Effects of Integrated Performance Assessment Modules on English Speaking Ability of Thai EFL Undergraduate Students

Benjawan Plengkham[a,*], Punchalee Wasanasomsithi[b]

[a] ajarnbenjawan@hotmail.com, English as an International Language Program, Graduate School, Chulalongkorn University, Thailand
[b] punchaleew@yahoo.com, Chulalongkorn University Language Institute, Thailand
* Corresponding author, ajarnbenjawan@hotmail.com

ABSTRACT

Integrating assessment into speaking instruction has been a challenge for English language instructors. The Integrated Performance Assessment (IPA), proposed by the American Council on the Teaching of Foreign Languages (ACTFL) (Adair-Hauck et al., 2015), is an assessment protocol featuring three modes of communication: interpretive, interpersonal, and presentational, to reinforce second and foreign language learning. This present mixed-method research aimed to investigate the effects of the IPA modules on Thai EFL undergraduate students' speaking ability. Three IPA modules were taught over the course of nine weeks to 33 Thai EFL undergraduate students enrolled in an English speaking course, and the students' performance was assessed using an adapted IPA rubric. The research instruments included a pretest, a posttest, unit tests, an IPA opinion questionnaire, and a semi-structured interview protocol. The findings indicated that the implementation of three IPA modules had a positive effect on students' speaking ability as evidenced by the posttest results. In addition, the findings obtained from the IPA opinion

questionnaire and the semi-structured interview protocol yielded further evidence of positive effects of the IPA, particularly on students' perceived speaking improvement. Based on the study findings, it is suggested that the three IPA modules can be used as an instructional and assessment model to promote students' speaking ability as well as to elicit insights into how English language teaching and assessment can be seamlessly integrated.

Keywords: integrated performance assessment, speaking ability, EFL, undergraduate students

Introduction

English language instruction has shifted away from grammar-based instruction toward communication and performance-based learning in recent years (Adair-Huack et al., 2006). According to the American Council on the Teaching of Foreign Languages (ACTFL) (2011), language educators have come to the belief that the objective of language teaching is to increase students' language competency through modes of communication that seem to mirror real-life situations. This is essential since language and communication are considered the most fundamental aspects of human experience. When it comes to communication skills, it is not questionable that conversation is a crucial part of education and teaching pedagogy. However, it has been discovered that EFL students are often less motivated to learn English speaking when compared to ESL students. More importantly, many of them are unable to communicate in English. This is because their opportunities to speak and practice English outside of the classroom are limited due to a lack of English-speaking environments (Nakhalah, 2016). It is worth noting that although many educational institutes in EFL settings provide English speaking courses, each institute's teaching pedagogy is different. In other words, despite the fact that Communicative Language Teaching (CLT) has been implemented in many countries, including Thailand, the English speaking ability of EFL students appears to be unsatisfactory (Atagi, 2011). Because of this, merely teaching students to speak may not be adequate to improve their ability to communicate; it is also necessary to actively involve them in the learning process at the same time.

Choosing an assessment technique is a crucial part of language instruction. This is because assessment reflects how teachers recognize students' needs, signal their progress, and establish their own educational plans (Frank, 2012); therefore, selecting an appropriate assessment technique can help students learn more effectively. Traditional assessment, on the one hand, is defined as a method of evaluation that is generally focused on pencil-

and-paper examinations and quizzes (Mertler, 2017). Traditional evaluation is extensively utilized in classroom assessment since the instructor may be required to examine a large number of students at once. In traditional assessment, the instruction and assessment are seen as two separate constructs, and assessment always occurs after a unit of instruction has been completed to determine whether and how well students have achieved the learning goals, which is signified in the form of letter grades or comments. It is often argued that such forms of assessment might not help students develop their language skills fully because they do not give sufficient information for them to improve themselves (Tunstall & Gipps, 1996). Alternative assessment, on the other hand, enables instructors to place their focus more on the students' learning process rather than just the final output. It is believed that alternative assessment may be more effective to help students develop mastery of the language skills as instructors, instead of having to wait until the end of the instruction period, have more chances to learn about their students' progress and to help them overcome challenges while they are still in the process of learning.

In terms of assessment types, summative and formative assessments have both been employed in language classrooms. Summative evaluation is described as an assessment strategy that intends to assess what students have learned and how well they have achieved the course goals at the end of the course (Brown, 2004). However, it is clear that summative evaluation alone is unlikely to be sufficient in assessing students' learning progress. Consequently, language instructors have increasingly turned to formative assessment to reflect how students learn the language. Formative assessment is a method of evaluating students' performance in terms of their progress in developing their language competence and abilities with the goal of continuing their learning development process (Brown, 2004). Formative assessment may not only assist language instructors in providing immediate assistance to students, but it may also enable students to develop themselves throughout the learning process. This is because formative evaluation focuses on students' performance as it occurs and actively incorporates both instructors and students in the learning process (Madison-Harris et al, 2012).

There are different types of assessments that may be used to assess students formatively (Brown, 2004). The Integrated Performance Assessment (henceforth IPA) is seen as one form of assessment that offers instructors and language learners a rich opportunity to focus on the learning progress. The IPA was first introduced by the American Council on the Teaching of Foreign Languages (ACTFL). The purpose of the implementation of the IPA is to seamlessly integrate assessment into second and foreign language instruction. As such, IPA changes the focus of assessment from pure memorization to students' actual performance by themselves or with their

peers. At present, IPA has been implemented in many foreign language classes including Spanish, Chinese, and French, in addition to English (Center for Advanced Research on Language Acquisition, 2019). However, it is noteworthy that even though the implementation of the IPA has been explored in different language learning contexts and with positive results (Adair-Hauck et al., 2015), to date, little research on the IPA with EFL students has been undertaken, particularly in English speaking courses. This study, therefore, aimed to investigate the effects of the IPA on Thai EFL students' speaking ability to help shed more light on the effectiveness of this assessment protocol when it is used in the Thai context. Accordingly, two research questions were formulated as follows:

1. What are the effects of IPA modules on undergraduate students' English speaking ability?
2. What are undergraduate students' attitudes towards the implementation of IPA modules?

Literature Review

Integrated Performance Assessment

The Integrated Performance Assessment (IPA) is defined as "a multi-task cluster assessment featuring three tasks, each of which reflects one of the three modes of communication, namely interpersonal, interpretive, and presentational" (Adair-Hauck et al., 2015, p. 22). The multi-task cluster assessment in the IPA is outlined in the Standards for Foreign Language Learning in the 21st century (National Standards for Foreign Language Education Project, 1996, 1999, 2006). The standard features five goal areas, known as the Five Cs, as outlined in the World-Readiness Standards for Foreign Languages: Communication, Cultures, Connections, Comparisons, and Communities (Adair-Hauck et al., 2015). The IPA can evaluate students' language proficiency at three levels: novice, intermediate, and advanced, and the use of standard-based and performance-based language evaluation is integral to the assessment process (Center for Advanced Research on Language Acquisition, 2019). The IPA is utilized in conjunction with a scoring rubric to assess students' performance in terms of whether they meet, exceed, or do not meet expectations upon completion of the assigned task.

According to the theoretical framework of the IPA, there are six important characteristics of which language instructors and course developers should be aware when wishing to implement the IPA in their classroom.

Figure 1

IPA theoretical framework

To begin with, the IPA is based on the three modes of communication, namely interpretive, interpersonal, and presentational, as outlined in the Standards for Foreign Language Learning in the 21st Century (National Standards for Foreign Language Education Project, 1996, 1999, 2009, as cited in Adair-Hauck et al., 2015). These three modes are integrated as a single unit of assessment and are all aligned with the same theme throughout the tasks in the unit. Each task reflects the necessary information as well as the linguistic interaction needed for students to perform the subsequent tasks in the same unit of assessment. Theoretically, the interpretive IPA refers to the interpretation of meaning from the written or spoken form. As for the interpersonal mode, it is defined as the active negotiation of meaning among individuals. The last IPA mode is presentational, which is the production of oral or written messages with the interpretation made by the audience. These principles serve as a key component of the design of an IPA unit (Adair-Hauck et al., 2015; Sandrock, 2008).

As for the second characteristic, the IPA is a performance-based assessment that enables students to perform tasks that are meaningful, motivating, and suitable for their backgrounds (Adair-Hauck et al., 2015; Sandrock, 2008). It has been found that a performance-based assessment shows great potential for students' performance. The principle of performance-based assessment has influenced the design of the IPA, as it could enable learners to apply their linguistic and cultural backgrounds to communicative tasks. (Sandrock, 2008).

Thirdly, the IPA is an authentic assessment as it mirrors real-world tasks (Adair-Hauck et al., 2015; Sandrock, 2008). Authentic assessment refers

to an assessment method that is based on real-world tasks involving the meaningful application of fundamental knowledge and skills (Mueller, 2005). Accordingly, IPA units are designed to feature authentic tasks, which are based on real-world situations. This authentic assessment helps to provide the real-world context necessary to assess students' competence based on the 5C goal areas.

Another characteristic of the IPA is its developmental nature. Development in the IPA refers to the developmental progress of language learners towards the specific levels of proficiency outlined in the ACTFL Proficiency Guidelines (2012), which describe what language learners can do with the language in terms of the four skills of speaking, writing, listening, and reading at different levels of proficiency based on real-world situations. The proficiency levels are ranged along a continuum of language learning, labeled Novice, Intermediate, and Advanced.

Additionally, the IPA is a standards-based assessment focusing on the integration of the standard goal areas outlined in the World-Readiness Standards for Learning Languages (5Cs). Apart from emphasizing the Communication standard in the three modes of communication, the IPA also integrates the other standard goal areas in the 5Cs, namely Cultures, Connections, Comparisons, and Communities. The incorporation of the Communication goal in the three modes of communication with other standard goal areas is reflected through the IPA assessment tasks (Adair-Hauck et al., 2015; Sandrock, 2008; Center for Advanced Research on Language Acquisition, 2019).

The last important characteristic of the IPA is its seamless connection between instruction and assessment. In other words, the IPA blends classroom instruction and experiences by rejecting the disconnection between instruction and assessment. In doing so, it redefines the practice of assessment. Recent research on language assessment has emphasized the use of integrated assessment for improving instruction and learning. Accordingly, it is believed that the IPA's combination of performance-based and authentic assessment, together with its model of performance descriptions and the role of feedback in assessment, can enable it to link instruction and assessment in a seamless manner (Adair-Hauck et al., 2015; Sandrock, 2008).

Speaking ability

Speaking ability is defined in various ways. Brown (2004) states that speaking is primarily related to listening since it is difficult to constrain speech alone without the interaction of aural awareness. Speaking ability is separated into two subskills: microskills and macroskills. Microskills stress smaller parts of the language, whereas macroskills emphasize larger elements of the

language such as discourse, fluency, and nonverbal communication (Brown, 2004).

According to Goh and Burns (2012), despite their awareness of the significance of speaking, language instructors are not often equipped with the skills they need to teach speaking due to a lack of a particular speaking instructional model. As such, it is imperative that instructors adopt a suitable speaking teaching model to effectively develop the speaking skills and ability of their students.

Figure 2

The Conceptual framework of the study
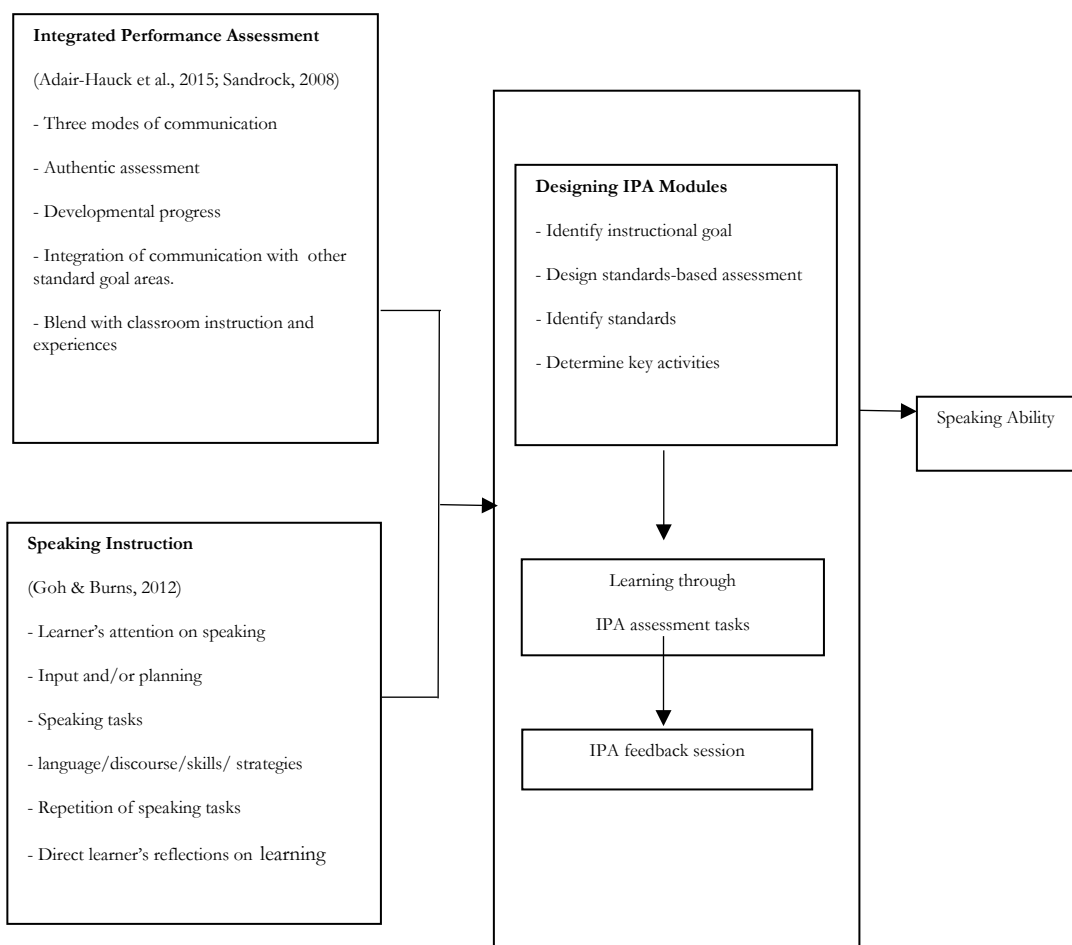
In this study, the IPA modules were developed based on two theoretical frameworks: the IPA framework proposed by Adair-Hauck et al., (2015) and Sandrock (2008) and the speaking instruction framework (Goh &

Burns, 2012). These frameworks were selected because the speaking instruction cycle and the IPA seemed to share some common concepts in terms of the characteristic of learning activities. Hence, the speaking tasks in the IPA modules were designed by using principles of both frameworks.

To design each IPA module, the instructional goal was first identified. After that, standards-based assessment was employed to ensure that the module included communication standards as well as integrating the other goal areas outlined in the 5Cs standards. The key activities were then designed and appropriate resources were provided for students to complete the tasks. The next component to be designed was the IPA assessment tasks, which offered opportunities for students to learn through assessment application. The last component was the IPA feedback session, which allowed students to receive co-constructed feedback on their performance and thus improve their speaking ability.

Methodology

The present study employed a mixed-methods research design to collect data on students' speaking ability. A one-group pretest-posttest design was used to investigate the effects of the IPA modules on Thai EFL undergraduate students' English speaking ability. An IPA opinion questionnaire was administered to explore students' opinions on the implementation of the IPA modules, particularly their effects on their speaking ability. Also, a semi-structured interview protocol was utilized to elicit in-depth information from students.

Participants

Thirty-three third-year undergraduate students majoring in English Education who were enrolled in an English speaking course at the Faculty of Humanities and Social Sciences at a government university in Nakhon Pathom Province, Thailand, participated in this study. There were both 14 male and 19 female participants. As for the English proficiency of the participants, most of them were at an intermediate level. Students were required to enroll in this course in their third year of study since it was a compulsory course.

Data Collection and Research Instruments

Data collection was conducted over the course of nine weeks in the first semester of the 2021 academic year. Four research instruments were pilot tested and then used in this study as follows:

Pretest and Posttest

A pretest was conducted before the implementation of the IPA modules. The pretest and posttest included three test tasks that reflected three IPA modes: interpretative, interpersonal, and presentational.

The first task tested students' ability to interpret meaning. In this task, they watched a video and completed four assessment activities. In the second task, students were tested on their ability to negotiate meaning with the interlocutor by talking together and exchanging ideas in the form of a conversation. In the final task, students were assigned to deliver a speech on the same topic as the first two tasks. After nine weeks of the implementation of the three IPA modules, students completed the posttest, which shared the same test constructs with the pretest but with a different topic. In terms of scoring, the ACTFL IPA rubric was adapted based on the context of this study. The students were assessed on whether they exceeded, met, or did not meet the performance expectations described in the rubric. The description of each task in the pretest and posttest is illustrated below.

**Table 1**

The element of pretest and posttest

Pretest - Posttest			
IPA Modes	**Assessment Criteria**	**Assessment Tasks**	**Time Allocation**
Interpretive	Keyword recognition	Students watched a movie trailer and provided 10 words they heard from the video.	15 minutes
	Main idea detection	Students watched the whole movie and wrote about the main idea of the story.	15 minutes (After watching movie)
	Supporting detail detection	Students wrote the supporting details of the story based on the words given.	15 minutes
	Cultural perspectives	Students answered questions related to the cultural perspectives found in the movie compared to their own culture.	15 minutes
Interpersonal	Language Function	Students had a conversation about the movie with the instructor in the form of an interview.	5 minutes
	Discourse Type		
	Communication Strategies		
	Comprehensibility		
	Language Control		
Presentational	Language Function	Students delivered a two-minute speech to review the movie.	2 minutes
	Discourse Type		
	Impact		
	Comprehensibility		
	Language Control		

Unit Tests

Three unit tests were conducted in each unit. The test construct was similar to that of the pretest and posttest. Each unit took three weeks to cover, and students completed the unit test upon completion of each IPA task in each unit. The IPA modules task cycle is illustrated below:

Figure 3

IPA Modules Task Cycle
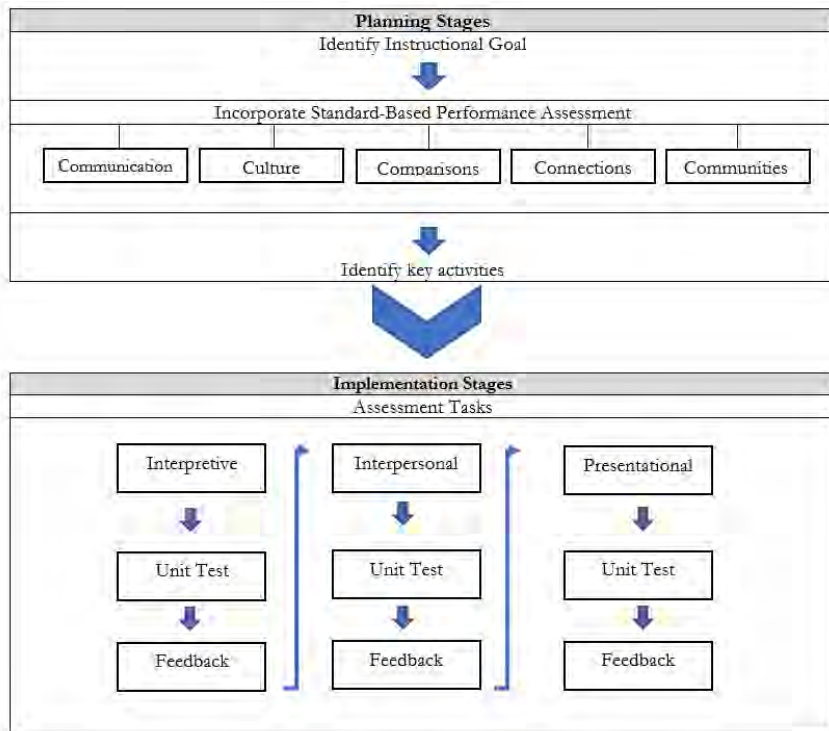
According to Figure 3, each IPA module was designed based on the aforementioned conceptual framework. The three modes of communication—interpretive, interpersonal, and presentational—were the main focus in designing the IPA modules. The 5Cs standards were also integrated into the assessment tasks. Moreover, the assessment tasks were constructed based on the students' proficiency levels, which were at the intermediate level in the current study.

As for the implementation stage, three IPA modules were taught over the nine weeks. Each IPA module employed a different theme, while the test constructs were the same. After each IPA module, students completed the

unit test, which featured three assessment tasks, each one corresponding to one of the three modes: interpretive, interpersonal, and presentational. The format of tasks was similar to those in the pretest and posttest, but with different themes. In terms of assessment criteria, the adapted IPA rubric was used to evaluate students' performance in each IPA task.

After each unit test, each student co-constructed IPA feedback with the instructor. In terms of the nature of the feedback loop, the instructor had a conversation with each student and discussed their performance on each IPA task. The exchange between the student and the instructor included questions from the teacher and self-reflection from the student. Through the co-constructed feedback, the instructor encouraged students to realize their strengths and weaknesses. The instructor avoided making judgmental statements in the feedback session.

IPA Opinion Questionnaire

The IPA opinion questionnaire was designed by the researcher to explore students' opinions on the implementation of the IPA modules and whether or not they perceived that the IPA modules had helped improve their speaking ability. To ensure the validity of the questionnaire, the Item-Objective Congruence Index (Henceforth IOC) was used by having three experts evaluate the questions. The Cronbach's alpha was employed to evaluate its reliability showing a high reliability of .911. The questionnaire was administered after the course ended.

Semi-structured interview protocol

To elicit more insightful information from the participants, a semi-structured interview protocol developed by the researcher was also used. Six students were chosen for the interview. The criteria for selecting the participants was based on the students' scores. Two students from each of the three score levels—high, middle, and low—were asked to take part in the interview. The interview was conducted one week after the students completed the questionnaire. There were ten questions, which had already been validated by using IOC, that investigated the effect of the IPA module. The theme of the interview questions was derived from the questions in the questionnaire in order to provide more insight from the data obtained from the questionnaire. The questions were based on the overall satisfaction, the development of English speaking ability through the IPA modules, the IPA instructional activities, and the IPA feedback session.

Data Analysis

As for quantitative data analysis, descriptive statistics of mean, standard deviation, and frequency were used to analyze the data obtained from the pretest and posttest. In addition, the data elicited from the semi-structured interviews were analyzed by means of content analysis.

Findings

Effectiveness of IPA Modules on Students' English Speaking Ability

The pretest and posttest were used to determine if the students' speaking ability had improved after the nine-week implementation of the three modules of the IPA.

Table 2 presents the findings from the pretest and posttest. The students' total posttest mean score ($M = 3.27$; $SD = .32$) was higher than the pretest mean score ($M = 2.62$; $SD = .31$). Moreover, the students' scores showed improvement in all three IPA modes, namely the interpretive, interpersonal, and presentational modes. When the paired sample T-test was used to compare the two sets of scores, the results revealed that there was a statistically significant difference ($p < .05$) between the pretest and posttest scores.

Table 2

Summary of the results of the pretest and posttest

IPA mode	Pretest		Posttest		t	p
	M	*SD*	*M*	*SD*		
Interpretive	2.12	.38	3.03	.41	-10.42	.000
Interpersonal	2.80	.43	3.34	.38	-14.66	.000
Presentational	2.94	.43	3.45	.34	-12.88	.000
Total	2.62	.31	3.27	.32	-18.76	.000

*$p < .05$

When further considering the test task aspects in each of the IPA modes, it was discovered that the students' posttest mean scores increased across all IPA aspects in each of the IPA modes.

Table 3 displays the findings from the pretest and posttest when it came to the interpretive task. It was found that the students' posttest mean score ($M = 3.03$; $SD = .41$) was higher than the pretest mean score ($M = 2.12$; $SD = .38$), and the performance of the students had improved in all aspects. To begin with, the posttest mean score of the keyword recognition aspect (*M*

= 2.65; *SD* = .59) was higher than the pretest mean score (*M* = 2.15; *SD* = .49). As for the main idea detection aspect, the posttest mean score (*M* = 2.71; *SD* = .59) was higher than the pretest mean score (*M* = 1.82; *SD* = .74). Furthermore, the posttest mean score of the supporting detail detection aspect (*M* = 3.48; *SD* = .63) was higher than the pretest mean score (*M* = 2.35; *SD* = .81). Finally, the posttest mean score of the cultural perspectives aspect (*M* = 3.27; *SD* = .49) was higher than the pretest mean score (*M* = 2.17; *SD* = .65). Of the four aspects of the interpretive test, the students' improvement in the supporting detail detection aspect was largest, while that in the keyword recognition aspect improved the least.

Table 3

Result of the pretest and posttest in the interpretive task

IPA mode	Pretest		Posttest		*t*	*p*
	M	*SD*	*M*	*SD*		
Keyword recognition	2.15	.49	2.65	.59	-4.00	.000
Main idea detection	1.82	.74	2.71	.59	-5.77	.000
Supporting detail detection	2.35	.81	3.48	.63	-6.62	.000
Cultural perspectives	2.17	.65	3.27	.49	-7.60	.000
Total	2.12	.38	3.03	.41	-10.42	.000

* *p* < .05

Table 4 illustrates the results of the pretest and posttest of the interpersonal task. The overall posttest mean score (*M* = 3.34; *SD* = .38) was significantly higher than the pretest mean score (*M* = 2.80; *SD* = .43) The students' improvement could be seen in all five aspects of the test. In terms of the language function aspect, the students' posttest mean score (*M* = 3.30; *SD* = .57) was higher than the pretest mean score (*M* = 2.59; *SD* = .52). The students also performed better in the discourse type aspect as the posttest mean score (*M* = 3.06; *SD* = .35) was higher than the pretest mean score (*M* = 2.64; *SD* = .49). As regards the communication strategies aspect, the students' posttest mean score (*M* = 3.42; *SD* = .56) increased and was higher than the pretest mean score (*M* = 2.73; *SD* = .56). Moreover, there was an increase in the posttest mean score of the comprehensibility aspect (*M* = 3.74; *SD* = .40) compared to the pretest mean score (*M* = 3.24; *SD* = .63). The last aspect was language control which also showed the students' improvement as the posttest mean score (*M* = 3.18; *SD* = .50) was higher than the pretest mean score (*M* = 2.79; *SD* = .43).

Table 4

Results of the pretest and posttest in the interpersonal task

IPA mode	Pretest		Posttest		t	p
	M	SD	M	SD		
Language Function	2.59	.52	3.30	.57	-9.43	.000
Discourse Type	2.64	.49	3.06	.35	-6.12	.000
Communication Strategies	2.73	.56	3.42	.56	-9.68	.000
Comprehensibility	3.24	.63	3.74	.40	-5.41	.000
Language Control	2.79	.43	3.18	.50	-5.28	.000
Total	2.80	.43	3.34	.38	-14.66	.000

*p < .05

Table 5 presents the findings from the pretest and posttest in the presentational task which was the final IPA test task. As can be seen, the overall posttest mean score ($M = 3.45$; $SD = .34$) was significantly higher than the pretest mean score ($M = 2.94$; $SD = .43$). and the students' performance improved significantly in all aspects. Firstly, the posttest mean score of the language function aspect ($M = 3.58$; $SD = .50$) was higher than the pretest mean score ($M = 3.14$; $SD = .79$). As for the discourse type aspect, the students had a better performance as evidenced by the increase in the posttest mean score ($M = 3.06$; $SD = .24$) compared to the pretest mean score ($M = 2.52$; $SD = .51$). In terms of the impact aspect, the posttest mean score ($M = 3.27$; $SD = .66$) was higher than the pretest mean score ($M = 2.67$; $SD = .49$). Similarly, the posttest mean score of the comprehensibility aspect ($M = 3.90$; $SD = .26$) was higher than the pretest mean score ($M = 3.50$; $SD = .48$). Finally, the students' performance in the language control aspect improved as the posttest mean score ($M = 3.40$; $SD = .52$) was higher than the pretest mean score ($M = 2.90$; $SD = .39$).

Table 5

Results of the pretest and posttest in the presentational task

IPA mode	Pretest		Posttest		t	p
	M	SD	M	SD		
Language Function	3.14	.79	3.58	.50	-4.93	.000
Discourse Type	2.52	.51	3.06	.24	-7.13	.000
Impact	2.67	.49	3.27	.66	-7.01	.000
Comprehensibility	3.50	.48	3.90	.26	-5.12	.000
Language Control	2.90	.39	3.40	.52	-6.45	.000
Total	2.94	.43	3.45	.34	-12.88	.000

*p < .05

In addition to the results of the pretest and posttest, the summary of the unit tests yielded further evidence of the effectiveness of the three IPA modules implemented in the present study. As can be seen in Table 6, the students' mean scores of the IPA unit tests progressively improved over the nine weeks during which the IPA modules were implemented (Unit 1: M = 2.91; SD = 0.8; Unit 2: M = 3.07; SD = .25; Unit 3: M = 3.21; SD =.19).

Table 6

Summary of the results of the unit tests

Unit test	M	SD
Unit 1	2.91	.08
Unit 2	3.07	.25
Unit 3	3.21	.19

To shed further light on the findings of the quantitative tests, semi-structured interviews were conducted. Although these interviews did not indicate the students' actual improvement in their speaking ability, the qualitative findings yielded support to the effectiveness of the IPA modules as perceived by the students. For example, the following students both explained that their speaking ability had improved because they had more confidence to speak English, something they were too worried to do in the past:

> I thought my speaking skill greatly improved. Before I studied this course, I did not dare to speak English, and I did not realize my weaknesses. [Student #1]
> My speaking skill has improved a lot. I am more confident and I am not afraid to speak in front of an audience any more. I used to be afraid that someone would make fun of me because of my poor speaking ability, but after taking this course, I no longer had that worry. I am quite pleased with all of my presentations. [Student #2]

Therefore, it could be seen that the implementation of the three IPA modules enabled the students to improve their speaking ability as they felt that they were now able to speak with more confidence.

Students' Opinions Towards the Implementation of the IPA Modules

To elicit students' opinions towards the implementation of the IPA modules, an IPA opinion questionnaire was administered. As presented in Table 7, the students were satisfied with the implementation of the IPA

modules (*M* = 4.23; *SD* = .12). Furthermore, they agreed that the IPA modules helped them develop their English speaking ability (*M* = 4.32; *SD* = .25), and they had positive opinions of the 5 Cs standards (*M* = 4.27; *SD* = .05), the IPA instructional activities (*M* = 4.26; *SD* = .16), and the IPA feedback session (*M* = 4.70; *SD* = .17).

Table 7

The mean score of the students' opinions of the implementation of the IPA modules

Aspects	M	SD	Interpretation of agreement level
Overall satisfaction with the implementation of the IPA modules	4.23	.12	Strongly agree
Opinions on the development of English speaking ability through the IPA modules	4.32	.25	Strongly agree
Opinions on the 5 Cs standards	4.27	.05	Strongly agree
Opinions on the IPA instructional activities	4.26	.16	Strongly agree
Opinions on the IPA feedback session	4.70	.17	Strongly agree

Furthermore, qualitative findings elicited from the interviews also reflected the students' positive opinions towards the implementation of the IPA modules. It could be seen that the students realized how the IPA modules helped enhance their speaking ability. Some students commented that the cycle of IPA tasks was one of the factors that improved their speaking ability in that they had the opportunity to develop ideas on the speaking topic in the interpretive task and to speak and exchange ideas with their classmates in the interpersonal task. After they completed these two tasks, they were able to generate adequate ideas and could apply them in their final speech performance. As a result, they could perform well in their speech delivery in the presentational task, as can be seen in the following excerpts:

> Overall, I was happy to study the way this course taught me. I thought it was a good thing when the instructor started by giving an overview of the topic as if no one had background knowledge about the topic. Moreover, I thought that watching a video in the first task was an effective way to lead students into the lesson. At first, I did not have any idea about the topic and the video was very helpful. Moreover, when I had a chance to talk with my friends about the topic in task 2, I could improve not only my speaking skill but also my ability to expand ideas. Finally, I could apply the knowledge obtained from the first two tasks in my final presentation. I assumed that once I had sufficient information and proper language structures, preparing my final oral presentation would be simple. [Student #3]

> In the first task, I listened to the idea that I could use, and then I discuss it with my friends in the second task. These two tasks assisted me to come up with my final presentation in the third task. The learning steps were organized in a way that I enjoyed them. Although it took a long time to complete all tasks, I believed this teaching modules was better than starting with explanations and then assigning the presentation task just like in a traditional class. [Student #4]

The students also mentioned that the feedback session was very important to help them improve their speaking ability. Many students agreed that the feedback session in the IPA modules enabled them to have a better speaking performance, as illustrated in the excerpts below:

> The feedback was very useful for me. I totally agreed with the feedback I received from the instructor. I could realize my weaknesses and improve my speaking. [Student #2]
> I really liked the feedback session as it was very useful for me. When the instructor gave me comments, it reflected my strengths and weaknesses and I could develop my speaking skill. [Student #5]
>
> The feedback from the instructors and classmates was very useful. For me, my heart was fulfilled again because I had never received this kind of feedback before. It wasn't a criticism for me, but rather a constructive suggestion. As a result, I didn't feel threatened when I received the feedback. [Student #4]

However, although the feedback sessions were beneficial for students, many of them agreed that it would be better if they could have an extra feedback session with the teacher so that they could ask more questions on performance, as demonstrated below:

> If it was possible, I would like to have more time on feedback session. I had more questions to ask. [Student#1]

Some students also mentioned that more time is needed for speaking tasks in interpersonal and presentational tasks. Although each IPA task builds upon the others, different emphasis should be put on them, as shown in the excerpt below:

> I thought that I needed more preparation time for my final presentation. Although I had a lot of ideas from the previous two tasks, I needed more time to complete the speaking task. [Student #1]

Discussion

The findings of the present study confirmed that the implementation of the IPA modules had positive effects on the students' development of speaking ability. The posttest scores revealed that the students improved after the implementation of the three IPA modules. Furthermore, the findings from the questionnaire and semi-structured interviews also yielded support to the claim that the students had positive opinions towards the implementation of the IPA modules in the English speaking course.

Such findings were consistent with the findings of Davin et al. (2011) that students performed better when they were assigned IPA speaking tests. Similarly, Kurniawan et al. (2022) also confirmed that the implementation of IPA in speaking classrooms helped enhance students' speaking ability. One plausible explanation of the positive effects of the IPA modules may come from the specific characteristics of the IPA that enabled assessment tasks to be integrated into instruction in a seamless manner. In so doing, the students did not have to wait until the end of the lessons to take a test to find out if they had learned what they were supposed to master. Instead, with the IPA, assessment was included in the learning process and the students were encouraged to think about what they were and were not yet able to do and try to improve. As such, as Glisan et al. (2007) have claimed, all three components of teaching, learning, and assessment took place simultaneously. As a result, the students had a better chance to learn and improve themselves based on the assessment results during the learning process rather than having to wait until the end of the course to realize what they needed to further develop.

Besides this, class activities were another key component of the IPA that may have contributed to the improvement of the students' speaking ability. According to the qualitative data, the students realized that the IPA modules provided them with a new experience in learning English and they were satisfied with the cycle of tasks sequenced in the three IPA modules. It could therefore be assumed that the IPA's cyclical approach enabled the students to learn more effectively. To explain, the IPA's interpretive task provided the students with the overall related content regarding the theme of the unit, and their understanding of the topic gave them the information they needed for the next task. In the IPA's interpersonal task, the students relied on the information from the previous task when exchanging ideas with their classmates. Finally, the information obtained from the previous two IPA tasks led to the completion of the presentational task, which was the final IPA task. This was an opportunity that they had rarely had in a traditional classroom since the IPA provided authentic class activities which were applied through the cycle of the three IPA tasks. As a result, students were more likely to be

interested in doing such activities, which were based on real-world situations. As Adair-Hauck et al. (2015) mentioned, the IPA is different from traditional speaking tasks, which are assessed mainly on linguistic accuracy; in fact, the IPA provides learners with authentic tasks that positively enable students to learn through a real-world situation. These findings were also in line with a previous study undertaken by Zapata (2006) in that the IPA class activities were beneficial for students as they helped them prepare for and complete assessment tasks.

The results of the pretest and the posttest confirmed that the students' speaking ability significantly improved after the IPA implementation. This is because students became familiar with learning through the IPA task cycles, which made them aware of what they had to improve in the subsequent tasks. The findings elicited from the semi-structured interviews also confirmed this since students mentioned that the interrelatedness of each task helped them when delivering their speeches. The connections among the three IPA tasks allowed students to have adequate information and increased their understanding of what was required in their speaking performance. Likewise, the Center for Advanced Research on Language Acquisition (2019) points out that the interconnectedness of the three IPA modes maximizes students' capacity in learning. In contrast, in traditional classrooms, the learning task cycles may not occur in three consecutive stages within one unit of instruction as the instructors may use a different number of activities which may not build upon one another. Hence, some students in traditional classrooms have neither adequate knowledge about the topic nor the language functions that are required to produce a successful speaking performance, whereas the IPA modules can increase students' knowledge of the topic, give them opportunities to practice speaking English, and help them to generate ideas for their speech performance based on the first two tasks. This is consistent with Adair-Hauck et al.'s (2015) finding that the interpersonal task encouraged students to negotiate meaning and obtain new information from their interlocutor. Consequently, it helped students to effectively create a speech performance by using necessary information they perceived from the interpersonal task. However, it is worth noting that this finding was contrary to a previous study carried out by Glisan et al. (2007) which suggested that students could not perform well in the interpersonal task since it required spontaneous communication. It could be assumed that if students could not perform well in the interpersonal tasks, it might affect their speaking performance in the presentational task as well.

Another possible explanation of the positive findings is that students had a better understanding of how each IPA task built upon the previous one. Understanding how IPA modes were interrelated contributed to students'

greater performance in using their language ability for communication, particularly in speaking. This assumption is congruent with Glisan et al.'s study (2007), which showed that when students understood the relationships among the three IPA modes of communication, they were more aware of how they could utilize the language for a communicative purpose. Kissau and Adams (2016) lent support to the claim that when students were familiar with learning through the three IPA modes and understood the interconnectedness among the modes, they would be able to use the language for communicative purposes more effectively. However, Davin et al. (2011) reported a contrasting finding indicating that there was a seeming discontinuity between the interpretive task and the two subsequent tasks.

Additionally, the assessment method used in the IPA modules may be considered another factor contributing to the students' development of their speaking ability. This is because the assessment method in the IPA modules dynamically assessed students throughout their learning based on each IPA mode (Glisan et al., 2007), while the assessment method in a traditional speaking classroom is more likely to assess students when they complete the final speaking performance. This dynamic nature of the IPA in the present study led to improvement in speaking ability, primarily due to its cyclical procedure. That is, students were firstly assessed on the interpretive task, and then they received feedback on this task. Similarly, students were assessed in the same way after completing the interpersonal and presentational tasks, respectively, so there were more chances to learn and improve. Such a finding is consistent with Madison (2019), who contended that the principle of dynamic assessment embedded in the IPA helped promote students' ability to learn the language more effectively as the instructor continually interacted with each student throughout the assessment cycle. Furthermore, in the assessment procedure of the IPA, learners were clearly informed of their performance expectations through the IPA rubrics that had been specifically designed for the implementation of the IPA in the classrooms. In this study, the IPA rubrics were given to students prior to undertaking the assessment and the instructor had to ensure that the students clearly understood them (Adair-Hauck et al., 2015). Having understood that the IPA rubrics were an important part of the assessment method in this study, students tended to improve their speaking performance inasmuch as they received feedback based on the assessment criteria in the IPA rubrics. This interpretation also agrees with that of Adair-Hauck and Troyan (2013) that the use of IPA rubrics potentially enabled students to become familiarized with rubric language, as well as the types of performance and feedback sessions that would occur during the assessment procedure. Based on the qualitative findings, students mentioned that they were satisfied with how each task was assessed. When it came to the assessment in the IPA

modules, the students were able to realize their strengths and weaknesses almost immediately after being assessed rather than having to wait until the end of the unit or course like in a traditional classroom. In particular, Sternberg and Grigorenko (2002) posited that, in traditional classrooms, students rarely received feedback after the test and were aware of only a test score; therefore, giving immediate descriptive feedback should allow students to better improve their performance.

The findings of the present study also contribute to a clearer understanding of the importance of the constructive feedback used in the IPA feedback loop. The qualitative data confirmed that the feedback session was beneficial for the students' improvement. Feedback was provided at the end of each IPA mode, and such feedback enabled the students to reflect on their past performance as well as prepare themselves to perform better in the next IPA task. According to Adair-Hauck et al. (2015), IPA feedback should be considered an ongoing feedback loop between the teacher and students. The fact that feedback was ongoing was vital to the development of the speaking ability of the students in this study because students not only received the feedback in the form of a letter grade but they also received dynamic feedback in the form of oral or written feedback which was given at the end of each IPA task. This feedback was responsive to students' needs as the teacher guided the students' performance through the feedback loop, thus fostering their ability to deliver a more effective speech in the final presentational task,. Adair-Hauck et al. (2013) explained that the IPA feedback loop was vital in raising students' awareness of how to strengthen their language performance. It is likely that the co-constructed feedback in the IPA feedback loop provided students with a rich opportunity to identify areas for self-improvement. This is because, in each IPA task, students received descriptive feedback based on detailed criteria rather than judgmental feedback. According to Adair-Hauck and Troyan (2013), providing descriptive feedback in IPA was more meaningful to students; therefore, it enhanced students' performance.

In summary, it was shown that the implementation of the IPA modules had positive effects on students' speaking ability. It can be concluded that IPA characteristics contributed to students' improvement in their speaking ability. These characteristics include the way in which the IPA bridged instruction and assessment by offering a dynamic cluster of assessment tasks along with a systematic cycle of class activities and effective feedback sessions. Therefore, integrating IPA modules into speaking instruction may be considered an effective instructional and assessment method in the English speaking classes of Thai EFL undergraduate students.

Implications

Pedagogically, the findings of the present study offer support to the claim that IPA modules can be implemented effectively in a speaking classroom. However, there are some issues that instructors should take into careful consideration when implementing IPA modules in their classes.

Firstly, instructors should consider which IPA mode should be given the most emphasis. Although the quantitative findings showed that students' performance on every IPA mode improved, it is suggested that interpersonal and presentational tasks should be mostly focused on. This is because these two tasks concern students' speaking ability and require more time for students to complete each task. Based on the findings obtained from the interview, many students agreed that interpersonal and presentational tasks were difficult for them and more time is needed to produce the final oral presentational tasks. Although an interpretive task is also necessary, the instructor could pay less attention to it than the interpersonal and presentational tasks since an interpretive task serves mainly to provide students with necessary information about the theme of the topic. This approach is in accordance with what Madison's (2019) position that it is not necessary to place an equal emphasis on each IPA task as it should depend on the skills that are being focused on.

Another important issue that instructors should be aware of when implementing the IPA in a speaking classroom is time constraints. Since each IPA task may take longer than the instructors expect, it is suggested that the instructors administer no more than one IPA task per class period, which gives adequate time for instructors and students to complete the feedback sessions. Madison (2019) explains that if instructors administer the three IPA tasks in the same class period, students may miss out on the feedback loop. Also, they may be too overwhelmed to make the most out of the feedback they receive.

All in all, the implementation of the IPA modules in this study helped confirm that integrating assessment into instruction is helpful to maximize students' capacity for learning as learners could monitor their learning performance. As stated by Viengsang and Wasanasomsithi (2022), integrating assessments into instruction has been shown to be a beneficial aid for language learning since it gives students the capability to identify their own learning ability; as a result, it brings about the development of students' performance.

Limitations and Recommendations

Some limitations of the present study need to be acknowledged. First of all, this study was conducted for only one semester by using three IPA modules to investigate the improvement of students' English speaking ability. Although the findings showed an improvement in students' speaking ability, more time might be required to examine the full development of students' speaking skills, which take time to develop. Moreover, the time to administer the feedback session was rather limited, so instructors need to carefully devise a teaching plan to make sure that there is sufficient time for feedback provision so that students can better realize their strengths and weaknesses to maximize their speaking ability development.

As for future research, since IPA feedback seemed to play a vital role in students' speaking ability, further studies on students' reactions to IPA feedback are recommended. This should shed light on how students apply feedback to correct, practice, and improve their speaking. Research can also be undertaken to explore different techniques of feedback provision to determine which is more effective for instructors to utilize. Moreover, longitudinal research should be carried out to explore if the IPA results in positive long-term effects on students' speaking improvement.

About the Authors

Benjawan Plengkham: A Ph.D. candidate in English as an International Language Program, Graduate School, Chulalongkorn University, Thailand. Her research interests include English language teaching, speaking skills, and language assessment and evaluation.

Punchalee Wasanasomsithi: An associate professor at Chulalongkorn University Language Institute and Director of the English as an International Language Program, Graduate School, Chulalongkorn University, Thailand. She holds a Ph.D. in Language Education from Indiana University, Indiana, USA. Her research interests include writing instruction, classroom-based assessment, assessment literacy, and learning-oriented assessment.

References

Adair-Hauck, B., Glisan, E. W., Koda, K., Swender, E. B., & Sandrock, P. (2006). The Integrated Performance Assessment (IPA): Connecting assessment to instruction and learning. *Foreign Language Annals*, *39*(3), 359–382. https://doi.org/10.1111/j.1944-9720.2006.tb02894.x

Adair-Hauck, B., & Troyan, F. J. (2013). A descriptive and co-constructive approach to integrated performance assessment feedback. *Foreign Language Annals, 46*(1), 23–44. https://doi.org/10.1111/flan.12017

Adair-Hauck, B., Glisan, E. W., & Troyan, F. J. (2015). *Implementing integrated performance assessment.* American Council on the Teaching of Foreign Languages.

American Council on the Teaching Foreign Language (ACTFL). (2011). *21st century skills map.* Retrieved December 12, 2019, from https://www.actfl.org/sites/default/files/resources/21st%20Century%20Skills%20Map-World%20Languages.pdf

American Council on Teaching Foreign Languages (ACTFL). (2012). *ACTFL Proficiency Guidelines 2012.* Retrieved October 3, 2019, from https://www.actfl.org/publications/guidelines-and-manuals/actfl-proficiency-guidelines-2012

Atagi, R. (2011). *Secondary teacher policy research: Secondary teachers in Thailand.* UNESCO.

Brown, H. D. (2004). *Language assessment: Principles and classroom practices.* Pearson Education.

Center for Advanced Research on Language Acquisition (2019). *Integrated Performance Assessment (IPA).* Retrieved December 10, 2019, from https://carla.umn.edu/assessment/vac/CreateUnit/p_2.html

Davin, K., Troyan, F. J., Donato, R., & Hellman, A. (2011). Research on the Integrated Performance Assessment in an early foreign language learning program. *Foreign Language Annals, 44*(4), 605–625. doi: 10.1111/j.1944-9720.2011.01153.x

Frank, J. (2012). The roles of assessment in language teaching. *English Teaching Forum, 50*(3), 32.

Goh, C. C. M., & Burns, A. (2012). *Teaching speaking: A holistic approach.* Cambridge University Press.

Glisan, E. W., Uribe, D., & Adair-Hauck, B. (2007). Research on Integrated Performance Assessment at the post-secondary level: Student performance across the modes of communication. *Canadian Modern Language Review, 64*(1), 39–67. doi: 10.3138/cmlr.64.1.039

Kissau, S., & Adams, M. J. (2016). Instructional decision making and IPAs: Assessing the modes of communication. *Foreign Language Annals, 49*(1), 105–123. https://doi.org/10.1111/flan.12184

Kurniawan, I., Sabaruddin, S., & Lubis, A. A. (2022). The implementation of the integrated performance assessment (IPA) model to improve English students' speaking skills. *JOALL (Journal of Applied Linguistics and Literature), 7*(2), 432–443. https://doi.org/10.33369/joall.v7i2.20860

Madison, S. (2019). Integrated Performance Assessments: A review of the literature and steps to move forward. *Spanish and Portuguese Review*, 5, 99-111.

Madison-Harris, R.D., Muoneke, A., & Times, C. (2012). *Using Formative Assessment to Improve Student Achievement in the Core Content Areas.* Briefing Paper.

Mertler, C. A. (2017). *Classroom assessment: A practical guide for educators.* Routledge.

Mueller, J. (2005). The authentic assessment toolbox: Enhancing student learning through online faculty development. *Journal of Online Learning and Teaching, 1*(1).

Nakhalah, D.A. (2016). Problems and difficulties of speaking that encounter English language students at Al Quds Open University. *International Journal of Humanities and Social Science Invention, 5*(12), 96-101.

National Standards in Foreign Language Education Project (NSFLEP). (1996). *Standards for foreign language learning in the 21st century (SFLL).* Allen Press.

National Standards in Foreign Language Education Project (NSFLEP). (1999). *Standards for foreign language learning in the 21st century (SFLL).* Allen Press.

National Standards in Foreign Language Education Project (NSFLEP). (2006). *Standards for foreign language learning in the 21st century (SFLL).* Allen Press.

Sandrock , P. (2008). Integrated Performance Assessment. Retrieved July 5, 2019, from http://depts.washington.edu/mellwa/Events/20081105/sandrock_ipa_handout.pdf

Sternberg, R.J., & Grigorenko, E.L. (2002). *Dynamic testing: The nature and measurement of learning potential.* Cambridge University Press.

Tunstall, P., & Gipps, C. (1996). Teacher feedback to young children in formative assessment: A typology. *British Educational Research Journal, 22*, 389–404.

Viengsang, R., & Wasanasomsithi, P. (2022). Effects of a learning-oriented reading assessment model on Thai undergraduate students' reading ability. *LEARN Journal: Language Education and Acquisition Research Network, 15*, 709–747.

Zapata, G. C. (2016). University students' perceptions of Integrated Performance Assessment and the connection between classroom learning and assessment. *Foreign Language Annals, 49*(1), 93–104. doi: 10.1111/flan.12176