

An Investigation of Global Englishes-Informed Assessment Perceptions and Practices Among Secondary School Teachers in Thailand

Pongpol Suansri and Supong Tangkiengsirisin*

Language Institute Thammasat University, Bangkok, Thailand

*Corresponding author: supong.t@litu.tu.ac.th

Article information	
Abstract	<p>The rising status of English as a global language prompts a call for a paradigm shift in English language teaching in order to correspond to the new sociolinguistic landscape. However, studies related to English language assessment, which can catalyze the transformation especially in a test-oriented context like Thailand, remain limited. This mixed-methods study aimed to explore the perceptions and practices of 61 secondary school teachers regarding adopting Global Englishes (GE)-informed assessment by using a questionnaire, documentation, and semi-structured interviews. It was found that the participants expressed positive perceptions towards adopting GE principles when designing their classroom-based assessment, especially in the domains of speaking, listening, and the involvement of non-native English-speaking teachers in test design and development. However, little were GE principles reflected in their assessment practices. Themes, such as extensive use of standardized proficiency tests and native speakerism in test constructs, which are considered major challenges for successful implementation of GE, emerged. The findings</p>

	contribute to a better understanding of the current status of GE-informed assessments in the context of Thailand and serve as a foundation for further developments in this field.
Keywords	Global Englishes, perceptions, assessment practices, secondary school teachers, Thailand
APA citation:	Suansri, P. & Tangkiengsirisin, S. (2024). An investigation of Global Englishes-informed assessment perceptions and practices among secondary school teachers in Thailand. <i>PASAA</i> , 69, 232–276.

1. Introduction

The widespread use of English has led to the unprecedented increase in English speakers worldwide. English is currently used by an estimated two billion people, and it is an official language in more than 70 nations and territories (Selvi & Yazan, 2021). By 2050, there could be over four billion English users all over the world, of which slightly over half a billion will be classified as L1 English speakers based on current growth rates (Rose & Galloway, 2019). As a result, the 1:3 ratio of L1 to L2 users that Crystal (2003) first suggested will rise to a 1:8 ratio. With this rising number, there are clearly far more non-native English speakers (NNES) than native English speakers (NES).

Such changes in the sociolinguistic landscape result in calls for the reconceptualization of the current status of English and a pedagogical change in English language teaching (ELT) from English as a foreign language (EFL) towards Global Englishes (GE) orientation. This is evident in the abundance of publications relating to GE. Books exploring GE-related fields (Galloway & Rose, 2015; Jenkins, 2014; Matsuda, 2012; Rose & Galloway, 2019) have been written, while theoretical and conceptual frameworks, including World Englishes (WE)-informed ELT (Kachru, 1992; Matsuda, 2020), English as an international language (EIL) curriculum blueprint (Matsuda & Friedrich, 2011), and Global Englishes Language

Teaching (Galloway & Rose, 2015) have been proposed for this pedagogical transformation.

However, a body of research still revolves around the awareness-raising stage, with most focusing on exploring teachers' and learners' perceptions as well as attitudes towards GE (Ahn et al., 2021; Aydin & Karakas, 2021; Lee & Drajadi, 2019; Sifakis, 2014; Vettorel, 2016), GE-aware teacher education (Anderson, 2018; Blair, 2017; Chen et al., 2021; Illes and Akcan, 2016), and the development and evaluation of GE-based classroom materials in various contexts, such as Italy (Caleffi, 2016; Vettorel, 2018) and Japan (Rose & Galloway, 2019).

In Thailand, GE research has explored various areas including teachers' and students' perceptions and attitudes toward GE (Ambele, 2022; Buripakdi, 2012; Jindapitak et al., 2022; Kirkpatrick, 2014; Ploywattanawong & Trakulkasemsuk, 2014; Snodin & Young, 2015; Thienthong & Uthai Korn, 2023), teacher identity and education (Montakantiwong, 2024; Prabjandee, 2020; Prabjandee & Fang, 2022), language policy following the rise of GE (Baker & Jarunthawatchai, 2017; Bennui & Hashim, 2014), and GE classroom implementation (Boonsuk et al., 2021; Rajprasit, 2024). However, research studies pertaining to GE language assessment remain understudied and seemingly unresponsive to ongoing changes. Researchers have even been accused of hesitating to adopt aspects derived from GE studies (Dimova, 2017), as Smaoui & Rahal (2022, p. 2) has pointed out: "No clear picture has emerged about the educational implications of this phenomenon for example in relation to teaching...and most importantly language assessment." Exploring secondary school teachers' perceptions of GE in assessment design and their current practices would pave the way for more openness towards GE, especially in a test-oriented context such as secondary schools in Thailand where large-scale English proficiency tests based on prescriptive language standards have long been adopted among Thai, bilingual, international, and English Medium Instruction (EMI) programs as language proficiency tests, achievement tests, and exit exams. Without any changes in assessment practices, little can GE principles

be adopted in language policies, curriculum designs, instructional materials, or learning activities due to the washback effect that tests generally used in secondary schools in Thailand have on teaching. Therefore, the present study aimed to answer the following research questions:

1. What are the perceptions of secondary school teachers in Thailand towards adopting GE-informed assessment?
2. To what extent is GE-informed assessment incorporated into Thai secondary schools' assessment practices?

2. Literature Review

2.1 GE-informed Assessment

GE-informed assessment is based on the underlying principles of the GE paradigm that shifts away from the traditional way of prioritizing linguistic accuracy over communicative competence to assessing students using the NES varieties/norms as a standard to design test constructs. The purpose of such assessment is to assess success in communication within the new sociolinguistic realities that value multilingualism over monolingualism and native speakerism, as well as intelligibility over accuracy.

In order to design assessment based on GE principles, the term 'proficiency' needs to be reconceptualized. Canagarajah (2006, p. 233) defines proficiency as "the ability to shuttle between different varieties of English and different speech communities." Therefore, language proficiency should no longer be universally defined as a uniform system of language with prescriptive grammatical rules. In order to measure proficiency based on the new definition, a number of guiding frameworks and principles have been proposed.

Hu (2012, p. 134) suggests five principles of test designs when assessing English as an international language as follows: 1) choosing the varieties for a test based on its intended use; 2) selecting a standard variety if more than one variety is adequate for the intended test use in a society; 3) providing test takers with

exposure to several NES and NNES varieties; 4) expanding the test construct in order to incorporate intercultural strategic competence; and 5) allowing individual aspirations to inner-circle norms.

Moreover, two additional approaches including the weak approach and the strong approach have been proposed (Hu, 2012). The former, which is more accommodating, suggests five aspects of test delivery and scoring method. It encompasses 1) checking the NES texts in the test in order to ensure that the contents are likely to be encountered by NNES; 2) avoiding lexical items or providing glosses that might be unfamiliar to NNES; 3) using proficient NNES examiners for oral assessment; 4) using NNES raters that are trained to penalize test takers only on unintelligible errors; and 5) employing NNES in standard settings in order to determine cutoffs that enable proficient NNES to pass the test. The latter, which is more radical, proposes a necessity to incorporate forms and functions of NNES varieties in the test construct.

Moreover, Canagarajah (2006) suggests that test tasks should be performance-based and include social negotiation that requires pragmatic competence. Similarly, Pill and Harding (2013) recommend that the ability to shuttle, negotiate, and use ELF pragmatics to accommodate listeners should be the area of focus when designing GE-informed tests. This corresponds to Newbold's (2014) suggestion that the pragmatics of GE interaction should serve as the foundation for GE-informed assessment, which should be norm-defocused and user-centered. In addition, scholars such as Canagarajah (2006), Hynninen (2014), and Jenkins (2020) agree that language assessment should be locally developed and contextualized, thus resulting in a call for a more context-based and culturally-aware assessment. Similarly, Leung and Scarino (2016) also suggest adopting small-scale assessments that reflect local demands and usage rather than using standardized large-scale tests. Based on the newly proposed definition of proficiency, proficiency tests should be locally designed for specific purposes and functions with specific discourse for specific communities.

2.2 Assessment Practices

Previous studies exploring teachers' assessment practices have revealed that teachers' assessment practices have evolved from traditional approaches of rote memorization and prescriptive grammar rules to more communicative and technology-integrated approaches. Although teachers may adopt performance-based assessment in higher education (Imsa-ard & Tangkiengsirisin, 2023), "teaching the test" remains a common assessment practice among many secondary school teachers in Thailand. It has been reported that secondary school teachers focus more on helping their students pass the university entrance examination (Jianrattanapong, 2011), while students also are more likely to pay attention to their grades instead of their learning (Padermprach, 2017). As pointed out by Watson Todd (2019), Thai EFL teachers tend to adopt summative assessments such as midterm exams, final exams, and quizzes rather than formative or alternative assessment in their class.

The use of standardized English proficiency tests to assess test takers' language competency has raised concerns about the issues of test fairness and inclusivity among GE scholars (Canagarajah 2006; Shohamy, 2011). The widespread use of such tests, especially in assessing international students from non-English speaking backgrounds in the expanding circle, is considered a major hurdle for successful introduction of GE in language classrooms (Deterding & Gardiner, 2019). Such practices have been criticized for prioritizing linguistic accuracy over pragmatic effectiveness (Hu, 2012), positioning English as a language system rather than social practice (Canagarajah, 2006), and emphasizing competence instead of performance (Hu, 2018) and the context of language use (Davidson, 2006).

A shift from formal accuracy to communicative competence or fluency has been witnessed in the assessment practices of standardized proficiency tests. There have been attempts to incorporate NNES varieties into test constructs. The TOEFL iBT speaking scale, for example, emphasizes intelligibility,

comprehensibility, and successful delivery of meaning, while the IELTS has included regional language variations and recruited proficient L2 speakers as oral examiners and raters. The speaking criteria for the Test of English Proficiency of Academic Staff (TOEPAS) have been redefined, deviating from NES norms to allow more accent variation. Moreover, the Cambridge English: Advanced test paper uses a paired-speaking test format that characterizes GE communication, allowing test takers to activate the communicative competency necessary for NNES-NNES interactions. While references to NES competence in the assessment criteria of some standardized proficiency tests are no longer as predominant as they used to be (Kling & Dimova, 2015), criticisms regarding the lack of reference to self-repair, repetition, or paraphrase, all of which refer to communicative competence that facilitates communication success, still exist.

2.3 Criticisms Against Current Use of Standardized English Proficiency Tests

One of the major challenges for the successful implementation of GE lies in the prevalent use of standardized proficiency tests. Test authenticity, construct validity, and washback are three qualities of Bachman and Palmer's test usefulness evaluation (1996), which have been subjects of criticism.

2.3.1 Authenticity

The main criticism against the current use of standardized English proficiency tests lies in the fact that these tests fail to reflect authenticity, which is defined as "the degree of correspondence of characteristics of a given language test task to the features of a Target Language Use (TLU) task" (Bachman & Palmer, 1996, p. 23). The language use in such tests has been shown to underrepresent authenticity due to an over-reliance on NES standards and limited NNES varieties in test tasks.

The absence of authenticity in standardized English proficiency tests in relation to TLU domains, for which proficiency tests are designed, has led to

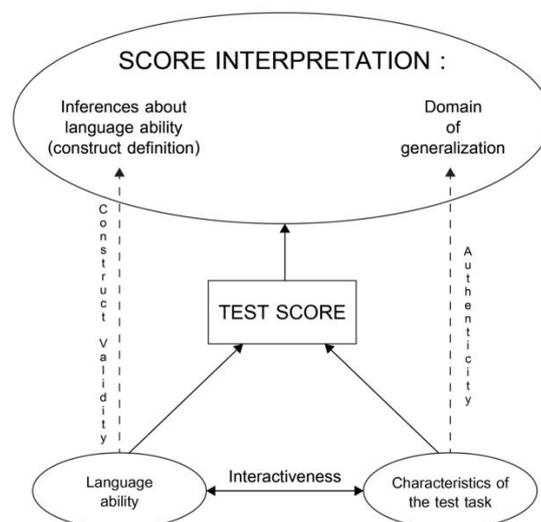
concerns regarding test usefulness and test purposes, especially as a gatekeeper for university admissions in the expanding circle countries. Dimova and Kling (2015) illustrate the discrepancy between the language characteristics of the test tasks used in standardized English proficiency tests and those of the TLU tasks in the EMI context of a Danish university where students are more likely to have Danish speakers of English as their instructors than any of the NES varieties represented in the listening components of the tests. The assessment based on native norms fails to reflect the actual use of English in classrooms, thus tarnishing test authenticity.

2.3.2 Construct Validity

The construct validity of standardized English language tests has also been subject to criticism, with the argument being based on the notion that construct validity and authenticity are interrelated, and authenticity is a necessary condition for construct validity that relates to the domain of generalization. In order to explain the relationship between construct validity and authenticity, Bachman and Palmer's conceptualization (1996) is presented below.

Figure 1

Construct Validity and Domain of Generalization (Bachman & Palmer, 1996, p. 22)



As shown in figure 1, construct validity is related to the domain of generalization of score interpretations. In order to attain the domain of generalization, test tasks have to correspond to the set of tasks in the TLU domain. Therefore, it is important that the interpretations of test takers' language ability be generalized beyond the test condition to a particular TLU. However, the degree of generalizability is impacted by the discrepancy between the language used in the test tasks of standardized proficiency tests and that in the authentic TLU domain. Due to the absence of test authenticity, which is a prerequisite for construct validity, standardized English proficiency tests, whose items are scored in relation to certain varieties, are considered invalid when such tests are used in a context where different varieties are used.

2.3.3 Washback

It has been argued that native norms continue to be prevalent in teaching and learning partly due to the power of washback that the standardized English proficiency tests yield. Brown and Abeywickrama (2019) point out that the widespread use of such tests as a gatekeeper and a model for classroom test designs can result in negative washback effects as students tend to prioritize obtaining acceptable scores over language development. Brown (2012) further highlights that these exam papers serve as the 'templates' for a number of English proficiency tests created and used in diverse NNEs contexts. As a result, the conventional idea about legitimate varieties of English persists due to the constraints imposed by language assessment (Kubota, 2019). In addition, Jenkins (2014, p. 125) also shares a similar remark: "No matter how much effort is put into making English language teaching more appropriate to the context of teaching, if the examination boards continue to measure students' success in English against native norms, then little is likely to change." In short, the usefulness of standardized English proficiency tests, especially in terms of authenticity, construct validity, and washback, has been challenged, and this results in calls for assessment incorporated with GE principles in order to ensure the correspondence

between the current sociolinguistic realities of the 21st century English and language teaching, learning, and testing.

2.4 Related Research Studies

To the best of the researchers' knowledge, only a few previous studies have explored teachers' and learners' perceptions of GE in relation to assessment. For example, Hamid (2014) conducted a study on IELTS test takers' perceptions of different English varieties, and the findings revealed mixed views and complexities. The majority of the test takers denied that there was a linguistic hierarchy and argued for equal status for all varieties. However, when presented with examples of GE features, an overwhelming majority judged those items unacceptable and opposed the incorporation of GE in IELTS in order to maintain test fairness, equality, and standards and to avoid complexity in test design. However, a study by Suebsook (2018), who investigated the perceptions of stakeholders of the inclusion of Asian English accents in listening proficiency tests for Thai healthcare professionals, reveals contradictory findings. The results showed that the stakeholders accepted the incorporation of NNES accents in the English proficiency test as it was believed to yield beneficial effects.

Regarding GE assessment practices, the use of NNES raters in assessing the productive skills of writing and speaking has been the area of focus in previous studies. Concerning writing assessment, Johnson and Lim (2009), who explored NNES rater language background-related bias, found that there was no significant pattern of bias, thus having an insignificant effect on the examinees' scores. Zhang and Elder (2011) and Bogorevich (2018) reported similar findings when assessing speaking performance. With the objectives of comparing the potential differences between NES and NNES raters, both studies reported no significant difference between NES and NNES raters' judgement of the speech samples produced by L2 test takers. These studies show that NES and NNES raters are indistinguishable from one another, so NES should not be exclusively recruited as raters.

With regards to studies related to listening assessment, Abeywickrama (2013) examined the impact of NNES varieties on NNES test takers and the results showed that the performance of the test takers was not affected by the diverse kinds of English utilized as audio input. Similarly, Kang et al. (2018) explored the effect of integrating international English accents into a simulated TOEFL listening comprehension test. Six different English varieties were recorded as speech samples for a simulated TOEFL iBT listening exam; the findings revealed that the listening test scores for stimuli based on outer and expanding circle varieties were not significantly lower than those based on inner circle varieties. These studies suggest that test takers did not perform differently in their listening comprehension tests when the listening inputs contained NNES varieties.

As discussed above, the findings of previous studies on the perceptions of incorporating GE principles into test tasks remain inconclusive. However, the findings concerning assessment practices incorporated with GE principles seem to be more conclusive. From the GE standpoint, studies relating to the assessment practices of both productive and receptive skills appear promising. It has been reported that NNES raters could rate productive skills as accurately as their NES counterparts, and NNES varieties used in listening test tasks had minimal impact on the test takers' performance. However, more studies, especially in the assessment practices of reading skills, are still needed in order to provide a comprehensive view of GE in relation to assessment practices.

3. Methodology

3.1 Research Design

The present study used a concurrent mixed-methods research design. This design allows the researchers to compare the results of quantitative and qualitative data to see if the findings confirm or disconfirm each other (Creswell & Creswell, 2018). A questionnaire and semi-structured interviews were employed to answer RQ 1, while documentation and semi-structured interviews were adopted to yield data for RQ 2.

3.2 Research Participants

A total of 61 non-native English-speaking teachers (NNESTs) from nine private secondary schools in different regions of Thailand participated in the study. Purposive sampling was used to recruit participants for both the quantitative and qualitative phases of the study. Teachers that had at least one year of teaching experience took part in the study. Those with less than one year of teaching experience were excluded from the study as they had limited exposure to the schools' assessment practices.

All of the participants worked for nine schools that belong to an educational foundation known as Saint Gabriel's Foundation of Thailand, which operates a number of schools nationwide. The member schools shared common policies and practices imposed by the foundation. These institutions have long been valued as a leading part of the Thai educational establishment. Their English language teaching is highly reputable, and this is reflected in one of their current mission statements to be qualified as world-class standard schools. Founded and run by a group of French missionary priests, English has been used as the medium of instruction since their first school was established in 1885. With their long history of bilingual heritage, five core subjects, including English, mathematics, science, social studies, and computer and technology, continue to be taught by using English as the medium of instruction. In addition to the bilingual programs, the schools also provide English programs (EP) in which all subjects are taught in English. Students are also encouraged to use English while studying and participating in different co-curricular and extra-curricular activities organized by Thai and non-Thai teachers. These NNEST and NNEST interactions using English as a medium characterize GE communication styles. Moreover, as private schools, they possess certain degree of academic freedom in order to implement any additional international curriculum in addition to the compulsory national Basic Education Core Curriculum B.E. 2551 (Office of the Basic Education Commission, 2008). Due to this academic freedom, the schools use renowned international test papers developed by Cambridge University Press and Assessment, such as B1

Preliminary, B2 First, and IELTS, together with locally designed tests, as their placement, achievement, and proficiency tests. Apart from the midterm and final exams, students are required to undergo a series of assessments throughout the academic year such as dictation, listening, essay writing, and interviews. It is worth noting that with the objectives to promote healthy competition and enthusiasm in learning, these assessments designed by the school teachers have opened up an opportunity for the researchers to explore their perceptions and assessment practices.

3.2 Research Instruments and Data Collection

This study employed three research instruments, including 1) a questionnaire, 2) assessment-related documents, and 3) semi-structured interviews. The questionnaire aimed to elicit data regarding the teachers' perceptions towards adopting GE-informed assessment. It consisted of three sections. The first section gathered the participants' demographic data. The second section, which explored the participants' perceptions of GE-informed assessment, was adapted from the instruments used in three previous studies. First, the instrument used in Grazi and Lopriore's (2020) study on teachers' awareness concerning GE was adopted because the study objective was to investigate the teaching and assessment practices of secondary school teachers. Secondly, the instrument used in Ramadhani and Muslim's (2021) study to explore the perceptions of school teachers' teaching objectives regarding assessment was selected. Thirdly, the instrument used in Thao et al.'s (2022) study was adopted to guide the main content of the questionnaire used in the present study. There was a total of 30 items in the questionnaire arranged on a five-point Likert scale ranging from 1 = strongly disagree to 5 = strongly agree. The questionnaire items could be divided into four areas of secondary school teachers' assessment practices: the perceptions of the current use of large-scale proficiency tests (items 1-2); the perceptions of assessing productive skills while assessing grammar embedded in the writing assessment section (items 3-13); the perceptions of assessing receptive skills while assessing vocabulary incorporated (items 14-27), and the

perceptions of the roles of NNESTs in test development (items 28-30). The final section of the questionnaire contained the open-ended part for short answers.

The second instrument included printed and electronic documents that provided evidence of assessment practices, including curricula, course syllabi, course assessment plans, and course assessment criteria, as well as the objectives and specifications of the examinations. In order to triangulate the data, document analysis was used in conjunction with the third research data collection instrument, which was a set of semi-structured interview questions. Interviews were conducted with six participants who were NNESTs. They were chosen by means of purposive sampling using information-rich cases (Patton, 2002, p. 46). The selection criterion for recruiting the participants was based on the number of responses from the open-ended section of the questionnaire. The interview questions were designed in order to probe further into the perceptions and practices of the secondary school teachers, with the questions being categorized into three parts: the perceptions of the test designs (five questions); the perceptions of standardized English proficiency tests (two questions); and the perceptions of the impact of GE on test designs (four questions). The interviews were conducted online via the Zoom application in English and recorded for subsequent analysis. Since the semi-structured interviews took place after the data were collected by means of the questionnaire and the document analysis, the researchers were able to see if the actual assessment practices in the classrooms were consistent with those indicated in the documents. This also allowed the researchers to elicit additional information if the obtained documents were incomplete or not specific enough. All of the participants were asked to confirm their consent before data collection commenced.

In order to ensure the validity of the instruments, three experts in the field of language assessment quantitatively and qualitatively evaluated the items of the questionnaire. The index of item-objective congruence (IOC) was used to establish the content validity. The items demonstrated an acceptable IOC value (0.5 or

higher). The questionnaire was subsequently piloted with 13 teachers that had characteristics similar to those of the participants of the main study. In terms of reliability, internal consistency was estimated using Cronbach's alpha. The perceptions of the current use of large-scale proficiency tests had an alpha value of .83, while productive skills had an alpha value of .70. The coefficients of .73 and .77 belonged to the perceptions of receptive skills and the role of NNESTs in test development, respectively. For documentation, evaluating and verifying the sources of the obtained documents and confirming the authenticity with the participants were conducted; and in order to ensure the representativeness of the selected documents, the original purpose of each document, the context in which it was produced, and the intended audience (Bowen, 2009) were all considered. Similarly, the interview questions were refined and confirmed by three experts in the field of language assessment before being subsequently tried out with the participants in the pilot study.

3.3 Data Analysis

The data analysis for the first research question consisted of three phases as suggested by Creswell and Creswell (2018). First, the questionnaire results were analyzed using SPSS version 24 to obtain the descriptive statistics of mean scores, standard deviations, and percentages for each questionnaire item. Second, the qualitative data from the semi-structured interviews and open-ended sections of the questionnaire were analyzed using thematic content analysis. Both sets of data were integrated in the final data analysis phase. To answer the second research question, the analytic procedure of documentation, which involved finding, selecting, assessing, and synthesizing the data (Bowen, 2009), was used. After obtaining the documents, the researchers superficially examined the documents in order to determine their relevance to the second research question. This initial screening process was followed by thorough examination and interpretation of the obtained documents. The documentary data were analyzed together with the qualitative data from the semi-structured interviews and open-ended sections of the questionnaire so that themes would emerge across the three

sets of data. For both research questions, thematic analysis was employed for coding and sub-coding construction, which were subsequently categorized into emerging themes. In order to establish the trustworthiness of the qualitative study, the multiple validity procedure (Creswell & Creswell, 2018) was adopted to ensure the accuracy of the findings. Apart from triangulating different datasets to ensure coherent justification for emerging themes, member checking was used. The verification procedure also included intra-coder and inter-coder agreement.

4. Findings

4.1 Secondary School Teachers' Perceptions of Adopting GE-informed Assessment

The majority of the participants reported mixed views towards GE-informed assessment in various areas. Table 2 summarizes the findings into four areas, including the perceptions of the use of standardized proficiency tests, the perceptions of assessing productive skills, the perceptions of assessing receptive skills, and the perceptions of the roles of NNESTs in test design and development.

Table 2

Participants' Perceptions of GE-informed Assessment

Areas	M	SD
Use of standardized proficiency tests	3.63	0.99
Productive skills	3.31	0.44
Speaking skills	3.50	0.60
Writing skills	3.12	0.49
Receptive skills	3.24	0.49
Reading skills	2.97	0.51
Listening skills	3.50	0.68
Roles of NNESTs in Test Design and Development	4.03	0.91

Regarding the use of standardized proficiency tests, the participants agreed that the tests were still useful in assessing students ($M = 3.63$, $SD = 0.99$). The majority of the participants were more likely to adopt GE principles when assessing students and designing test materials for productive skills ($M = 3.31$, $SD = 0.44$). Similarly, the majority of the participants were prone to adopting GE principles when assessing students and designing test materials for overall receptive skills ($M = 3.24$, $SD = 0.49$). Out of the four subscales, the participants were most likely to adopt GE principles in the speaking domain ($M = 3.50$, $SD = 0.60$) and listening domain ($M = 3.50$, $SD = 0.68$), but least likely in the reading domain ($M = 2.97$, $SD = 0.51$). Lastly, the participants were positive about NNESTs playing a role in developing assessment and test materials ($M = 4.03$, $SD = 0.91$).

The following sections further address the quantitative findings in greater detail. In order to complement the questionnaire results, the findings from the semi-structured interview are also presented.

4.1.1 Perceptions of the Use of Standardized Proficiency Tests

The first area addresses the use of standardized English proficiency tests. Table 3 illustrates the results regarding the perceptions towards the use of such tests.

Table 3

Frequency of Participants' Responses by Items Towards the Use of Standardized Proficiency Tests

Items	M (SD)	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
Q1. English proficiency tests such as the TOEFL and IELTS should be used to assess high school students' language ability.	3.72 (1.07)	1 1.64%	8 13.11%	15 24.59%	20 32.79%	17 27.87%
Q2. I use English proficiency tests such as the TOEFL and IELTS as a model for designing my own English test papers for students.	3.54 (1.12)	4 6.56%	6 9.84%	16 26.23%	23 37.70%	12 19.67%

According to Table 3, the majority of the participants showed positive perceptions of the use of large-scale proficiency tests in assessing students' language abilities ($M = 3.72$, $SD = 1.07$) and as a model for designing their own assessments ($M = 3.54$, $SD = 1.12$), with more than half of the participants (60.66% and 57.37% respectively) either agreeing or strongly agreeing with using standardized proficiency test papers to assess students and to be a model for designing their classroom English test papers.

These results aligned with the qualitative data from the semi-structured interviews, as the participants reported that well-accepted English proficiency test papers possessed a standard that allowed them to assess students in a uniform manner, as illustrated in the excerpts below (pseudonyms are used).

Non-native speakers come from many locations, so they have different accents and different ways of using English. When it comes to testing, there has to be a uniformed standard. (Alex)

I used IELTS test samples as a main assessment because I can assess my students in a standardizable way. (Roberto)

4.1.2 Perceptions of Productive Skills

As Table 4 indicates, the majority of the participants had favorable perceptions of GE in assessing the speaking domain.

Table 4

Frequency of Participants' Responses by Items Towards the Speaking Domain

Items	M (SD)	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
Q5. I don't mind students who use a Thai accent during the speaking tests.	3.84 (1.02)	2 3.28%	3 4.92%	16 26.23%	22 36.07%	18 29.51%
Q7. Speaking with grammatical correctness is more important than maintaining the flow of the conversation.*	2.54 (1.12)	10 16.39%	24 39.34%	15 24.59%	8 13.11%	4 6.56%

*Note * indicates reverse-scoring items.*

To illustrate, in Q5, almost two-thirds of the participants (65.58%) reported that they did not mind the Thai accent during speaking tests. This was elaborated during the semi-structured interviews:

I myself do not encourage them (students) to imitate the British or American accent or expect them to speak like a native speaker. What matters more than the accent is correct pronunciation which includes stress patterns and intonation. (Jim)

In terms of fluency and accuracy in Q7, most of the participants (55.73%) viewed that the flow of the conversation was more important than grammatical correctness. An excerpt from the semi-structured interviews reflected this sentiment:

Some of my students were afraid to speak because they were worried that their points would be deducted if they made any grammatical mistakes. So, during a speaking test, their responses were limited and not sufficient for examiners to evaluate. I would prefer students to speak more even with some grammatical errors. (Saheed)

Compared to the speaking domain, the participants were more ambivalent about adopting GE when assessing writing, as illustrated in Table 5.

Table 5

Frequency of Participants' Responses by Items Towards the Writing Domain

Items	M (SD)	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
Q9. Intelligibility is more important than grammatical accuracy when assessing writing.	2.84 (1.08)	8 3.11%	14 22.95%	22 36.07%	14 22.95%	3 4.92%
Q10. When designing rubrics for writing tasks, grammatical accuracy is the most important component.*	3.46 (1.12)	3 4.92%	9 14.75%	18 29.51%	19 31.15%	12 19.67%
Q12. The use of non-native grammatical forms in writing tasks is not considered an error as long as it promotes comprehension.	3.08 (1.16)	6 9.84%	13 21.31%	19 31.15%	16 26.23%	7 11.48%
Q13. As long as the minor grammatical errors do not cause misunderstanding, I do not deduct students' scores.	3.52 (1.11)	2 3.28%	11 18.03%	13 21.31%	23 37.70%	12 19.67%

*Note * indicates reverse-scoring items.*

In Q9 and Q12, most of the participants were ambiguous concerning the importance of intelligibility over grammatical accuracy (approximately 30% neutral,

while another 45% equally agreed or disagreed). In Q10, most of the participants still reported that grammatical accuracy was the most important component when designing a writing rubric (50.82% agreed and strongly agreed). Likewise, the data from the semi-structured interviews were congruent with the quantitative data, as the participants reported the importance of accuracy, as shown in the following excerpt:

Writing is closely related to grammar, and it encompasses rules and convention. I would still adhere to grammatical rules when it comes to assessing students' writing abilities. (Lisa)

However, in Q13, the participants were more relaxed towards the grammatical errors as long as they did not cause misunderstanding (57.37% of responses were 'agree' and 'strongly agree'). This was mentioned by the participants who addressed the concept of major and minor errors in relation to intelligibility, as illustrated in the excerpt below:

When I grade my students' writing tasks, they will only be penalized when major errors are made but not with minor errors like preposition and articles that do not affect understanding. (Alex)

4.1.3 Perceptions of Receptive Skills

As shown in Table 6, the majority of the participants were ambivalent regarding incorporating GE in reading assessment.

Table 6*Frequency of Participants' Responses by Items Towards the Reading Domain*

Items	M (SD)	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
Q14. When designing a reading test, materials from British and American publishers are preferred over local materials that contain non-native language features.*	3.87 (1.06)	4 6.56%	1 1.6%	11 18.03%	28 45.90%	17 27.87%
Q17. I try to provide my students with diverse reading texts so they understand different non-native varieties of English.	3.98 (1.06)	2 3.28%	4 6.56%	10 16.39%	22 36.07%	23 37.70%
Q19. One criterion for selecting a passage in reading tests is that students have background knowledge about that topic.	3.52 (1.12)	1 1.64%	12 19.67%	17 27.87%	16 26.23%	15 24.59%
Q20. I consider standardized tests, such as TOEFL and IELTS, as a model for designing my reading tests.*	3.72 (0.86)	0	5 8.20%	18 29.51%	27 44.26%	11 18.03%

*Note * indicates reverse-scoring items.*

For example, in Q17 and Q19, the majority of the participants (73.77% and 50.82%, respectively) reported that they would provide diverse reading texts or select reading texts whose topics their students had knowledge about. However, in Q14 and Q20, most of the participants (73.77% and 62.29%, respectively) still preferred reading texts from native publishers and would design their assessment according to standardized tests such as TOEFL or IELTS.

The code “standardized proficiency tests” frequently emerged when the participants were asked to justify the design of reading tests, as exemplified below.

I would consider including non-native varieties, but I need to model my test according to the standardized proficiency test papers like IELTS and Cambridge ESOL because we need to familiarize our students and prepare them for the tests at the end of the academic year. (Rose)

Likewise, this corresponds to the following excerpt:

Both seen and unseen passages are taken from the commercial textbooks and the teachers' resources provided by UK publishers to prepare students for Cambridge tests and other standardized proficiency tests like IELTS. (Jim)

Compared to assessing reading, the participants were more likely to adopt GE principles for the listening domain, as can be seen in Table 7.

Table 7

Frequency of Participants' Responses by Items Towards the Listening Domain

Items	M (SD)	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
Q23. I include different sources of listening materials to get my students familiar with different non-native accents.	3.87 (1.04)	2 3.28%	4 6.56%	13 21.31%	23 37.70%	19 31.15%
4. I would include a simulated interaction between a Thai businessperson negotiating a sale with a Filipino customer in my listening test.	3.75 (1.14)	4 6.56%	3 4.92%	15 24.59%	21 34.43%	18 29.51%
Q26. Test designers should use authentic voice recordings of speakers of various geographical and linguistic backgrounds on a listening test.	3.82 (1.09)	3 4.92%	3 4.92%	15 24.59%	21 34.43%	19 31.15%

In Q23 and Q26, approximately two-thirds of the participants (68.85% and 65.58%, respectively) had favorable perceptions of utilizing different accents from diverse cultural and geographical backgrounds. Likewise, in Q24, the majority of the participants (63.94%) expressed positive perceptions of the inclusion of NNEs interactions in listening tests. These findings were confirmed by the interview data, as the participants reported that exposure to NNEs varieties resembled real-life communication, as shown in the following excerpts:

My students should be prepared to understand different accents because, in the real world, they may need to interact with non-native speakers. (Roberto)

I think understanding various accents would be an advantage for students. Knowing the British or American accents would not be enough in this globalized world. (Rose)

4.1.4 Perceptions of the Role of NNEs in Test Design and Development

The majority of the participants showed favorable perceptions of NNEs in test design and development. They reportedly agreed with NNEs being involved in test design and development ($M = 4.11$, $SD = 1.03$), writing test items ($M = 4.03$, $SD = 1.03$), and setting cut-off scores ($M = 3.95$, $SD = 1.04$). Most of the participants (75.41%, 72.13%, and 67.21%, respectively) either agreed or strongly agreed with these statements.

This was confirmed by the data from the open-ended sections of the questionnaire, as the participants shared favorable comments regarding the involvement of NNEs in test design:

From the perspective of a NEAR-native English speaker, I believe that test designs follow a specific structure and are not dependent on who designed them. It is not imperative that if English test

papers were designed by a native speaker, then it will not be subject to scrutiny, errors, or revisions, in the same way if it was designed by a near-native English speaker. What I strongly agree with is that examiners and test designers should have a clear understanding of how tests are designed and how an exam paper is drafted, edited, and approved by reliable experts in the field of English, regardless of their cultural backgrounds. (Saheed)

A similar sentiment can be seen in the following excerpt:

English is a global language. Any professional English teacher, regardless of ethnicity, can design an English test paper that is valid and reliable. (Alex)

4.2 The Incorporation of GE-informed Assessment in Thai Secondary Schools' Assessment Practices

Based on the analyses of the documents and semi-structured interviews, four themes were identified.

4.2.1 The Extensive Use of Standardized English Proficiency Tests

The use of large-scale standardized English proficiency tests based on prescriptive standards was prevalent in secondary schools' assessment practices. As part of summative assessment, the Common European Framework of Reference for Languages (CEFR)-based scores of such tests were requirements for students at certain levels at the end of the academic year. For example, the A2 Key English Test (KET) and B1 Preliminary English Test (PET) developed by Cambridge University Press and Assessment were used to assess Secondary 1 and 3 students' language proficiency, respectively. Moreover, standardized English proficiency tests were also used as an exit exam. As part of graduation criteria for students in the EP programs, prospective Secondary 6 graduates were required to submit test scores of locally developed tests or any well-received English proficiency test papers, such as IELTS, TOEFL iBT, TOEIC, CU-TEP, and TU-GET

at an equivalence of the B2 level of the CEFR. It can be noticed that these assessment practices aligned with teachers' perceptions towards the use of standardized English proficiency tests.

4.2.2 Adoption of Inner Circle-oriented Curricula

It was found that all of the schools in this study adopted the curriculum and its accompanying assessment practices developed by an inner-circle country, especially the United Kingdom. Documentary evidence suggested that British-based English curriculum, namely the Bell English Program, which was implemented at the lower secondary level of the two participating schools, added to the dominance of NES varieties in testing. In order to adopt the curriculum, several specifications, such as recruiting British teachers as teaching staff and using Cambridge English's B1 Preliminary tests to measure the students' language proficiency, were mandatory. Adopting inner-circle-oriented curricula contributed to the prevalence of native norms in the assessment practices. Moreover, three schools also adopted another inner-circle curriculum accredited by Cambridge Assessment International Education (CAIE). With this curriculum, the students' achievement in these three core subjects—mathematics, science, and English—was assessed by using the inner-circle test papers entitled Cambridge Lower Secondary Checkpoint Assessment.

Furthermore, the course syllabi showed evidence of curricula whose contents were based on standardized English proficiency tests. To illustrate, Secondary 4 English courses were titled *English for International Standardized Tests* and *English for Examination*. The course descriptions and objectives of the former included “the course is designed to help them become familiar with international standardized testing procedures...the final exam will be a simulation of an IELTS speaking exam,” while the latter stated that “students will have to analyze content to figure out the answers in any top examinations, for example, IELTS, CU-TEP, and TOEIC.”

Similarly, the Secondary 5 English curriculum also focused on standardized English proficiency tests. In an English course entitled *English for Examination*, one of its learning objectives was “to demonstrate advanced English language proficiency, enabling success in a diverse range of standardized English tests.” In terms of its course description, the following was stated: “This comprehensive programme equips students with the skills and knowledge necessary to navigate a wide spectrum of English standardized tests, encompassing national and international assessments, such as TGAT, A- level, IELTS, and more.” Additionally, the course description for a Secondary 6 English course indicated the following:

This course is to equip grammar topics by means of learning mainly through a variety of real practices by Cambridge and international standardized examinations, including IELTS and SAT...Besides, according to the optimum goal of the school, prior to their graduation, achieving an overall IELTS band score of 6.0 or higher, they (students) are highly expected to excel in this course.

While these curricula based on standardized English proficiency tests were understood to prepare learners for the competitive university admissions tests, they were not limited to the higher secondary levels. The analysis of the course syllabi for the lower secondary levels also showed evidence of inner circle-oriented course design, which led to assessment practices that were based on inner-circle varieties. For instance, the course description of a Secondary 1 English language course stated the following: “This course also focuses on the grammatical principles covered on Cambridge’s Key English Test (KET).” In short, NES varieties were dominant in the secondary schools’ assessment practices partly due to the adoption of inner circle-oriented curricula.

4.2.3 Native Speakerism in Test Constructs

In this study, it was found that NES varieties widely existed in the Thai secondary schools’ assessment practices. Every semester, teachers in each grade level were required to conduct three forms of tests: dictation, interviews, and

listening tests. The dictation tests required pre-recorded audios of the target words selected from each learning unit; semi-structured interviews suggested that only native English-speaking teachers (NESTs) were invited to record the audio input used for the tests. These practices contradicted their perceptions, as the majority of the teachers reported that they were likely to include various accents from diverse cultural and linguistic backgrounds in listening assessment in order to familiarize the students with different NNEs varieties. In addition, the criterion for selecting the target words was also based on the NES varieties. When asked to justify the target word selection, the participants stated that the textbooks developed by inner-circle publishers were the sources of the target words selected for the tests, as the objective of this assessment was to familiarize the students with the key words from each learning unit.

Similar to the dictation tests, NES varieties were used as a basis for selecting oral examiners and test materials for the interview tests. During the test, students were required to attend this oral assessment with examiners to check the students' comprehension of the prescribed graded readers from major inner-circle publishers, such as Macmillan Readers, Penguin Readers, and Cambridge Readers. The semi-structured interviews revealed that the role of the oral examiners was reserved for NESTs only. With regards to the listening tests, since part of the test tasks required students to listen to audio inputs readily available on textbooks and teacher resources provided by the inner-circle publishers, it could be seen that pronunciation in the audio inputs had little relevance to the local varieties. In sum, native speakerism remained predominant in secondary schools' assessment practices, especially in the construct of direct tests.

4.2.4 The Role of NNESTs in Test Design and Development

Despite the dominance of NESTs in conducting direct tests, NNESTs were responsible for the administrative duties of test design and development, especially in test preparation, operationalization, and administration. It was found that NNESTs played a key role in the initial stage of test design and development.

The head teachers of both divisions, including the English language Department and the school's English Program (EP) whose L1 was Thai, were responsible for setting test specifications which also included format, timing, and scoring methods. Apart from specifying the test specifications, they also defined test constructs and identified resources, including individuals (number of invigilators, examiners, and raters required for different tests), materials (audio recordings, scripts for dictation, a stopwatch timer, etc.), and time. For test operationalization, each subject teacher, regardless of their L1, was responsible for writing their own test items. This also encompassed writing scoring keys and scoring criteria. However, proofreading was still conducted by NESTs in order to ensure that the items were free of any kinds of errors and had the same standard and style of language use. With regards to administering the tests, both NESTs and NNESTs were assigned to assume different roles that aligned with the test purposes, such as invigilators, oral examiners, and raters. However, some test administering tasks were reserved for certain groups of teachers only. For paper-based tests, teachers, regardless of their L1, were responsible for proctoring and marking the students' test papers. This was the only test type where both NESTs and NNESTs were involved in the process. As stated earlier, for the direct tests that assessed students' oral and listening abilities, only NESTs were involved in conducting interview tests and recording listening inputs. In sum, the role that NESTs and NNESTs played at the stage of test design and development varied depending on the purposes of the tests. While NESTs were mainly responsible for administering direct tests, NNESTs were involved at the initial stage of test development.

5. Discussion

The objectives of the current study were to explore the perceptions and practices of secondary school teachers in Thailand towards adopting GE-informed assessment. The study identified three key points that have yielded insights into the teachers' perceptions and practices with regards to GE-informed assessment.

First, secondary school teachers are aware of the changing sociolinguistic landscape of English and its implications for assessment, as the findings suggest some positive views towards incorporating GE principles in certain assessment practices. This is evident in the speaking and listening domains, as the participants favored fluency and pragmatic effectiveness over accuracy when assessing oral proficiency. For the listening tests, the majority of the participants were likely to include different non-native accents and authentic voice recordings of speakers of various geographical and linguistic backgrounds. This finding presents a new insight into the current situation of GE in the area of language assessment. Previous studies have reported that teachers acknowledged this linguistic phenomenon (Ambele, 2022; Galloway & Numajiri, 2019; Aydın & Karakas, 2021; Vettorel, 2016) and expressed the changing attitudes from NES orientation toward GE, especially in the areas of NES norms and a variety of Englishes (Prabjandee & Fang, 2022). This study added that the teachers were aware of GE principles when designing their classroom-based assessment tasks. This increased awareness can influence teachers and administrators in terms of redesigning their teaching and formulating language policies that correspond to current linguistic diversity.

Second, despite their acknowledgement of incorporating GE into assessment, teachers still adopted large-scale standardized proficiency tests based on prescriptive standards, both in direct and indirect ways. For their direct utilization, the evidence suggests that these test papers functioned beyond their primary purpose of measuring students' English proficiency. They were also used as achievement tests, placement tests, screening tests, and exit exams by the schools. As for indirect use, the teachers adopted these well-received standardized tests of English proficiency as a prototype for the design of classroom-based assessment and English language courses. Apparently, these direct and indirect uses of such tests contribute to the hegemony of native norms that exist not only in the area of testing, but also in teaching and learning through the power of washback. This adds complexity to GE-related innovations and serves as a major hurdle for the successful implementation of the GE paradigm in the

classroom. The findings in this study support the ideas of Galloway and Numajiri (2019) and Jenkins (2014), who suggest that the traditional EFL paradigm will remain dominant as long as students' English proficiency is still measured against native norms.

Third, while teachers' assessment practices were heavily influenced by the EFL paradigm, some evidence of GE principles were identified. It was found that NNESTs took part in three initial stages of test development: preparation, operationalization, and administration. To illustrate, tasks that included specifying test specification, defining test construct, writing test items, developing scoring keys and rubrics, and marking test papers were performed by NNESTs. Kling and Dimova (2015) and Taylor (2006) also addressed similar attempts to incorporate NNESTs in the design and development of the language proficiency tests. However, the involvement of NNESTs was limited to certain areas as some tasks, such as administering direct tests to measure students' oral and listening abilities, were reserved exclusively for NESTs. The findings by Prabjandee (2020) regarding the perceptions of accented speeches were in line with the assessment practices of the present study. It was reported that the British accent and the American accent were rated at a higher level of acceptability than most accents from the inner circle and the outer circle. Although some developments towards GE were observed in their practices, they are still minimal and have limited potential to catalyze pedagogical changes partly due to stereotypical views of the native varieties and dominance of large-scale standardized English proficiency tests. Introducing GE elements into test constructs and administration, such as incorporating NNEST varieties into listening and reading tests, inviting NNESTs as examiners for speaking and writing assessment, and redesigning rubrics to include intelligibility and communicative competence, can facilitate a shift towards assessment that corresponds to the current sociolinguistic realities of English.

6. Limitations

There are some limitations of the present study. First, since the participants were drawn from nine different schools located in different regions of Thailand, it is very likely that their exposure and familiarity with the concept of GE are different. The level of exposure to GE might have influenced these teachers' perceptions and decisions regarding assessment practices. Second, the study did not account for other variables that may have affected the results, such as the schools' assessment policies, as these regulations could impact teachers' assessment practices. Future studies might expand the scope of investigation to include more variables.

7. Implications

The findings from this study provide the following pedagogical implications. First, a series of tests, including scoring schemes and rubrics that contain elements of GE developed by local language assessment researchers and specialists, might be provided to teachers as practical resources and references. This locally designed assessment can raise awareness and create more interest among teachers and administrators to opt for more culturally-aware assessment. Such tests can also minimize heavy reliance on standardized proficiency tests, which is considered a major obstacle for implementing GE in classrooms. Secondly, classroom materials or textbooks should be designed by incorporating GE principles in order to be resources or references for practitioners. NNES varieties and authentic materials with culturally appropriate content relevant to the local context should be integrated into the contents of commercial textbooks. Thirdly, professional development programs on GE test design and development might be necessary for secondary school teachers and related educators in order to ensure a clear understanding of GE principles and to provide them with assistance to develop their classroom-based assessment incorporated with GE elements. Such trainings on language assessment could contribute to a more successful operationalization of GE-informed assessment.

8. Conclusion

While the participants in the present study demonstrated positive perceptions regarding the adoption of GE principles when designing certain test tasks, it should be noted that GE principles were little reflected in the teacher participants' assessment practices. It appears that the teachers still favored large-scale standardized proficiency tests based on prescriptive standards. A plausible explanation can be that these test papers are associated with accountability, prestige, and the commercial viability of the school. Unless there are changes in assessment from the EFL paradigm to greater incorporation of GE, native norms, which have been deeply ingrained in the educational context, will continue to be the predominant paradigm in teaching, learning, and testing. To correspond to the changing landscape of English, more studies that provide frameworks and guiding principles for operationalizing assessment tasks based on GE principles are needed in order to pave the way for more context-specific, locally developed, and in-house assessment. To date, some attempts, such as Jenkin's *Lingua Franca* core (2000) and the THAI Test of English Proficiency (THAI-TEP), can serve as points of reference. This would not only provide sufficient opportunities for local language assessment researchers to develop their capacity and improve their expertise (Wudthayagorn, 2022), but would also serve as impetus for the paradigm shift that has been called for.

8. About the Authors

Pongpol Suansri is a Ph.D. candidate in English Language Teaching at the Language Institute, Thammasat University, Thailand. He has an M.Ed. in TESOL from the University of Wollongong, Australia. His research interests include language assessment, Global Englishes, and English for specific purposes. He can be reached at pongpol-suansri@gmail.com.

Supong Tangkiengsirisin is an associate professor at the Language Institute, Thammasat University, Thailand. He holds a Ph.D. from the University of Nottingham, U.K. His research interests include second language writing, written

discourse analysis, and English for specific purposes. He can be reached at supong.t@litu.tu.ac.th.

9. Reference

- Abeywickrama, P. (2013). Why not non-native varieties of English as listening comprehension test input? *RELC Journal*, 44(1), 59–74.
<https://doi.org/10.1177/0033688212473270>
- Ahn, H., Choi, N. & Kiaer, J. (2021). Multilingual “native” speakers of the English language: The perceptions of university students from the United Kingdom, Singapore and South Korea. *Journal of Language, Identity and Education*, 22(5), 463–477.
<https://doi.org/10.1080/15348458.2021.1893174>
- Ambele, E. A. (2022). Thai English? Non-Thai English lecturers’ perceptions of Thai English and World Englishes. *LEARN Journal: Language Education and Acquisition Research Network*, 15(2), 724–750. <https://so04.tci-thaijo.org/index.php/LEARN/article/view/259949>
- Anderson, J. (2018). Reimagining English language learners from a translingual perspective. *ELT Journal*, 72(1), 26–37. <https://doi.org/10.1093/elt/ccx029>
- Aydın, F. & Karakas, A. (2021). Turkish EFL teachers’ awareness about ELF and the underlying ideologies behind their (non)-awareness. *Reading Matrix: An International Online Journal*, 21(1), 72–88.
- Bachman, L. F. & Palmer, A. S. (1996). *Language testing in practice: Designing and developing useful language tests*. Oxford University Press.
- Baker, W. & Jarunthawatchai, W. (2017). English language policy in Thailand. *European Journal of Language Policy*, 9(1), 27–44.
<https://doi.org/10.3828/ejlp.2017.3>
- Bennui, P. & Hashim, A. (2014). English in Thailand: Development of English in a non-postcolonial context. *Asian Englishes*, 16(3), 209–228.
<https://doi.org/10.1080/13488678.2014.939810>

- Blair, A. (2017). Standard language models, variable lingua franca goals: How can ELF-aware teacher education square the circle? *Journal of English as a Lingua Franca*, 6(2), 345–366. <https://doi.org/10.1515/jelf-2017-0016>
- Bogorevich, V. (2018). *Native and non-native raters of L2 speaking performance: accent familiarity and cognitive processes* [Doctoral dissertation, Northern Arizona University]. Northern Arizona University. <http://openknowledge.nau.edu/id/eprint/5406>
- Boonsuk, Y., Ambele, E. A., & McKinley, J. (2021), Developing awareness of Global Englishes: Moving away from ‘native standards’ for Thai university ELT, *System*, 99, Article 102511. <https://doi.org/10.1016/j.system.2021.102511>
- Bowen, G. A. (2009), Document analysis as a qualitative research method. *Qualitative Research Journal*, 9(2), 27–40. <https://doi.org/10.3316/QRJ0902027>
- Brown, H. D. & Abeywickrama, P. (2019). *Language assessment: Principles and classroom practices*. Pearson Education.
- Brown, J. D. (2012). Choosing the right type of assessment. In C. Coombe, P. Davidson, B. O'Sullivan, & S. Stoyhoff (Eds.), *The Cambridge guide to second language assessment* (pp. 133–139). Cambridge University Press.
- Buripakdi, A. (2012). On professional writing: Thai writers' views on their English. *International Journal of Applied Linguistics*, 22(2), 245–264. <https://doi.org/10.1111/j.1473-4192.2012.00311.x>
- Caleffi, P. (2016). ELF in the speaking and listening activities of recently published English-language Coursebooks. In L. Lopriore, & E. Grazi (Eds.), *Intercultural communication: New perspectives from EFL* (pp. 63–82). RomaTrE-Press.
- Canagarajah, S. (2006). Changing communicative needs, revised assessment objectives: Testing English as an international language. *Language Assessment Quarterly*, 3(3), 229–242. https://doi.org/10.1207/s15434311laq0303_1

- Chen, Z., Chen, X. & Fang, F. (2021). Global Englishes and teacher education: Present cases and future directions. *RELC Journal*, 54(3), 873–880.
<https://doi.org/10.1177/00336882211044872>
- Creswell, J. W. & Creswell, J. D. (2018). *Research Design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). Sage.
- Crystal, D. (2003). *English as a global language* (2nd ed.). Cambridge University Press.
- Davidson, F. (2006). World Englishes and test construction. In B. B. Kachru, Y. Kachru, & C. L. Nelson (eds.), *The handbook of World Englishes* (pp. 709–717). Blackwell.
- Deterding, D. & Gardiner, I. (2019). Language selection and assessment in Brunei Darussalam. In F. Fang & H. Widodo (Eds.), *Critical perspectives on Global Englishes in Asia: Language policy, curriculum, pedagogy and assessment* (pp. 45–63). Multilingual Matters.
<https://doi.org/10.21832/9781788924108-006>
- Dimova, S. (2017). Pronunciation assessment in the context of World Englishes. In O. Kang, & A. Ginther (Eds.), *Assessment in second language pronunciation* (pp. 49–66). Routledge.
- Dimova, S., & Kling, J. (2015). Lecturers' English proficiency and university language policies for quality assurance. In R. Wilkinson, & M. L. Walsh (Eds.), *Integrating content and language in higher education: From theory to practice selected papers from the 2013 ICLHE Conference* (pp. 50–65). Peter Lang.
- Galloway, N. & Numajiri, T. (2019). Global Englishes language teaching: Bottom-up curriculum implementation. *TESOL Quarterly*, 54(1), 118–145.
<https://doi.org/10.1002/tesq.547>
- Galloway, N. & Rose, H. (2015). *Introducing Global Englishes*, Routledge.
<https://doi.org/10.4324/9781315734347>

- Grazzi, E., & Lopriore, L. (2020). ELF awareness for teacher education in Italy: Attitudes and actions. *Estudos Linguísticos E Literários*, 1(65), 69–89. <https://doi.org/10.9771/ell.v1i65.36469>
- Hamid, M. O. (2014). World Englishes in international proficiency tests. *World Englishes*, 33(2), 263–277. <https://doi.org/10.1111/weng.12084>
- Hu, G. (2012). Assessing English as an international language. In L. Alsagoff, S. L. McKay, G. Hu, & W. A. Renadya (Eds.), *Principles and practices for teaching English as an International Language* (pp. 123–144). Routledge.
- Hu, G. (2018). The challenges of World Englishes for assessing English proficiency. In E. L. Low & A. Pakir (eds.), *World Englishes: Rethinking paradigms* (pp. 120–138). Routledge.
- Hynninen, N. (2014). The Common European Framework of Reference from the perspective of English as a lingua franca: What can we learn from a focus on language regulation? *Journal of English as a Lingua Franca*, 3(2), 293–316. <https://doi.org/10.1515/jelf-2014-0018>
- Illes, E. & Akcan, S. (2016). Bringing real-life language use into EFL classrooms. *ELT Journal*, 71(1), 3–12. <https://doi.org/10.1093/elt/ccw049>
- Imsa-ard, P. & Tangkiengsirisin, S. (2023). Language assessment literacy: Thai EFL lecturers' language assessment knowledge, beliefs, and practices. *PASAA*, 66, 37–80. <https://doi.org/10.58837/CHULA.PASAA.66.1.2>
- Jenkins, J. (2000). *The phonology of English as an international language*. Oxford University Press.
- Jenkins, J. (2014). *Global Englishes: A resource book for students* (3rd ed.). Routledge. <https://doi.org/10.4324/9781315761596>
- Jenkins, J. (2020). Where are we with ELF and language testing? An opinion piece. *ELT Journal*, 74(4), 473–479. <https://doi.org/10.1093/elt/ccaa045>
- Jianrattanapong, A. (2011). Positive washback from Thai university entrance examinations. *Language Testing in Asia*, 1(1), Article 50. <https://doi.org/10.1186/2229-0443-1-1-50>

- Jindapitak, N., Teo, A. & Savski, K. (2022). Bringing Global Englishes to the ELT classroom: English language learners' reflections. *Asian Englishes*, 24(2), 279–293. <https://doi.org/10.1080/13488678.2022.2033910>
- Johnson, J. S. & Lim, G. S. (2009). The influence of rater language background on writing performance assessment. *Language Testing*, 26(4), 485–505. <https://doi.org/10.1177/0265532209340186>
- Kachru, B. B. (1992). World Englishes: Approaches, issues and resources. *Language Teaching*, 25(1), 1–14. <https://doi.org/10.1017/S0261444800006583>
- Kang, O., Thomson, R., & Moran, M. (2018). The effects of international accents and shared first language on listening comprehension tests. *TESOL Quarterly*, 53(1), 56–81. <http://doi.org/10.1002/tesq.463>
- Kirkpatrick, A. (2014). English as a medium of instruction in East and Southeast Asian universities. In N. Murrey & A. Scarino (Eds.), *Dynamic Ecologies* (pp. 15 – 29). Springer. https://doi.org/10.1007/978-94-007-7972-3_2
- Kling, J. & Dimova, S. (2015). The test of oral English for academic staff (TOEPAS): Validation of standards and scoring procedures. In A. Knapp & K. Aguado (Eds.), *Fremdsprachen in Studium und Lehre—Chancen und Herausforderungen für den Wissenserwerb* (pp. 247–268). Peter Lang.
- Kubota, R. (2019). A critical examination of common beliefs about language teaching: From research insights to professional engagement. In F. Fang & H. Widodo (Eds.), *Critical perspectives on Global Englishes in Asia: Language policy, curriculum, pedagogy and assessment* (pp. 10–26). Multilingual Matters. <https://doi.org/10.21832/9781788924108-004>
- Lee, J. S., & Drajadi, N. A. (2019). English as an international language beyond the ELT classroom. *ELT Journal*, 73(4), 419–427. <https://doi.org/10.1093/elt/ccz018>
- Lueng, C. & Scarino, A. (2016). Reconceptualizing the nature of goals and outcomes in language/s education. *The Modern Language Journal*, 100(S1), 81–95. <https://doi.org/10.1111/modl.12300>

- Matsuda, A. & Friedrich, P. (2011). English as an international language: A curriculum blueprint. *World Englishes*, 30(3), 332–344.
<https://doi.org/10.1111/j.1467-971X.2011.01717.x>
- Matsuda, A. (2020). World Englishes and pedagogy. In C. L. Nelson, Z. G. Proshina & D. R. Davies, *The handbook of World Englishes* (pp. 686–702). Willey-Blackwell.
- Matsuda, A. (Ed.). (2012). *Principles and practices of teaching English as an international language*. Multilingual Matters.
<https://doi.org/10.21832/9781847697042>
- Montakantiwong, A. (2023). Bridging conceptual gaps in Global Englishes Language Teaching: Ethnographic insights from Thai higher education. *Asian Englishes*, 26(2), 391–409.
<https://doi.org/10.1080/13488678.2023.2251736>
- Newbold, D. (2014). Engaging with ELF in an entrance test for European university students. In Y. Bayyurt & A. Sumru (Eds.), *Current perspectives on pedagogy for English as a lingua franca* (pp. 205–221). De Gruyter.
- Office of the Basic Education Commission. (2008). *The basic education core curriculum B.E. 2551 (A.D. 2008)*. Ministry of Education.
http://academic.obec.go.th/images/document/1525235513_d_1.pdf
- Padermprach, N. (2017). The problems of English language education at the upper secondary level in Thailand: the perceptions of Thai EFL teachers and students in Bangkok. [Doctoral dissertation, University of Warwick]. University of Warwick. <http://wrap.warwick.ac.uk/98408>
- Patton, M. Q. (2002). *Qualitative evaluation and research methods* (3rd ed.). SAGE.
- Pill, J. & Harding, L. (2013). Defining the language assessment literacy gap: Evidence from a parliamentary inquiry. *Language Testing*, 30(3), 381–402.
<https://doi.org/10.1177/0265532213480337>
- Ploywattanawong, P. & Trakulkasemsuk, W. (2014). Attitudes of Thai graduates toward English as a Lingua Franca of ASEAN. *Asian Englishes*, 16(2), 141–156. <https://doi.org/10.1080/13488678.2014.910902>

- Prabjandee, D. & Fang, F. (2022): 'I was like, just wow!': Insights from Global Englishes teacher professional development, *Asian Englishes*, 24(3), 294–311. <https://doi.org/10.1080/13488678.2021.1989546>
- Prabjandee, D. (2020). Teacher professional development to implement Global Englishes language teaching, *Asian Englishes*, 22(1), 52–67. <https://doi.org/10.1080/13488678.2019.1624931>
- Rajprasit, K. (2024). MOOC-ing around 'Englishes' by Thai students: Applying the GELT framework. *Asian Englishes*, 26(2), 428–443. <https://doi.org/10.1080/13488678.2023.2251739>
- Ramadhani, S. A. F. & Muslim, A. B. (2021). Investigating teachers' attitudes towards teaching and learning of English as a Lingua Franca (ELF) in Indonesian EFL context. *Jurnal Penelitian Pendidikan*, 21(2), 66–79. <https://doi.org/10.17509/jpp.v21i2.35131>
- Rose, H. & Galloway, N. (2019). *Global Englishes for language teaching*. Cambridge University Press.
- Selvi, A. F. & Yazan, B. (2021). *Language teacher education for Global Englishes*. Routledge.
- Shohamy, E. (2011). Assessing multilingual competencies: Adopting construct valid assessment policies. *The Modern Language Journal*, 95(3), 418–429. <https://doi.org/10.1111/j.1540-4781.2011.01210.x>
- Sifakis, N. C. (2014). ELF awareness as an opportunity for change: A transformative perspective for ESOL teacher education. *Journal of English as a Lingua Franca*, 3(2), 317–335. <https://doi.org/10.1515/jelf-2014-0019>
- Smaoui, C. & Rahal, A. (Eds.) (2022). *World Englishes and language assessment*. Cambridge Scholars Publishing.
- Snodin, N. & Young, T. (2015). 'Native-speaker' varieties of English: Thai perceptions and attitudes. *Asian Englishes*, 17(3), 248–260. <https://doi.org/10.1080/13488678.2015.1083354>
- Suebsook, K. (2018). *The inclusion of Asian English accents in the listening proficiency test for healthcare personnel working in the Thai context*.

- [Doctoral dissertation, Thammasat University]. TU Digital Collection.
https://digital.library.tu.ac.th/tu_dc/frontend/Info/item/dc:126214.
- Taylor, L. (2006). The changing landscape of English: Implications for language assessment. *ELT Journal*, 60(1), 51–60. <https://doi.org/10.1093/elt/cci081>
- Thao, T. Q., Phoung, N. L. & Tham, D. M. (2022). Vietnamese teachers' attitudes toward English as a Lingua Franca (ELF) and classroom EFL teaching practices. *VNU Journal of Science: Education Research*, 38(4), 90–102. <https://doi.org/10.25073/2588-1159/vnuer.4641>
- Thienthong, A., & Uthaiorn, K. (2023). Thai learners' attitudes toward English accents: Fields and stages of study. *LEARN Journal: Language Education and Acquisition Research Network*, 16(1), 676–702. <https://so04.tci-thaijo.org/index.php/LEARN/article/view/263461>
- Vettorel, P. (2016). WE- and ELF-informed classroom practices: Proposals from a pre-service teacher education programme in Italy. *Journal of English as a Lingua Franca*, 5(1), 107–133. <https://doi.org/10.1515/jelf-2016-0005>
- Vettorel, P. (2018). ELF and communication strategies: Are they taken into account in ELT materials? *RELC Journal*, 49(1), 58–73. <https://doi.org/10.1177/0033688217746204>
- Watson Todd, R. (2019). How is English assessed at Thai schools? *THAITESOL Journal*, 32(1), 1-15.
- Wudthayagorn, J. (2022). An exploration of the English exit examination policy in Thai public universities. *Language Assessment Quarterly*, 19(2), 107–123. <https://doi.org/10.1080/15434303.2021.1937174>
- Zhang, Y. & Elder, C. (2011). Judgments of oral proficiency by non-native and native English speaking teacher raters: Competing or complementary constructs? *Language Testing*, 28(1), 31–50. <https://doi.org/10.1177/0265532209360671>

10. Appendix**Appendix A****Teachers' Perceptions towards GE-informed Assessment Questionnaire****SECTION 1: Demographic data**

Gender: _____

Age: _____

Level of Education: _____

Level of Teaching: _____

Teaching Experience (in years): _____

On a scale of 1 to 5, please indicate what applies to you: (1) Strongly disagree, (2) Disagree, (3) Neutral, (4) Agree, (5) Strongly agree.

SECTION 2

The Current Use of Large-scale Proficiency Tests					
1. English proficiency tests like TOEFL and IELTS should be used to assess high school students' language ability.	1	2	3	4	5
2. I use English proficiency tests like TOEFL and IELTS as a model for designing my own English test papers for students.	1	2	3	4	5
Assessing Productive Skills					
3. I would give higher scores to the students whose accent is closer to native speakers	1	2	3	4	5
4. As long as my students get the meaning across, how they pronounce English words does not matter.	1	2	3	4	5
5. I don't mind students who use a Thai accent during the speaking tests.	1	2	3	4	5
6. One criterion of oral examiners is being a native speaker.	1	2	3	4	5

7. Speaking with grammatical correctness is more important than maintaining the flow of the conversation.	1	2	3	4	5
8. Students' points should not be deducted for causing misunderstanding if they show the ability to repair the conversation.	1	2	3	4	5
9. Intelligibility is more important than grammatical accuracy when assessing writing.	1	2	3	4	5
10. When designing rubrics for writing tasks, grammatical accuracy is the most important component.	1	2	3	4	5
11. Native English raters assess writing tasks better than non-native English raters.	1	2	3	4	5
12. The use of non-native grammatical forms on writing tasks is not considered an error as long as it promotes comprehension.	1	2	3	4	5
13. As long as the minor grammatical errors do not cause misunderstanding, I do not deduct students' scores.	1	2	3	4	5
Assessing receptive skills					
14. When designing a reading test, materials from British and American publishers are preferred over local materials that contain non-native language features.	1	2	3	4	5
15. Test papers should include some local/non-native words.	1	2	3	4	5
16. In the test, teachers should provide a gloss (a brief explanation of an unfamiliar word or expression) to students when assessing reading comprehension.	1	2	3	4	5

17. I try to provide my students with diverse reading texts so they understand different non-native varieties of English.	1	2	3	4	5
18. It is important that students know the native English idioms, even though they would rarely use them.	1	2	3	4	5
19. One criterion for selecting a passage in reading tests is that students have background knowledge about that topic.	1	2	3	4	5
20. I consider standardized tests, such as TOEFL and IELTS as a model for designing my reading tests.	1	2	3	4	5
21. The accents used for the listening tests should be native speakers' standard only.	1	2	3	4	5
22. Contents for the audio clips used for listening tests should be selected only from native speakers' sources.	1	2	3	4	5
23. I include different sources of listening materials to get my students familiar with different non-native accents.	1	2	3	4	5
24. I would include a simulated interaction between a Thai businessman negotiating a sale with a Filipino customer in my listening test.	1	2	3	4	5
25. The language in the listening test should be similar to students' daily life interactions, even though it is not similar to that of native speakers.	1	2	3	4	5
26. Test designers should use authentic voice recordings of speakers of various geographical and linguistic backgrounds on a listening test.	1	2	3	4	5
27. It is better to include conversations between native and non-native speakers (NES vs. NNES) in a listening	1	2	3	4	5

test rather than between native and native speakers (NES vs. NES).					
Roles of Non-native Speakers in Test Development					
28. Non-native speakers should be involved in test design and development.	1	2	3	4	5
29. Non-native speakers should be involved in writing test items.	1	2	3	4	5
30. Non-native speakers should involve in setting cut-off scores.	1	2	3	4	5

SECTION 3: OPEN-ENDED QUESTION

As non-native speakers outnumber native speakers, how would that affect the design of your English test papers? (สามารถตอบเป็นภาษาไทยได้)

Appendix B

Interview Questions

1. How do you design the reading test?
2. How do you design the listening test?
3. How do you design the speaking test?
4. How do you design the writing test?
5. How do you design the grammar test?
6. What is your opinion about the design of the proficiency tests?
7. What do you think about the use of English Proficiency Tests like TOEFL and IELTS to assess students
8. As there are more non-native speakers than native speakers, how do you think it would affect language assessment?
9. How do you think the concepts of GE or ELF could be incorporated into language assessment?
10. What is your idea about incorporating local varieties into English test papers?
11. What do you think are the constraints in incorporating GE into your test papers?