# Tailoring CEFR-Based Writing Assessment: A Localised Framework for Evaluating International Students' Academic Readiness in a Malaysian Technical University

**Yuziana Yasin[1,2,*], Harshita Aini Haroon[1,2], Sharmini Abdullah[1,2]**

[1]Faculty of Business and Communication, Universiti Malaysia Perlis (UniMAP), Perlis, Malaysia.
[2]Languages and General Studies Department, Universiti Malaysia Perlis (UniMAP), Perlis, Malaysia.

| KEYWORDS | ABSTRACT |
|---|---|
| Thesis<br>Recycled aggregate<br>Alkali activated material<br>Aerodynamic<br>Coefficient of friction | This study explores the development and validation of a localised CEFR-based writing test tailored for international students at Universiti Malaysia Perlis (UniMAP). The test addresses gaps in standardised assessments by aligning language evaluation with the specific academic and cultural contexts of Malaysian universities. The research employs a qualitative methodology across two phases—test development and validation. Findings indicate the test's potential to enhance the alignment between curriculum, teaching, and assessment while providing insights into students' readiness for academic challenges. This localised framework has broader implications for language assessment and international student integration in higher education. |

## 1. INTRODUCTION

The use of standardised tests like IELTS, TOEFL, and MUET offers a set of general benchmarks for English proficiency but often fail to address the specific academic contexts. Many students with high scores still struggle with academic writing in Malaysian universities. This misalignment between test outcomes and actual academic demands creates challenges for higher institution in supporting international students' academic success.

### 1.1 CEFR in Language Assessment: Overview of Its Role and Global Application

The Common European Framework of Reference for Languages (CEFR) is widely recognised as a comprehensive framework for assessing language proficiency. Initially developed to standardise language education in Europe, CEFR has gained global adoption due to its transparent descriptors and proficiency levels ranging from A1 (beginner) to C2 (proficient). In academic settings, CEFR provides a structured approach to aligning language instruction and evaluation, offering a common reference point for learners, educators, and institutions. However, cultural and linguistic differences require adaptation to ensure the framework's relevance and effectiveness in local educational environments (Foley, 2019). These challenges highlight the need for localised approaches to CEFR implementation, particularly in higher education institutions serving diverse linguistic populations.

### 1.2 Localisation in Language Testing: Adapting Frameworks to Institutional Needs

Localisation in language testing refers to tailoring assessments to reflect the unique linguistic, cultural, and academic needs of specific contexts. O'Sullivan (2011) emphasises the importance of contextual relevance in test development, suggesting for adjustments that align assessments with local curricula, cultural norms, and institutional goals. While IELTS and TOEFL provide a broad measure of English proficiency, they often fail to account for the specific writing demands faced by students in Malaysian universities (Mohd Ali et al., 2018). Localisation efforts in Malaysia, such as aligning CEFR descriptors with national education standards, have demonstrated the potential to bridge this gap, enhancing the relevance of language assessments to the local academic context (Rahman et al., 2021). However, developing localised assessments requires significant resources, expertise, and collaboration among stakeholders, including educators, policymakers, and language assessment experts. Furthermore, balancing localisation with the need for standardisation poses a critical challenge, as overly localised tests may affect the comparability and recognition of students' performance on an international scale.

### 1.3 Challenges in Writing Assessments: Issues with Task Alignment, Rater Consistency, and Standardisation

Writing assessments involves cognitive, linguistic, and sociocultural dimensions. One major challenge lies in aligning writing tasks with proficiency descriptors, such as those in CEFR. Studies by Harsch & Rupp (2011) highlight the importance of level-specific tasks, but this

*Corresponding author: yuziana@unimap.edu.my

approach can limit flexibility to cover multiple levels of language proficiency. In addition, variability in how raters interpret and apply scoring rubrics can lead to unreliable results, thus affecting the validity of writing assessments (Deygers & Van Gorp, 2015). Moderation and the use of statistical tools, such as the Intraclass Correlation Coefficient (ICC), have been proposed as solutions to enhance inter-rater reliability, particularly in high-stakes assessments. Standardisation further affect writing assessments, especially in diverse educational contexts. While standardised frameworks like CEFR offer a common foundation, their application in culturally distinct settings often requires adaptation. Excessive standardisation can also overlook the unique linguistic and academic needs of learners, leading to assessments that are misaligned with their educational goals (Fox & Artemeva, 2022). Hence, balancing standardisation with contextual adaptation is essential to creating fair and effective writing assessments.

Building on the identified challenges of standardised tests and the need for a localised approach in assessing writing proficiency, this study adopts a thorough methodology to develop and validate a CEFR-based writing test tailored to a technical Malaysian university.

## 2. METHODOLOGY

To address the gaps highlighted in the introduction, a systematic approach is employed to design and validate a localised CEFR-based writing test that aligns with the specific academic and cultural needs of international students at UniMAP.

### 2.1 Research Design

This study employs a mixed-methods approach structured into two phases: test development and validation. In the test development phase, writing tasks and scoring rubrics were designed to align with CEFR descriptors. The validation phase ensured that the test accurately measured students' writing proficiency and addressed their specific linguistic and academic needs. Qualitative content analysis was conducted on focus group discussion data and to identify recurring patterns and themes. Descriptive statistics were applied to task difficulty ratings and inter-rater reliability scores, to assess consistency. Statistical measures, such as Intraclass Correlation Coefficients (ICC), were used to evaluate scoring reliability and ensure consistency among raters. This methodological approach provided a framework for developing and validating a localised CEFR-based writing test, ensuring its relevance, fairness, and reliability in assessing international students' academic readiness as in Figure 1 below.
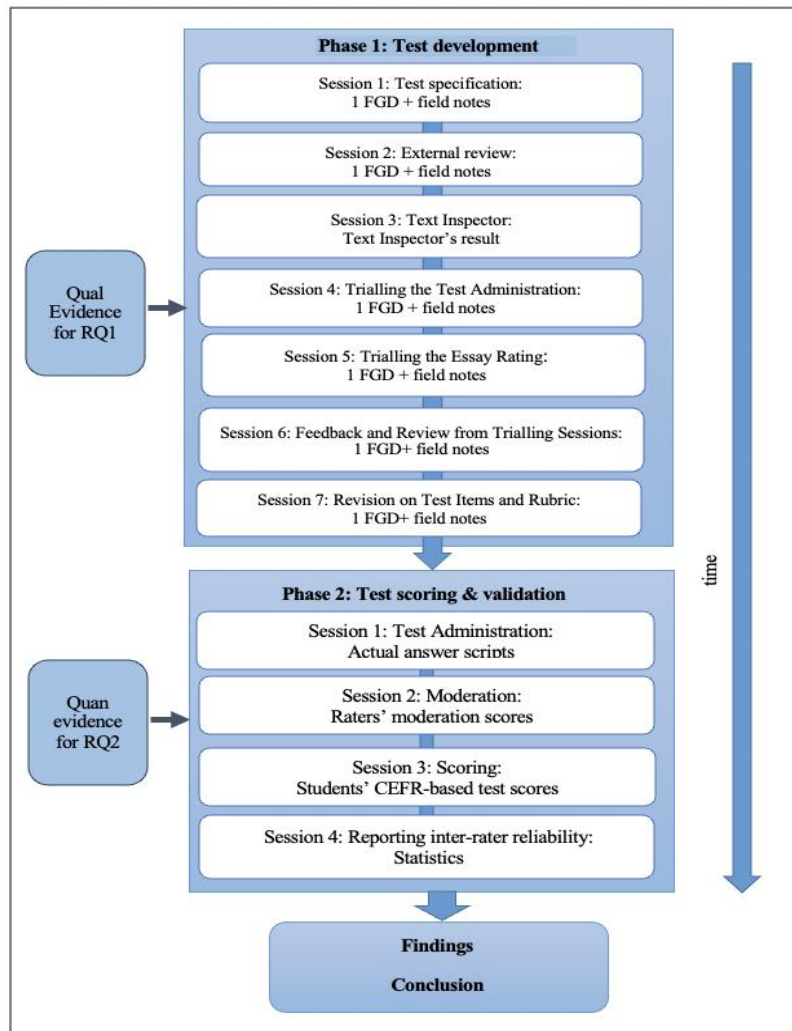
**Figure 1.** Framework for developing and validating a localised CEFR-based writing test.

## 2.2 Participants

To ensure the test's effectiveness and relevance, the study involved a diverse group of participants:

i. 10 UniMAP's teachers, provided insights into task design, alignment with CEFR descriptors, and contextual appropriateness.
ii. 5 Raters: scoring writing tasks, offered feedback on rubric clarity and inter-rater reliability.
iii. 3 External experts from other institution, reviewed the test specifications and ensured both theoretical and practical alignment.
iv. 32 International students valuated the tasks' relevance, clarity, and perceived fairness.

## 3. RESULTS AND DISCUSSION

### 3.1 Insights from Teachers, Raters, and External Experts

i. Teachers highlighted the importance of aligning writing tasks with the academic challenges commonly faced by international students. Additionally, teachers stressed the need for clear and concise instructions to avoid ambiguity among test-takers.
ii. Raters showed inconsistencies in the initial rubrics. It was recommended to revise the descriptors to enhance inter-rater reliability.
iii. External experts validated the alignment of tasks with CEFR descriptors and offered refinements to better address the Malaysian educational context. It was suggested to refine test instructions and improve task clarity.

**3.2 Adjustments to Writing Prompts and Scoring Descriptors**

i.   Writing Prompts were revised to ensure cultural neutrality and relevance, focusing on topics applicable to international students' academic and personal experiences.

ii.  Scoring Descriptors were to provide clearer distinctions between proficiency levels. Adjustments ensured the rubrics reflected CEFR descriptors more effectively, supporting fair and consistent evaluation.

**3.2  Statistical Reliability of Test Scores**

Inter-rater reliability, measured using the Intraclass Correlation Coefficient (ICC), indicated strong consistency among raters. The internal consistency of the test, assessed through Cronbach's alpha, confirming the reliability of the test across its components.

| Components | Cronbach's Alpha | ICC Value |
|---|---|---|
| 1. Part A (Language) | 0.951 | 0.866 |
| 2. Part A (Task Fulfilment) | 0.951 | 0.866 |
| 3. Part A (Total Score) | 0.970 | 0.914 |
| 4. Part B (Language) | 0.983 | 0.950 |
| 5. Part B (Task Fulfilment) | 0.978 | 0.936 |
| 6. Part B (Total Score) | 0.988 | 0.964 |
| 7. Overall score | 0.989 | 0.967 |

**3.3 Perceptions of Test-Takers Regarding Relevance and Clarity**

i.   Relevance: The participants agreed that the writing tasks were applicable to their academic contexts and effectively mirrored real-world academic requirements.

ii.  Clarity: Participants reported that the test instructions were clear and easy to follow, contributing to a positive test-taking experience. Some test-takers highlighted difficulties with specific vocabulary in prompts, suggesting further refinement in language complexity.

**4. CONCLUSION**

This study developed and validated a localised CEFR-based writing test tailored to the linguistic and academic needs of international students at a technical university in Malaysia. Strong inter-rater reliability and positive feedback reflected the effectiveness of the localised test, highlighting its potential to evaluate writing proficiency. These findings are particularly significant to support international students, contributing to a more inclusive and effective admissions process.

**REFERENCE**

[1]  Council of Europe. (2001). *Common European Framework of Reference for Languages: Learning, teaching, assessment*. Cambridge University Press.

[2]  Deygers, B., & Van Gorp, K. (2015). Rater effects in writing assessment: A multilevel analysis of variability in rater severity and consistency. *Language Testing, 32*(2), 179–200. https://doi.org/10.1177/0265532214562097.

[3]  Foley, J. A. (2019). Language learning and teaching: Global frameworks and local realities. *Asian Journal of Language Teaching, 12*(3), 10–22.

[4]  Fox, J., & Artemeva, N. (2022). Standardization and adaptation: Balancing global frameworks and local contexts in language testing. *Journal of Language Assessment and Evaluation, 20*(1), 45–62.

[5]  Harsch, C., & Rupp, A. A. (2011). Designing and scaling level-specific writing tasks aligned with the CEFR. *Language Assessment Quarterly, 8*(1), 1–25. https://doi.org/10.1080/15434303.2010.538725

[6]  Mohd Ali, Z., Hassan, R., & Sulaiman, M. (2018). Challenges in adapting standardized language tests to local contexts: The Malaysian perspective. *International Journal of Language Education and Assessment, 6*(2), 90–112.

[7]  O'Sullivan, B. (2011). Localisation and language testing: Matching test design and use. In B. O'Sullivan (Ed.), *Language Testing: Theories and Practices* (pp. 51–75). Palgrave Macmillan.

[8]  Rahman, N. A., Aziz, M. N. A., & Karim, S. A. (2021). Aligning CEFR with the Malaysian education system: Issues and challenges in localized test development. *Education in Malaysia: Policies and Practices, 9*(2), 120–138.