

Exploring the connection between the testing of reading and literacy: The case of the MUET

Lee King Siong

**School of Language Studies and Linguistics
Faculty of Social Sciences and Humanities
Universiti Kebangsaan Malaysia
bking@pkriscc.ukm.my**

Abstract

The Malaysian University English Test (MUET), introduced in 1999, is a potentially high-stakes test because of the fact that it is a mandatory requirement for admission into public universities. It was introduced with the aim of bringing about a higher level of English proficiency, especially crucial for those entering university. The paper focuses on the impact of the Reading component on literacy, and examines four factors that are perceived to affect the impact of the MUET, namely: 1) the perceived status of the MUET, 2) the sociopolitical situation, 3) the teacher/teaching, and 4) the test construct. It concludes that the reading construct in the MUET Reading component is inadequately operationalized and suggests that the construct be reviewed in the light of long-term literacy goals.

Introduction

This paper is based partly on an on-going study on the impact of a high-stakes standardized university entrance English exam on reading instruction in the Sixth form class. The small study involves 9 sixth form English teachers of the Malaysian University English Test (popularly referred to as MUET) and 230 lower six form students. The teachers were from five urban schools in Kuala Lumpur and Petaling Jaya selected for their comparatively better performance on the MUET over the past few years. Although it was not possible to carry out classroom observations, what teachers said about their teaching activities seemed to be confirmed by students' reports in their focus group discussions. Besides this study, I also draw from a paper I presented at the Malaysian English Language Teaching Association (MELTA) Conference on Reading in 2002 (Lee, 2004), in which I analysed the construct operationalized in two Reading tests in two major English examinations in the Malaysian school system, the Sijil Pelajaran Malaysia or SPM, and the MUET, and showed how reading tests perpetuate a narrow view of reading and influence the way reading is taught.

Although the MUET is made up of 4 components, namely, Listening, Speaking, Reading and Writing, I have chosen to look at the impact on reading instruction for a start. The importance of reading for university education is reflected in the weightage given to the reading comprehension component in the MUET: it is 45% of the total marks.

Let me first define the notions of literacy and reading that form part of the framework of this paper. “Reading is clearly primary to any definition of literacy” (Venezky,1990:9), but the view of literacy taken in this paper is not the skill-based notion of literacy as consisting of the ability to read and write according to prescriptive, normative standards. I am not looking at literacy as an end product of instruction but literacy that involves “not only the ability to produce and interpret texts, but also a critical awareness of the relationships between texts, discourse conventions, and social and cultural contexts” (Kern,2000:6) The current view of reading has moved beyond reading as a psycholinguistic process; reading is also a socially and culturally embedded practice. It is not just schema theory but also reader-response theory. Kern describes it as “a dynamic, interactive process of deriving discourse from text” (Kern,2000)

I should also explain the choice of the word “impact” rather than “washback”. Washback is the term used to refer more specifically to the influences of a new test on teaching, teachers and learning, including curriculum and materials (Alderson and Wall,1993). To decide what constitutes washback, one would have to compare the situation before and after the new test was introduced. In the case of the MUET, there was no English being taught in the sixth form at the time the test was introduced, hence the problem of deciding what is washback. In any case, my concern is with the wider influences of the MUET, encompassing washback, or any kind of influence attributed to the test on teaching. The study of test impact is concerned with the consequences of a newly introduced, national test on society and the educational system. It is concerned with how individuals are affected by the test results, and how the general education process is affected by the use of the test.

Why study impact? In Bachman and Palmer’s framework of test usefulness (1996), impact is one of the six facets that need to be assessed in the design as well as the evaluation of a test. But impact studies serve different purposes: it can help to ensure that ethical language testing is achieved, and, from a validation perspective, gathering information on the social consequences of testing is “an aspect of construct validity” (Messick, 1989:18).

My aim in this paper is to share my perspectives on why a high-stakes test meant to induce greater interest in English and bring about a higher level of proficiency does not seem to be making any significant impact on the reading competence of the learners.

I shall first give a brief description of the MUET, followed by some main findings of research on the impact of high-stakes tests. The discussion that follows will be confined to just four factors that I perceive to affect the impact of the MUET, and the conclusion will be my reflection on the implications for literacy development among students entering university.

The MUET

The Malaysian University English Test is a potentially high-stakes test introduced into Malaysian schools in 1999. The stated aims are: 1) to bridge the gap in English language needs between secondary and tertiary education, and 2) to consolidate and enhance the English proficiency of students preparing to enter Malaysian public

universities. For some time prior to the introduction of the MUET, there had been no provision for English classes in the sixth form.

The MUET booklet on Regulations and Scheme of Test, Syllabus and Sample Questions published by the authority behind the test, the Malaysian Examinations Council, describes the test as a competency test designed to measure students' level of proficiency. The test is a battery of four tests: Listening (15%), Speaking (15%), Reading Comprehension (45%) and Writing (25%). The Reading component has a heavier weightage because it is perceived that students at university will be expected to read more than they would write. Performance on the test is reported in terms of an aggregated score with respect to six levels of achievement, referred to as Bands 1 to 6, Band 1 being the lowest and Band 6 the highest. Since it is a criterion-referenced test, the bands have descriptors of the expected performance at each level. Although it has not been validated by any foreign examination body, and not as well-known and widely accepted as the Sijil Tinggi Persekolahan Malaysia (STPM), which is the Malaysian version of the Higher School Certificate, endorsed by the University of Cambridge Local Examinations Syndicate (UCLES), "some foreign universities have allowed students to use MUET as proof of English proficiency (*sic*) in lieu of TOEFL and IELTS" (The Sunday Star, 27 July 2003).

The Reading component in the test battery is a two-hour paper comprising 50 multiple choice questions (MCQ). The first 15 questions are in a cloze passage, the next 7 questions are on information transfer, and the rest of the questions are based on four reading passages of varying length and difficulty level.

Since the introduction of the MUET, schools with form six classes have been directed to allocate 6 to 8 periods (4 to 5.3 hrs) per week for the teaching of English, or more specifically, for MUET coaching. On paper, the MUET is a very important test because it is a stipulated requirement for admission into public universities. However, as Madaus (1988) points out, whether a test is high stakes or not, it is the perception of the stakeholders that will influence their response to it.

What research tells us about high stakes tests and impact

The notion that testing has impact on society is not new. In general education, as far back as 1877, Latham (cited in Spolsky, 1995) referred to tests as "the encroaching power" that can revolutionise education. Frederiksen in his 1984 paper declared that the real bias was the influence that tests had on what was being taught in the classroom. Popham (1987), believing in the influence of high-stakes tests, was a strong advocate of using high-stakes tests to influence curriculum (measurement-driven instruction, or MDI). He was of the view that MDI could be "a potent force for educational improvement". It was also seen as having the potential for being "a particularly cost-effective way of improving public education" (Gipps, 1994 :33).

Popham believed that "if tests were properly conceived and implemented", then focusing on what they assessed was a positive activity. He did, however, lay down certain conditions: criterion referencing, defensible content, a manageable number of targets, "instructional illumination", and adequate support for teachers in the form of instructional support, useful teaching materials and advice on beneficial teaching activities. Madaus (1988), a fierce critic of measurement-driven instruction, pointed

out several negative effects of using high stakes test to drive the curriculum. Among his observations were that while test scores may rise, students' skills may not necessarily have improved, because the teaching has been aimed at the test item, and the format, rather than at the construct or the skill which it intends to measure. Other negative effects he observed were that high-stakes tests led to cramming and narrowing of the curriculum because teachers would naturally teach to the test, and tend to neglect those aspects of the curriculum that were not going to be tested.

In language education, impact studies actually took off only in the 1990s. A few of the more well-known studies are the landmark Sri Lanka Impact Study (Wall and Alderson, 1993), Shohamy's study comparing the impact of three different tests in Israel (1993), and Li's study of the impact of the MET in China (1990). Studies on the impact of language tests have so far been on national or large scale standardized tests. There have been several other smaller studies on washback in different countries too (Andrews, 1995; Watanabe, 1996; Lam, 1995; Lee and Wong, 2000). While there were those in language education who saw the potential in using tests as a lever for change (Pearson' 1988), there were others who viewed it as a simplistic notion, because there was insufficient empirical evidence to support it. The empirical studies on washback and impact have established a few important findings:

(1) it is a mistaken notion that tests in and of themselves will have an impact on teaching and learning,

(2) test impact is a complex phenomenon, and all the factors that work one way in one context may not work the same way in other contexts, nor will they necessarily lead to the same outcomes, (Wall and Alderson, 1993; Alderson and Hamp-Lyons, 1996; Shohamy, 1993; Shohamy et al, 1996) and

(3) other factors are important in affecting impact : the status of the test, the social and political situation, the teachers – their background, training, willingness to innovate etc, and, of course, the test itself .

Discussion

The introduction of the MUET represents an innovation in the Malaysian education system, because for the first time, a test was introduced with neither a pre-designed syllabus nor textbook. In fact, what is now referred to and used as the syllabus is actually the test specifications that appeared in the first edition of the MUET guide book in 1999. The MUET was meant to be a lever for change.

Although impact depends on several factors, I shall discuss only four which are relevant to the MUET: the status, the sociopolitical situation, the teachers' perceptions of the test, and the construct of the test.

1. The perceived status of the MUET

One important thing to establish is the perceptions of the stakeholders as to the status of the MUET. Both parties of stakeholders, the teachers and the students, acknowledged that the MUET, being a compulsory test for admission into public universities, was a high stakes test. However, the fact that there is yet no clear official

statement from the Education Ministry on specific minimum requirements for admission into public universities does lower the status somewhat. Although universities are expected to use the MUET result as either an entry or exit requirement, only one university so far has spelt out specific band requirements for acceptance into specific faculties; admission into the medical and TESL programmes requires Band 5, admission into Law requires Band 4, and Band 3 is the minimum requirement for acceptance into other faculties. A few universities are reportedly imposing a minimum exit requirement of Band 3; however, time will tell whether this is at all feasible.

Apparently, none of the students interviewed were taking tuition for the MUET. In our very examination-oriented culture, this is very telling on the status of the MUET – that it certainly is not in the same league as Bahasa Malaysia, or Maths and Science subjects which most students generally seek tuition for because good results in these are perceived to be crucial in deciding the type of courses they will pursue in the next stage of their education.. A few teachers commented that if only there was some sort of certificate to show for their effort, students might have a higher regard for the MUET.

A close look at job advertisements in the national newspapers reveal an absence of the mention of MUET among the requirements - this is again an indication that the MUET has not yet attained any symbolic importance in Malaysian society.

Despite the ambivalent status of the test, the teachers and students who participated in the study generally take it seriously and work very hard at achieving good results. A few students have been known to re-sit the paper (it is offered twice a year) for the sake of improving their results.

2. The Sociopolitical situation

The MUET was introduced at a time when the educational system was beginning to re-emphasise the importance of learning English. This national concern with improving students' proficiency in the English language came about with the awareness of the importance of English in the era of globalisation. Along with this was the realization that the achievement of national aspirations depended greatly on having a workforce that had access to current technical and scientific knowledge, and that the key to access was a high level of competency in English.

One major move was the introduction of a “new” English test , the SPM 1119, in 1997. This new test was reportedly introduced as a reaction to falling pass rates in the English paper in the national public examination administered at the end of secondary schooling. It was seen as a means of inducing greater interest and motivation to do well in the English paper. Things took a critical turn in 2003 when English was introduced as the language for teaching Science and Maths . Students in Primary one, Form four and Lower six were the first cohorts to be taught Maths and Science in English. Public universities have begun converting more of their courses to English.

The public is constantly reminded that English is an important language to master for the sake of the nation's development and progress. There is indeed a great deal of awareness about the need to master the English Language. But in our exam-oriented

culture, the focus is on the product, the examination results, not the process. Finally, although English is taught as a second language in all national schools in Malaysia, there has always been an ambivalence about its status in the education system : “compulsory to take, but not to pass”.

3. The Teacher/Teaching

In the literature, several factors that affect impact are related to the teachers: teacher ability, teacher understanding of the test (Wall and Alderson,1993), teacher-style, commitment and willingness to innovate (Alderson and Hamp-Lyons,1996), teacher background (Watanabe,1996), and teacher training (Cheng,1997).

All the nine teachers that I interviewed were very experienced English teachers, and all but one had a Masters degree. They were very familiar with the test specifications that served as the syllabus, and had a positive regard for the test. They believe that the MUET reading test is a good test because it is challenging. While they felt that the level of difficulty was appropriate considering that this was a university entrance requirement, they pointed out, however, that many of the students who enter sixth form these days do not seem to cope well with its demands.

They had no problem with the test consisting entirely of multiple choice questions (MCQs) – because, they said, the distractors are “tricky” and so make students think carefully before selecting the right answer. Only a few teachers felt that there should be some open-ended questions. Not one of the teachers expressed doubt about the efficacy of teaching reading skills ala MUET. They talked about teaching reading only in terms of skills. There was no mention of teaching reading strategies, or awareness of text or critical reading.

I have no doubt as to the capability of the teachers, their interest and motivation, and their efforts at coaching and coaxing. However, when I look at the descriptions of what goes on in the reading class, I find a very traditional approach to reading, what is described as “procedural display” – getting the lesson done. The routine of a typical reading class described by the students did not differ much from what teachers themselves reported of how they taught. It was basically the same in all the schools: teacher provides a text, with or without a set induction, students read (silently, or take turns to read aloud), teacher carries out a discussion with students, students attempt the questions, teacher corrects answers with the class. There were variations: in one case, students reported that the teacher would get them to read aloud to correct their pronunciation, in another case, the teacher would have a discussion on the topic before students read the text. Integration with other skills was a common practice, usually with speaking and writing, and some teachers reported introducing other fun and useful activities into the reading class. For example, a few teachers organized their students to work on projects which required research and an oral presentation, while yet another said that she taught her class basic library research skills, including writing bibliographical data.

Since the MUET was in fact designed as a tool to bring about improvement in the English proficiency of students, there is no syllabus as such and therefore no textbooks. The teachers I interviewed said that they sourced their teaching materials from various sources, or selected a particular commercially-produced book for use as

a basic text, and selected passages from the same sources as those used for the MUET (non-specialist magazines such as Time, the Malayan Naturalist, the Far Eastern Economic Review, in-flight magazines, and of course, the newspapers.), but much of the time was spent on practising with commercially-prepared MUET exam papers.

By and large, however, it would appear that what goes on in the MUET reading class is in fact strictly coaching for the exam. It seems to be a very much skills oriented approach, as the syllabus suggests. The training in reading skills, according to most of the teachers and students, was not much different from what had gone on in their upper secondary years, except that the texts were now definitely more challenging (expository texts rather than fictional narratives or descriptions) and some higher level skills such as inferencing and critical thinking were involved. The problems that most students had were simply the result of a previous lack of training in basic skills and lack of interest in reading in English. As one teacher lamented, some of the students have such a weak grasp of English that “there just isn’t enough time to catch up in form six”.

4. The Test Construct

A perusal of the test specifications/syllabus for the Reading Comprehension component revealed a very modest aim: “The syllabus seeks to enable students to comprehend linear and non-linear texts”. The list of skills range from basic reading skills such as extracting information and distinguishing main ideas from supporting details to higher level skills such as inferencing, paraphrasing, summarizing and even reading critically. However, the problem with a skills-based test is that the results can only tell us how well a reader has done on the test; it does not tell us about his reading ability or reading behaviour beyond the test situation. Indeed, as someone has pointed out, if a test can be prepared for, then the test no longer can be said to measure proficiency, but instead, how well people have studied for the test.

The construct of reading that seems to underly the MUET reading component is that reading is a product, that it consists of a list of discrete skills, that readers have to extract everything that a writer has put into a text. This is a very limited view of reading, and is disconnected from the social /cultural view of literacy. Indeed, the types of texts used in the MUET Reading Comprehension component are not contextualised, and the reader has no purpose for reading other than to extract the information from the text. There is no choice of responding to the text individually.

The use of MCQ items implies that there’s always only one right answer, one interpretation of the meaning of anything in the text. In fact, some researchers argue that the ability to answer MCQ is by itself a separate ability, different from the reading ability. There may be some truth in that since many teachers do resort to teaching test-wisness strategies where MCQ is concerned. Kern (2000:275) notes too that the MCQ (and true/false item types) often is problematic among the *best* readers, who find some logical flaw or are confused by the only partial adequacy of the possible answers provided.

The MUET and Literacy

The MUET is perceived, albeit with some ambivalence, to be a high-stakes test. The teachers have a positive regard for the test and faithfully follow the “syllabus” (which is really the test specifications). Indeed, the tail does wag the dog. And in the context of Malaysia at this time in history, when there is political encouragement and support for the learning of English, why is literacy seen to be on the wane among undergraduates? Shouldn’t the coaching for the MUET have prepared students adequately for reading at the university?

A test may be a good and reliable test, but it will have little impact on learning and improving learners’ competence in real and meaningful ways if the construct underlying it is a mere shadow of the real thing. Grabe (cited in Alderson, 2000:206) notes that “although our understanding of reading has advanced considerably over the past fifteen years, this has not affected the assessment of reading”. Testing authorities and test designers “should be aware that their tests reflect their model of the nature of reading, and they should thus seek to ensure that they reflect and build upon what recent research suggests about the process and the product of reading” (Alderson 2000:120).

I believe that the main problem lies in the way the reading construct is operationalized in the MUET. The choice of assessment is based on the choice of the model of reading.

Education authorities and policy makers are overly concerned with psychometric reliability and statistics. If the MUET is to be used as a university admissions test, then the starting point surely, at least for Reading, should be an examination of the kinds of reading (text types, genre awareness, as well as purposes for reading) that are required in the university, and the specific kinds of skills needed to read in these ways. Most importantly, all those involved in test design and construction and in teaching should operationalize reading as a literacy goal. Literacy should be the long-term goal.

At present, the use of the MUET as a lever for change is achieving nothing more than mere compliance. Indeed we should learn from examples elsewhere that using high-stakes tests to drive the curriculum could be “dangerous driving”(Gregory Anrig, President of ETS till 1993). The MUET , as it is at present, cannot be expected to improve the proficiency of students who, for the past eleven to thirteen years, have not learnt the language in order to *use* it but merely to take an exam. For the MUET to have any impact on the way reading is taught, the testing authorities would have to review the construct of reading underlying the test and decide on a more dynamic way of operationalizing that construct.

References

Alderson, J.C. (2000) *Assessing Reading*. Cambridge: Cambridge University Press

Alderson, J.C. and Hamp-Lyons,L. (1996) TOEFL preparation courses: a study of washback. *Language Testing*13 (3), 280-297

Alderson, J.C. and Wall, D. (1993) Does washback exist? *Applied Linguistics* 14 (2). 115-129

Bachman, L.F. and Palmer, A.S. (1996) *Language Testing in practice* Oxford: Oxford University Press

Cheng, L. (1997) How does washback influence teaching? Implications for Hong Kong. *Language and Education* 235 11(1), 38-54

Frederiksen, N. (1984) The Real Test Bias *American Psychologist*, March, 193-202

Gipps, Caroline V. (1994) *Beyond Testing: Towards a theory of educational assessment*. London: The Falmer Press

Heath, S.B. (1991) 'The Sense of Being Literate : Historical and Cross-cultural Features' in P.D. Pearson, R.Barr, M.L. Kamil, and P.B.Mosenthal (Eds) *Handbook of Reading Research Vol II*. New York: Longman, 3-25 Kern, Richard (2000) *Literacy and Language Teaching*. Oxford University Press.

Latham, H. (1877) *On the Action of Examinations Considered as a Means of Selection*. Cambridge: Deighton, Bell

Lee, King Siong (2004) Reading into our Reading Tests in Wijayasuriya, Basil and Fatimah Hashim (Eds) *Reading Connections* Petaling Jaya: Sasbadi Sdn Bhd ,128-140

Lee, King Siong and Wong, Fook Fei (2001) Making washback work. *The English Teacher (MELTA) Vol XXIX March 2001*, 12-22

Li, X. (1990) How powerful can a language test be? *Journal of Multilingual and Multicultural Development* 11 (5), 393-404

Madaus, G.F. (1988) The influence of testing on the curriculum. In Travers L. (ed) *Critical Issues in curriculum (87 th yearbook of the Society for the study of Education) Part I*. Chicago, IL: Chicago University Press, 83-121

Messick, S. (1989) Validity. In Linn, R.L. (ed) *Educational Measurement* 3 rd edition, American Council on Education/Macmillan, New York, 13-103

Popham, J. (1987) *The merits of measurement driven instruction*. Phi Delta Kappa May, 679-682

Shohamy, E. (1993) The power of tests:the impact of language tests on teaching and learning (*NFLC Occasional Papers*) The National Foreign Language Centre, Washington D.C.

Shohamy, E., Donitsa-Schmidt, S., Ferman, I. (1996) Test impact revisited: washback effect over time. *Language Testing* 13 (3) 298-317

Street, B.V. (1984) *Literacy in Theory and Practice*. Cambridge: Cambridge University Press

Spolsky, Bernard (1995) *Measured Words: The Development of Objective Language Testing*. Oxford: OUP

Sunday Star, Malaysia 27 July 2003

Wall, D. and Alderson, J.C. (1993) Examining washback: the Sri Lankan Impact Study. *Language Testing* 10 (1), 41-69

Venezky, R.L. (1990) Definitions of Literacy. In Venezky, R.L. et al (Eds) 1990 *Towards Defining Literacy*. Newark, D.E. International Reading Association

Urquhart, A.H. and Weir, C.J. (1998) *Reading in a second Language: Process, Product and Practice*. Longman.

Wallace, C. (1992) *Reading*. Oxford: Oxford University Press

Watanabe, Y. (1996) Does grammar-translation come from the entrance examination? Preliminary findings from classroom-based research. *Language Testing* 13 (3), 319-333

Biodata

Lee King Siong is an English lecturer with Universiti Kebangsaan Malaysia. Her main research interests are in language testing particularly testing as social practice, and its impact on teaching and learning), and in reading (particularly critical reading and literacy issues).