



The Development of a Self-Efficacy Measurement Tool and an Assessment Framework for Academic English-Speaking Ability among EFL Pre-Service Teachers

Ranya Jitardhan^{1*}, Phitcharie Koanamsai²

^{1,2}Faculty of Education, Valaya Alongkorn Rajabhat University under the Royal Patronage, Thailand

✉: * ranya.jit@vru.ac.th

(*Author Corresponding Email)

Received: 08 May 2026; Revised: 22 May 2026; Accepted: 21 June 2026

© The Author(s) 2026

Abstract

Academic English-speaking competence and strong self-efficacy are essential for EFL pre-service teachers, yet context-specific instruments that assess both performance and self-beliefs remain limited. Therefore, this study aims to (1) develop a performance-based Academic English-Speaking Assessment (AESA) and (2) develop and preliminarily validate the Self-Efficacy in Academic English-Speaking Scale (SE-AESS) for EFL pre-service teachers. The AESA was adapted from the authentic assessment framework proposed by O'Malley and Pierce (1996) and the analytic speaking rubric developed by Syahidah and Umasugi (2021). The SE-AESS was constructed based on Bandura's Self-Efficacy Theory, encompassing both perceived academic English-speaking self-efficacy and its four sources: mastery experience, vicarious experience, verbal persuasion, and somatic and emotional states. Content validity of both instruments was evaluated by three experts in educational psychology, English language teaching, and applied linguistics. The SE-AESS demonstrated excellent content validity, with item-objective congruence (IOC) values ranging from 0.80 to 1.00, while the AESA showed IOC values between 0.88 and 0.93. Reliability analysis revealed good internal consistency for the academic English-speaking self-efficacy dimension ($\alpha = 0.851$) and moderate reliability for the sources of self-efficacy dimension ($\alpha = 0.554$). Both instruments achieved the minimum psychometric requirements for early-stage instrument development, even though the latter reflected the multidimensional and experiential nature of the construct. The AESA and SE-AESS together provide a comprehensive framework for assessing both observable academic English-speaking performance and learners' self-efficacy beliefs. In order to improve academic English-speaking competency among EFL pre-service teachers in ASEAN and multicultural contexts, these tools may be crucial for assessment and instructional design.

Keywords: Self-efficacy, Academic English speaking, Instrument development, EFL pre-service teachers

1. Introduction

In the 21st century, academic English-speaking ability has become an essential competence for EFL pre-service teachers since they are increasingly required academic presentation to present ideas including participating in group discussion in classrooms and communicating pedagogical knowledge using English. Effective academic speaking is therefore not only a linguistic skill but also a critical professional competence that supports teaching practicum, academic success, and the development of teacher identity. However,

academic English speaking is not only learning language itself, but also profoundly related to the learner's beliefs in their ability or it's commonly referred to as self-efficacy. (Bandura, 1997)

Self-efficacy is a psychological factor that affects learners' learning outcomes and an individual's different behaviors. Research in educational psychology has consistently shown that learners' performance is strongly influenced by their beliefs about their own capabilities when they face challenging tasks. Self-efficacy also affects individuals' motivation, persistence, effort, and emotional responses when confronting demanding tasks. Learners with high self-efficacy are more likely to engage actively, persist through difficulties, and interpret challenges as opportunities for growth, whereas those with low self-efficacy tend to avoid tasks they perceive as exceeding their abilities. Additionally, research in various fields reflected the importance of self-efficacy towards effort, facing obstacles, and preparation for undertaking challenging mission. Jaruwat Prada et al. (2010) found that increasing of self-efficacy in terms of taking care of family in diverse dimensions helping caregivers feel confident and able to cope with new problems or difficulties about challenges related to the care of children with developmental and intellectual disabilities. These beliefs make a person have an effort and encouragement to overcome challenging circumstances during upbringing. Similarly, Siriporn Chudjeujeen et al. (2017) investigated the program to promote self-efficacy in building life skills related to sexual health behaviors in early adolescents. The result showed that self-efficacy helps adolescents understand their sexual behavior, life skills, and sexual health behavior which has been a significant increase. It is also emphasized that student teachers' perceived self-efficacy influences their level of participation in academic discourse, classroom communication, and the development of professional identity (Tschannen-Moran & Hoy, 2007) (Schunk et al., 2020)

In the field of English as a Foreign Language (EFL), numerous studies have demonstrated that self-efficacy is positively associated with language achievement and learning experiences. Learners with higher self-efficacy generally report greater confidence, stronger engagement, and better performance in English-speaking tasks, while learners with lower self-efficacy often experience anxiety, avoidance, and negative learning outcomes. Empirical studies have also indicated that self-efficacy can be a stronger predictor of language performance than affective variables such as anxiety. When considered in the context of learning the English language, several studies have reported consistent results that self-efficacy is a key variable that influences learners' language achievement and English language experience. Rahimi (2007) found that learners with low self-efficacy tend to have lower English language achievement and negative experiences in language learning. While Tilfarlioglu and Cinkara (2009) reported that highly proficient English learners tend to have higher self-efficacy scores, Mills, Pajares, and Herron (2006) pointed out that self-efficacy is a better predictor of language achievement than language anxiety.

Although numerous studies confirm the importance of self-efficacy in English language learning, the development of past measurement tools has been challenging mostly focusing on study of general English speaking or everyday English communication rather than academic English speaking. Academic speaking involves discipline-specific vocabulary, complex grammatical structures, coherent argumentation, and precise expression of ideas, which differ substantially from conversational English. Moreover, many studies rely on self-report questionnaires or academic achievement scores as proxies for speaking ability, with limited use of performance-based assessments evaluated by trained raters using analytic rubrics.

Another limitation of prior research is the lack of integration between the assessment of speaking performance and the psychological factors underlying it. While self-efficacy theory emphasizes the importance of mastery experience, vicarious experience, verbal persuasion, and emotional states, few instruments measure these sources of self-efficacy specifically within

academic English-speaking contexts. This gap is particularly evident in EFL teacher education, where pre-service teachers must develop both academic speaking competence and the confidence to use English professionally.

To address these gaps, the present study aims to develop two complementary instruments: (1) the Academic English-Speaking Assessment (AESA), a performance-based assessment using an analytic rubric to evaluate academic speaking ability, and (2) the Self-Efficacy in Academic English-Speaking Scale (SE-AESS), designed to measure learners' perceived self-efficacy and its sources within academic speaking contexts. Together, these instruments provide an integrated framework for assessing both academic English-speaking performance and the self-efficacy beliefs that influence it among EFL pre-service teachers, and can serve as essential tools for assessment and for the design of learning activities aimed at effectively promoting academic English-speaking development among pre-service teachers in ASEAN and multicultural contexts.

2. Literature Review

1. Self-Efficacy

Albert Bandura (1997), a well-known educational psychologist, proposed the perspective of an individual's expectation of their own ability and it was defined as "self-efficacy" which refers to an individual's beliefs in their ability to successfully complete a particular task and to execute a plan as expected. These beliefs influence whether individuals initiate tasks, the level of effort they expend, their persistence in the face of difficulty, and their emotional reactions during task performance. Rather than reflecting actual ability alone, self-efficacy represents individuals' judgments of what they can do under particular conditions.

Bandura conceptualized human functioning as the product of reciprocal interactions among personal factors, behavior, and environmental influences. Within this framework, self-efficacy plays a pivotal role in self-regulation and adaptive functioning, particularly in complex and demanding tasks such as academic communication in a second language. Learners with high self-efficacy are more likely to approach difficult tasks with confidence, view challenges as manageable, and recover quickly from setbacks, whereas those with low self-efficacy are prone to avoidance, anxiety, and self-doubt (Bandura, 1994).

In present studies, several scholars have provided similar definitions, both in other countries such as Dagmer El-Hmoudova (2015) and James E. Maddux (2018), and in Thailand e.g., Siriporn Chudjeujeen et al. (2017) and Pasit Sirithes et al., (2019). summarized that "Self-efficacy" refers to an assessment of one's own ability to perform a specific task within a given individual situation which is based on self-confidence derived from considering and evaluating one's own performance capabilities and skills for managing and coping with various challenging situations. This arises from social learning based on past experiences. Beliefs about one's self-efficacy influence decisions about whether or not to exhibit certain behaviors to achieve goals in various situations.

According to Bandura (1997), self-efficacy beliefs are shaped by four primary sources. Mastery experience is the most influential source, as successful task performance provides direct evidence of capability. Vicarious experience involves observing others with similar characteristics succeed, which can strengthen beliefs in one's own ability. Verbal persuasion refers to encouragement and feedback from significant others, which can enhance confidence when realistic and supportive. Somatic and emotional states involve individuals' interpretations of physiological arousal and emotional reactions; anxiety and stress may undermine self-efficacy, while positive emotional states can enhance it.

These sources operate dynamically and are influenced by contextual and personal factors. In academic settings, particularly in second language learning, learners' self-efficacy is shaped

by prior learning experiences, instructional practices, feedback from teachers and peers, and emotional responses to performance situations. Because self-efficacy is task- and context-specific, its measurement must align closely with the nature of the target task.

2. Academic English Speaking and Self-Efficacy

Academic English speaking refers to the use of spoken English in formal educational and professional contexts, such as presenting research, engaging in academic discussions, and responding to content-based questions. Compared with everyday conversational English, academic speaking requires greater linguistic accuracy, advanced vocabulary, coherent organization of ideas, and clarity of expression. These demands make academic speaking particularly challenging for EFL learners and increase the importance of psychological factors such as self-efficacy.

Self-efficacy plays a significant role in English language learning, and there have been numerous studies related to the relationship between self-efficacy and English language learning. For example, Rahimi (2007) investigated the relationship between self-efficacy and students' academic performance. The results showed a significant relationship between self-efficacy and success in learning English. Students with low self-efficacy performed poorly in English language learning because they had negative attitudes and bad experiences in learning English. In addition, the results of the study by Tilfarlioglu and Cinkara (2009) showed that students with different levels of self-efficacy exhibited different levels of self-efficacy. Upper-intermediate students had higher perceived self-efficacy than pre-intermediate students.

Previous research has consistently shown a positive relationship between self-efficacy and English language performance. Learners with higher self-efficacy tend to demonstrate greater persistence, employ more effective learning strategies, and perform better in speaking tasks. In contrast, low self-efficacy is often associated with avoidance of speaking opportunities, heightened anxiety, and weaker performance. However, most existing studies examine self-efficacy in general English learning or public speaking contexts, with limited focus on academic English-speaking tasks.

Furthermore, many instruments used in prior research assess self-efficacy without examining its underlying sources or without linking self-efficacy beliefs to actual performance-based measures. This limitation reduces the explanatory power of self-efficacy research and constrains its application to instructional design. In response, the present study adopts Bandura's theoretical framework to develop a context-specific self-efficacy scale for academic English speaking and integrates it with a performance-based assessment. This approach enables a more comprehensive understanding of both learners' academic speaking competence and the psychological factors that support or hinder its development.

3. Related Research

To develop valid and reliable instruments for assessing academic English-speaking ability (AESA) and self-efficacy in academic English speaking (SE-AESS), it was essential to examine previous studies related to language assessment, self-efficacy theory, and instrument development in both international and Asian EFL contexts. A review of the literature indicates that most prior research has focused either on general English proficiency or on learners' perceived self-efficacy, with limited integration of performance-based speaking assessment and psychological constructs within academic contexts.

Several studies have demonstrated a positive relationship between self-efficacy and language achievement. Mills, Pajares, and Herron (2006) found that self-efficacy was a stronger predictor of foreign language performance than anxiety, while Rahemi (2007) reported that learners with low self-efficacy tended to experience lower English achievement and more negative learning experiences. Similar findings were reported by Tilfarlioglu and Cinkara

(2009), who observed that learners with higher English proficiency levels also exhibited higher self-efficacy beliefs. These studies, however, primarily measured achievement through course grades or standardized tests and did not assess speaking performance using expert-rated, performance-based instruments.

Research conducted in Asian EFL contexts has yielded comparable results. Kitikanan and Sasimonton (2017) reported a significant positive relationship between English self-efficacy and English achievement among Thai learners, while Siritaratn (2013) identified attitudes, prior learning experiences, and instructional practices as important factors shaping English self-efficacy among low-proficiency Thai graduate students. Although these studies highlight the importance of self-efficacy in English learning, they generally focus on general English use rather than academic English speaking and do not examine the specific sources of self-efficacy proposed by Bandura.

With regard to instrument development, several studies have constructed English self-efficacy scales grounded in Bandura's framework. Wang et al. (2013) developed the Questionnaire of English Self-Efficacy (QESE) to assess learners' perceived self-efficacy across four language skills and reported strong psychometric properties. Subsequent studies adapted the QESE for different contexts and proficiency levels, such as Nguyen and Habók (2022) for CEFR A2 learners and Kitikanan and Sasimonton (2017) for Thai EFL learners. While these instruments provide valuable insights into learners' perceived abilities, they rely primarily on self-report measures and do not incorporate performance-based assessments of speaking ability.

Only a limited number of studies have examined the relationship between speaking self-efficacy and actual speaking performance. Suharja (2022) reported a positive association between speaking self-efficacy and speaking performance among EFL university students, and Chang, Zhou, and Zhang (2024) explored the dynamic relationship between speaking self-efficacy and enjoyment during oral tasks. However, these studies were conducted in general speaking or public speaking contexts and did not focus specifically on academic English speaking. Moreover, the sources of self-efficacy were not systematically examined in most of these studies.

In teacher education contexts, research by Tschannen-Moran and Hoy (2007) emphasized the role of self-efficacy in teachers' professional communication and instructional engagement. More recently, Lin, Mohamad, and Mahmud (2025) examined the relationship between teachers' classroom English proficiency and teacher self-efficacy using self-reported measures. These studies suggest that self-efficacy plays an important role in professional communication; however, they do not provide performance-based assessments of academic English speaking or examine the sources of self-efficacy within this specific context (Tschannen-Moran et al., 2001).

Overall, the review of related research reveals three major gaps. First, there is a lack of performance-based instruments designed to assess academic English-speaking ability using expert-rated analytic rubrics. Second, existing self-efficacy instruments largely focus on general English use and do not capture the specific demands of academic English speaking. Third, few studies integrate the assessment of academic speaking performance with the measurement of self-efficacy and its underlying sources. These gaps directly informed the development of the AESA and SE-AESS in the present study, which aim to provide a comprehensive and theoretically grounded framework for assessing academic English-speaking competence and self-efficacy among EFL pre-service teachers.

As summarized in **Table 1**, most previous studies examined self-efficacy in general English contexts or relied on self-report measures without integrating performance-based assessments of academic English speaking. These limitations highlight the need for instruments that simultaneously assess academic speaking performance and context-specific self-efficacy.

Table 1: Summary of Related Research on Self-Efficacy and English-Speaking Ability

Researchers	Participants & Context	Instruments	Key Findings	Gaps Relevant to the Present Study
Mills, Pajares, & Herron (2006)	University EFL learners (French)	Language proficiency tests; Self-efficacy questionnaire	Self-efficacy was a stronger predictor of language performance than anxiety.	Focused on reading/listening; no speaking or academic speaking assessment; sources of self-efficacy not examined.
Rahemi (2007)	Iranian high-school EFL students	English achievement scores; self-efficacy questionnaire	Low self-efficacy associated with low English achievement and negative learning experiences.	Used achievement scores rather than performance-based speaking assessment; general English context.
Tschannen-Moran & Hoy (2007)	Pre- and in-service teachers	Teacher Sense of Efficacy Scale (TSES)	Teacher self-efficacy linked to professional engagement and communication.	Did not examine English-speaking performance or academic English contexts.
Tilfarlioglu & Cinkara (2009)	Turkish EFL learners (varied proficiency)	GPA; EFL self-efficacy questionnaire	Higher proficiency learners reported higher self-efficacy.	No expert-rated speaking assessment; sources of self-efficacy not investigated.
Siritaratn (2013)	Thai EFL graduate students (low proficiency)	Placement tests; self-efficacy questionnaire; interviews	Attitudes and prior experiences influenced English self-efficacy.	Focused on general English; lacked academic speaking performance measures.
Kitikanan & Sasimonton (2017)	Thai EFL English majors	English achievement scores; QESE	Positive relationship between English self-efficacy and achievement.	Speaking assessed indirectly via achievement scores; sources of self-efficacy not examined.
Wang et al. (2013); Nguyen & Habók (2022)	University EFL learners (Asia)	Questionnaire of English Self-Efficacy (QESE)	QESE showed good reliability and construct validity across four language skills.	Relied on self-report only; no linkage to actual speaking performance or academic speaking context.

Suharja (2022)	Indonesian EFL university students	Speaking performance test; self- efficacy questionnaire	Speaking self- efficacy positively correlated with speaking performance.	Conducted in general speaking context; academic speaking and sources of self-efficacy not addressed.
Chang, Zhou, & Zhang (2024)	Chinese EFL university students	Speaking tasks; speaking self- efficacy scale	Dynamic relationship between self- efficacy and task enjoyment.	Focused on public/general speaking; sources of self-efficacy not examined.
Lin, Mohamad, & Mahmud (2025)	Polytechnic EFL teachers	CEPS; TSES (self-report)	Classroom English proficiency related to teacher self- efficacy.	English proficiency measured by self- report; no expert- rated academic speaking assessment.

3. Research Methodology

This study was conducted with EFL pre-service teachers in a multicultural academic context. The participants consisted of 40 students enrolled in a Master of Education program in English Language Teaching who possess the characteristics that align with the objectives of developing measurement tools in this study. As EFL learners, they frequently use academic English for communication in linguistically and culturally diverse classroom settings and are preparing to become future English teachers.

1. Development and Validation of Research Instruments

Two instruments were systematically developed and validated in this study: (1) the Academic English-Speaking Assessment (AESAs), and (2) the Self-Efficacy in Academic English-Speaking Scale (SE-AESS). The development process involved multiple steps, including: literature review, item construction, content validity evaluation, pilot testing, reliability analysis, and expert-based speaking assessment (inter-rater scoring). The procedures are detailed below.

1.1 Literature Review

A comprehensive review of research and theoretical literature on English-speaking assessment and self-efficacy was conducted. Key frameworks included: O'Malley and Pierce's (1996) model for authentic speaking assessment, Syahidah and Umasugi's (2021) analytic speaking rubric, and Bandura's (1997) Self-Efficacy Theory, covering four major sources: Mastery Experience, Vicarious Experience, Verbal Persuasion, and Somatic and Emotional States. Previous studies on self-efficacy in English language learning and teacher education were also reviewed to identify gaps and determine the essential constructs for both instruments (Bandura, 1986; Bandura, 2001).

1.2 Item Generation and Indicator Development

Items for both instruments were generated based on theoretical and empirical foundations. The process consisted of two parts:

(1) Development of the Academic English-Speaking Assessment (AESA)

The AESA was designed to assess pre-service teachers' academic speaking competence through tasks that reflect real classroom situations such as, academic presentations, academic topic discussions, and responding to discipline-related questions an analytic rubric was developed for performance-based assessment, consisting of six components: Grammar, Vocabulary, Comprehensibility, Fluency, Pronunciation, and Accuracy, each rated on a 5-point scale (1–5).

The AESA was created by adapting the assessment frameworks of O'Malley and Pierce (1996) and Syahidah and Umasugi (2021) due to their strong relevance in EFL performance assessment. O'Malley and Pierce (1996) emphasize authentic communicative performance, while Syahidah and Umasugi (2021) provide updated descriptors for academic discourse; organization, coherence, complexity, and topic development that aligned with contemporary EFL academic speaking tasks.

An inter-rater reliability procedure was implemented to ensure consistency in scoring, audio recordings of all speaking tasks were evaluated by two expert raters experienced in language assessment. Before scoring, the researcher conducted a training session to familiarize the raters with the AESA rubric, clarify construct definitions, and review sample recordings for calibration.

After training, both raters independently scored all recordings. Inter-rater reliability was calculated using the Intraclass Correlation Coefficient (ICC), which is suitable for continuous scale ratings. An ICC value of 0.75 or above was considered indicative of good reliability. Large discrepancies, if present, were resolved through discussion to reach a consensus score. This ensured that the assessment process was rigorous, aligned, and methodologically reliable.

(2) Development of the Self-Efficacy in Academic English-Speaking Scale (SE-AESS)

The development of the Self-Efficacy in Academic English-Speaking Scale (SE-AESS) resulted in a 39-item instrument comprising two dimensions: self-efficacy in academic English speaking (19 items) and the sources of self-efficacy (20 items). All items were systematically constructed based on Bandura's theoretical framework and supported by evidence from relevant empirical studies. A five-point Likert scale was employed for all items to capture varying levels of agreement and perceived confidence among participants.

1.3 Content Validity Evaluation

Three experts in educational psychology, English Language Teaching, and Applied linguistics evaluated the correspondence between each item, the theoretical constructs, and the research objectives using the Index of Item–Objective Congruence (IOC). The results indicated that the SE-AESS demonstrated excellent content validity, with IOC values ranging from 0.80 to 1.00, while the AESA also showed high content validity, with IOC values between 0.88 and 0.93. All items for which the experts' provided recommendations were subsequently revised and refined to ensure clarity, alignment, and accuracy before the instruments were administered in the pilot phase.

1.4 Pilot Testing

Pilot testing was conducted for both instruments to evaluate the clarity of questionnaire items and speaking task instructions, the completeness and usability of the analytic rubric, and the appropriateness of the response time required for each task. Based on the pilot results, several questionnaire items were refined to improve clarity, and minor wording adjustments

were made to the rubric descriptors to enhance precision and usability before proceeding to the full administration phase.

1.5 Reliability Analysis

The reliability of the SE-AESS was examined using Cronbach's alpha to assess internal consistency. The results showed that the self-efficacy dimension demonstrated good reliability ($\alpha = 0.851$), while the sources of self-efficacy dimension yielded a moderate reliability level ($\alpha = 0.554$). Although the latter value indicates room for improvement, it remains acceptable for an instrument in its early developmental stage, with further refinement of item clarity and internal coherence recommended before large-scale implementation. Overall, both instruments met and exceeded the minimum reliability thresholds suggested for preliminary research (Nunnally, 1978, as cited in Tavakol & Dennick, 2011; Hair et al., 2019), supporting their suitability for use in the present study.

2. Data collection and scoring procedures

The data collection process and Scoring Procedures in this study was carried out in two main phases to obtain comprehensive information on participants' academic English-speaking ability and their perceived self-efficacy.

2.1 The first phase involved the performance-based assessment of Academic English-Speaking ability

Data were collected during the participants' regular course schedule to ensure ecological validity and naturalistic performance conditions, following recommendations that assessments should be administered in authentic learning environments to enhance validity (Bachman & Palmer, 2010). Before data collection began, the researcher clarified the study purpose, procedures, rights of participants, and confidentiality measures. Informed consent was obtained in accordance with ethical guidelines for educational research (Creswell, 2018). Each participant received a coded identifier to ensure anonymity during scoring and analysis.

The first component of data collection involved administering the Academic English-Speaking Assessment (AESA). Each participant completed an academic speaking task consisting of a short presentation and responses to academic questions designed to elicit discourse-level academic language. Performance-based speaking tasks are widely recognized as effective means of assessing communicative ability because they generate authentic, observable language use (O'Malley & Pierce, 1996; Luoma, 2004). The researcher provided standardized instructions to all participants to reduce administration bias and enhance score comparability, as recommended in language assessment best practices (Fulcher, 2010).

All speaking performances were audio-recorded using a digital device with high audio quality to allow repeated listening and precise scoring, which is essential for reliability (McNamara, 1996). The recordings were labeled only with participant codes. After collecting, the recordings were submitted to two expert raters with extensive experience in EFL oral proficiency assessment. Prior to scoring, the two raters participated in a structured rater training session. Rater training is essential for ensuring shared understanding of rating criteria, reducing variability, and improving inter-rater reliability (Weigle, 2002; Eckes, 2015). The training involved reviewing the analytic rubric, discussing performance descriptors, and practicing scoring with benchmark samples. Such calibration procedures enhance rating consistency and reduce subjective interpretation (Brown, 2012).

The raters then independently evaluated all recordings using the AESA analytic rubric. Independent scoring without discussion helps avoid rater contamination and ensures that reliability metrics reflect true inter-rater agreement (Eckes, 2015). After initial scoring, the two

sets of scores were compared. For performances where scores differed substantially, raters participated in a moderation session to discuss discrepancies and reach a consensus score, following established procedures for maintaining fairness and consistency in performance assessments (McNamara, 1996; Brown & Abeywickrama, 2019). Inter-rater reliability was calculated using the Intraclass Correlation Coefficient (ICC), a solid statistical measure recommended for continuous rating-scale data (Koo & Li, 2016).

2.2 The second phase involved the administration of the Self-Efficacy in Academic English-Speaking Scale (SE-AESS)

The second phase of data collection involved the administration of the Self-Efficacy in Academic English-Speaking Scale (SE-AESS). After completing the speaking task, participants completed the SE-AESS questionnaire during class time. The 15–20-minute administration period aligns with typical recommendations for self-report questionnaires in educational research (Dörnyei & Taguchi, 2010). The researcher remained available to clarify instructions but refrained from influencing responses, adhering to best practices for minimizing researcher bias (Creswell, 2018).

Importantly, the speaking assessment was administered before the self-efficacy questionnaire to avoid priming effects or altered performance based on self-reflection. This sequencing reflects Bandura's (1997) recommendation that self-efficacy assessments should not artificially inflate or depress actual task performance by preceding the performance task.

All completed questionnaires were collected immediately, checked for completeness, and coded for analysis. The AESA performance scores and SE-AESS responses were then entered into a secure database using participant codes. This systematic approach supports the reliability, validity, and methodological rigor standardized administration, rater training, independent scoring, and validated instruments which supports the reliability, validity, and overall methodological rigor of the data collection process (Bachman & Palmer, 2010; McNamara, 1996).

4. Results

The analysis of the quality of the two research instruments indicated that both tools possessed satisfactory psychometric properties and were appropriate for use in achieving the research objectives. For the AESA (Academic English-Speaking Assessment), the content validity was found to be high, as experts had examined the alignment between the speaking tasks, scoring criteria, and research objectives prior to implementation. The analytic rubric used in AESA assesses five-level analytic rubric comprising six components; Grammar, Vocabulary, Comprehensibility, Fluency, Pronunciation, and Accuracy and was adapted from O'Malley and Pierce (1996) and Syahidah and Umasugi (2021). This adaptation resulted in a clearer and more systematic set of criteria for evaluating learners' academic English-speaking competence. Furthermore, the scores assigned by two independent raters showed high consistency, reflecting a satisfactory level of inter-rater reliability, which confirms that the AESA can reliably and consistently measure learners' speaking competence across evaluators.

For the SE-AESS (Self-Efficacy in Academic English-Speaking Scale), which measures learners' confidence in their academic English-speaking ability, content validity was evaluated by three experts. The item-objective congruence (IOC) indices ranged from 0.80 to 1.00, indicating excellent content validity. After revising the items based on expert feedback, the finalized SE-AESS was pilot-tested and subjected to reliability analysis. The internal consistency of the self-efficacy dimension yielded a Cronbach's alpha coefficient of 0.851, indicating good internal reliability. In contrast, the sources of self-efficacy dimension resulted in a Cronbach's alpha of 0.554, reflecting a moderate level of internal consistency. This

suggests certain items particularly those related to wording clarity or conceptual alignment. It should be refined in future iterations to strengthen the internal coherence of this dimension.

Although the reliability for the sources dimension was moderate, the overall reliability of SE-AESS remained within the acceptable range for preliminary research. This aligns with the recommendations of Nunnally (1978), as cited in Tavakol and Dennick (2011) and Hair et al. (2019), which state that newly developed instruments may be considered acceptable for initial research use when Cronbach's alpha exceeds 0.50.

In summary, both AESA and SE-AESS demonstrated satisfactory content validity and reliability. The instruments are suitable for assessing learners' academic English-speaking performance and self-efficacy, and they effectively support the data collection objectives of this study. These tools can be used to evaluate both academic speaking ability and self-efficacy among EFL pre-service teachers and hold potential for informing instructional design aimed at enhancing academic English-speaking competence in multicultural educational contexts.

5. Conclusion

This study successfully developed and preliminarily validated two complementary instruments for use in EFL teacher education: the Academic English-Speaking Assessment (AESA) and the Self-Efficacy in Academic English-Speaking Scale (SE-AESS). Together, these instruments address important gaps in previous research by integrating a performance-based assessment of academic English speaking with a context-specific measure of self-efficacy grounded in Bandura's theoretical framework.

The AESA provides a systematic and linguistically grounded approach to assessing academic English-speaking competence through an analytic rubric adapted from established assessment frameworks. Expert review confirmed strong content validity, indicating that the assessment tasks and scoring criteria appropriately represent the construct of academic English speaking. The use of trained raters and structured scoring procedures further supports the reliability and rigor of the assessment, making AESA a valuable tool for evaluating pre-service teachers' academic speaking performance in formal educational contexts.

The SE-AESS demonstrated excellent content validity and good internal consistency for the academic English-speaking self-efficacy dimension. Although the reliability of the sources of self-efficacy dimension was moderate, this result is consistent with the multidimensional and experiential nature of the construct and is acceptable for an instrument at an early stage of development. Importantly, the SE-AESS operationalizes Bandura's four sources of self-efficacy specifically within academic English-speaking contexts, which has been largely absent from prior research.

5.1 Theoretical Implications

From a theoretical perspective, this study reinforces the principle that self-efficacy measurement must be task- and context-specific. By focusing on academic English speaking rather than general language use, the SE-AESS extends self-efficacy theory into a domain that is critical for EFL pre-service teachers' academic and professional development. The integration of performance-based assessment with psychological measurement also contributes to a more comprehensive understanding of how self-efficacy beliefs relate to observable language performance.

5.2 Practical Implications

From a practical standpoint, the combined use of AESA and SE-AESS offers valuable diagnostic information for teacher education programs. The AESA enables instructors to identify specific strengths and weaknesses in learners' academic speaking skills, while the SE-

AESS provides insights into learners' confidence levels and the sources influencing their self-efficacy. This information can inform the design of targeted instructional interventions, such as creating structured mastery experiences, incorporating peer modeling, providing constructive feedback, and fostering supportive learning environments that reduce anxiety and enhance learners' readiness to speak in academic contexts.

Overall, the AESA and SE-AESS provide a coherent and evidence-based framework for assessing and supporting the development of academic English-speaking competence among EFL pre-service teachers. Their application has the potential to enhance instructional design, learner reflection, and professional preparation in multilingual and multicultural educational settings.

6. Discussion

The findings from the development and preliminary validation of the AESA and SE-AESS indicate that both instruments address critical limitations in previous research on English-speaking assessment and self-efficacy in EFL contexts. Unlike many earlier studies that relied on self-report measures or academic achievement scores as proxies for speaking ability, the AESA employs a performance-based assessment evaluated by expert raters using an analytic rubric. This approach allows for a more precise and linguistically grounded evaluation of academic English-speaking competence, particularly for pre-service teachers who must use English in formal academic and professional contexts.

The strong content validity of the AESA confirms that the speaking tasks and rubric components appropriately represent the construct of academic English speaking. By incorporating dimensions such as grammar, vocabulary, fluency, pronunciation, comprehensibility, and accuracy, the AESA aligns with established frameworks for authentic and analytic language assessment. The use of trained raters and inter-rater reliability procedures further strengthens the methodological rigor of the assessment and enhances confidence in the consistency of scoring.

The SE-AESS also demonstrated excellent content validity and acceptable reliability for early-stage instrument development. The high internal consistency of the academic English-speaking self-efficacy dimension indicates that the items effectively capture learners' perceived confidence in performing academic speaking tasks. The moderate reliability coefficient observed for the sources of self-efficacy dimension reflects the complex and experiential nature of this construct. Because mastery experience, vicarious experience, verbal persuasion, and emotional states are influenced by diverse personal histories and contextual factors, variability in responses is expected. Similar challenges have been reported in previous studies examining multidimensional psychological constructs, particularly during initial stages of scale development.

Importantly, the development of SE-AESS extends prior self-efficacy research by operationalizing Bandura's four sources of self-efficacy within academic English-speaking contexts. Most existing instruments assess self-efficacy beliefs without examining their underlying sources or without linking them to actual performance outcomes. By integrating SE-AESS with the AESA, the present study provides a dual-dimensional assessment framework that captures both observable speaking performance and the psychological factors that influence it. This integrated approach offers greater explanatory power than using either performance assessment or self-report measures alone.

From an instructional perspective, the combined use of AESA and SE-AESS has practical implications for teacher education programs. Performance data from the AESA can help instructors identify specific areas of strength and weakness in learners' academic speaking skills, while SE-AESS results can reveal learners' confidence levels and the sources shaping their self-efficacy. This information enables the design of targeted instructional interventions,

such as providing structured mastery experiences, modeling effective academic speaking, offering constructive feedback, and creating supportive learning environments that reduce anxiety and enhance emotional readiness.

Overall, the findings support the usefulness of AESA and SE-AESS as complementary tools for assessing and supporting the development of academic English-speaking competence among EFL pre-service teachers. The instruments contribute to both research and practice by providing a theoretically grounded and methodologically robust framework that addresses the limitations of prior studies in this area.

7. Suggestions

Future research should focus on refining the SE-AESS, particularly the sources of self-efficacy dimension, by revising item wording and increasing the number of items within each source to enhance internal consistency. Further validation studies using larger and more diverse samples across different institutions and proficiency levels are also recommended to strengthen the generalizability of the instruments.

In addition, future studies may examine the relationships among sources of self-efficacy, perceived academic English-speaking self-efficacy, and actual speaking performance using advanced statistical techniques such as structural equation modeling or path analysis. Longitudinal research designs could also provide insights into how self-efficacy and academic speaking competence develop over time in response to instructional interventions.

References

- Bachman, L. F., & Palmer, A. S. (2010). *Language assessment in practice: Developing language assessments and justifying their use in the real world*. Oxford University Press.
- Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Prentice Hall.
- Bandura, A. (1994). *Self-efficacy*. In V. S. Ramachandran (Ed.), *Encyclopedia of human behavior* (Vol. 4, pp. 71–81). Academic Press.
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. W. H. Freeman.
- Bandura, A. (2001). Social cognitive theory: An agentic perspective. *Annual Review of Psychology*, 52, 1–26.
- Brown, H. D. (2012). *Principles of language learning and teaching (6th ed.)*. Pearson Education.
- Brown, H. D., & Abeywickrama, P. (2019). *Language assessment: Principles and classroom practices (3rd ed.)*. Pearson.
- Chang, P., Zhou, L., & Zhang, L. J. (2024). *The roles of task-specific self-efficacy and enjoyment in EFL learners' development of speaking CALF: A longitudinal study*. *System*, 127, Article 103539.
- Creswell, J. W. (2018). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research (6th ed.)*. Pearson.
- Dörnyei, Z., & Taguchi, T. (2010). *Questionnaires in second language research: Construction, administration, and processing (2nd ed.)*. Routledge.
- Eckes, T. (2015). *Introduction to many-facet Rasch measurement: Analyzing and evaluating rater-mediated assessments*. Peter Lang.
- El-Hmoudová, D. (2015). *Developing speaking skills through task-based learning*. *Procedia – Social and Behavioral Sciences*, 197, 219–223.
- Fulcher, G. (2010). *Practical language testing*. Hodder Education.
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). *Multivariate data analysis (8th ed.)*. Cengage Learning.
- Maddux, J. E. (2018). *Self-efficacy: The power of believing you can*. In J. E. Maddux (Ed.), *Self-efficacy, adaptation, and adjustment* (pp. 3–35). Routledge.
- Kitikanan, P., & Sasimonton, P. (2017). The relationship between English self-efficacy and English learning achievement of L2 Thai learners. *LEARN Journal*, 10(1), 149–164.
- Koo, T. K., & Li, M. Y. (2016). A guideline for selecting and reporting intraclass correlation coefficients

- for reliability research. *Journal of Chiropractic Medicine*, 15(2), 155–163.
- Lin, Y., Mohamad, M. B. T., & Mahmud, M. I. (2025). Exploring the relationship between classroom English proficiency and teacher self-efficacy among polytechnic EFL teachers. *Arab World English Journal*, 16(2), 80–95.
- Luoma, S. (2004). *Assessing speaking*. Cambridge University Press.
- McNamara, T. (1996). *Measuring second language performance*. Longman.
- Mills, K., Pajares, F., & Herron, C. (2006). *A reevaluation of the role of anxiety: Self-efficacy, anxiety, and their relation to reading and listening proficiency in foreign language learning*. *Foreign Language Annals*, 39(2), 276–295.
- Nguyen, T. T. M., & Habók, A. (2022). Developing and validating a self-efficacy questionnaire for CEFR A2 level EFL learners. *MEXTESOL Journal*, 46(2), 1–17.
- O'Malley, J. M., & Pierce, L. V. (1996). *Authentic assessment for English language learners: Practical approaches for teachers*. Addison-Wesley.
- Prada, J. (2010). *Effects of a family program on perceived caregiving behavioral self-efficacy (Master's thesis)*. Chulalongkorn University.
- Rahimi, M. (2007). Self-efficacy, attribution, and achievement in English as a foreign language. *Journal of Applied Linguistics*, 10(2), 119–137.
- Schunk, D. H., & DiBenedetto, M. K. (2020). *Self-efficacy and human motivation*. In *Advances in motivation science* (Vol. 7, pp. 153–179). Elsevier.
- Siriporn, C., Intaraprasert, C., & Akkakoson, S. (2017). Factors affecting English speaking ability of Thai undergraduate students. *Journal of Asia TEFL*, 14(3), 507–521.
- Siritaratn, N. (2013). English self-efficacy beliefs of EFL low proficiency graduate students. *Academic Journal of Interdisciplinary Studies*, 2(3), 461–468.
- Sirithes, P., Intaraprasert, C., & Akkakoson, S. (2019). *Thai EFL learners' willingness to communicate in English: The roles of speaking self-efficacy and anxiety*. *PASAA*, 56, 1–30.
- Suharja. (2022). The relationship between self-efficacy and speaking performance of EFL students. *Journal of English Language Teaching*, 11(1), 45–54.
- Syahidah, W., & Umasugi, R. (2021). Developing an English-speaking assessment rubric for academic purposes. *Journal of English for Academic Purposes*, 54, 101041.
- Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education*, 2, 53–55.
- Tilfarlioglu, F. Y., & Cinkara, E. (2009). *Self-efficacy in EFL: Differences among proficiency groups and relationship with success*. *Novitas-ROYAL*, 3(2), 129–142.
- Tschannen-Moran, M., & Hoy, A. W. (2001). *Teacher efficacy: Capturing an elusive construct*. *Teaching and Teacher Education*, 17(7), 783–805.
- Tschannen-Moran, M., & Hoy, A. W. (2007). *The differential antecedents of self-efficacy beliefs of novice and experienced teachers*. *Teaching and Teacher Education*, 23(6), 944–956.
- Wang, C., Kim, D.-H., Bong, M., & Ahn, H. S. (2013). Examining measurement properties of an English self-efficacy scale for English language learners. *International Journal of Educational Research*, 59, 68–79.
- Weigle, S. C. (2002). *Assessing writing*. Cambridge University Press.