

Editorial

Introduction to the special issue: Exploring language assessment issues and challenges in ASEAN contexts

This special issue of PASAA is dedicated to research into issues and challenges in language assessment from classroom-based assessment to large-scale assessment in terms of principles, policies, and practices in ASEAN regions. This issue focused on the exploration of different perspectives on how principles and policies are transformed into actual practices in language assessment, ranging from the application of small-scale assessment in traditional classroom settings to the administration of large-scale and standardized assessment on a regional and national level. This special issue featured seven papers that focused on language assessment in various contexts related to the ways in which language tests are designed, interpreted, and used to make decisions about test takers.

In the first paper, Khamboonruang explored differential rater functioning, also known as rater bias or rater interaction in high-stakes EFL classroom writing assessment using a many-facets Rasch measurement (MFRM) approach. Khamboonruang found that the less experienced teachers rated their own classroom students more leniently but graded each other's classroom students more harshly than expected. This finding underscores the crucial need for well-designed rating rubrics and comprehensive rater training to ensure the validity and fairness of rater-mediated assessments in the classroom. It is clear that fairness in cross-classroom rating cannot be achieved without the presence of high-quality rating rubrics and well-prepared raters.

In a similar vein, Imsa-ard and Tangkiengsirisin have highlighted the significance of language assessment knowledge as a fundamental requirement for

effective assessment practices in language classrooms. Their research reveals that Thai EFL university lecturers generally possess a satisfactory level of language assessment knowledge and show a preference for performance assessment over traditional methods. However, they still require support in terms of scoring, providing feedback, and assisting students in performance-based assessments.

This need for additional support in the context of Thai tertiary EFL education is echoed in Piamsai's study, which examined the perceptions of Thai undergraduate students regarding their English speaking, listening, reading, and writing skills. Piamsai used the Thai version of the CEFR-based self-assessment grid and also explored the correlations between this self-assessment and a university English proficiency test. The findings indicate that Thai students exhibited more confidence in their receptive language skills compared to their productive language skills. Furthermore, the study identified significant but relatively weak positive correlations between their self-assessment scores and their proficiency test results. Piamsai argued that enhancing the accuracy of self-assessment can be achieved through training and aligning the test constructs measured by the self-assessment grids with the actual test tasks.

Two additional empirical studies in this special edition delve into the broader aspects of implementing English language assessment in K-12 contexts. Nguyen et al. explore the intricate relationship between the macro and micro contexts of English language assessment in Vietnam, examining various conditions across different educational levels, including primary, lower-secondary, upper-secondary, and higher education. Their research also extends to different sectors, encompassing both the public and private domains. This investigation is based on intensive interviews with five specialists and educational leaders, focusing on nationwide English language assessment policies and the provision of guidelines for teacher training (the macro context), as well as the institutional conditions and practices (the micro context). Their findings reveal that standardization, initiated

at the national level, is evident in both the public and private sectors. Moreover, a shift towards assessment for learning, going beyond traditional assessment methods, has been observed. However, they identify a certain level of tension at the micro level, as effective policy implementation necessitates consistent quality management. This can be achieved by ensuring that human resources possess the requisite language competence and assessment literacy, along with the necessary financial resources. Their study emphasizes the significance of understanding the context and the interplay among its various factors in facilitating the intended policy implementation to enhance English education overall.

Mohd Salleh et al. present a study on the washback effects of Malaysia's school-based assessment (SBA) on the English language learning of secondary school students. Utilizing various data collection methods, including questionnaires, interviews, and classroom observations, their research identifies a positive washback of the test, particularly regarding the students' perceptions and attitudes toward SBA. The study highlights that engaging activities and constructive feedback provided through SBA significantly contribute to these positive perceptions. However, in alignment with the findings of Nguyen et al. on this issue, it is worth noting that a significant number of students expressed concerns about the adequacy of resources and class time for the effective implementation of SBA in schools.

In this special issue, we also present a test review and idea-sharing papers. Lu conducts a comprehensive review of the Duolingo English Test (DET), an AI-powered online English proficiency test that has been gaining significant attention in the field of language assessment. Lu's critical analysis of the DET is based on Bachman and Palmer's (1996) test usefulness framework, with a particular focus on reliability, content validity, and impact. Lu highlights that the DET demonstrates a high level of reliability, effectively eliminating inconsistencies and biases associated with human raters. However, Lu also identifies certain limitations, primarily related to the test's lack of authenticity and absence of human

interaction. These limitations may restrict the assessment of academic English abilities for which the DET is primarily used. To address these concerns, Lu offers several recommendations aimed at improving the DET to enhance its intended score interpretation and applications.

Finally, Shaw shares his insights on crafting item difficulty in the TOEFL iBT listening test, drawing from his extensive experience as a test developer at the Educational Testing Center (ETS). Shaw identifies three crucial factors that determine item difficulty in three main areas: 1) text characteristics, including topic familiarity, vocabulary frequency, saliency, location, repetition of tested information, and information density; 2) recording characteristics of the text, such as speech rate and the presence of a pause after tested information; 3) characteristics of test questions, with a particular emphasis on creating distractors in multiple-choice questions. Shaw provides illustrative examples that clarify these aspects, offering valuable guidance for English teachers seeking to enhance their understanding of how to manage item difficulty when creating their own listening tests and interpreting standardized English listening test results.

This special issue has addressed a variety of assessment issues and challenges within the ASEAN region. It has provided solutions and shared insights derived from diverse contexts, spanning large-scale and classroom-based assessments at both regional and national levels. While the issues discussed in the articles are specific to their respective local settings, we believe that by integrating these findings, a clearer and more comprehensive understanding can be achieved regarding the complex relationships among the various principles and policies that underpin language teaching and assessment practices in ASEAN contexts.

Sun-Young Shin, Senyung Lee, and Jirada Wudthayagorn

Guest Editors

PASAA