Toward an English Proficiency Test for Postgraduates in Malaysia

Saidatul Akmar Zainal Abidin\(^1\) and Asiah Jamil\(^1\)

**Abstract**

Malaysia is fast becoming a major attraction for candidates from all over the world to pursue their higher education. Currently students (local and international) who pursue postgraduate (hereafter, PG) education in Malaysia use the Test of English as a Foreign Language (TOEFL) or International English Language Testing System (IELTS) scores as indicators of their English ability. These are tests from the United States and the United Kingdom, respectively, tailor-made for university education in those countries. Recent literature in testing and evaluation describes the need for more localized tests, developed for the “local” context of a particular country. Thus, the need for a test that could be utilized and customized to the needs of the students studying in Malaysia is foreseeable. This is in line with the concept of test localization. It stipulates that for a test to be valid, its design and development must take into consideration the population, context, and the domain in which the test is used. A project was undertaken where a new English test named Graduate Admission Test of English (GATE) was developed for PG admission into universities in Malaysia. This article describes the process of developing a new test that measures English language competency of PG students who intend to pursue their studies in Malaysia. It includes the use of a test specification/blueprint that contains validity elements adopted from a test validation framework developed by Weir. The article emphasizes the rigor of developing such a test, which includes aspects of test development, operation, analysis, and validation.

**Keywords**

postgraduate education, test localization, high-stakes tests, GATE, CEFR, English language competence.

**Introduction**

This research was undertaken with the hope of improving the postgraduate (PG) intakes at local universities in Malaysia in terms of candidates’ language ability, which would in turn enhance skills and competencies required to succeed in the respective programs. At present, candidates who intend to take up a PG course in the country are required to obtain a score from tests such as the Test of English as a Foreign Language (TOEFL) and International English Language Testing System (IELTS), which shows their English language ability at entry point into a university. In fact, since 2010, even local candidates were subject to the same English language requirement. However, there is evidence to show a mismatch between a test score and a candidate’s actual writing and speaking performance in the PG classrooms, albeit reports from foundation courses offered at various universities for foreign candidates who do not make the mark, and others who are exempted from such courses.

In their article on processing strategies in reading, Ponniah and Tay (1992) stated that in Malaysian tertiary institutions, it is essential for students to have a near-native English competency in reading to be able to comprehend/read academic texts in various disciplines. This will subsequently produce graduates who are competent and able to demonstrate the required skills at the workplace as well as for other endeavors.

One of the main reasons for the development of a new test for PGs is the influx of foreign nationals into Malaysia, mostly for higher education. Their lack of English language ability to pursue PG education and the requirement to secure an IELTS or TOEFL score to enter universities in Malaysia are contributing factors. In 2008, the Council of Deans of Post Graduate Studies, Science University of Malaysia (USM; minutes of meeting MDPS17/18/19/20 2008) was of the opinion that it is time to introduce an English test developed in Malaysia for these international candidates, just as the Malaysian University English Test (MUET) is a measure of Malaysian students’ proficiency at the undergraduate level. As reported in a research article by Juliana Othman and Nordin (2013), it has been argued that to cope with the linguistic demands of the courses, a certain level of proficiency in the English language is required. In most...
if not all applications for universities, jobs, and even government agencies, the English language is obligatory; indeed, even Malaysian graduates have at times demonstrated below-average levels of English in university tasks and on tests such as the MUET, IELTS, and TOEFL.

Along these lines, if candidates need to take IELTS or TOEFL to study in the United Kingdom and the United States respectively, it is only appropriate that they take Graduate Admission Test of English (GATE) to study in Malaysia. The requirement of taking a foreign test as opposed to a local test to study locally in itself warrants justification. This is so because IELTS and TOEFL were designed for those who want to pursue their studies in the English-speaking countries. Based on the justification provided earlier on the need for a localized test, no one test or two tests in this context fit all. If all elements and principles of testing can be incorporated in the development of a localized test, that should provide a sound basis for its use locally.

The proposal for developing a localized test is inevitable as it can only be beneficial for those who intend to pursue higher education in Malaysia, and where their English language abilities require filtering. This is to prevent the mismatch between test scores and true performance of candidates in the classroom, and cases where even Malaysians take a foreign test to study in Malaysia.

Thus, the significance of this project will be evident in terms of saving cost for many candidates and the universities, encouraging candidates to enhance their English language ability, the potential for universities to venture into the testing field, and more importantly, achieving standardization of PG intake and quality among Malaysian universities.

To develop a test that would be acceptable, it has to have the qualities of validity, reliability, test usefulness, and test fairness. To produce a test that has value, the test development process needs to be multifaceted and rigorous, and incorporate many parties. Indeed, experience in the current project has shown that these test qualities are not achievable when considerable time, resources, administrative support, and budget are not made available.

Objective

The main objective of the new English test for postgraduates is to ascertain their ability in the English language to be able to cope academically in universities in Malaysia. Because the literature shows that Malaysia is currently a favorite destination for higher education, and the number of students entering the country for this purpose has increased drastically, it is essential that the language level be ensured before a candidate can continue to study at PG level. In addition, this is more than ever vital seeing that the mode of instruction in most universities (public and private) in Malaysia is English. With this objective in mind, the thrust of this article is a description of the process involved in developing the proposed test of English for PG purpose.

Literature Review

This section provides a brief overview of test processes and the comparison of some major tests in the literature.

Test development begins with the concern that a test can be shown to produce scores that are an accurate reflection of a candidate’s ability in a particular area, such as reading for specific ideas, writing a research proposal, breadth of vocabulary knowledge, or speaking in a class presentation. We need to understand the trait (underlying construct) we wish to measure and the method (instruments we need to develop) we would use to provide us with the information about these constructs (Weir, 2005). In addition, a key part of this process is test validation in which evidence is gathered to corroborate the inferences we make regarding these traits from test scores obtained. Finally, testing also has an ethical dimension, which many in the testing field have referred to as consequential validity. This is the impact that the test has on individuals, institutions, and society as a whole, that is, the stakeholders (IELTS Handbook, 2007).

In the case of the proposed new English test for PG students, an early analysis of the literature found similarities between the MUET and IELTS. These similarities are presented in the comparison table (refer to Table 1).

Because there are close similarities between the MUET and the IELTS in terms of exam type, language components, and marking on a band scale, a new test should be fashioned after these tests, targeted at PG level. However, major changes will take place in terms of topics, levels of difficulty of tasks, coverage, and more importantly, taking into consideration the local context and test domain. Like the IELTS and TOEFL, which are developed by major exam boards, the MUET by the Malaysian Examination Council (MPM), the proposed new test would be developed by a group of academic professionals, and supported by the university for its operations. The introduction of such a test would eventually involve more parties (such as the Malaysian Examination Council and the Ministry of Higher Education) and could be implemented on a wider scale throughout the country and beyond.

The models below outline the test development process by some major exam boards (Cambridge English for Speakers of Other Languages [ESOL] Test Development Cycle, 2002, and Pearson Test Development Process, 2007). These models are used as references for test development in the present study (Figures 1 and 2).

It is evident that in both models, the processes include test design and development, test operations, data analysis, and an evaluation or validation, although these are labeled differently in each model.

In the context of the present study, the development process follows a similar outline as presented in Figure 3. In addition, there is an attempt at placing candidates according to levels in the Common European Framework of Reference (CEFR; refer to the appendix). The CEFR is a framework used widely in language programs, curriculum, schools, and other language contexts.
practitioners in Europe (see Cambridge ESOL Main Suite exams, 2002; Pearson Test of English, 2007; Saville, 2006, for various uses of the CEFR). It describes in a comprehensive way what language learners have to learn to use a language for communication, and what knowledge and skills they have to develop to be able to perform effectively. Little (2006) discussed the impact that the CEFR has had in European schools and higher institutions where it was used as a major reference for developing courses and tests in their respective programs.

Thus, the proposed new test in the study follows a model that has two major considerations:

i. the test development process and

ii. the test levels, based on existing models of language learning and testing.

Table 1. Comparison Between IELTS and MUET.

| Objective                          | IELTS                                                                 | MUET                                                                 |
|------------------------------------|----------------------------------------------------------------------|----------------------------------------------------------------------|
| Assess a candidate’s ability to study at pre-degree or postgraduate levels in an English-speaking country | Assess a candidate’s English language ability for admission into university at pre-degree |
| Type of exam                        | Criterion-referenced                                                | Criterion-referenced                                                |
| Components                          | Reading, writing, listening, speaking                               | Reading, writing, listening, speaking                               |
| Place offered and recognized         | International                                                      | Malaysia and two universities in Singapore                           |
| Administration                      | British Council and IDP Australia                                   | Malaysian Examination Council                                        |
| Marking scheme                      | 9-band scale of 1–9                                                 | 6-band scale of 1–6                                                 |
| Mode                                | Paper and pencil                                                    | Paper and pencil                                                    |

Note. IELTS = International English Language Testing System; MUET = Malaysian University English Test; IDP = International Development Programme.

This model ensures test validity in terms of its development and reliability in terms of its scoring outcome.

The GATE (Graduate Admission Test of English)

Before a test can be developed, pertinent questions need to be put forward such as the following:

- Why must we develop a new test when there are such English tests in the market?
- What would the test contain, and what would it look like?
- Who would be developing the test?
- How will the test be graded?
- What about issues of cost, operations and administration, and security of the test?

A possible solution to address these questions is the use of a framework for test development, which can ensure a systematic and effective process, and essentially a valid test. The GATE will use a framework proposed by Weir (2005) labeled Socio-Cognitive Framework for Developing and Validating tests of reading, writing, listening, and speaking. This framework will be the backbone for the test from the initial development process, scoring, and to its final evaluation process. We need clear...
objectives, expressed in a set of terms of reference; what is the test used for, for which target group of test takers, and what structure will the test take in terms of its format, content, and levels. In addition, the new test would need to have qualities of sustainability if it is to be useful not just for its immediate purpose, but also for a longer shelf life when it can be extended to a more global level and application. For this purpose, it is proposed that the test would eventually be administered via computers, or a computer-based test (CBT).

To generate a test of this significance and scale, it has to be given a name, or branding, that reflects its purpose and impact. The proposed test, GATE, is a test that measures a candidate’s language ability and competency in that it measures beyond English language proficiency, it ascertains the candidate’s ability to demonstrate competence required to partake and succeed in PG education. As indicated by Moon and Siew (2004), the level of proficiency in the English language is a contributing factor for a better academic performance. This includes the ability “. . . to participate in scholarly discussions, to be able to defend their arguments, explain their opinions, develop hypotheses, all of the things they need to write and defend their dissertation” (Redden, 2008).

Similarly, in the context of the present study, the skills and the level of proficiency, which are required at PG level are measured to ascertain if graduates have achieved the proficiency needed in postgraduate education. Thus, it is labeled a test of language competence administered at the point of a candidate’s entry into a PG university program. The GATE consists of three sections (refer to Table 2).

The test consists of three sections. They are

i. Reading
ii. Grammar, and
iii. Writing.

In each of the sections, different skills are tested. As indicated in the table, the skills tested in the Reading section are making inferences, determining author’s purpose/point of view, determining author’s argument, and evaluating the author’s argument. In this section, students are required to respond to 18 items. The items are tested in the multiple-choice question format.

The second section, which is Grammar, is divided into two types: error correction and transformation items. The skills tested are identifying errors, making corrections, and paraphrasing sentences. Students are required to respond to 23 items in this section.

The last section is Writing, and the skills tested are summarizing a text and writing an argumentative essay. Students are required to respond to one item for each of these skills.

Development of GATE

This section discusses the stages of developing the GATE. The study utilized the test process (refer to Figure 3) as a basis of the development of the GATE. The three phases in the process are further elaborated and explained.

Phase I—Test Development

Developing the GATE. The GATE was developed based on the following principles:

1. Literature on high-stakes standardized tests such as TOEFL and IELTS, the framework for validating language tests (Weir, 2005), and the CEFR.
2. The test specifications focus on the elements of “context validity” or the test tasks or questions, such as purpose, text length, lexical and structural range, topic familiarity, and so on. The objectives of the test were spelled out according to the major sections, Reading, Writing, and Grammar, and the subsections for each component. The Listening and Speaking sections were left out as they required a lot more time
and resources to develop, and at PG level, candidates are expected to write extensively to produce a thesis. They need to be able to read extensively and intensively, be critical on their readings and discussions, and be able to summarize and paraphrase ideas that subsequently need to be organized in the thesis. In addition to the conventional methods of testing Reading and Writing, the test was creative in that a different technique for testing Grammar called “transformation items” was introduced; it tests a candidates’ ability to paraphrase sentences and very short paragraphs using appropriate and accurate language. Thus, not only did the test contain the necessary skills that PG candidates are expected to demonstrate at entry level, it also had innovative ideas for a test of this nature. In fact, the test takers should be the starting point of a test development process; we need to ask the following questions:

1. Who are the candidates?
2. What is required of them in the test task?
3. Do the tasks match the candidates’ levels of ability, content knowledge, and language knowledge?
4. Are they well prepared for the test tasks?
5. What processes would the candidates need to use to fulfill the test tasks?

In the case of the GATE test, the factors were spelled out in the test specifications, for example, the topics selected for the Reading passages were fairly academic yet general enough and were within candidates’ scope of knowledge and ability. The topics ranged from education, international trade to technology and social ills. The skills tested include applying inductive and deductive reasoning to determine author’s argument, making inferences and drawing conclusions and evaluating the author’s argument. The Writing tasks require that the candidates demonstrate their ability to discuss a topic in an argumentative manner using appropriate and accurate language. The Grammar tasks require error recognition, correction, and paraphrasing skills.

These are skills that are expected of PG students in their programs; they need to be able to read critically, write in a cohesive and coherent manner, and display language knowledge and proficiency that is up to mark.

### Table 2. GATE: Components Tested.

| Section | Item type/format | Skill tested                                                                 | Number of items |
|---------|------------------|-------------------------------------------------------------------------------|-----------------|
| Reading | MCQ              | Making inferences and drawing conclusion based on evidence or information found in text | 3               |
|         |                  | Determining author’s purpose and point of view                               | 4               |
|         |                  | Using deductive or inductive reason to determine author’s argument/ interpreting graphic aids, which authors use to present information | 7               |
|         |                  | Evaluating author’s argument to determine objectivity, completeness, validity, and credibility | 4               |
| Grammar | Error correction | Identifying errors and making corrections                                    | 15              |
|         | Transformation items | Paraphrasing sentences given first words                                    | 8               |
| Writing | Summary          | Summarizing a text                                                           | 1               |
|         | Essay            | Writing an argumentative essay                                               | 1               |

*Note. GATE = Graduate Admission Test of English; MCQ = multiple-choice question.*

In the beginning stages of the test development process, the participants were randomly sampled from PG candidates (Malaysian and international) currently enrolled in Universiti Teknologi Mara (UiTM) and some selected from other universities. When a candidate has completed the test and obtains a score, she was placed on a band, in line with the levels of the CEFR (although it was anticipated that candidates may reach
a maximum of B2; the highest level in the CEFR is C2 = proficient user).

Phase II—Test Operations (Implementing the GATE)

After many months of developing, vetting, and refining the GATE, it was administered to PG candidates who were already enrolled in various university programs and faculties.

The test was administered in two stages as it was dependent on the availability of the test takers.

Stage 1—October 23, 2010
Stage 2—May 14, 2011

Test examiners were assigned, and the scores were reported (refer to the test analysis).

Phase III—Test Analysis

Before test scores can be analyzed for further validation and revision of test items and the test as a whole,

i. items were calibrated accordingly so they match the objectives and
ii. bias was detected to find items where one group performs much better than the other group: Such items function differentially for the two groups, and this is known as Differential Item Functioning (DIF). Bias is “systematic error that disadvantages the test performance of one group” (Shepard, 1981) or “systematic under- or overestimation of a population parameter by a statistic” (Jensen, 1980).

Results

This section provides the findings of the GATE taken by students in the two stages above.

A. The distribution of scores is reported as follows:

- According to each section—Reading, Writing, Grammar (refer to Table 3)
- The total score for each candidate (refer to Table 3)
- Data on descriptive statistics—mean, median, mode, range (refer to Table 3)
- Distribution of scores (refer to Figure 4)
- Distribution of final scores (refer to Figure 5)
- Placement of student scores against the CEFR (refer to Table 4)

B. Summary of distribution of scores:

- The scores on the GATE showed a distribution of scores that are erratic and fluctuating
- The mean scores for individual sections are just average or below the average expected

Reading > 20.45/40 Grammar > 10.37/30 Writing > 12.7/30

- The candidates performed best in the Reading section and scored lowest in the Grammar section
- The range between the highest and the lowest scores obtained is big (59.7)
- The distribution of final scores showed several peaks (5 scores in the 70s) and several scores in the low 20s (11 scores)

C. Placement of students against the CEFR

1. Before students are placed along the CEFR table, cut-off scores needed to be determined for the individual sections and the final scores of the test. In doing so, the following factors were taken into consideration:

- The GATE was developed solely on research, the team members’ expertise and experience, and its own specifications

| Table 3. Mean Distribution of Scores for Each Section, the Final Scores, and Descriptive Scores. |
|-----------------|-----------------|-----------------|-----------------|-----------------|
|                 | Reading         | Grammar         | Writing         | Final score     |
| Total           | 1,043           | 528.2           | 647             | 2,218.2         |
| M               | 20.45           | 10.37           | 12.7            | 43.5            |
| Mode            | 20              | 3.3, 8.4, 9.6, 11.1 | 20             | 33              |
| Maximum         | 34              | 25.2            | 26              | 79.3            |
| Minimum         | 10              | 1.5             | 3               | 19.6            |
| Median          | 20              | 8.4             | 13              | 40.3            |
| Range           | 24              | 23.7            | 23              | 59.7            |

Data analyzed in the project were based on several sources:

i. Test trials/pilot studies: Test scores from these trials were analyzed using the SPSS and results placed according to the CEFR band.
ii. The new English test for PG students: Test scores were rendered in SPSS and results placed according to the CEFR band.
Table 4. Placement of Students Against the CEFR.

| Score | Grade | Band | Levels | GATE |
|-------|-------|------|--------|------|
| 100-90 | A+    | 6    | C2     |      |
| 89-80  | A     |      |        |      |
| 79-75  | A−    | 5    | C1     | 5 candidates |
| 74-70  | B+    | 4    | B2     | 5 candidates |
| 69-65  | B     |      | B1     |      |
| 64-60  | B−    |      |        |      |
| 59-55  | C+    | 3    | A2     | 7 candidates |
| 54-50  | C     |      |        |      |
| 49-47  | C−    | 2    | A1     | 9 candidates |
| 46-44  | D+    |      |        |      |
| 43-40  | D     |      |        |      |
| 39-30  | E     | 1    |        | 25 candidates |
| 29-0   | F     |      |        |      |

Note. CEFR = Common European Framework of Reference for Languages; MUET = Malaysian University English Test; GATE = Graduate Admission Test of English.
• The GATE did not have an equivalent; however, it benchmarked items against the MUET, IELTS, and other internal tests.
• Results from the test indicated that the test performance for the sample of candidates ranged from good to below average and weak; in fact, 50% of the candidates scored below the proposed cutoff scores, and 50% scored above the proposed cutoff scores.

2. Thus, it was decided that we could take a central point to determine the cutoff scores as follows:

| Sections   | Final score |
|------------|-------------|
| Reading: 20/40 | 40/100*     |
| Grammar: 15/30   |             |
| Writing: 15/30    |             |

*The cutoff for the final scores is the same as the median of the distribution.

3. Thus, referring to the GATE results, we derive the following details:

- 5 scored in the 70s
- 14 scored in the 30s
- 5 scored in the 60s
- 11 scored in the 20s
- 7 scored in the 50s
- 9 scored in the 40s

= 26 scores above 40
= 25 scores below 40

These results were transferred on to the following table to highlight the levels attained by the candidates in the GATE.

It appears that almost 10% of the candidates managed placements in C1, 17% in the B1 to B2 range, 23% in the A1 to A2 range, and the others scored below the range. Although the candidates demonstrated performance that were below the expectations of PG candidacy, especially because many were already in the respective courses, some inferences could be made of the situation.

As mentioned earlier, there seems to be a mismatch in the performance of the candidates in the PG classes although they were able to produce an IELTS or TOEFL score when applying to the university. One possible explanation is in the nature of these tests; many have questioned whether they are tests for academic purpose or tests of proficiency. Test takers may have been able to succeed in the tests, but the demands of PG tasks, especially in critical reading and writing, are far different from the tasks in these standardized tests. Aspects of bias have often been one of the criticisms of these tests (Hawkey, 2004; Jaschik, 2010; TOEFL Research Reports (RR), 2005).

Last, the GATE may have been a rather overwhelming experience for the candidates; they may not have taken such tests for some time, and results of the GATE trial showed their weakest performance in the grammar followed by the Writing section. Overall, despite the test trials and results, the likelihood of developing such a test within the local context in Malaysia is now evident. Given time, resources, and the wealth of test development “experts” in the country, the GATE can be realized as a promising option to the foreign and costly tests.

**Discussion**

The main reasons for developing a “localized,” “home-grown” test for entry into PG education in the country is clear: Recent literature suggests it, and a clear mismatch is found in language ability at PG intake into universities and candidates’ performance in the classroom.

Thus, it is proposed that the GATE is used to measure a candidate’s ability to use the English language at entry level into a PG program in Malaysia. This in itself encourages potential candidates to develop and improve their English language competency so that they are able to communicate and function effectively and successfully in PG education and in a global world where English is the lingua franca of trade, international business, political governance, mass communication, and so on. More importantly, we have reiterated the significance of standardization in the PG enrollment across faculties in Mara University of Technology and across universities in Malaysia.

The GATE was developed systematically according to knowledge and literature in test development, and details in its Test Specification document are based on elements described according to components in the framework for test validation (Weir, 2005). Accordingly, several raters carried out the marking while the Writing section was graded based on the Graduate Record Examination (GRE) scale for writing.

Results of the test (detailed in the section “Results”) indicate that candidates performed fairly well in spite of constraints such as time, exposure to test format and level, and little or no preparation leading to the test. More importantly, the results draw attention to the mismatch between student ability and actual performance within a course, and the concerns of teaching, learning, and testing at PG level in Malaysia.

**Recommendations**

The following recommendations are made based on factors that need to be considered to enhance and facilitate the test development process, and the research project as a whole, leading to the production of a more reliable and valid GATE test.

The test development process requires many hours of research into best practices of test development from organizations such as Cambridge ESOL, English Testing Services (ETS), the Malaysian Examination Board, and other more localized exam boards. The process and procedures need to
be examined and discussed, and test specifications need to be drawn up before the test development process begins.

If the project is to be a success, time, resources in terms of personnel, and money are essential. Teams of human resource are needed for the following:

a. test development: writing, vetting, refining, and collating
b. administering the test, analyzing the outcome, and reviewing the test
c. test operations: administrating the test, monitoring, and collecting
d. grading: marking/rating the test/task using grading/rating/markings schemes, and moderating
e. analyzing test scores and reporting test results, and
f. test validation, which incorporates all of the above to ensure construct validity and reliability.

Furthermore, infrastructure needs to be made available for the project to run smoothly, and this includes facilities such as office space, basic office equipment such as computers, printers, scanners, photocopying facility, statistical software, and other applications to assist in item analysis, score analysis, organization, and reporting of test results.

**Conclusion**

The topic of *localization* in testing is new in the field (O’Sullivan, 2011a), yet one of the major considerations of the test development process is the purpose or needs of the test we are developing. Before we can consider the type of task, test content, scoring, and test outcomes, we have to first determine its purpose for the target test takers. Given this important aspect in testing, it can be concluded that developing the GATE is a viable project as it bears significance and could have a big impact on the quality of PG candidates as well as the PG system in universities in Malaysia. Given much more consideration in terms of time, budget, human resource assistance, and proper management, a test such as the GATE could help raise the status of local academic institutions by having quality PG education, students, and system, which will ultimately produce competent graduates equipped to face a challenging and globalized world.

**Appendix**

Common Reference Levels: Global Scale.

| Level         | Description                                                                 |
|---------------|-----------------------------------------------------------------------------|
| Proficient user | C2 Can understand with ease virtually everything heard or read. Can summarize information from different spoken and written sources, reconstructing arguments, and accounts in a coherent presentation. Can express himself or herself spontaneously, very fluently, and precisely, differentiating finer shades of meaning even in more complex situations. |
|               | C1 Can understand a wide range of demanding, longer texts, and recognize implicit meaning. Can express himself or herself fluently and spontaneously without much obvious searching for expressions. Can use language flexibly and effectively for social, academic and professional purposes. Can produce clear, well-structured, detailed text on complex subjects, showing controlled use of organizational patterns, connectors, and cohesive devices. |
| Independent user | B1 Can understand the main ideas of complex text on both concrete and abstract topics, including technical discussions in his or her field of specialization. Can interact with a degree of fluency and spontaneity that makes regular interaction with native speakers quite possible without strain for either party. Can produce clear, detailed text on a wide range of subjects and explain a viewpoint on a topical issue giving the advantages and disadvantages of various options. |
|               | B2 Can understand the main points of clear standard input on familiar matters regularly encountered in work, school, leisure, etc. Can deal with most situations likely to arise whilst travelling in an area where the language is spoken. Can produce simple connected text on topics that are familiar or of personal interest. Can describe experiences and events, dreams, hopes, and ambitions, and briefly give reasons and explanations for opinions and plans. |
| Basic user    | A2 Can understand sentences and frequently used expressions related to areas of most immediate relevance (e.g., very basic personal and family information, shopping, local geography, employment). Can communicate in simple and routine tasks requiring a simple and direct exchange of information on familiar and routine matters. Can describe in simple terms aspects of his or her background, immediate environment, and matters in areas of immediate need. |
|               | A1 Can understand and use familiar everyday expressions and very basic phrases aimed at the satisfaction of needs of a concrete type. Can introduce himself or herself and others, and can ask and answer questions about personal details such as where he or she lives, people he or she knows, and things he or she has. Can interact in a simple way provided the other person talks slowly and clearly and is prepared to help. |
Declaration of Conflicting Interests
The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding
The author(s) received no financial support for the research and/or authorship of this article.

References
Hawkey, R. (2004, December 8-10). A study of the impacts of IELTS, especially on candidates and teachers. Paper presented at Going Global. The UK International Education Conference, Edinburgh, Scotland.
IELTS Handbook. (2007). University of Cambridge web publication. Available from www.ielts.org
Jaschik, S. (2010, August 2). New questions on test bias. Inside Higher Ed. Available from http://www.insidehighered.com/
Jensen, A. R. (1980). Methods for identifying biased test items. In G. Camilli & L. A. Shepard (Eds.). Thousand Oaks, CA: Sage.
Little, D. (2006). The Common European Framework of Reference for Languages: Content, purpose, origin, reception and impact. Language Teaching, 39, 167-190.
Moon, T. S., & Siew, H. O. (2004). A study on the factors that impact on academic performance of the computer science and the information technology students in University of Malaya. CMU Journal, 3(2), 169-184.
O’Sullivan, B. (2011a). Language testing: Theories and practices (Advances in linguistics). London, UK: Palgrave Macmillan.
Othman, A. B., & Nordin, J. (2013). MUET as a predictor of academic achievement in ESL teacher education. GEMA Online Journal of Language Studies, 13(1), 99-111.
Ponniah, K. S., & Tay, B. (1992). Processing strategies in reading. Suara Pendidik, 15(4), 27-40.
Redden, E. (2008). English for graduate students. Inside Higher Ed. Retrieved from http://www.insidehighered.com/news/2008/06/05/english
Saville, N. (2006). The CEFR and common standards for international language examinations a QMS approach. Paper presented at ALTE conference, Copenhagen, Denmark.
Shepard, L. A. (1981). Identifying bias in test items. In B. F. Green (Ed.), Issues in testing: Coaching, disclosure and ethnic bias (pp. 79-104). San Francisco, CA: Jossey-Bass. Retrieved from http://www.eric.ed.gov
Tannenbaum, R. J., Wylie, E. C. (2005) Mapping English language proficiency test scores onto the Common European Framework (ETS Research Rep. No. RR-05-18; TOEFL Research Rep. No. RR–80). Princeton, NJ: ETS.
Weir, C. J. (2005). Language testing and validation: An Evidence-based approach. New York, UK: Palgrave Macmillan.

Author Biographies
Saidatul Akmar Zainal Abidin is associate professor and senior lecturer in the field of English language and Linguistics. She has more than 28 years experience in teaching English to second language speakers (TESL), and the programme ‘English for Professional Communication’. Dr Saidatul’s research interest is in the fields of language testing & evaluation, professional communication and curriculum development.
Asiah Jamil is associate professor and senior lecturer in the field of English language and Linguistics. She has more than 25 years experience in teaching English to second language speakers (TESL), and the programme ‘English for Professional Communication’. Dr Asiah’s research interest is in the fields of language testing & evaluation, test taking strategies and qualitative research.