Effects of Formative Self-Assessment on Thai Students' English Language Proficiency

Patcharee Imsri
School of Education and Liberal Arts
Walailak University
222 Thaiburi, Thasala
Nakhon Si Thammarat, Thailand, 80160
Tel: +660617041814
Email: ipatchar@wu.ac.th

Suttida Sangpoom
School of Science
Walailak University
222 Thaiburi, Thasala
Nakhon Si Thammarat, Thailand, 80160
Tel: +660653905453
Email: suttida.sa@mail.wu.ac.th

Abstract: This study aimed to gain more understanding of Thai EFL students' self-assessment during practice test engagement. The participants were 95 Thai undergraduate students majoring in English at a public university in Thailand. The correlation and multiple regression analyses were employed to study the relationship between the practice test and English language proficiency test (TOEIC) scores. The analysis of the confusion matrix was also performed to measure the accuracy of students' self-assessment from the practice tests. The findings revealed that while Thai EFL students perceived themselves as "poor performers" for grammar and listening skill, their self-perception for vocabulary improvement was positive. A significant correlation was also found between the vocabulary practice test and the English proficiency test scores. The completion of multiple practice tests enhanced the students' vocabulary knowledge and enabled them to improve their English proficiency test score. Finally, the confusion matrix results revealed the accuracy for students' self-perception of grammar and listening skill. They did not overestimate or underestimate their abilities, but accurately predicted themselves as "low achievers" for these two skills. As demonstrated in the findings, students are the key factors for their learning success. It is essential to engage them with their own learning process, so that they will be able to identify their strengths and weaknesses in learning.

Keywords: English language proficiency, Thai EFL students, formative self-assessment, practice tests

Introduction

One of the major aspects of self-assessment is assessment for learning. According to Andrade and Valtcheva (2009), assessment for learning is "a process of formative assessment during which students reflect on the quality of their work, judge the degree to which it reflects explicitly stated goals or criteria, and revise accordingly." During the process of formative self-assessment, students have the opportunity to monitor, change, adapt, and improve their learning (Bullock, 2011; Kavaliauskienė, 2004; Zimmerman & Schunk, 2011). When the focus of self-assessment is on the

process, students will gain benefits, especially in terms of self-regulation of learning (SRL) (Panadero, Brown, & Strijbos, 2015). Students will be empowered to take ownership of their learning, and able to identify their errors and what areas they need to improve (Black & William, 1988; Killimangalam, 2020; Kostons, van Gog, & Paas, 2012; Ramdass & Zimmerman, 2008). Formative self-assessment does not only improve learning performance but also enables students to be independent in their learning (Killimangalam, 2020). Students will directly benefit from formative self-assessment and are able to execute activities that can improve their learning (Rea, 1981). As Boud (1995) and Murakami, Valvona, and Broudy (2012) point out, self-assessment creates an autonomous learning environment, because students are engaged in making decisions and judging their learning quality. In-depth analysis of their strengths and weaknesses enables the students to develop effective learning strategies and achieve their learning goals (Butler & Lee, 2010).

There are several methods in the implementation of self-assessment. The present study used practice tests to investigate regular formative self-assessment among Thai EFL students. A wide range of research shows that undertaking practice tests results in learning improvement (Carpenter, Pashler, & Vul, 2006; Roediger & Karpicke, 2006). Practice tests function as retrieval tools that are highly effective for academic improvement (Karpicke, Butler, & Roediger III, 2009; Martin, English, Morley, O'Keefe-Quinn, & Whitfield, 2020; Naujoks, Harder, & Händel, 2022). Formative self-assessment of practice tests can support self-regulation among students (Cogliano, Kardash, & Bernacki, 2019).

Despite the clear benefits of formative self-assessment, one of the concerns is whether students are able to accurately evaluate   their learning abilities. Many studies have found that low achievers tend to overestimate themselves more than the high achievers (e.g., Blue, 1988; Killimangalam, 2020; Wangsotorn, 1981). According to Nunan (2004), "While self-assessment has been criticized on the grounds that not all learners are accurate judges of their own ability, this criticism misses the point to some extent, which is to involve learners in their own learning process" (p. 149). The present study attempted to investigate how formative self-assessment could be used to enhance academic achievement among Thai EFL students. It explored the process of formative self-assessment while students completed English practice tests.

The study also examined the correlation between these practice tests and the Test of English for International Communication (TOEIC) scores. Hence, the research questions are:

RQ1. How do Thai EFL students' practice test scores correlate with TOEIC score?

RQ2. How do Thai EFL students self-assess their English vocabulary, grammar and listening skills during practice test engagement?

RQ3. How accurate is Thai EFL students' self-assessment on their actual performance in TOEIC?

Formative Self-Assessment as a Learning Tool

Formative self-assessment has a significant impact on students' learning improvement. Butler and Lee (2010) found that formative self-assessment positively affected EFL students' English performance and their confidence in learning the language. In a reviewed study that investigated educational factors on students' academic achievement, formative assessment was found to be one of the most influential factors for students' achievement (Hattie, 2009). Alderson and Huhta (2005) adopted formative self-assessment in their study by assigning students to regularly record their learning by keeping a journal. Research findings suggested that self-assessment encouraged

students to be more responsible for their own learning (Lopez & Kossack, 2007; Bourke, 2014; Ndoye, 2017). A positive association was also found between self-assessment and achievement (Graham, Hebert, & Harris, 2015; Sanchez, Atkinson, Koenka, Moshontz, & Cooper, 2017).

Mazloomi and Khabiri (2016) investigated the formative aspect of self-assessment and found that when students practiced their writing skills using the assessment criteria, students were able to improve their writing skills. Karaman (2021) conducted the meta-analysis study from 16 empirical studies to investigate the effects of self-assessment on academic performance from primary to university levels and found a positive impact on students' academic performance. Mahdavinia and Ahmadi (2011) explored how Iranian university students used portfolios to self-assess their essay writing. The interview results showed that students were able to assess themselves and reflect on how their essay writing improved during the process. They also gained responsibility in their own learning. In addition to writing assessment, Shelton-Strong (2018) investigated the effects of self-assessment on Japanese university students' speaking tasks. Students reported that they developed both metacognitive skills and awareness. Andrade (2019) suggested that these positive results were possible due to the formative nature of self-assessment. Students developed both knowledge and skills during the self-assessment processes. During practice, conducting self-assessment allows students to get feedback and revise their performance accordingly.

Formative Self-Assessment Studies in the Thai Context

Self-assessment has also been a topic of interest in the Thai EFL context. Research has sought to measure the accuracy of self-assessment and investigated its effects upon Thai students' learning behavior. Sapsirin (2014) compared the relationship between Thai EFL students' self-assessment ratings on writing ability and writing achievement test performance. The findings revealed a slight , positive correlation between the self-assessment ratings and the writing test scores. The result indicated that the students were not able to accurately judge their English writing ability. In contrast, Suwanarak (2018) explored Thai students' self-evaluation of their English writing performance compared with the evaluation of their teacher. The findings from closed and open-ended questionnaires, semi-structured interviews and English writing assignments showed that the students were able to accurately evaluate their writing performance.

Recent findings have shown how Thai EFL students use self-assessment to improve their English performance. Dhanarattigannon and Thienpermpool (2022) found that the process of self-assessment helped to improve students' writing performance, confidence and develop a more positive attitude towards the self-assessment process. Konchiab and Munpanya (2021) explored Thai EFL students' self-assessment on their oral presentation and found that the students were not confident with their pronunciation, vocabulary and grammar. They did not rate themselves highly for these language features and reflected their needs for improving these skills. Seenak and Adunyarittigun (2019) investigated the effects of Thai EFL students' self-assessment on intonation improvement. The findings revealed that the students were able to recognize their strengths and weaknesses and improve their intonation ability. In addition, the students showed positive attitudes towards self-assessment.

While self-assessment has increasingly been investigated in the Thai context, there is a lack in theoretical and empirical knowledge on the effects of Thai EFL students' self-assessment of practice tests. The present study aimed to advance the understanding of Thai EFL students' self-assessment during practice test engagement.

Methodology

Participants

Participants in this study were 95 second year English major students at a public university in Thailand. Their native language was Thai. Since they were English major students, most courses were conducted using English as a medium of instruction.

Materials and Instruments

The data-gathering tools in this study consisted of (1) practice tests that aimed to assess the three skills: Grammar, listening and vocabulary, (2) English proficiency test, the Test of English for International Communication (TOEIC), and (3) students' self-reflection journals.

Procedure

During the 10-week study period, the participants completed multiple practice tests from the website: https://www.examenglish.com. These online practice tests aim to improve students' grammar, listening and vocabulary skills. The selected CEFR level for the practice tests was B1, the minimum target level of university students' proficiency in Thailand (Piamsai, 2023).

　　1. https://www.examenglish.com/B1/B1_listening.htm

　　2. https://www.examenglish.com/grammar/b1_grammar.html

　　3. https://www.examenglish.com/vocabulary/B1_vocabulary_topics.htm

The students could access these websites at their convenience and as frequently as they preferred. After completing each formative practice test, the students recorded their scores, studied the explanations for the correct/incorrect answers, and noted down their self-reflection of their performance. The students focused their self-assessment on their performance during the practice tests. The skills that had and had not been executed well were classified as "Good" and "Not Good," respectively.

After the students completed the 10-week practice period, two types of tests were administered: The summative practice test and the Test of English for International Communication (TOEIC). The summative practice test was the Oxford Online Placement Test, which was composed of two sections: Use of English (Grammar and Vocabulary) and Listening. The test was administered to assess the students' performance after completing the regular self-assessment of the practice tests. The test of English for International Communication (TOEIC) was also administered to assess students' English performance.

Data Analysis

The scores from the summative practice test, the English language proficiency (TOEIC) test, and the students' reflections of their grammar, listening and vocabulary skills after completing the practice tests were analyzed using descriptive statistics.

The correlation analysis was used to describe the relationship between the practice test and English language proficiency test (TOEIC). The correlation coefficient ranges from -1 to 1. If the value is close to 1 or 1, it is shown that two variables are a very strong association. On the contrary, the value close to 0 indicates 'weak association' or 'no association.' The null hypothesis of the analysis is that one variable and the other variable do not correlate. The relationship between two continuous variables: Practice tests score and English language proficiency score (TOEIC) were examined.

Because some variables or both variables did not have normal distribution, the Spearman correlation coefficient (Spearman, 1904) was used to measure the correlation coefficient.

The multiple regression analysis was employed to study the association between English practice tests and the English language proficiency (TOEIC). Finally, the analysis of the confusion matrix was employed to measure the accuracy of students' self-assessment during practice test engagement. A confusion matrix is a matrix of numbers that tells us where a model gets confused. It is an organized way of mapping the predictions to the original classes to which the data belong (Luque, Carrasco, Martin, & de Las Heras, 2019) Click here to enter text.

Results

Descriptive Statistics for Practice Tests and English Language Proficiency (TOEIC)

Although the participants in this study were majoring in English, they did not perform well on the Test of English for International Communication (TOEIC). The test results, as shown in Table 2, revealed that these students achieved the levels of Elementary Proficiency Plus and Elementary Proficiency (See Table 1) with the average scores of 523 and 272, respectively. They were then divided into two groups according to their TOEIC test scores: Elementary Proficiency Plus (Group A) and Elementary Proficiency (Group B) in order to investigate whether any differences could be found between these two groups of students.

Table 1

TOEIC Score Description

Score Level	Proficiency Description
905 – 990 (91% - 100%)	International Professional Proficiency
785 – 900 (79% - 90%)	Working Proficiency Plus
605 – 780 (61% - 78%)	Limited Working Proficiency
405 – 600 (41% - 60%)	Elementary Proficiency Plus
255 – 400 (26% - 40%)	Elementary Proficiency
0 – 250 (0 – 25%)	Basic Proficiency

Note. Waikato Institute of Education (2013, p. 1)

As shown in Table 2, the average score from the practice tests was 212.74. The skills practiced were grammar, vocabulary, and listening. The average scores for grammar, vocabulary and listening skill practice were 64.65, 76.85 and 71.24 respectively. The English language proficiency was measured by Test of English for International Communication (TOEIC) with 990 total score. The results showed that the average TOEIC score was 329.68 (Elementary Proficiency). The average scores for the students in Group A (Elementary Proficiency Plus) and those in Group B (Elementary Proficiency) were 523 and 272, respectively.

Table 2

Mean and Deviation of English Language Proficiency Score (TOEIC) and Practice Test Score

Score	N	Minimum	Maximum	Mean	Std. Deviation
TOEIC	95	100.00	695.00	329.68	134.8025
Group A	22	405.00	695.00	522.73	79.8917
Group B	73	100.00	395.00	271.51	83.9299
Practice Tests	95	100.00	296.96	212.74	40.6437
Grammar Practice	95	15.00	100.00	64.65	17.3343
Vocabulary Practice	95	41.00	98.25	76.85	13.9510
Listening Practice	95	10.00	98.71	71.24	16.9144

RQ1: How do Thai EFL students' practice test scores correlate with TOEIC score?

In the section, the correlation analysis was employed to describe the relationship between the practice tests and English language proficiency (TOEIC). The correlation coefficient ranges from -1 to 1. If the value is close to -1 or 1, it is shown that two variables have a very strong association. On the contrary, the value close to 0 indicates 'weak association' or 'no association.' The null hypothesis of the analysis is that one variable and the other variable do not correlate.

The relationship between two continuous variables: Practice test score and English language proficiency score (TOEIC) were examined. Because some variables or both variables did not have normal distribution, the Spearman correlation coefficient (Spearman, 1904) was used to measure the correlation coefficient. The result is shown in Table 3.

The practice tests which were self-analyzed as 1, Good, or 0, Not Good, is the value of a dichotomous variable. Therefore, the correlation between the practice test (grammar, vocabulary and listening) and English language proficiency scores (TOEIC) was analyzed with the point Biserial correlation coefficient. The methodology was a special case of Pearson correlation coefficient (Lev, 1949). The result is shown in Table 3.

Table 3

Correlation (Spearman, 1904) between Practice Test Score and English Language Proficiency Score (TOEIC)

Score	English Language Proficiency (TOEIC)	English Practice Tests	Grammar Practice	Vocabulary Practice	Listening Practice
TOEIC	1	0.366[*]	0.145	0.529[*]	0.245[*]
Group A	1	0.233	0.308	-0.231	0.192
Group B	1	0.177	0.009	0.364[*]	0.072
English Practice Tests	0.366[*]	1	0.823[*]	0.731[*]	0.855[*]
Grammar Practice	0.145	0.823[*]	1	0.383[*]	0.632[*]
Vocabulary Practice	0.529[*]	0.731[*]	0.383[*]	1	0.504[*]
Listening Practice	0.245[*]	0.855[*]	0.632[*]	0.504[*]	1

Note. *. Correlation is significant at the 0.05 level (2-tailed).

As shown in Table 3 at the 0.05 level of significance, English practice test score was found to be significantly correlated with TOEIC score. The scores for the grammar, vocabulary and listening skills were found to be significantly correlated with TOEIC scores. The correlation between

vocabulary practice score from the students in Group B (Elementary Proficiency) was found to be statistically significant. No significant correlation was found for the students in Group A (Elementary Proficiency Plus). This result demonstrates that for students in the lower proficiency group, their vocabulary performance in the practice tests correlates with their English performance.

We further examined whether the vocabulary and listening practice scores were associated with English language proficiency score (TOEIC) by employing the multiple regression analysis, which is a statistical technique used to study the association between two or more independent variables that affect changes in a dependent variable. In Table 4, adjusted R-squared value of 0.2533 means that 25.33% of the variance in the English language proficiency score (TOEIC) can be explained by the vocabulary and listening practice score included in the model. As shown in Table 4, the result revealed that the vocabulary practice score correlated with the English language proficiency score (TOEIC) at 0.05 significance level. Moreover, simple linear regression was used to analyze the association between both variables. It was found that if the vocabulary practice score increased by 1, the English language proficiency score (TOEIC) also increased by approximately 4.94 scores. R-squared value was 0.2612. This means that 26.12% of variance in English language proficiency score (TOEIC) was explained by the vocabulary practice scores. Vocabulary knowledge is a good predictor of English language performance.

Table 4

Multiple Regression Result of English Practice Test Score and English Language Proficiency Score (TOEIC)

	Unstandardized Coefficients		t	p-value
	ß	Std. Error		
(Constant)	-39.59	69.307	-0.571	0.569
Vocabulary Practice Score	5.323	1.050	5.071[*]	0.000
Listening Practice Score	-0.558	0.866	-0.644	0.521

Note. Dependent Variable: English language Proficiency sore
Adjusted R-squared = 0.2533
*. Correlation is significant at the 0.05 level (2-tailed)

RQ2: How do Thai EFL students self-assess their English vocabulary, grammar and listening skills during practice test engagement?

In this section, Table 5 shows students' self-assessment of their grammar, listening and vocabulary skills after completing the practice tests. They judged whether they were "Good" or "Not Good" for each skill based on their reflections of their own performance during the practice tests.

Table 5

Students' Analysis of their Grammar, Listening and Vocabulary Skills after Completing the Practice Tests

Vocabulary			Listening		Total
			Not Good	**Good**	
Not good	Grammar	Not good	1 (3.45%)	17 (58.62%)	18 (62.07%)
		Good	10 (34.48%)	1 (3.45%)	11 (37.93%)
	Total		11 (37.93%)	18 (62.07%)	29 (30.53%)
Good	Grammar	Not Good	60 (90.90%)	3 (4.55%)	63 (95.45%)
		Good	3 (4.55%)	0 (0%)	3 (4.55%)
	Total		63 (95.45%)	3 (4.55%)	66 (69.47%)
Total	Grammar	Not Good	61 (64.21%)	20 (21.05%)	81 (85.26%)
		Good	13 (13.68%)	1 (1.05%)	14 (14.74%)
	Total		74 (77.89%)	21 (22.11%)	95 (100%)

Table 5 reports the results from students' self-analysis after completing the vocabulary, grammar and listening practice tests. As shown in Table 5, nine students (30.53%) judged themselves that they were "Not Good" for their vocabulary skills. Among these students, 18 students (62.07%) and 10 students (34.48%) viewed themselves as having good listening and grammar skills, respectively. There was 1 student (3.45%) who judged himself as "Good" for both grammar and listening skills. The highest number of the students who perceived themselves as "Not Good" for grammar but "Good" for listening skills after completing the practice tests was 17 students (58.62%). These results revealed that the students who performed poorly for vocabulary skills in the practice tests, also analyzed themselves as "poor" for their grammar skills, but not for listening skills.

As also shown in Table 5, most of the students (69.47%) assessed themselves as "Good" for their vocabulary skills during the practice tests. Among these students, 60 of them (90.90%) viewed themselves as "Not Good" for both listening and grammar skills. None of them saw themselves as competent for listening and grammar skills. The results from this group showed that the students with positive perception on their vocabulary skills still felt that they had not been doing well for their grammar and listening skills.

When excluding students' analysis on their vocabulary skills, only one student (1.05%) viewed himself as "Good" for both listening and grammar skills. 13 students (13.68%) assessed themselves as "Good" for grammar but "Not Good" for listening skills. 61 students (64.21%) viewed themselves as "Not Good" for both grammar and listening skills.

From the students' self-analysis of their English vocabulary, grammar and listening practice tests, while most of the students perceived themselves as having "good" performance for the vocabulary practice tests, they did not perform well for the grammar and listening skills in the practice tests.

RQ3: How accurate is Thai EFL students' self-assessment on their actual performance in TOEIC?

The analysis of the confusion matrix, as shown in Table 6, was used to measure the accuracy of students' assessment from the practice tests. A confusion matrix is a matrix of numbers that indicates where a model gets confused. It is an organized way of mapping the predictions to the original classes to which the data belong (Luque et al., 2019).

Table 6

Confusion Matrix

	Actually Positive (1)	Actually Negative (0)
Predicted Positive (1)	True Positives (TPs)	False Positives (FPs)
Predicted Negative (0)	False Negatives (FNs)	True Negatives (TNs)

True Positive (TP) and True Negative (TN) refer to the accurate prediction of the actual performance. For example, if something is predicted to happen, it actually happens (TP), or if something is predicted not to happen, it actually does not happen (TN).

False Positive (FP) and False Negative (FN) refer to the inaccurate prediction of the actual performance. For example, if something is predicted to happen, it actually does not happen (FP), or if something is predicted not to happen, it actually happens (FN).

Since Parts 1 and 5 of the TOEIC test examine students' listening and grammar skills, in the current study, students' prediction from their grammar and listening practice tests was measured using the confusion matrix to investigate whether they were accurate for their self-assessment of their grammar and listening skills. The results for these skills are shown in Tables 7 and 8, respectively.

Table 7

Confusion Matrix of the Self-Analysis on Grammar Performance from Practice Tests

The Self-Analysis on Grammar Proficiency (Practice Tests)	Actual Grammar Performance (TOEIC)	
	Good	Not Good
Good	4 (4.21%)	10 (10.53%)
Not good	7 (7.37%)	74 (77.89%)

The results in Table 7 showed that only 4 students (4.21%) correctly predicted their actual grammar performance to belong to "Good" group (True Positive). Most of the students (77.89%) accurately predicted their actual grammar performance to belong to the "Not Good" group (True Negative). While the number of incorrect predictions was 17.90% (False), most of the students (82.10%) correctly predicted their actual grammar performance (True).

Table 8

Confusion Matrix of the Self-Analysis on Listening Performance from Practice Tests

The Self-Analysis on Listening Proficiency (Practice Tests)	Actual Listening Performance (TOEIC)	
	Good	Not Good
Good	0	21(22.11%)
Not Good	2 (2.11%)	72 (75.79%)

Table 8 showed that none of the students correctly predicted their actual listening performance to belong to the "Good" group (True Positive). Most of the students (75.79%) accurately predicted their actual listening performance to belong to the "Not Good" group (True Negative). While the number of incorrect predictions was 24.21% (False), most of the students (75.79%) correctly predicted their actual listening performance (True).

Discussion

Andrade (2019) underscores the value of formative self-assessment in allowing students to monitor their learning and provide feedback to themselves, which is crucial for academic improvement. Furthermore, the process-oriented nature of self-assessment, as supported by Panadero et al., (2015), facilitates self-regulation, empowering students to independently navigate their learning journey by recognizing and rectifying errors. As Boston (2002) points out, self-assessment enables lower achieving students to improve their learning performance because they are able to focus on the areas that need improvement.

Since students are the key factors for their learning success, the present study investigated the impact of Thai EFL students' formative self-assessment on their English performance. As the results presented , most of the Thai EFL students had a positive perception for their vocabulary skills, but they viewed themselves as poor performers for grammar and listening skills. The findings align with Wang and Rajprasit's (2015) self-perception study. They found that Thai EFL students believed that vocabulary played an important role in English learning.

The findings demonstrated that these students benefited from the practice tests and were able to see their strengths in vocabulary improvement but weaknesses in their attempts for grammar and listening sections. The students' accurate perception for their vocabulary performance was confirmed by the correlation analysis, the multiple regression and the confusion matrix results. A positive correlation was found between vocabulary practice and English proficiency (TOEIC) scores. The findings were in line with Karakoç and Köse's (2017) study. They found a positive correlation between vocabulary knowledge and English ability. Milton (2013) concluded that improving vocabulary knowledge results in good performance in language skills.

In terms of academic improvement, engagement in vocabulary practice tests enabled Thai EFL students to perform better in the English performance test (TOEIC). Findings are in line with a number of previous studies (e.g., Alqahtani, 2015; Laufer & Nation, 1995; Meebangsai, Pongtin, Kitipoontanakorn, & Laosrirattanachai, 2023). For the benefits of practice tests, Martin et al. (2020) reported on the effectiveness of practice test engagement on academic performance. As explained by Roediger and Karpicke (2006), practice tests enhance retention of previously learned information even if no feedback was provided. Practice tests provide the opportunity for students to retrieve and recall information from memory and significantly benefit learning (Carpenter, 2023). Therefore, engaging students with practice tests is clearly beneficial for their academic improvement. However, since these students' level of English proficiency was very low, their background knowledge for listening and grammar was not sufficient for them to benefit from the practice tests of listening and grammar skills. Future research is needed for factors that can enhance listening and grammar skills when these students engage in the practice tests on their own. As pointed out by Pan, Dunlosky, Xu, and Ouwehand (2024), "Over the past decade-and-half, researchers have uncovered a great deal about test-enhanced learning. As the articles in this special issue indicate, however, that work is far from over. There remain many unanswered research questions, under-explored approaches to practice testing, and other dimensions of test-enhanced learning that have yet to be thoroughly investigated" (p.20).

For the accuracy of self-assessment, the results from the confusion matrix confirmed that these Thai EFL students correctly self-perceived their grammar and listening skills. During the practice tests, the students perceived themselves as "Not Good" for both grammar and listening skills. Their negative self-perception for these skills was confirmed by their poor performance shown in the actual TOEIC score. These students did not overestimate or underestimate their English abilities, as reported in some of the previous studies (e.g., Kuncel et al., 2005). The results also lend support to the self-assessment research by Konchiab and Munpanya (2021). They found that Thai EFL students were not confident with their pronunciation, vocabulary, and grammar for their oral presentation. They did not rate themselves highly for these language features and reflected their needs for improving these skills.

Conclusion and Implications

Research has shown that formative self-assessment is one of the highly regarded educational interventions for academic enhancement (Black & Wiliam, 1998). One of the key factors that enables teachers to better serve students' varieties in terms of needs and levels of achievements is by using formative assessment to identify these varieties and adjust their teaching styles and techniques that can produce positive learning outcomes (OECD/CERI, 2008). Formative self-assessment cultivates learning-to-learn skills and actively engages students in the learning process. This enables students to better understand their learning processes, and effectively adjusts and develops strategies that can yield positive outcomes for their learning achievements (OECD/CERI, 2008).

This study hopes to contribute to a growing body of research on self-assessment. The findings revealed that the process of formative self-assessment during the completion of multiple practice tests enabled Thai EFL students to be more accurate in their self-perception of their English abilities. They were able to correctly judge their strengths and weaknesses for each of the English skills. The completion of multiple practice tests enhanced EFL students' vocabulary skills and enabled them to improve their English proficiency test score. However, the findings revealed that these students were not able to improve their grammar and listening skills during and after the completion of the practice tests. Future research should be taken to further investigate the different results found among different language skills. While students benefited from the practice tests for vocabulary skills improvement, further investigation should be conducted for the different results found for grammar and listening skills.

Pedagogical Implications

As demonstrated in the findings, self-assessment empowers EFL students with the awareness of their strengths and weaknesses and enables them to be independent and autonomous. Blue (1988) suggests that formative self-assessment should be part of the lessons learned in class along with other aspects of learning such as goal setting or selecting learning materials. As Andrade (2019) points out, "self-assessments of competence are only useful if students have opportunities to do something about their perceived low competence - that is, it serves the purpose of formative feedback for the learner" (p. 87). Students can fully benefit from self-assessment when they have the opportunity to relearn and revise (Andrade, 2010).

Self-assessment allows students to be more responsible for their own learning (Dickinson, 1978). This is echoed by Zimmerman and Schunk (2011) and Boud (2003), who note that self-assessment is linked to enhanced academic achievement and encourages students to take responsibility for their learning. It is essential to engage EFL students with their own learning process. When students are involved in their assessment process, this can lead to effective learning. They can link and apply what was previously learned for their future learning (Hosseini & Nimehchisalem, 2021; Little,

2011). As Yan and Brown (2017) state, students benefit from self-assessment processes because they actively engage in this learning process.

Limitation of the Study

Since the participants in this study were English major students, the generalizability of the findings may not extend to other groups of EFL students whose majors are not English.

Acknowledgements

References

Alderson, C. J., & Huhta, A. (2005). The development of a suite of computer-based diagnostic tests based on the Common European Framework. *Language Testing, 22*(3), 301-320.

Alqahtani, M. (2015). The importance of vocabulary in language learning and how to be taught. *International Journal of Teaching and Education, III*(3), 21-34.

Andrade, H. (2010). Students as the definitive source of formative assessment: Academic self-assessment and the self-regulation of learning. In H. Andrade & G. Cizek (Eds.), *Handbook of Formative Assessment* (pp. 90-105). New York, NY: Routledge.

Andrade, H. (2019). A critical review of research on student self-assessment. *Frontiers in Education, 4.* https://doi.org/10.3389/feduc.2019.00087

Andrade, H., & Valtcheva, A. (2009). Promoting learning and achievement through self-assessment. *Theory into Practice, 48*(1), 12–19. https://doi.org/10.1080/00405840802577544

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *International Journal of Phytoremediation, 21(*1), 7-74. https://doi.org/http://dx.doi.org/10.1080/0969595980050102

Blue, G.M. (1988). Self-assessment: The limits of learner independence. In A. Brooks & P. Grundy (Eds.), *Individualization and autonomy in language learning: ELT document 131* (pp. 100-118). Hong Kong: Modern English Publications / British Council.

Boston, C. (2002). The concept of formative assessment. *Practical Assessment, Research & Evaluation, 8*(1),9. https://doi.org/10.7275/kmcq-dj31

Boud, D. (1995). Enhancing learning through self-assessment. London: Routledge. https://doi.org/https://doi.org/10.4324/9781315041520

Boud, D. (2003). Enhancing learning through self-assessment. London: Routledge Falmer, Taylor & Francis Group.

Bourke, R. (2014). Self-assessment in professional programmes within tertiary institutions. *Teaching in Higher Education, 19*(8), 908-918.

Bullock, D. (2011). Learner self-assessment: An investigation into teachers' beliefs. *ELT Journal, 65(2),* 114-125. https://doi.org/https://doi.org/10.1093/elt/ccq041

Butler, Y. G., & Lee, J. (2010). The effects of self-assessment among young learners of *English. Language Testing, 27*(1), 5–31. https://doi.org/10.1177/0265532209346370

Carpenter, S. K. (2023). Encouraging students to use retrieval practice: A review of emerging research from five types of interventions. *Educational Psychology Review, 35.* https://doi.org/10.1007/s10648-023-09811-8

Carpenter, S.K., Pashler, H., & Vul, E. (2006). What types of learning are enhanced by a cued recall test? *Psychonomic Bulletin & Review, 13*(5), 826-830. https://doi.org/10.3758/BF03194004

Cogliano, M. C., Kardash, C. A. M., & Bernacki, M. L., (2019). The effects of retrieval practice and prior topic knowledge on test performance and confidence judgments, 56, 117- 129. http://dx.doi.org/10.1016/j.cedpsych.2018.12.001

Dhanarattigannon, J., & Thienpermpool, P. (2022). EFL tertiary learners' perceptions of self-assessment on writing in English. *LEARN Journal: Language Education and Acquisition Research Network, 15*(2), 521-545. https://so04.tci-thaijo.org/index.php/LEARN/index

Dickinson, L. (1978). Autonomy, self-directed learning and individualization. In A. Brooks & P. Grundy (Eds.), *Individualization and autonomy in language learning: ELT document 131* (pp. 7-28). Hong Kong: Modern English Publications and the British Council.

Graham, S., Hebert, M., & Harris, K. R. (2015). Formative assessment and writing. *Elementary School Journal, 115*(4), 523-547. https://doi.org/https://doi.org/10.1086/681947

Hattie, J. (2009). Visible learning: A synthesis of over 800 meta-analyses relating to achievement. New York, NY: Routledge.

Hosseini, M., & Nimehchisalem, V. (2021). Self-assessment in English language teaching and learning in the current decade (2010-2020): A systematic review. *Open Journal of Modern Linguistics, 11*(6), 854-872. https://doi.org/10.4236/ojml.2021.116066

Karakoç, D., & Köse, G. D. (2017). The impact of vocabulary knowledge on reading, writing and proficiency scores of EFL learners. *Journal of Language and Linguistic Studies, 13*(1), 352-378.

Karaman, P. (2021). The impact of self-assessment on academic performance: A meta-analysis study. *International Journal of Research in Education and Science, 7*(4), 1151–1166. https://doi.org/10.46328/ijres.2344

Karpicke, J. D., Butler, A. C., & Roediger III, H. L. (2009). Metacognitive strategies in student learning: Do students practise retrieval when they study on their own? *Memory, 17*(4), 471–479. https://doi.org/10.1080/09658210802647009

Kavaliauskienė, G. (2004). Quality assessment in teaching English for specific purposes. *English for Specific Purposes World, 3*, 8-17.

Killimangalam, A. (2020). Student self-assessment: boon or bane? *Teacher, 14*(2), 30-33.

Konchiab, S., & Munpanya, P. (2021). Investigating Thai EFL undergraduates' oral presentation performances and experiences, using teacher and student self-assessments. *THAITESOL JOURNAL, 34*(1), 96–117.

Kostons, D., van Gog. T., & Paas, F. (2012). Training self-assessment and task-selection skills: A cognitive approach to improving self-regulated learning. *Learning and Instruction, 22*(2), 121-132. https://doi.org/https://doi.org/10.1016/j.learninstruc.2011.08.004

Kuncel, N. R., Crede, M., & Thomas, L. L. (2005). The validity of self-reported grade point averages, class ranks, and test scores: A meta-analysis and review of the literature. *Review of Educational Research, 75*(1), 63-82. http://dx.doi.org/10.3102/00346543075001063

Laufer, B., & Nation, P. (1995). Vocabulary size and use: Lexical richness in L2 written production. *Applied Linguistics, 16*(3), 307-322.

Lev, J. (1949). The point biserial coefficient of correlation. *Annals of Mathematical Statistics, 20*(1), 125-126.

Little, D. (2011). Learner autonomy, self-assessment and language tests: Towards a new assessment culture. In B. B. Morrison (Ed.), *Independent language learning: Building on experience, seeking new perspectives* (pp. 25-39). Hong Kong: Hong Kong University Press.

Lopez, R., & Kossack, S. (2007). Effects of recurring use of self-assessment in university courses. *International Journal of Learning, 14*(4), 203-216.

Luque, A., Carrasco, A., Martin, A., & de Las Heras, A. (2019). The impact of class imbalance in classification performance metrics based on the binary confusion matrix. *Pattern Recognition, 91*, 216-231.

Mahdavinia, M., & Ahmadi, L. N. (2011). Portfolio Assessment: A tool for self-directed learning at post-secondary level. In D. Gardner (Ed.), Fostering Learner Autonomy in Language Learning (pp. 76-89). Gaziantep: Zirve University.

Martin, S., English, T., Morley, L., O'Keefe-Quinn, T., & Whitfield, P. (2020). Practice tests improve performance, increase engagement and protect from psychological distress. Proceedings from *International Conference on Higher Education Advances, 2020-June*, 811–818. https://doi.org/10.4995/HEAd20.2020.11151

Mazloomi, S., & Khabiri, M. (2016). The impact of self-assessment on language learners' writing skill. *Innovations in Education and Teaching International, 55*(1), 91-100. https://doi.org/10.1080/14703297.2016.1214078

Meebangsai, D., Pongtin, P., Kitipoontanakorn, P., & Laosrirattanachai, P. (2023). Investigating proficiency of academic English in student writing: A comparative case study on vocabulary utilization in student research article writing visà-vis national and international research. *PASAA, 67*(1), 66-100.

Milton, J. (2013). Measuring the contribution of vocabulary knowledge to proficiency in the four skills. In C. Bardel, C. Lindquist, & B. Laufer (Eds), *L2 vocabulary acquisition, knowledge and use: New perspectives on assessment and corpus analysis* (pp. 57-78). Eurosla Monographs Series, 2.

Murakami, C., Valvona, C., & Broudy, D. (2012). Turning apathy into activeness in oral communication classes: Regular self- and peer-assessment in a TBLT programme. *System, 40*(3), 407-420.

Naujoks, N., Harder, B. & Händel, M. (2022). Testing pays off twice: Potentials of practice tests and feedback regarding exam performance and judgment accuracy. *Metacognition Learning, 17*, 479-498. https://doi.org/10.1007/s11409-022-09295-x

Ndoye, A. (2017). Peer/self-assessment and student learning. *International Journal of Teaching and Learning in Higher Education, 29*(2), 255-269.

Nunan, D. (2004). *Task-based language teaching*. Cambridge: Cambridge University Press.

OECD/ CERI. (2008). Proceedings from *International Conference Learning in the 21st Century: Research, Innovation and Policy*. Retrieved from https:// www. csuchico. edu/ futurepossibilities/ _assets/ documents/ oecd- 21st- century-learning-40554299.pdf

Pan, S. C., Dunlosky, J., Xu, K. M., & Ouwehand, K. (2024). Emerging and future directions in test-enhanced learning research. *Educational Psychology Review, 36*(1). https://doi.org/10.1007/s10648-024-09857-2

Panadero, E., Brown, G. T. L., & Strijbos, J. W. (2015). The future of student self-assessment: A review of known unknowns and potential directions. *Educational Psychology Review, 28*(4), 803-830. https://doi.org/10.1007/s10648-015-9350-2

Piamsai, C. (2023). Development and use of CEFR based self-assessment in a Thai tertiary context article information. *PASAA Journal, 66*(1), 81-126.

Ramdass, D., & Zimmerman, B. J. (2008). Effects of self-correction strategy training on middle school students' self-efficacy, self-evaluation, and mathematics division learning. *Journal of Advanced Academics, 20*(1), 18-41. https://doi.org/10.4219/jaa-2008-869

Rea, P. R. (1981). Formative assessment of student performance: The role of self-appraisal. *Indian Journal of Applied Linguistics, 7*, 66-68.

Roediger, H. L., & Karpicke, J. D. (2006). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science, 17*(3), 249-255. https://doi.org/https://doi.org/10.1111/j.1467-9280.2006.01693.x

Sanchez, C. E., Atkinson, K. M., Koenka, A. C., Moshontz, H., & Cooper, H. (2017). Self-grading and peer-grading for formative and summative assessments in 3rd through 12th grade classrooms: A meta-analysis. J*ournal of Educational Psychology, 109(8)*, 1049–1066.

Sapsirin S. (2014). Self-assessment of business English writing ability of Thai university students: Criterion-related validity and gender differences. *NIDA Journal of Language and Communication, 19*(22), 77–91.

Seenak, P., & Adunyarittigun, D. (2019). Effects of self- and peer-assessments on Thai EFL students' intonation learning. *Social Science Journal, 8*(1), 1–15.

Shelton-Strong, S. J. (2018). Fostering the development of language learner autonomy through Peer and self-Assessment. *Relay Journal, 1*(1), 21-46. https://doi.org/https://doi.org/10.37237/relay/010103

Spearman, C. (1904). The proof and measurement of association between two things. *The American Journal of Psychology, 15*(1), 72-101.

Suwanarak, K. (2018). Self-evaluation of Thai adult learners in English writing practice. *In 3L: Language, Linguistics, Literature, 24*(2), 95-111. https://doi.org/10.17576/3L-2018-2402-08

Waikato Institute of Education. (2013). TOEIC® test scores and what they mean. Retrieved from https://www.wie.ac.nz/TOEICconversion.htm

Wang, T., & Rajprasit, K. (2015). Identifying affirmative beliefs about English language learning: Self-perceptions of Thai learners with different language proficiency. *English Language Teaching, 8*(4), 1–13. https://doi.org/10.5539/elt.v8n4p1

Wangsotorn, A. (1981). Self-assessment in English skills by undergraduate and graduate students in Thai universities. In J. Read (Ed.), *Directions in language testing* (pp.240-260). Singapore: Singapore University Press.

Yan, Z., & Brown, G. T. L. (2017). A cyclical self-assessment process: Towards a model of how students engage in self-assessment. *Assessment and Evaluation in Higher Education, 42*(8), 1247-1262. https://doi.org/10.1080/02602938.2016.1260091

Zimmerman, B. J., & Schunk, D. H. (2011). Self-regulated learning and performance: An introduction and overview. In B. J. Zimmerman & D. H. Schunk (Eds.), *Handbook of Self-Regulation of Learning and Performance* (pp. 1-14). London: Routledge.