Potential Guidelines for Implementing Performance-Based Assessment for Thai EFL Secondary School Teachers

Pariwat Imsa-ard
Ramkhamhaeng University

Abstract

In education settings, it is widely accepted that teachers play a significant role in improving the students' performance. Despite the fact that teachers, especially in terms of their qualities, are so important in developing students' proficiency, previous literature demonstrated that a number of poorly trained teachers in classrooms have adverse effects on students' learning development. In addition, many teachers still concentrate on pencil-and-paper tests ascribed to the high-stakes tests such as the national examination, which leads to negative washback on teaching and learning. To yield positive washback, the performance-based assessment (PBA), which gathers students' performance over a period of time, has potential as an approach. In fact, teachers can elevate and support students' English language performance through the use of PBA which is designed to measure students' language skills in a real-life context. However, teachers' lack of practical assessment knowledge and insufficient support with effective training development programs can lead to avoidance of implementing PBA effectively. Thus, this paper aims to provide the potential guideline for Thai EFL secondary school teachers to implement PBA in classrooms.

Keywords: performance-based assessment, professional development, secondary school teachers

Introduction

Over the decades, English education policy in Thailand has focused on English language skills for communication and encouraged teachers to promote communicative competence which emphasizes the ability to use the language effectively in authentic contexts (Richards, 2006) with students. However, despite considerable effort and incentives devoted to English education development, the students' English-language proficiency still remains low and is far from satisfactory (Kaur, Young, & Kirkpatrick, 2016). In spite of the fact that Thai EFL students have spent at least nine years studying English in their basic education, the results from the national examination and some other high-stakes tests, for instance TOEIC (Educational Testing Service, 2019), are still in doubt and the students' ability to communicate and use English in their real-life context is still trailing far behind other countries in Asia where English is a foreign language (English First, 2020; Waluyo, 2019).

A possible explanation to such a phenomenon is that Thai EFL students have little exposure to English language and, worse still, there are a number of poorly trained teachers in classrooms who eventually start to fossilize their teaching practices into rote-learning, spoon-

feeding, and teacher-centered approaches (Dhanasobhon, 2006; Imsa-ard, 2020; Noom-ura, 2013). In fact, it is widely accepted that one of the most crucial factors to support student learning progress is unquestionably teachers, especially in the aspects of teacher quality and teachers' beliefs (Geringer, 2003). To support this claim, Borg (2011) demonstrated that teaches' beliefs and knowledge offer a basis for action that leads to teachers' decision-making in the classroom, that is to say, teachers' beliefs can direct teachers' choices of practices such as implementing lessons, selecting learning activities, and assessing students (Rios, 1996).

The aforementioned practices such as focusing on the rote-learning and teacher-centered approaches, which are likely the result of the importance placed on standardized tests such as the national examination (Imsa-ard, 2020), result in a wide gap between the current position and the expected English-language performance. Moreover, despite the fact that the emphasis in the Basic Education Core Curriculum has been placed on "building the communicative competence" (Ministry of Education, 2009, p. 252), teachers' assessment methods do not correspond to this emphasis. To elaborate, Thai EFL teachers have been encouraged to adopt communicative methods and make their teaching interactive and relevant to the students' real-world situations, but they still focus on traditional paper-and-pencil tests. By focusing on the traditional paper-and-pencil tests, it has often been noted that, in spite the fact that students may perform well on the test, they may have no actual ability to use the target language in real-life situations. However, students should be placed into a situation where they produce the target language to complete an authentic task (Lim & Griffith, 2011).

In the light of the discrepancy between the actual and expected performances, there has been a growing interest in performance-based assessment in language teaching and assessment to encourage students to improve their English-language performance (Koné, 2015). Some previous studies (e.g., Inprasit, 2016; Imsa-ard, 2020) indicated that Thai EFL teachers lack a clear concept of how to promote English communicative skills and have insufficient support with effective training development programs. Moreover, some teachers even expressed their concerns about their confidence in implementing performance-based assessments in their own classrooms due to the unpalatable fact that they did not know how to do it. Such reports are in line with Stoynoff and Coombe (2012) who claimed that teacher's lack of preparation and professional development can underrepresent appropriate language assessment. For these reasons, this paper aims to provide a potential guideline for Thai EFL secondary school teachers to implement performance-based assessment.

Performance-Based Assessment (PBA)

With the advent of the communicative language teaching (CLT) approach, there have been some shifts that raise strong objections to using merely standardized test scores by de-emphasizing large-scale standardized tests (Brown & Abeywickrama, 2019) and language teaching has been encouraged to emphasize students' performance rather than merely discrete-point test items in the traditional assessment. Concerning such standardized tests as the national examination and TOEIC, they comprise only multiple-choice questions without performance tests focusing on productive skills. This test format has had a considerable impact on stakeholders such as students, teachers, and even school principals. To illustrate, drilling and test-wiseness strategies are more concentrated on than practicing communicative language

(Prapphal, 2008). In addition, Brown and Abeywickrama (2019, p. 18) demonstrated the standardized tests cannot "elicit actual performance" of the students through discrete-point items, while performance can be systematically evaluated through direct observation.

While the traditional assessment concentrates on "the rank ordering of students, […] individual test performance', scholars have recently proposed alternative assessment which focuses on 'an investigation of developmental sequences in student learning, […] and the provision of an opportunity for further learning" (Lynch, 2001, pp. 228-229). To elaborate, the alternative assessment involves students' active participation as part of their learning development process. Alderson and Banerjee (2001) further illustrated that the alternative assessment is an on-going process which gathers students' performance over a period of time and is claimed to yield positive washback. Concerning positive washback, when the assessment practice focuses on authentic and direct ways of performing language skills rather than the activities involved in preparing for the test, the assessment is likely to yield optimal positive washback by allowing teachers to enhance students' learning process and motivate students to learn and apply their language skills and knowledge to real-world situations (Messick, 1996; Phongsirikul, 2018). To involve students in actually performing the tasks, performance-based assessment is designed to pose tasks "that are based directly on the learners' intended (or hypothesized) use of the target language" (Bailey, 1998, p. 215).

Having discussed the alternative assessment, it is necessary to discuss PBA, which falls under the blanket term of alternative assessments (Herman et al., 1992). The term "performance-based assessment," is not a novelty, as it has been in use for a few decades (McNamara, 1996; Norris et al., 1998). McNamara (1996) defines PBA as "the assessment of the actual performances of relevant tasks are required of candidates, rather than the more abstract demonstration of knowledge" such as paper-and-pencil tests (p. 6). To illustrate, PBA requires test-takers to perform particular tasks designed to measure their language skills presumably required in a real-life context (Wigglesworth & Foster, 2008). Wiggins (1990) indicated four key elements of the PBA: 1) having collaborative elements, 2) being contextualized and complex, 3) measuring real-world tasks, and 4) having standards that are authentic and clear to students. Later, O'Malley and Valdez Pierce (1996, p. 5) and Norris et al. (1998, pp. 9-10) provided some characteristics of PBA as follows:

1. Constructed Responses: students need to produce a response, engage in a performance, or create a product, which is aligned with the daily actions in the language classroom.
2. Higher-order Thinking: students engage higher levels of thinking in constructing responses to open-ended questions given based on needs analysis including student input in terms of rating criteria, content, and contexts.
3. Authenticity: tasks are meaningful and engaging that reflect real-world contexts. In this matter, the performance-based tasks should be as authentic as possible with the goal of measuring real-world activities.
4. Integrative: tasks should call for integration of language skills and knowledge and skills across other content areas and include collaborative elements that stimulate communicative interactions.

5. Process and Product: procedures and strategies used to achieve the tasks are assessed. This should be appropriate in terms of timing and frequency of assessment.
6. Depth VS. Breath: performance-based assessments provide information in depth about students' performance. Unlike standardized testing which cannot elicit actual performance of the students, the information from performance-based assessment can yield more detailed interpretation and higher content validity.

As PBA focuses on authenticity and the tasks should be "authentic," it is crucial to understand the criteria and the characteristics of authentic activities. The National Capital Language Resource Center (n.d., as cited in Griffith & Lim, 2012, pp. 2-3) noted the criteria of authentic assessment as follows:

- Be relevant to students' interests;
- Reflect real-world contexts;
- Involve creative use of language rather than memorization and repetition;
- Require students to produce a quality performance;
- Students know the evaluation criteria and standards prior to the tasks;
- Involve interactions;
- Allow self-evaluation as students proceed.

Moving on now to consider the several washback effects of PBA, there are some potential advantages of PBA. First, compared to traditional assessments, PBA provides more valid measures of students' abilities responding to real-life situations. To elaborate, it offers more effective estimates and more reliable predictions of students' performances in real-life contexts (Quyên, 2019). Furthermore, by using PBA in classrooms, teachers can directly observe how well students apply desired skills and knowledge and diagnose the students' strengths and weaknesses in support of their learning (Espinosa, 2015; McTighe & Ferrara, 2011; Pierce, 2002).

Bailey (1998) and Hamp-Lyons (1997) further contrast PBA and standardized tests, as shown in Table 1, while specific examples of these two types of tests are shown in Table 2. PBAs have demonstrated some positive practices in comparison to standardized tests. For example, as PBA is continuous, a teacher can monitor students' progress so as to promote their learning and enhance their learning motivation. Moreover, with its characteristics of authenticity and contextualization, the performance-based tasks are relevant to students' life so students can relate themselves to contexts given. In addition, the focus of standardized testing rests on developed uniform exams, meaning that the standardized test is administered and marked according to a uniform standard. Unlike standardized tests, PBA allows teachers to plan and modify their practices according to the students' needs in order to boost and encourage students' learning progress (Espinosa, 2015). In addition, PBA is criterion-referenced assessment; a student's performance is assessed against a fixed set of predetermined criteria or learning standards. PBA is therefore less discouraging or stressful for students when compared norm-referenced assessment, in which a student's performance is compared against the performance of his or her peers.

Table 1. Characteristics of standardized tests and performance-based assessment
(adapted from Bailey, 1998, p. 207; Hamp-Lyons, 1997, p. 300)

Standardized tests	Performance-based assessment
Norm referenced	Criterion referenced
Decontextualized objectives	Contextual objective
Uniform	Modifiable
Restricted dimensions	Multi-dimensional
Pre/post "snapshots"	Continuous assessment
Inauthentic tests	Authentic and more real-world assessment
Static view of achievement	Monitors progress
Restricted behavior sampling	Extensive behaviors sampling
Reflects speed and accuracy	Reflects quality of work
Promotes skill in test-taking	Promotes student learning
Promotes student anxiety	Enhances student motivation
Negative washback	Positive washback

Table 2. Example of standardized tests and performance-based assessment

Standardized Test (National Examination)	Performance-Based Test Tasks

Listening and Speaking Part
Directions: Choose the best answer.
#1: A foreign tourist wants to take a bus to Kanchanaburi, so he asks a man for directions to the bus terminal. The man says: _____
  a) Let's take a taxi to Kanchanaburi.
  b) There are many buses going to the bus terminal.
  c) You need to go by train because it's faster.
  d) It's on your left, just after you pass the next intersection.
  e) Kanchanaburi's about 120 kilometers from Bangkok.
	Oral Performance and Listening Tasks
	Direction: You want to visit a tourist attraction in Bangkok, but you do not know how to get there. In pairs, you have five minutes to prepare a brief conversation about asking the directions. In one or two minutes, state in complete sentences:
	  a) Where do you want to go?
	  b) How to ask for help (directions)?
	  c) How to get there?

Writing Part
Directions: Read each sentence and choose the alternative that best completes it.
#1: Montri, along with his friends, _____ from Chiang Mai to Bangkok at the end of the year.
  a) is planning to cycle
  b) are planning to cycle
  c) have been planning for cycling
  d) plans for cycling
  e) plan to cycle
	Writing task
	Directions: In a paragraph, prepare a brief plan for your upcoming trip on the weekend.

Despite the aforementioned beneficial aspects, PBA has some limitations that should be taken into consideration. One obvious limitation is that PBA takes time and is energy-consuming to administer (Brown & Hudson, 1998). Another restriction is that it requires the teacher's previous experiences and knowledge to implement it as they lack sufficient knowledge to implement PBA and assess students (Kirmizi & Komec, 2016). To outweigh the advantages and transcend the limitations, proper and adequate professional development training in PBA practices should be provided.

To sum up, PBA refers to testing where test-takers are tested on what they can do in a target language in real-life situations. Moreover, it represents a set of strategies for assessing students' knowledge and skills through the performance of tasks that are meaningful and engaging. This PBA provides teachers with useful information about how well a student can apply knowledge in real-world situations, which goes beyond rote memorization of language rules. Simply put, students are required to perform the tasks that correspond to their real-world contexts and PBA allows a teacher to assess students as they perform actual or simulated real-world tasks.

A Call for the Implementation of PBA in a Thai Context

Previous research (e.g., Darling-Hammond, 2000; Hanushek & Rivkin, 2006) strongly demonstrated that the quality of teachers is one of the key factors that determine students' achievement as the teachers are positioned at the heart of the teaching and learning process that intends to bring about improved learning. Previous literature on assessment practices, however, indicated that assessment practices were problematic because of badly-prepared and inadequate training (Stiggins, 2002; Thong-Iam, 2017). Worse still, some teachers failed to align their practices with their learning objectives and had some misconceptions regarding assessment practices (Black & Wiliam, 1998). For instance, some teachers indirectly assessed students' speaking and listening skills through a paper-and-pencil test because it is simple to administer and tested on the national examination. From such practices, they can reaffirm that effective professional development on PBAs is necessary.

Richards and Farrell (2005) indicated that teacher professional development refers to training activities that focus directly on teachers' current duties and aim at improving teachers' understanding of particular knowledge, skills, and concepts in the hope that teachers can apply what they have learned to their practices in actual classrooms. To promote assessment practices among EFL teachers, some studies (e.g., Chinda, 2009; Koh, 2011; Thong-Iam, 2017) also examined the link between teacher assessment practices and teacher professional development and found that there is a positive link between them indicating that effective teacher professional development can result in effective teacher assessment practices.

Looking at a Thai context, despite the fact that PBAs have emerged over the past few decades aiming to support and increase students' performance, Thai EFL teachers' assessment practices remain limited to traditional methods focused on a paper-and-pencil testing due to the O-NET according to recent studies (Imsa-ard, 2020; Narathakoon et al., 2020). In addition, Imsa-ard (2020) asserted that teachers believe that teaching in accordance with the test format could help raise students' test scores, meaning that they were highly driven by the test. Worse still, Sinwongsuwat (2012, p. 76) affirms that "students' communicative abilities are still assessed by means of paper-and-pencil multiple-choice tests, particularly in large-scale school and university admission exams." Moreover, this finding also suggests that teachers' lack of sufficient and practical assessment knowledge is one of the major problems. Furthermore, Noom-Ura (2013) indicated that teachers needed professional development, particularly in instructional/pedagogical strategies for teaching and assessing English-language skills.

However, to date, there is scant research on the implementation of PBA for Thai EFL secondary school teachers despite some studies (e.g., Chinda, 2013, 2014) focusing on

university lecturers. Since teaching and learning English in secondary schools and teaching and learning English at tertiary level are not the same in terms of students' maturity, learning goals, and learning contexts, developing some potential guidelines for implementing PDA in secondary school levels is worthwhile.

General Guidelines to Implement PBA in Secondary School Classrooms

After teachers understand the characteristics of PBA, it is important for them to understand potential ways to create PBA for their students in classrooms. Generally speaking, there are five main steps of planning PBA for classroom units, which are adapted from Airasian (2000), Brualdi (1998), and Hilliard (2015).

1. Define the purposes of performance-based assessment.

It is necessary to establish a clearly defined purpose of the assessment. Teachers may need to ask themselves some key questions: (1) What skill or knowledge am I trying to assess? (2) What type of knowledge is assessed? (3) What should the students know? After having answered these questions, teachers can better select the scenario that best suits their assessment.

2. Design the scenario/theme.

To make the assessment authentic, teachers should select the scenario that is relevant to the defined purposes as well as students' life-life contexts. For instance, the teacher wants to assess speaking skills on how to give directions. In this case, the teacher should expect students to know particular language functions on how to give directions. To establish the scenario, the teacher might establish the theme of the "Amusement Park Navigation" where students are posited in the amusement park and they have to produce language to find the various places in the amusement park.

- *Choose an activity*. In this stage, there are three major factors that teachers should take into account.
- *Time constraints.* Teachers should ask themselves how long this activity can last in the classroom session. This may help the teacher plan and select the appropriate activity within the time constraints.
- *Availability of resources*. Teachers should ask themselves whether students have access to required resources or not. This aspect can cause the failure of the activity if the activity requires some resources (e.g., speakers, or certain learning materials) that are not available.
- *Amount of data.* How much information collected from the assessment practice is sufficient to make an informed decision about the student's performance. This aspect leads to the next stage which refers to setting up the criteria, meaning that what aspects or elements determine the success of students' performance. Also, this aspect refers to the purpose of the assessment in what teachers are looking for in judging students' performance.

Another important point to keep in mind is that not all hands-on activities can be considered as PBA (Wiggins, 1993). PBA requires students to demonstrate and apply their

knowledge and skills in the specified context to complete the task, and the teacher has to ensure that the students do not merely complete a task given by passively completing the task without any attempts to apply their knowledge and skills to do so.

There are a number of PBAs which allow students to perform meaningful and engaging tasks and demonstrate authentic learning. For example, teachers can have students conduct a presentation. This activity requires students to practice their teaching and reporting skills, and to learn how to persuade an audience using the target language. In addition to a presentation, teachers can have students create student portfolios which allow them to collect tasks or assignments over the semester. Self-reflection in the target language may be included which requires them to reflect and make notes of their learning growth. At the end of semester, teachers and students can see the students' progress. Furthermore, doing a project is another activity that is commonly used by many teachers. In this activity, students are required to apply what they have learned and skills by undertaking a project.

3. Establish the criteria and scoring rubric.

When the activity has been selected and designed, teachers need to define which aspects or elements determine the success of students' performance. These criteria may be drawn from the Basic Education Core Curriculum and pre-defined purposes or based on their own decisions. For instance, the indicator in the Basic Education Core Curriculum indicating "Use requests and give instructions and clarifications according to the situation" (Ministry of Education, 2009, p. 265) may be drawn on and partly adapted to be such a criterion for the activity as "Students will be able to give instructions in the given situation." Importantly, the number of criteria should be bounded and finite, as teachers may have some trouble observing every criteria if there are too many.

Since PBA does not have clear-cut right or wrong answers as is the case in such traditional assessments as standardized tests, PBA focuses on how successful the students are at completing the tasks. Importantly, the observable behaviors produced by the students are integral to PBA, and the student are required to demonstrate good performance which masters the underlying competence to perform the task (Fastré et al., 2010). To evaluate the success, moreover, scoring rubrics are crucial. According to Moskal (2000), scoring rubrics refer to descriptive scoring schemes or rating systems that aim at guiding the analysis and determining the processes or products of students' performance. Having scoring rubrics, teachers can define and differentiate each level of proficiency in each criterion. For instance, if the teacher wants to evaluate the aspect of "content delivery," the teacher may adopt such scales as "excellent," "proficient," "apprentice," and "needs for improvement."

It should be noted that there are two main types of scoring rubrics: (1) holistic rubrics and (2) analytic rubrics (Brown & Abeywickrama, 2019). First, holistic rubrics, or overall evaluation, are used to judge the overall quality and performance of the students. This type of the scoring rubric is useful when it is not practical to break down students' performance. Second, analytic rubrics are used to judge specific features of the ability. The decision on the types of scoring rubrics depends on the purpose of the assessment. To create and develop a scoring rubric, based on Brookhart (2013) and Brown and Abeywickrama (2019), there are four potential steps of developing scoring rubrics as follows:

- **Step 1: Define the construct or ability measured in the task.** In this step, teachers may consider the purpose of the assessment to help them establish the ability to be measured. For example, the teachers want to measure the speaking ability of students.
- **Step 2: Indicate observable behavior that demonstrates students' performance.** In this step, teachers specify the skills or behavior that they want to measure. For instance, in the speaking ability, teachers may specify such behavior as (1) pronunciation, (2) delivery, and (3) content.
- **Step 3: Specify levels of performance in each scale.** In this step, teachers describe each criterion at different levels of ability. For example, there are three levels: excellent, proficient, and apprentice.
- **Step 4: Write narrative descriptions for each criterion.** In this step, teachers have to describe what each criterion means using the descriptors. To illustrate, what does it mean by "excellent" and to what degree do students need to perform to reach this level.

Once teachers have finished developing the scoring rubric, the scoring rubric may look similar to the example shown in Table 3.

Table 3: Example of a scoring rubric

	Excellent	**Proficient**	**Apprentice**
Pronunciation	Description of pronunciation at excellent level	Description of pronunciation at proficient level	Description of pronunciation at apprentice level
Delivery	Description of delivery at excellent level	Description of delivery at proficient level	Description of delivery at apprentice level
Content	Description of content at excellent level	Description of content at proficient level	Description of content at apprentice level

Significantly, teachers should share those assessed criteria and scoring rubrics with the students before they complete the task or project, so that they know how they can improve themselves to meet the expected goal.

4. Assess students' performance

In this stage, teachers can assess students' performance based on the criteria and scoring rubrics so as to give feedback in the form of narrative report or numerical data. Moreover, teachers can have students assess themselves (self-assessment) or assess their peers (peer-assessment), which allows students to have the opportunity to reflect on their own work so that they can monitor their own learning progress as well.

Conclusion

It is undeniable that, in improving the students' English language proficiency, teachers play a pivotal role in raising students' performance. Despite a large number of efforts and incentives implemented through English language education policy, student English language

proficiency remains questionable. Although teachers, especially in terms of their qualities, are crucial in developing students' proficiency, previous literature reveals that there are a number of poorly trained teachers in classrooms. Moreover, a number of previous studies indicate that many teachers still focus on pencil-and-paper tests due to continuation of national high-stakes testing. Such practices can yield negative washback which has a negative impact on teaching and learning.

To avoid negative washback, PBA, which gathers students' performance over a period of time, is claimed to yield positive washback. When the PBA requires test-takers to perform particular tasks which are designed to measure their language skills presumably required in a real-life context, teachers can elevate and support students' English-language performance. However, previous pertinent literature indicates some problems in implementing PBA can occur when teachers lack the practical assessment knowledge and have insufficient support with effective training development programs. In light of this, this paper aims to provide a potential guideline for Thai EFL secondary school teachers to implement PBA in classrooms. Taken together, this guideline is hoped to provide some practical guidance for educators and teacher trainers to design effective training on PBA for secondary school teachers, and can be used to develop targeted interventions aimed at developing teachers' assessment knowledge on PBA.

The Author

Pariwat Imsa-ard is a lecturer of ELT at Ramkhamhaeng University, Thailand. He is pursuing a Ph.D. in ELT at Thammasat University. He holds M.A. in TESOL with merit from the University of York, UK. His research interests include language assessment, ELT Methodology, and teacher education. He can be reached at pariwat.i@ru.ac.th.

References

Airasian, P. W. (2000). *Assessment in the classroom: A concise approach*. New York, NY: McGraw-Hill.

Alderson, J. C., & Banerjee, J. (2001). Language testing and assessment (Part I). *Language teaching*, *34*, 213–236.

Bailey, K. M. (1998). *Learning about language assessment: Dilemmas, decisions, and directions*. Boston, MA: Heinle and Heinle.

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, *5*, 7–74.

Borg, S. (2011). The impact of in-service teacher education on language teachers' beliefs. *System*, *39*, 370–380.

Brown, H. D., & Abeywickrama, P. (2019). *Language assessment: Principles and classroom practices* (3rd ed.). Hoboken NJ: Pearson Education.

Brown, J. D., & Hudson, T. (1998). The alternatives in language assessment. *TESOL Quarterly*, *32*, 653–675.

Brualdi, A. (1998). *Implementing performance assessment in the classroom.* Washington, DC: ERIC Clearinghouse on Assessment and Evaluation.

Chinda, B. (2009). *Professional development in language testing and assessment: a case study of supporting change in assessment practice in in-service EFL teachers in Thailand* [Dissertation, University of Nottingham].

Chinda, B. (2013). Consideration in performance-based language assessment: Rating scales and rater training. *PASAA: Journal of Language Teaching and Learning*, *46*, 141–159.

Chinda, B. (2014). Teachers' reactions towards performance-based language assessment. *PASAA: Journal of Language Teaching and Learning*, *48*, 57–88.

Darling-Hammond, L. (2000). Teacher quality and student achievement. *Education policy analysis archives*, *8*, 1–44.

Dhanasobhon, S. (2006). *English language teaching dilemma in Thailand*. http://www.curriculumandinstruction.org/index.php?lay=show&ac=article&Id=539134523&Ntype=7.

Education First. (2020). *EF English proficiency index: A ranking of 100 countries and regions by English skills*. https://www.ef.com/wwen/epi/downloads/

Educational Testing Service. (2019). *TOEIC L&R 2018 report on test takers worldwide.* https://www.ets.org/s/toeic/pdf/2018-report-on-test-takers-worldwide.pdf.

Espinosa, L. F. (2015). Effective use of performance-based assessments to identify English knowledge and skills of EFL students in Ecuador. *Theory and Practice in Language Studies*, *5*, 2441–2447.

Fastré, G. M. J., Van der Klink, M. R., & Van Merriënboer, J. J. (2010). The effects of performance-based assessment criteria on student performance and self-assessment skills. *Advances in Health Sciences Education*, *15*, 517–532.

Geringer, J. (2003). Reflections on professional development: Toward high-quality teaching and learning. *Phi Delta Kappan*, *84*, 373–380.

Griffith, W. I., & Lim, H. Y. (2012). Performance-based assessment: Rubrics, web 2.0 tools and language competencies. *MEXTESOL Journal*, *36*, 1–12.

Hamp-Lyons, L. (1997). Washback, impact and validity: Ethical concerns. *Language Testing*, *14*, 295–303.

Hanushek, E. A., & Rivkin, S. G. (2006). Teacher quality. *Handbook of the economics of education*, *2*, 1051–1078.

Herman, J. L., Aschbacher, P. R., & Winters, L. (1992). *A practical guide to alternative assessment.* Alexandria, VA: Association for Supervision and Curriculum Development.

Hilliard, P. (2015). *Performance-based assessment: Reviewing the basics.* https://www.edutopia.org/.

Imsa-ard, P. (2020). Voices from Thai EFL teachers: Perceptions and beliefs towards the English test in the national examination in Thailand. *LEARN Journal: Language Education and Acquisition Research Network*, *13*, 269–287.

Inprasit, C. (2016). *Thai native teachers' experiences of using communicative language teaching in extended education schools* [Master's Thesis, Thammasat University].

Kaur, A., Young, D., & Kirkpatrick, R. (2016). English education policy in Thailand: Why the poor results? In R. Kirkpatrick (Ed.), *English language education policy in Asia* (pp. 345–361). Cham, Switzerland: Springer International Publishing.

Kirmizi, O., & Komec, F. (2016). An investigation of performance-based assessment at high schools. *Üniversitepark Bülten*, *5*, 53–65.

Koh, K. H. (2011). Improving teachers' assessment literacy through professional development. *Teaching Education*, *22*, 255–276.

Koné, K. (2015). *The impact of performance-based assessment on university ESL learners' motivation* [Master's thesis, Minnesota State University].

Lim, H-Y., & Griffith, W. I. (2011). Practice doesn't make perfect. *MEXTESOL Journal*, *35*, 1–9.

Lynch, B. (2001). The ethical potential of alternative language assessment. In C. Elder, A. Brown, E. Grove, K. Hill, N. Iwashita, T. Lumley, T. McNamara, & K. O'Loughlin (Eds.), *Experimenting with uncertainty: Essays in honour of Alan Davies* (pp. 228–239). Cambridge, UK: Cambridge University Press.

McNamara, T. F. (1996). *Measuring second language performance*. New York, NY: Addison Wesley Longman.

McTighe, J., & Ferrara, S. (2011). *Performance-based assessment in the classroom.* https://www.jaymctighe.com/wp-content/uploads/2011/04/Performance-Based-Assessment-in-the-Classroom.pdf

Messick, S. (1996). Validity and washback in language testing. *Language Testing*, *13*, 241–256.

Ministry of Education. (2009). *The Basic Education Core Curriculum B.E. 2551 (A.D. 2008).* http://academic.obec.go.th/images/document/1525235513_d_1.pdf

Moskal, B. M. (2000). *Practical Assessment, Research, and Evaluation.* https://scholarworks.umass.edu/cgi/viewcontent.cgi?article=1086&context=pare

Narathakoon, A., Sapsirin, S., & Subphadoongchone, P. (2020). Beliefs and classroom assessment practices of English teachers in primary schools in Thailand. *International Journal of Instruction*, *13*, 137–156.

Noom-Ura, S. (2013). English-teaching problems in Thailand and Thai teachers' professional development needs. *English Language Teaching*, *6*, 139–147.

Norris, J. M., Norris, J. M., Brown, J. D., Hudson, T., & Yoshioka, J. (1998). *Designing second language performance assessments* (No. 18). Honolulu, HI: University of Hawaii Press.

O'Malley J. M., & Valdez Pierce, L. (1996). *Authentic assessment for English language learners: Practical approaches for teachers*. London, UK: Longman.

Phongsirikul, M. (2018). Traditional and alternative assessments in ELT: Students' and teachers' perceptions. *REFLections*, *25*, 61–84.

Pierce, L. V. (2002). Performance-based assessment: Promoting achievement for English language learners. *Clearing House on Languages and Linguistics*, *26*, 1–3.

Prapphal, K. (2008). Issues and trends in language testing and assessment in Thailand. *Language Testing*, *25*, 127–143.

Quyên, V. P. (2019). Performance-based assessment: Implications to English language classrooms. In P. H. Yen & T. V. S. Nguyen (Eds.), *Using alternative assessment to improve EFL learners' learning achievement: From theory to practice* (pp. 63–79). Hauppauge, NY: Nova Science Publishers.

Richards, J. C. (2006). *Communicative language teaching today*. New York, NY: Cambridge University Press.

Richards, J., & Farrell, T. (2005). *Professional development for language teachers: Strategies for teacher learning*. New York, NY: Cambridge University Press.

Rios, F. A. (Ed.). (1996). *Teacher thinking in cultural contexts*. Albany, NY: State University of New York Press.

Sinwongsuwat, K. (2012). Rethinking assessment of Thai EFL learners' speaking skills. *Language Testing in Asia*, *2*, 75–85.

Stiggins, R. J. (2002). Assessment crisis: The absence of assessment for learning. *Phi Delta Kappan*, *83*, 758–765.

Stoynoff, S., & Coombe, C. (2012). Professional development in language assessment. In C. Coombe, P. Davidson, B. O'Sullivan, & S. Stoynoff (Eds.), *The Cambridge guide to second language assessment* (pp. 122–129). New York, NY: Cambridge University Press.

Thong-Iam, N. (2017). *The impact of language assessment literacy training on EFL university lecturers' language assessment literacy, efficacy and practice of student self-assessment* [Dissertation, Chulalongkorn University].

Waluyo, B. (2019). Examining Thai first-year university students' English proficiency on CEFR levels. *The New English Teacher*, *13*, 51–71.

Wiggins, G. (1990). The case for authentic assessment. *Practical Assessment, Research, and Evaluation*, *2*, 1-3.

Wiggins, G. (1993). Assessment, authenticity, context, and validity. *Phi Delta Kappan*, *75*, 200214.

Wigglesworth, G., & Foster, P. (2008, April). *Measuring accuracy in second language performance*. Paper presented at the TESOL Convention. New York City, New York.