

How is English Assessed at Thai Schools?

Richard Watson Todd
Faculty of Liberal Arts
King Mongkut University of Technology Thonburi
Thailand

Abstract

In the test-centric Thai education system, results on national exams are often viewed as indicators of educational success. These exams use multiple-choice which can have detrimental effects on students' attitudes and learning. If school assessments also rely on multiple-choice exams, the situation would be worrying, yet there is little data available on how English language courses at Thai schools are evaluated. This paper presents the initial results of a large-scale survey of the assessment practices of several hundred Thai teachers. On average, exams and quizzes account for 64% of marks for a course, and 60% of exam scores come from multiple-choice (meaning that multiple-choice accounts for 27% of marks). Furthermore, a high proportion of teachers use some innovative and creative assessment practices, providing chances for further positive developments. To promote more positive impacts from assessment, a greater proportion of the overall scores should be assigned to innovative formative assessment.

Keywords: assessment, Thai schools, exams, continuous assessment, multiple-choice

Introduction

The consequences of testing and grading students are immense. They can determine the colleges students attend, the careers open to them, and the lifestyles they ultimately maintain.

Arends (2007, p. 247)

As this quotation shows, assessment can have a massive impact on students, yet it is surprising how little is known about how students are assessed by educational institutions in Thailand. In the most recent UNESCO report on Thai education (OECD/UNESCO, 2016), there are numerous criticisms of current assessment practices in Thai schools, but the report also

acknowledges that it has no data on the actual assessment practices of Thai teachers and identifies this as a useful area for investigation. This paper aims to provide some preliminary information about the assessment of English language courses at Thai schools. To understand teachers' assessment practices, we first need to understand the context.

Assessment policy in Thailand

The key policy document governing Thai education is the National Education Act of 1999 (Office of the National Education Commission, 1999). Section 26 of the Act specifically focuses on assessment:

Educational institutions shall assess learners' performance through observation of their development; personal conduct; learning behaviour; participation in activities and results of the tests accompanying the teaching-learning process commensurate with the different levels and types of education.

A key implication of this section of the Act is that multiple methods should be used to assess students with five specific methods suggested.

Further guidance in how to assess students can be found in the Basic Education Core Curriculum (Basic Education Commission, 2008). This document lists five main ways to use assessment results:

1. To inform students of their progress
2. As information for improvement and development
3. As information for setting study plans and career goals
4. To issue documents for educational qualifications
5. To inform authorities to aid in developing policies

These purposes of assessment appear to be highly idealistic, and whether the actual assessment practices used in schools allows these purposes to be achieved warrants examination.

UNESCO in their report on Thai education (OECD/UNESCO, 2016) suggest a different list of purposes of assessment:

1. To inform decisions about student retention and promotion and about grouping students
2. To compare individual schools against other schools or national standards
3. To monitor schools' progress
4. To make judgments about teachers' effectiveness

5. To identify areas of the curriculum to improve

In contrast to the purposes given in the Core Curriculum, these purposes appear to be more grounded in reality, but again this needs investigation.

The main Thai educational policy documents, then, give guidance on how to conduct assessment and on how assessment should be used. Without knowing how assessment is actually practised in Thai schools, the value of this guidance is unclear.

National-level assessments in Thailand

The Basic Education Core Curriculum of 2008 identifies four levels at which assessment is conducted in Thailand: the national level, the district level, the school level, and the classroom level. Although the main focus of this paper is on how individual Thai teachers assess their students' English (i.e. the classroom level), we need to first consider national-level assessment in Thailand for two reasons. First, the Thai national-level assessments have been the focus of a substantial amount of previous research and thus much more is known about this level than the other levels. Second, since Thailand has a test-centric education system (Watson Todd & Shih, 2014), the main national exams have a large influence on how assessment is conducted at other levels.

School students in Thailand are assessed at the national level every three years. The main assessment instrument is a set of exams termed the Ordinary National Educational Test (ONET) which aims to measure achievement from school learning. The ONET exams for English consist of multiple-choice items only, a fact that has wide-reaching implications for English language education at Thai schools.

Tests, and especially national-level tests such as ONET, influence the behaviour of teachers and students so that they do things that they would not do if the tests did not exist (Brown, 2005). Such changes due to tests are called washback (see Cheng, 2014 for an overview). Since national-level tests have the greatest impact on students' lives, these tests have the greatest washback.

The massive washback effects of tests such as ONET can be seen in the results of several studies. In a survey of the problems that 156 secondary school teachers faced, the most

frequently mentioned problem was the influence of ONET which ranked above other apparently serious problems such as large class sizes, students' low proficiency, and excessive workload (Thongsri, Charumanee & Chatupote, 2006). In a study into what actually happens in Thai classrooms, Fitzpatrick (2011) found that much of the class time in the final year of secondary school was devoted to preparing students for ONET which took precedence over meeting the expected curriculum objectives.

The fact that national-level tests in Thailand create so much washback is not in itself a cause for concern since washback can be either positive or negative; in other words, washback can have either beneficial or detrimental impacts on the teaching-learning process. Unfortunately, in the case of Thailand, the washback from the national-level tests is overwhelmingly negative (OECD/UNESCO, 2016) and this is largely because of the excessive reliance on multiple-choice items in the tests.

The dominance of multiple-choice testing and directions for change

Multiple-choice tests have the benefits of usually being reliable and practical, and it is possible, if difficult, to design multiple-choice items that assess higher-order thinking (see e.g. Scully, 2017) and so potentially have positive washback effects. Nevertheless, most multiple-choice tests are severely limited in what they assess. Because of this, the use of multiple-choice as an assessment tool is generally associated with negative washback effects (Brown, 1997, 2005; Brown, Bull & Pendlebury, 1997; Burke, 1999; Forsyth, Jolliffe & Stevens, 1999), including the following:

- Teaching focused on vocabulary and grammar only
- Heavy emphasis on receptive skills
- Promotion of rote learning
- Minimal focus on higher-order thinking skills
- Teaching of simplistic knowledge only
- Encouraging students to see themselves as knowledge seekers, not understanding seekers

There is evidence that validity and design problems with the multiple-choice national-level tests cause these negative washback effects (Goodman, 2012; Nonthaisong, 2015; OECD/UNESCO, 2016). If these negative impacts were restricted to the national-level tests while classroom-level assessment promoted beneficial washback, the challenges facing Thai

education would not be too serious. However, there is some evidence that schools may use national-level tests such as ONET as a model when designing their own assessments (Watson Todd, 2008), and it is unlikely that teachers who have generally received little training in assessment would be able to design multiple-choice tests with positive washback. For these reasons, some large-scale evaluations of the Thai educational system have pinpointed the wide prevalence of multiple-choice testing as a major problem.

In 2000, Chulalongkorn University was commissioned to conduct a large-scale evaluation of the Thai educational system. One of the eight key recommendations of the report was that there was excessive use of multiple-choice testing. This recommendation does not appear to have led to any meaningful change, since in its 2016 report on Thai education (OECD/UNESCO, 2016) UNESCO highlighted a continuing “overreliance on summative testing” (p. 174) which focuses on “the reproduction of discrete knowledge” (p. 165) through the use of multiple-choice testing. They strongly recommend a shift away from traditional approaches “towards a broad mix of assessments that measure – and thus value – the application of knowledge and the development of a broad set of skills” (p. 165) and argue for “the innovative use of formative assessment techniques” (p. 174).

While such recommendations appear sound and progressive, they are largely based on evidence from the national-level tests. Information about the content and item types used at the national level is readily available, but how assessment is conducted at the classroom level is far less known. The UNESCO report admits that there is little information available about classroom-level assessment practices and calls for research into the area.

The only previous research into the assessment practices for English language courses at Thai schools that I am aware of is Piboonkanarax (2007). A questionnaire was sent to English language teachers around Thailand to elicit the proportions of overall course scores assigned to various assessment methods such as final exams, assignments and homework, and to find out the item types frequently used. The main findings from 78 responses are that test-style assessments (final exams, mid-term exams and quizzes) accounted for 59% of marks with continuous assessment contributing 41%. Multiple-choice was the most frequent item type used in the tests.

While the Piboonkanarax study provides useful information, the number of schools investigated was low and the situation may have changed in the past 12 years. Conducting a similar study with a larger number of schools would go some way to answering the UNESCO call for research into assessment practices at Thai schools. This paper therefore aims to find out how classroom-level assessment is conducted for English language courses at Thai schools. As part of a large research project, a survey was conducted to collect data on the methods used to assess English. The results from the quantitative section of the survey are presented in this paper.

Methodology

To collect information about classroom-level assessment practices for English language courses at Thai schools, a survey was conducted. To obtain a large sample of English language teachers, permission was sought to contact English language teachers undergoing the boot camp training run by the British Council (in total, this project trained over 15,000 English teachers). Consent was obtained allowing us to contact all teachers in the last three cohorts of the boot camp project. E-mails were sent to 4,020 English language teachers asking them to complete an online survey into their assessment practices.

The questionnaire used is available at <https://goo.gl/forms/TS72vPULq96B3hFE2>. It consists of three main sections. First, some personal information including location of school and years of experience was collected. Second, for one course taught, the teachers were asked to indicate what types of assessment they used and the proportions of overall score for each type. For this section, several choices were given: final exam, mid-term exam, assignment, quiz, portfolio, homework, participation, and attendance, with teachers free to insert other assessment types if they needed to. For the first three of these choices, teachers were also asked to give percentages of overall scores for each of the item types they used, and choices of item types were provided (for the exams, these include multiple-choice and matching; for assignments, these include reports and role-plays). Third, teachers were asked to describe in detail one assessment practice that they were proud of and would like to share with other teachers. This paper focuses primarily on the results from the second section, but also investigates whether certain variables from the first section are related to the responses given in the second section.

In total, 347 teachers from 337 schools responded to the survey. 152 of these teachers worked at primary schools, and 195 at secondary schools. On average, they had 12 years of teaching

experience with a minimum of 1 year and a maximum of 42 years. Of these, 320 responses were usable (for other teachers, scores for courses did not add up to 100%). For the questions in the second section of the questionnaire concerning proportions of scores assigned to each type of assessment, frequency of teachers using the assessment type, mean proportions and standard deviations were calculated. In addition, assessment types were categorised in various ways, such as whether they were open or closed assessment types, to allow us to see broader trends. The information from the first section was compared to the proportions given in the second section using point biserial correlation to see if there were differences in approaches to assessment between schools in Bangkok and surrounding provinces and schools upcountry and between primary and secondary level courses.

Results

From the survey, the most basic finding concerns the proportions assigned to the main assessment types. Table 1 presents these results in terms of the percentage of teachers who used each main assessment type and the average percentage of the overall score for a course assigned to each type.

Table 1 Teachers' use of the main assessment types

<i>Assessment type</i>	<i>% of teachers who use this type</i>	<i>% of total score assigned to type</i>	<i>Standard deviation of score assigned</i>
Final exam	100%	26.6%	6.2%
Mid-term exam	92%	17.8%	8.0%
Quizzes	91%	20.0%	16.3%
Continuous assessment	100%	35.6%	10.5%

From Table 1, we can see that all courses include marks from a final exam and from continuous assessment. The combined contribution of the three test-style assessments (final exam, mid-term exam, quizzes) makes up 64.4% of total scores with continuous assessment contributing 35.6%. This emphasis on test-style assessments is slightly higher than that reported by Piboonkanarax (2007) but comparable, suggesting that little has changed in the last 12 years. From the standard deviations, there is much greater variation in the extent to which different teachers used quizzes than there is for the other main assessment types.

Focusing on the mid-term and final exams, Table 2 shows the proportions of the various item types used (there is no information about item types used in quizzes). The findings show that multiple-choice is the most common item type on exams and accounts for the majority of marks given and for at least 27% of the total scores, suggesting that the national exams may have some influence on how teachers design exams at school.

Table 2 Item types used in exams

<i>Item type</i>	<i>Mid-term exams</i>		<i>Final exams</i>	
	<i>% of teachers who use this type</i>	<i>% of exam score assigned to type</i>	<i>% of teachers who use this type</i>	<i>% of exam score assigned to type</i>
Multiple-choice	95%	57.9%	97%	60.8%
Gap-fill	71%	13.8%	74%	13.4%
Matching	59%	9.5%	63%	9.3%
Short answer	59%	10.0%	62%	9.1%
Sequencing	34%	3.9%	35%	3.8%
Other	21%	4.8%	17%	3.6%

Turning to continuous assessment, Table 3 shows that most teachers use a wide variety of continuous assessment practices. However, in most cases each type of continuous assessment accounts for only a small proportion of the overall score. Focusing on the proportions of scores allocated to each continuous assessment type by those teachers who used them (Table 3 gives proportions for all teachers irrespective of whether they used the assessment type or not), only reports/projects account on average for more than 15% of the total score for a course.

Table 3 Types of continuous assessment used

<i>Continuous assessment type</i>	<i>% of teachers who use this type</i>	<i>% of total score assigned to type</i>
Reports/projects	75%	12.0%
Presentations	85%	11.9%
Role-plays	75%	9.1%
External reading	61%	6.0%
Portfolio	77%	8.5%
Homework	86%	10.8%
Participation	87%	10.9%
Attendance	73%	7.0%
Other types	30%	3.8%

Looking at broader patterns, item types can be categorised as open or closed. Open assessment types are usually productive, and there is no single correct answer. Open types include reports, presentations and some short-answer items. Closed assessment types are usually receptive, and there is a single correct answer. These include multiple-choice items, matching items and homework involving answering closed exercises from a coursebook. Open and closed assessment types evaluate students' language ability. Some assessment, however, such as participation and attendance, evaluates non-linguistic behaviour. Table 4 shows the extent to which open and closed assessment types and measures of non-linguistic behaviour are used. It should be noted that some items cannot be categorised into these types. The findings show that the majority of assessment on English language courses at Thai schools uses closed assessment types.

Table 4 Broader patterns of assessment

<i>Broad assessment type</i>	<i>% of teachers who use this type</i>	<i>% of total score assigned to type</i>
Closed assessment	100%	58.1%
Open assessment	94%	27.1%
Non-linguistic behaviour	90%	5.3%

The findings presented so far have looked at the results from all responses combined. However, some variables from the first section of the questionnaire may influence teachers' assessment practices. For example, we might expect some differences in assessment practices between primary and secondary schools, and between schools in Bangkok (and surrounding provinces) and schools upcountry. Table 5 presents the correlations between education level (primary or secondary) and assessment types. Although none of the correlation values are notably high, it appears that summative assessments in the form of final exams are emphasised at secondary schools. Primary schools, on the other hand, place a greater emphasis on mid-term exams and also use a greater proportion of multiple-choice items.

Table 5 Relationships between education level and assessment types

<i>Assessment type compared to education level</i>	<i>r_{pb} (point biserial correlation value)</i>	<i>p (probability value)</i>	<i>Interpretation</i>
% from final exam	0.13	0.02	Secondary schools rely more on final exams
% from mid-term exam	-0.23	0.00	Primary schools rely more on mid-term exams
% from multiple-choice items	-0.19	0.01	Primary schools rely more on multiple-choice items
% from continuous assessment	0.10	0.07	Not significant

Turning to school location, Table 6 shows the correlations between school location and assessment types. In this case, the only significant relationship is that schools in Bangkok and the surrounding provinces tend to place a greater emphasis on final exams (with upcountry schools using more continuous assessment, although this is not significant).

Table 6 Relationships between school location and assessment types

<i>Assessment type compared to school location</i>	<i>r_{pb} (point biserial correlation value)</i>	<i>p (probability value)</i>	<i>Interpretation</i>
% from final exam	0.14	0.01	Schools in Bangkok rely more on final exams
% from mid-term exam	-0.01	0.86	Not significant
% from multiple-choice items	0.01	0.89	Not significant
% from continuous assessment	-0.07	0.21	Not significant

The comparisons reported in Tables 5 and 6 examine the relationship between characteristics of schools and assessment types. We can also examine whether individual teacher characteristics influence the use of assessment types by comparing teachers' years of experience with the percentages of scores they assign to different assessment types using product-moment correlation. All of the correlation values are very low (for final exams, $r = -0.02$; for mid-term exams, $r = 0.03$; for continuous assessment, $r = -0.07$) and not significant. It therefore appears that the assessment types teachers use are not related to their years of teaching experience.

Discussion

The findings show that, at Thai schools, test-style assessments (exams and quizzes) and closed-ended assessments dominate how the English language ability of students is assessed. This key finding suggests that little has changed since the evaluation conducted by Chulalongkorn University in 2000 or the survey of Piboonkanarax in 2007. However, all the teachers surveyed assign at least a few marks to some form of open-ended continuous assessment. It is not the case, then, that teachers reject open-ended forms of assessment or that they are unfamiliar with this approach; rather, the emphasis on closed-ended tests seems more likely to be due to practicality, social expectations or the influence of national exams. Research into the reasons underlying teachers' choices of assessment practices is needed, and this is the next stage of the current research project.

Turning to the extent to which the results match policy documents, the National Education Act of 1999 calls for multiple methods of assessment to be used. The results show that teachers do use multiple methods, but also that the extent to which different methods are used is imbalanced. There is an overemphasis on “results of the tests” with most other methods given few marks and “observation of their development” apparently ignored. For the purposes of testing, the progressive purposes in the Core Curriculum of using assessment to inform students of their progress or as information for improvement and for setting study plans are very unlikely to be achieved, given that the test-style assessments used for the majority of marks do not usually provide detailed feedback for students to learn from. The Core Curriculum purpose that is more likely to be met is that of issuing documents of qualifications. Similarly, the UNESCO purpose of informing retention, promotion and grouping decisions is likely to be achieved. The preference for using test-style assessments shown in the survey results appears to prioritise assessment for administrative purposes rather than assessment for learning.

With the assessment practices currently used not meeting the recommendations and goals of the policies, what directions are needed to make it more likely that assessment has positive impacts on English language teaching in Thailand? The UNESCO report highlights the need to replace the current emphasis on summative testing using multiple-choice items with a broader mix of assessments, specifically including innovative formative assessments. The survey results show that there is an overreliance on summative testing, but that at the classroom level multiple-choice might be less of an issue since it accounts for only a quarter of marks. The call for a mix of assessments mirrors the National Education Act suggesting that there has been little improvement in the variety of assessment techniques used in the last twenty years. The survey shows that teachers do use formative assessment, perhaps especially at upcountry schools, but that the proportion of marks given to this is still low. To see whether the formative assessment practices are innovative, we need to turn to the third section of the questionnaire where teachers were asked to describe one assessment practice that they were proud of. Although it is beyond the scope of this article to provide detailed results from this section, there is evidence that a high proportion of teachers are producing and using innovative formative assessment practices (307 of the teachers gave details of such practices) and that much of this work is being done at upcountry secondary schools. Overall, the findings show that beneficial assessment practices are being used; the problem appears to be that the use of beneficial practices is far outweighed by the emphasis on test-style assessments. Future efforts

to improve the impacts of assessment on ELT in Thailand, then, need to focus on attaining a more beneficial balance between the currently dominant test-style assessment and the underused formative assessments. To see how this could be done, we need information about the reasons why teachers use the current imbalanced proportions and, as noted above, this is the next stage of the research project. Once these reasons are known, clear directions for improvement can be suggested.

Conclusion

The survey of English language teachers' assessment practices at Thai schools shows that, at the classroom level, assessment is largely test-centric with the majority of scores coming from test-style assessments. Such an assessment system has the potential to have negative impacts on students' learning and Thai education in general, and there appear to have been no clear changes in these assessment practices in the last twenty years. The picture, however, is not totally bleak. All teachers are familiar with and use, albeit to a lesser extent than is desirable, open-ended continuous assessment, and there is some evidence of the use of potentially beneficial innovative formative assessment practices. Such practices can have positive washback, and they may provide a foundation on which a more promising future for English language teaching in Thailand can be built.

Note: This research is part of a project funded by the National Research Council of Thailand entitled 'A survey of English language assessment practices at schools throughout Thailand focusing on practices with positive washback' (Thailand Research Fund Project RDG61H0028). In addition to the author of this paper (the project leader), the research team comprises Asst. Prof. Dr. Saowaluck Tepsuriwong, Asst. Prof. Dr. Wannapa Trakulkasemsuk, Asst. Prof. Dr. Natjiree Jaturapitakkul, Dr. Punjaporn Pojanapunya, Ms. Nawiya Chanchula, Ms. Duangjaichanok Pansa and Mr. Stuart Towns. The contributions of all research team members are gratefully acknowledged.

References

- Arends, R. I. (2007). *Learning to teach* (7th ed.). Singapore: McGraw-Hill.
- Basic Education Commission (2008). *Basic Education Core Curriculum BE 2551 (AD 2008)*. Bangkok: Ministry of Education.
- Brown, J. D. (1997). Testing washback in language education. *PASAA*, 27, 64-79.
- Brown, J. D. (2005) *Testing in language programs*. Singapore: McGraw-Hill.
- Brown, G., Bull, J. & Pendlebury, M. (1997). *Assessing student learning in higher education*. London: Routledge.
- Burke, K. (1999). *How to assess authentic learning*. Arlington Heights, IL: Skylight Professional Development.
- Cheng, L. (2014). Consequences, impact, and washback. In A. J. Kunnan (Ed.). *The companion to language assessment* (chapter 68). New York: John Wiley.
- Chulalongkorn University (2000). *Report on the project to evaluate the development of education at the primary and secondary levels in government and private sectors—science, mathematics and English*. Bangkok: Chulalongkorn University Academic Service Centre.
- Fitzpatrick, D. (2011). *Making sense of the English language policy in Thailand: An exploration of teachers' practices and dispositions* (Unpublished PhD thesis). University of Exeter.
- Forsyth, I., Jolliffe, A. & Stevens, D. (1999). *Planning a course*, 2nd edition. London: Kogan Page.
- Goodman, J. E. (2012). *Testing Thais: Establishing the ordinary in Thai national exams* (Unpublished MA thesis). Cornell University.
- OECD/UNESCO (2016). *Reviews of national policies for education: Education in Thailand*. Paris: OECD Publishing.
- Office of the National Education Commission (1999). *National Education Act of B.E.2542 (1999)*. Bangkok: Office of the National Education Commission.
- Nonthaisong, K. (2015). *The impact of the English language policy in the Basic Education Core Curriculum B.E. 2551(A.D. 2008) on English instruction of Thai secondary EFL teachers: A multiple case study in northeast Thailand* (Unpublished PhD thesis). University of Alabama.

- Piboonkanarax, K. (2007). *A survey of secondary school evaluation procedures focusing on continuous assessment* (Unpublished MA thesis). King Mongkut's University of Technology Thonburi, Bangkok.
- Scully, D. (2017). Constructing multiple-choice items to measure higher-order thinking. *Practical Assessment, Research & Evaluation*, 22(4), 1-13.
- Thongsri, M., Charumanee, N. & Chatupote, M. (2006). The implementation of 2001 English language curriculum in government secondary schools in Songkhla. *ThaiTESOL Bulletin*, 19(1), 60-94.
- Watson Todd, R. (2008). The impact of evaluation on Thai ELT. *Selected Proceedings of the 12th English in South – East Asia International Conference: Trends and Directions* (pp. 118 – 127). Bangkok: King Mongkut's University of Technology Thonburi.
- Watson Todd, R. & Shih, C. (2014). Assessing English in Southeast Asia. In A. J. Kunnan (Ed.) *The companion to language assessment* (chapter 102). New York: John Wiley.

Author's biography

Richard Watson Todd is Associate Professor in Applied Linguistics at King Mongkut's University of Technology Thonburi. He has a PhD from the University of Liverpool. He is the author of over 100 articles and several books, including *Classroom Teaching Strategies* (Prentice Hall, 1997), *Much Ado about English* (Nicholas Brealey Publishing, 2006), and *Discourse Topics* (John Benjamins, 2016).

Contact information

Name: Richard Watson Todd

Institution: King Mongkut's University of Technology Thonburi

Mailing address: School of Liberal Arts, King Mongkut's University of Technology Thonburi, 126 Pracha Uthit Rd., Bang Mod, Thung Khru, Bangkok 10140, Thailand

E-mail address: irictodd@kmutt.ac.th