?* Retrieved from http://www.ielts.org/institutions/test_format_and_results/how_is_ielts_marked.aspx.
- IELTS. (2012). *Test taker performance*. Retrieved from <http://www.ielts.org/researchers/analysis-of-test-data/test-taker-performance-2012.aspx>.
- IELTS. (2013). *IELTS guide for educational institutions, governments, professional bodies and commercial organisations*. Retrieved from http://www.ielts.org/PDF/Guide_Edu-%20Inst_Gov_2013.pdf.
- Lewthwaite, M. (2007). Teacher and student attitudes to IELTS writing tasks: Positive or negative washback? *UGRU Journal*, 5, 1–16.
- McLaren, P. (2011). *English medium in the United Arab Emirates: Serving local or global needs?* Unpublished doctoral dissertation, (Ed.D.). University of Exeter, Exeter, UK.
- McLean, M., Murdoch-Eaton, D., & Shaban, S. (2013). Poor English language proficiency hinders generic skills development: A qualitative study of the perspectives of first-year medical students. *Journal of Further and Higher Education*, 37, 462–481.
- Moore, T., & Morton, J. (2005). Dimensions of difference: A comparison of university writing and IELTS writing. *Journal of English for Academic Purposes*, 4, 43–66.
- Morrow, C. (2005). *An analysis of the course sequence in UGRU English* (Unpublished Report). Al Ain, UAE: Undergraduate General Requirements Unit, United Arab Emirates University.
- Moussly, R. (2012, May 27). Majority of Emirati students are unprepared for university. *Gulf News* (United Arab Emirates). Retrieved from <http://gulfnews.com/news/uae/education/majority-not-prepared-for-university-1.690616>.
- Nation, P. (2007). The four strands. *Innovation in Language Learning and Teaching*, 1, 1–12.
- Newbold, D. (2012). Global institution, local examination. In D. Tsagari & I. Csépes (Eds.), *Language testing and evaluation, Volume 26: Collaboration in language testing and assessment* (pp. 127–142). Frankfurt am Main: Peter Lang AG.
- O’Connell, S. (2007). *Focus on IELTS foundation*. London: Pearson Longman.
- O’Sullivan, A. (2004, September). Reading and Arab college students: Issues in the United Arab Emirates’ Higher Colleges of Technology. In *Proceedings of the First International Online Conference on Second and Foreign Language Teaching and Research: Beyond Borders*. Symposium conducted at First International Online Conference on Second and Foreign Language Teaching and Research.
- Pilgreen, J. (2007). Teaching the language of school to secondary English learners. In J. Lewis & G. Moorman (Eds.), *Adolescent literacy instruction: Policies and promising practices* (pp. 238–262). Newark, DE: International Reading Association.
- Rogier, D. (2012). The effects of English-medium instruction on language proficiency of students enrolled in higher education in the UAE (Ed.D.). Exeter, UK. Retrieved from <https://ore.exeter.ac.uk/repository/handle/10036/4482>.
- Troudi, S. (2009). The Effects of English as a medium of instruction on Arabic as a language of science and academia. In P. Wachob (Ed.), *Power in the EFL classroom: Critical pedagogy in the Middle East*. Newcastle upon Tyne: Cambridge Scholars Publishing.
- Weir, C. (2005). *Language testing and validation*. Houndmills, Basingstoke: Palgrave Macmillan.
- Weir, C., Hawkey, R., Green, A., & Devi, S. (2012). The cognitive processes underlying the academic reading construct as measured by IELTS. In L. Taylor & C. Weir (Eds.), *IELTS collected papers 2: Research in reading and listening assessment* (Vol. 2, pp. 157–189). New York: Cambridge University Press.