

Refining English Language Tests for University Admission: A Malaysian Example

Arshad Abd Samad
Sharifah Zainab Syed Abd Rahman
Siti Norbaiti
Universiti Putra Malaysia
Email: arshad@educ.upm.edu.my

ABSTRACT

English has now become the lingua franca of much of technological, business and academic endeavours. Consequently, learning the English language is now seen as vital, especially at the university level where proficiency in the language has become a selection criterion. At present, the Malaysian University English Test (MUET) has been adopted by Malaysian public universities as an indicator of English language proficiency. A student's overall result depends on all the four language components of the MUET and often determines the number and nature of the English language courses he or she has to attend at university. This study seeks to examine whether MUET is an accurate predictor of performance and success at university and how the MUET can be finetuned as an entry level English language test. It was carried out among 52 third year undergraduates of the Faculty of Educational Studies, Universiti Putra Malaysia, Serdang, Malaysia, admitted into the Teaching English as a Second Language (TESL) programme. The findings of the study do not offer conclusive evidence about the validity of MUET as a predictor of academic success. However, six models of various combinations of scores on language components on the MUET scores are examined in terms of their effectiveness in increasing the accuracy in selecting students for the TESL programme. The correlations obtained using these models indicate that the combination of various components of the MUET can be used to more accurately predict student achievement at tertiary level than the cumulative MUET score itself. The results of

these correlations and their implications in using language tests as admission requirements in general are also discussed.

Introduction

The English language is an important language in academe and Malaysian public universities have now taken English language proficiency seriously into consideration when making decisions on admission. One of the main measures of English language proficiency for admission into Malaysian universities is the Malaysian University English Test (MUET). According to Khatijah Mohd Tahir (2004), the MUET was first administered in 1999 for the purpose of bridging the gap in English Language proficiency between the final year of secondary school and university level study and providing a context for continuous teaching and learning of English.

There are four components graded in the MUET, namely Listening, Speaking, Reading Comprehension and Writing. The maximum score is 45 each for Listening and Speaking, 135 for Reading Comprehension and 75 for Writing. The total possible aggregate score is 300 and a score is reported as a Band score of 1 to 6, with the former being the lowest and the latter, the highest. The listening component requires candidates to listen to three recorded texts and answer 15 multiple-choice questions (MCQ) in half an hour. The component tests the ability of candidates to understand and listen to messages contained in long, serious, social and academic talk. In the speaking component, candidates are given two minutes for an individual presentation and 10 minutes for a group discussion. Marks are allocated for task fulfillment, language use and communicative ability. As for the writing component, candidates are required to write a summary and a composition in one and a half hours. The largest part of the MUET test is the two-hour reading comprehension component which consists of 50 questions based on different types of passages and tasks including cloze passages, information transfer, and interpretation of diagrams, tables and graphs.

Most candidates who sit for the MUET do so as part of their admission requirements into public universities and colleges in Malaysia. Universities generally use the MUET to determine the English language proficiency courses the students have to attend. In Universiti Putra Malaysia (UPM), for example, students who obtain Bands 1 or 2 in the MUET are required to take 3 English language proficiency courses; those who obtain Bands 3 or 4 need to take 2 English language proficiency courses; while those with either a Band 5 or 6 are exempted from any such courses.

It is not always the case that the students' English proficiency determines their academic success and the importance of performance in English varies according to the student's academic programme. For example, the MUET score may not be as important a predictor of success in programmes that have more emphasis on the application of scientific and content knowledge. Certain programmes, however, expect English language proficiency to have an influence in academic success and therefore specify a particular band in the MUET as an entrance requirement. The Bachelor of Education in Teaching English as a Second Language (TESL) programme at UPM, for example, requires a minimum MUET band score of 4 or an equivalent indicator of English language ability. About seventy percent of the courses in the TESL programme are taught in English and hence a higher MUET requirement is justified. There is a shortage of English language teachers in the country and hence many places are available in the programme. However, there is an equally large number of applications for the programme. At UPM, on average about 60 students have been accepted into the programme every year for the past three years, although applications for the programme are often triple this number. In future, the number of students admitted into the programme at UPM is expected to be slightly lower as the emphasis of the university changes towards graduate studies. With the expected reduction in available places in the programme, a more stringent but just method of student selection is required. In such a situation, it is necessary to examine how accurate the MUET is in predicting academic success and hence how valid it is as an entrance requirement into the Bachelor of Education (TESL) programme.

Examining Predictive Validity

Two major questions were addressed in this study. The first question asks what the relationship is between English language proficiency, as measured by the MUET and the student's academic performance at university level, as measured by the student's CGPA early (semester 01) and late (semester 06) in the programme. The purpose of this question was to examine the validity of the MUET in predicting academic success and consequently its appropriateness in determining the English language proficiency courses required. This question will also examine whether the use of individual component scores on the MUET (e.g. scores on reading or listening) would provide better predictors compared to the overall Band score.

Although the most common method for determining the validity of a test is examining its construct validity, when a test is used as a predictor of future performance the predictive validity of the test is often referred to. Predictive validity involves a time interval. A test is first administered and, after a period of time, a behaviour which the test is intended to predict is measured (Kubiszyn & Borich, 2003). Performance on the test is then correlated against the behaviour. Predictive validity is therefore concerned with the use of the test performance to predict future performance on some other valued measure or criterion. In many cases, this predicted measure is academic performance. Unlike construct validity, which is determined through analysis and logical reasoning, predictive validity involves correlating test performance to future performance and hence yields a correlation value. Various studies have examined the predictive validity of different tests as a predictor of academic success. These include the International English Language Testing System or IELTS (Dooley, 1999; Feast, 2002) and the Test of English as a Foreign Language or TOEFL (Ayers & Peters, 1977).

The second question examines how students and their CGPA are affected if various score combinations of the language components in the MUET are adopted as entry requirements into the TESL programme. These effects were tested using various models of score combinations devised according to specific language expectations of university level students. Six models were proposed as illustrated in Table 1 below.

The first model highlights the importance of speaking ability especially for teachers of English and hence the single criterion used is that the students had to achieve a score of 30 or higher in the speaking component of the MUET. The second model emphasizes an overall proficiency model with cut off scores for each language component in the MUET set at a minimum of 65% of the possible score for the component. The third model considers the result of the first question in this study and uses the reading component as the main criterion. Model number four examines the notion that most academic work involves reading and writing and hence uses both these components from the MUET as its criteria. In the fifth model, the speaking component is once again included in order to emphasise its importance to language teachers. In this model, reading is included as an important component for academic university level work. Finally, model six focuses on the productive skills of speaking and writing, both of which are the most commonly assessed skills due to their productive and hence observable nature. Models that used a single

Table 1: Alternative Models to Band Score as Criteria for Selection into the TESL Programme

Model	MUET based criteria	Defining feature and rationale
Model 1:	Speaking > 30	Importance of speaking for TESL teachers
Model 2:	Speaking > 25; listening > 25; reading > 90; writing > 50	Overall language proficiency
Model 3:	Reading > 100	Reading and aggregate score as highest correlated components.
Model 4:	Writing > 50, Reading > 90	Writing and reading as traditional academic necessities
Model 5:	Reading > 90, Speaking > 25	Speaking required of TESL teachers as well as reading as an academic necessity
Model 6:	Speaking > 25, Writing > 50	Speaking and writing as 'productive skills' and easily assessed.

skill as a criterion had a higher cut-off score for the skill compared to when the skill was used in combination with another skill. For example, in Model 1, the cut-off score for speaking was 30 whereas it was 25 when speaking was assessed together with other skills such as in Models 2, 5 and 6.

A total of 52 students who entered the TESL undergraduate course at Universiti Putra Malaysia (UPM) in 2004 were involved in the study. The overall MUET aggregate scores, the four component scores and the corresponding CGPA for the first and sixth semesters were obtained and analysed.

Results and Discussion

The two questions posed earlier concerning the predictive validity of the MUET as well as the feasibility of various alternative models as admission criteria are discussed as follows based on data collected in the study.

Examining the Predictive Validity of the MUET

The correlation between components of the MUET and the students' academic performance are presented in Table 2 below. We note that the

Table 2: Correlation between MUET and Cumulative Grade Point Average (CGPA) at the First and Sixth Semesters

	CGPA at Semester 01	CGPA at Semester 06
Listening	0.19	0.29*
Speaking	-0.08	0.125
Reading	0.35*	0.58**
Writing	0.19	0.32*
Aggregate	0.29*	0.55**

* statistically significant at $p < 0.05$,

** statistically significant at $p < 0.01$

correlation coefficient between results obtained for various components and the CGPA for both semesters 1 and 6 are low. For the CGPA during the first semester, low coefficients were observed for listening (0.19), speaking (-0.08), and listening (0.19). The correlations between CGPA during semester 6 and the various MUET components also yielded low coefficients, but only for speaking at 0.125. The two highest correlations in the table are the correlations between CGPA at semester 06 and reading ($r = 0.58$) and between CGPA at semester 06 and the aggregated MUET score at $r = 0.55$.

Based on the results obtained in Table 3, the following three observations need to be further examined and discussed.

- a. *The speaking component has a low correlation with CGPA.* The speaking component of the MUET does not correlate well with academic performance during either the first or the sixth semester. In fact, speaking and the CGPA for semester 1 has a negative correlation ($r = -0.08$). This indicates that a high score in the MUET speaking component does not indicate a high CGPA for semester 1. It was anticipated that since the TESL programme was a teaching programme that focused on language, the MUET speaking component would have a strong correlation with grades obtained in the programme.

There could be many reasons for the low correlation between scores on the speaking component and the students' academic performance. Assuming that both the scores are valid measures, the low correlation may imply that the task in the MUET speaking component is not similar to the kind of speaking that is assessed in courses in the TESL programme. In the MUET, scores are assigned

based on an individual and group task involving a single situation that is usually of general interest or social in nature. As for assessment at the tertiary level, there is no specific session to test speaking ability because the assessment is done simultaneously during various specialized tasks such as microteaching or presentations on academic topics. These tasks differ quite significantly from the types of tasks as well as the content used in the MUET. At the same time, the more formal testing conditions in the MUET differ from the in-class oral presentations given in front of peers and may also contribute to the low correlation observed between the two measures.

- b. *The correlations improve with time.* Compared to the first semester, the students' CGPA for semester 06 has a higher correlation with all the MUET's testing components and aggregate score. This seems to imply that the MUET score is more indicative of the students' academic performance after several semesters in the university. The longer time interval between the MUET test and the sixth semester or the third year of the student's study may have enabled students to more accurately show their language abilities. While some believe that "lifestyle, personality, attitudinal or experiential changes may alter an individual's rank on criterion measure 2 years from now from what it was an initial testing" (Kubiszyn and Borich, 2003: 305), it is also possible that English language demands during the first year in the university differed from what the students were used to. The indications here are that the students had to make adjustments to their language use during the first year and only achieved more stability and consistency in the later semesters. Another possible explanation for the improvement in correlations after time is that in the TESL programme at UPM, many of the more general education courses that do not need to be taught in English are presented in the first semester. Hence, performance in the English language in the MUET will not correlate with performance in subjects that are not taught in English.
- c. *The reading component shows the highest correlation with CGPA.* A positive and significant correlation between the MUET reading component test scores and students' CGPA for both semesters was observed. Since the correlation that was obtained apparently shows that the scores on the MUET reading component are linked to the students' academic performance in university studies, the reading component can be a good predictor of students' success at tertiary level. It is difficult to explain the higher correlation for the reading

component as opposed to the other three language components. A possible explanation, however, is that reading forms the basis of academic ability and hence the relatively higher correlation between the reading component score and academic performance at the two time periods. Alternatively, however, it may be possible that the Malaysian education system is largely text or reading based and hence scores on reading represent the most stable indicator of language ability and consequently the highest correlation with CGPA.

The Effect of Various Models of MUET Component Score Combinations

Six models, as described earlier in Table 1, are examined in terms of their effect on CGPA, the number of students retained from the original number, as well as its accuracy in selecting and not excluding high achieving students. Table 3 below illustrates the effects of the five models.

Based on the data in Table 3, all six models yielded an increase in the average CGPA. However, Model 2 results in the highest average CGPA at 3.38 which is an increase of 0.23 points compared to the original average CGPA. This increase represents a 7.30% increase in the CGPA. While Model 2 was the most effective in raising the average CGPA of the students, it was the least effective in retaining students from the original group of 52 students in the study. Only 8 of the 52 were retained as the 44 others failed to meet the minimum requirements set by the model. The low number of students retained is a definite drawback of the model.

Model 1 was the model that was able to retain the most number of students. This model used the speaking score as its sole criterion for selection. As many as 46 students or 88.46% of the original number of students were retained. Model 1 can also be considered a fair and just model as only 3 students or 11.11% of students who scored higher than the average CGPA of the retained students were excluded. However, Model 1 did not result in a significant positive change in the average CGPA as the CGPA average of the retained students was similar to that of the original set of students.

A third model that shows promise is Model 5 which specifies a reading score of 90 or more and a speaking score of at least 30 as its selection criteria. This model produces a 0.08 point increase in CGPA, is able to retain 56% of the original students, and excludes a relatively low

Table 3: Effect of Alternative Models on Student Selection and Predicting Academic Performance

	Original average CGPA	Average CGPA of remaining students	Change in CGPA	% increase in CGPA	No of students retained	% of students retained	No. students with CGPA > av. CGPA of remaining students	No of students excluded with CGPA > av. CGPA of remaining students	% of students excluded
Model 1: Speaking	3.15	3.15	0.00	0	46	88.46	27	3	1.11
Model 2: Overall	3.15	3.38	0.23	7.30	8	15.38	12	8	6.66
Model 3: Reading	3.15	3.37	0.22	6.98	20	38.46	12	3	25
Model 4: Writing and Reading	3.15	3.37	0.22	6.98	9	17.31	12	8	6.66
Model 5: Reading and Speaking	3.15	3.23	0.08	2.54	29	55.77	23	6	6.09
Model 6: Speaking and Writing	3.15	3.29	0.14	4.44	11	21.15	20	15	75

26% of students who had a CGPA higher than the average CGPA of the retained students. Overall, this model is able to provide fairly good results in the three categories of increasing average CGPA, retaining students, and not rejecting high CGPA students.

Conclusion

This study was initiated in order to examine whether English language examinations used as admission requirements can be accurate predictors of academic success among undergraduate students. Specifically, it examines the MUET used in Malaysia. The scores of the four components of the MUET as well as the MUET aggregate score were correlated with the academic performance of the students during their first and sixth semesters. The findings of this study suggest that among the four components of the MUET, the Reading component seems to be the

most valid as a predictor of early as well as late academic success. The correlation between the reading component scores and the students' CGPA is not only consistently positive for both the first and sixth semesters, but also yields the strongest correlation values.

Conversely, the MUET speaking component is a poor indicator of tertiary success. Since all of the students in this study are prospective English language teachers, it would seem that English language speaking ability is needed to ensure academic success. However, scores obtained on the MUET speaking component do not correlate well with academic performance. It was suggested that a major cause of this low correlation was the lack of congruence between the MUET speaking test tasks and the assessment of speaking related tasks in the TESL programme. Based on this observation, other forms of assessment either in addition to or in place of the MUET should be used if speaking ability is to be used as a criterion for admission. These assessment formats could take the form of presentations or micro teaching sessions which are reflective of the speaking tasks of prospective teachers during training at the tertiary level. Alternatively, the speaking component of the MUET could be revised in order to better reflect university level tasks and speaking related assessments.

It is also suggested that combinations of scores on the language components in the MUET be considered as admission criteria, especially into the TESL programme, rather than the single band score or aggregate score. In Malaysia, the MUET offers a clear advantage over other locally developed measures of language proficiency such as the secondary school English language exam or even the GCE O level English subject as, unlike in these examinations, scores on each individual language skill in the MUET are reported. These individual scores, however, are largely ignored in favour of the Band or aggregate score. It is suggested that selection decisions based on the MUET pay greater attention to the scores for each language skill or some form of combination of these scores. The use of an appropriate combination of scores can lead to a more just and accurate measure in making admission decisions. Of the six models examined in this study, model 5 which focuses on reading and speaking seems the most appropriate as an admission criterion instead of the more generalized aggregate or Band score. Instead of specifying a Band 4 as the only admission criteria, either an additional or alternative criterion of a combined score of at least 90 for reading and 25 for speaking could be used.

This paper has examined the possibility of using a combination of component scores to increase the precision of the MUET as a tool for predicting academic success at tertiary level in Malaysian universities. While the paper describes a Malaysian scenario, the results indicate that the same techniques may be used with other English language examinations used for university admission purposes. The MUET remains a useful tool in estimating student preparedness for using English at the tertiary level. However, when it is to be used for admission decisions, greater care should be taken as the decision can be highly sensitive and politicised. Admission into a programme or even determining the number of English language proficiency courses a student needs to take can hinge on not only how the students perform on the test but also how the scores on the MUET are interpreted and used. As presently the MUET is used largely for university entrance purposes, it may also be a useful idea to tailor the tasks in the MUET according to the types of language related activities that occur in universities.

References

- Abedi Abdullah, Sarah (2004). MUET & IELTS preparation: Can one size fit both? In Sarah Abedi Abdullah and Malachi Edwin Vethamani (Eds.) *Preparing students for the Malaysian University English Test*. SASBADI MELTA ELT Series.
- Awang Mat Dahan, H.B. (2002). *Language testing: Construction and validation*. Kuala Lumpur: University of Malaya Press.
- Ayers, J.B. & Peters, R.M. (1977). Predictive validity of the Test of English as a Foreign Language for Asian graduate students in engineering, chemistry and mathematics. *Educational and Psychological Measurement*, 37(2), 461-463.
- Dooey, P. (1999). An investigation into the predictive validity of the IELTS Test as an indicator of future success. In K. Martin, N. Stanley, and N. Davison (Eds.). *Teaching in the Disciplines/Learning in contexts, Proceedings of the 8th Annual teaching learning forum*, The University of Western Australia, February 1999. 114-118.

Feast, V. (2002). The impact of IELTS scores on performance at university. *International Education Journal*, 3(4), 70-85.

Hayashi Keiko (1999). Reading strategies and extensive reading in EFL classes. *RELC Journal*, 30(2), 114-129.

Kubiszyn T. & Borich G. (2003). *Educational testing and measurement: Classroom application and practice*. (7th ed) New York: Wiley.

Mohd Tahir, Khatijah (2004). MUET overview and examiner's feedback. In Sarah Abedi Abdullah and Malachi Edwin Vethamani (Eds). *Preparing Students for The Malaysian University English Test*. SASBADI-MELTA ELT Series, 1-12.

Pandian, A. (2002). *Practical guide to Malaysian University English Test*, University Putra Malaysia Press.

Stansfield, C.W. & Hewit, W.E. (2005). Examining the predictive validity of a screening test for Court Interpreters, *Language Testing*, 22(4), 438-465.