

Development and Validation of Writing Assessment Tools for Grade 7 Second Language Learners of the Division of Abra

Jelly P. Galinte ¹
jellygalinte21@gmail.com/ 0009-0004-9564-9116

1 – Abra State Institute of Sciences and Technology

Publication Date: August 16, 2026

DOI: [10.5281/zenodo.21966410](https://doi.org/10.5281/zenodo.21966410)

Abstract

This study developed and validated writing assessment tools for Grade 7 second-language learners in the Division of Abra, addressing the gap in systematically validated teacher-made instruments for assessing writing proficiency. Anchored on Constructivist Learning Theory (Vygotsky, 1978) and Messick's Unified Theory of Validity (1989), the research employed a developmental research design involving 27 Grade 7 learners from Abra High School–Calot Extension and three expert validators in English language teaching, curriculum development, and language assessment. The researcher-developed instruments comprised five contextualized writing tasks aligned with the Most Essential Learning Competencies (MELCs), evaluated using a holistic rubric assessing content, organization, language use, grammar, and mechanics on a four-point rating scale. Data analysis utilized Weighted Mean for validity levels, Fleiss' Kappa for inter-rater agreement, and Cronbach's Alpha for reliability. Results demonstrated substantial improvement from initial to final validation, with weighted mean scores increasing from 4.50–4.70 to 4.75–4.96, indicating Very Highly Valid ratings. Fleiss' Kappa coefficients improved from –0.154–0.375 (weak agreement) to 0.391–1.000 (moderate to almost perfect agreement) following expert revisions. Reliability analysis revealed acceptable internal consistency for Tool 1 ($\alpha = 0.770$), Tool 2 ($\alpha = 0.700$), and Tool 5 ($\alpha = 0.710$), while Tool 3 ($\alpha = 0.521$) and Tool 4 ($\alpha = 0.631$) require further refinement. The validated instruments provide English teachers with curriculum-aligned, evidence-based assessment tools that promote objective and consistent classroom assessment. The study recommends full implementation of the validated tools in Grade 7 classrooms, enhanced validation procedures with orientation sessions for experts, and further refinement of lower-reliability instruments through clearer rubric descriptions and rater training. Future research should involve larger and more diverse samples and additional expert validators to strengthen the generalizability and applicability of the developed writing assessment tools.

Keywords: *Writing Assessment, Validation, Inter-Rater Reliability, Second Language Learners, Rubric Development*

1. Introduction

Writing is an essential skill that enables learners to express ideas, develop critical thinking, and communicate effectively in academic settings. For second-language learners, writing presents greater challenges because it requires mastery of grammar, vocabulary, organization, and content while expressing ideas in a language that is not their first language. Effective assessment is therefore necessary to accurately measure learners' writing proficiency and identify areas for improvement.

The study is anchored on **Constructivist Learning Theory (Vygotsky, 1978)**, which emphasizes that learners acquire knowledge through interaction, experience, and guided learning, and on **Messick's Unified Theory of Validity (1989)**, which highlights the importance of validity and reliability in assessment instruments. Despite the importance of writing assessment, many teachers experience difficulties in developing validated assessment tools, resulting in inconsistent evaluation of learners' writing performance. In the local context, Grade 7 second-language learners in Abra may be assessed using teacher-made instruments that have not undergone systematic validation. Hence, this study developed and validated writing assessment tools to ensure that they are valid, reliable, and appropriate for assessing the writing proficiency of Grade 7 second-language learners in the Division of Abra.

2. Materials and Methods

Research Design

The study employed a **developmental research design** to develop and validate writing assessment tools for Grade 7 second-language learners. The design involved the systematic development, expert validation, revision, pilot testing, and statistical evaluation of the assessment instruments to establish their validity and reliability.

Participants

The study involved **27 Grade 7 second-language learners** from **Abra High School–Calot Extension** who participated in the pilot testing of the developed writing assessment tools. It also involved **three expert validators** with expertise in English language teaching, curriculum development, and language assessment who evaluated the instruments during the validation process.

Instruments

The primary instrument was a **researcher-developed Writing Assessment Tool for Second-Language Learners in English 7**, consisting of **five contextualized writing tasks**: Expository Writing, Making Inferences and Drawing Conclusions, Using Textual Evidence in an Argumentative Essay, Sending Letters as Response, and Letter of Request. Each task was evaluated using a **researcher-developed holistic rubric** assessing **content, organization, language use, grammar, and mechanics** on a four-point rating scale. A **validation checklist**

was also used by expert validators to evaluate the instruments in terms of content validity, clarity and organization, and language relevance and appropriateness.

Procedure

Permission to conduct the study was obtained from the Schools Division Superintendent and the school head. The researcher developed the writing assessment tasks and holistic rubrics based on the Most Essential Learning Competencies (MELCs). The instruments were submitted for expert validation, revised based on the validators' recommendations, and subjected to final validation. The validated instruments were then pilot-tested with Grade 7 learners, after which the collected data were analyzed to determine the validity, inter-rater agreement, and reliability of the developed assessment tools.

Data Analysis

Data were analyzed using **Weighted Mean** to determine the levels of validity, **Fleiss' Kappa** to measure inter-rater agreement among expert validators, and **Cronbach's Alpha** to determine the reliability of the developed writing assessment tools. These statistical analyses were used to establish the validity, consistency, and suitability of the assessment instruments for classroom use.

3. Results

The study established the validity, inter-rater agreement, and reliability of the developed writing assessment tools for Grade 7 second-language learners. Results showed that all five assessment tools achieved **Very Highly Valid** ratings during the initial validation, with weighted mean scores ranging from **4.50 to 4.70** across content validity, clarity and organization, and language relevance and appropriateness. After incorporating the recommendations of the expert validators, the final validation yielded improved weighted mean scores ranging from **4.75 to 4.96**, confirming the enhanced quality and suitability of the assessment tools for classroom use.

The level of agreement among the expert validators also improved substantially after revision. During the initial validation, **Fleiss' Kappa coefficients ranged from -0.154 to 0.375**, indicating weak agreement among validators. Following the refinement of the instruments, the final validation produced **Kappa values ranging from 0.391 to 1.000**, demonstrating moderate to almost perfect agreement. These findings indicate that the revisions enhanced the consistency and objectivity of expert evaluations.

Reliability analysis using **Cronbach's Alpha** further revealed that three assessment tools demonstrated acceptable internal consistency: **Tool 1 ($\alpha = 0.770$)**, **Tool 2 ($\alpha = 0.700$)**, and **Tool 5 ($\alpha = 0.710$)**. Meanwhile, **Tool 3 ($\alpha = 0.521$)** and **Tool 4 ($\alpha = 0.631$)** obtained lower reliability coefficients, indicating the need for further refinement to improve scoring consistency.

Overall, the findings demonstrate that the developed writing assessment tools possess strong evidence of validity, improved inter-rater agreement, and generally acceptable reliability.

These results support their use as curriculum-aligned assessment instruments for evaluating the writing proficiency of Grade 7 second-language learners while highlighting opportunities for the continued enhancement of selected tools.

4. Discussion

Interpretation of Findings

The findings demonstrate that the developed writing assessment tools are valid, reliable, and appropriate for assessing the writing proficiency of Grade 7 second-language learners. The improvement in weighted mean scores from the initial to the final validation indicates that incorporating expert recommendations enhanced the quality of the assessment tools. Likewise, the increase in Fleiss' Kappa values from weak to moderate and almost perfect agreement suggests that revisions improved the consistency of expert evaluations. Although three assessment tools demonstrated acceptable reliability, two instruments require further refinement to strengthen their internal consistency before broader implementation. Overall, the results support the effectiveness of a systematic development and validation process in producing quality writing assessment tools.

Comparison to Existing Studies

The findings are consistent with the literature reviewed in the study, which emphasizes that valid writing assessment instruments should undergo expert validation and reliability testing before classroom implementation. The high validity ratings and improved inter-rater agreement support **Messick's Unified Theory of Validity**, highlighting the importance of systematic validation in ensuring meaningful assessment. Likewise, the development of contextualized writing tasks and holistic rubrics aligns with **Constructivist Learning Theory**, which advocates authentic assessment that reflects learners' experiences and supports meaningful learning. These findings also reinforce previous studies cited in the review of related literature that reported the effectiveness of validated rubrics and expert-reviewed assessment tools in improving writing assessment practices.

Implications for Practice and Policy

The validated writing assessment tools provide English teachers with curriculum-aligned and evidence-based instruments for assessing the writing proficiency of Grade 7 second-language learners. Their use can promote more objective, consistent, and meaningful classroom assessment while supporting competency-based instruction aligned with the Most Essential Learning Competencies (MELCs). The findings also highlight the importance of continuous validation, expert review, and rubric refinement in developing assessment tools. School administrators and curriculum developers may use these findings to strengthen assessment practices and support the development of standardized writing assessment resources in basic education.

Study Limitations

The study was conducted using a relatively small sample of **27 Grade 7 learners** from a single public school in the Division of Abra, which may limit the generalizability of the findings to other educational settings. Additionally, only **three expert validators** participated in the validation process, and two assessment tools demonstrated lower reliability coefficients, indicating the need for further refinement. Future studies are encouraged to involve larger and more diverse samples, additional expert validators, and further validation across different schools to strengthen the reliability and applicability of the developed writing assessment tools.

5. Conclusion

1. The writing assessment instruments were already good enough during the first phase of the validation since they are considered well-crafted regarding their contents, organization, and appropriateness. However, after going through the second phase of the validation, the instruments become even more improved due to the increase in clarity, structure, and connection with Grade 7 competencies. Therefore, it seems that the validation procedure has been successful in improving the instruments' quality and making them fit for use in a classroom setting, especially when assessing second language writers.

2. Initially, expert agreement was low due to some degree of disagreement on how the assessment instruments were used in their original forms. Nonetheless, the agreement levels increased remarkably after the instruments had undergone modification. The significant improvement shows that there was an increase in the agreement of the experts on how the assessments should be done. Thus, the modification helped address issues that could have been responsible for the inconsistencies in the application of the tools.

3. There exists variability among the writing assessment tools in relation to their reliability. In this regard, while some of the tools exhibit high levels of reliability and consistent scoring, others have shown poor reliability scores, which might be attributed to ambiguous assessment rubrics. Based on the findings, it can be concluded that, despite the fact that the tools may be useful for assessments, they require further improvements in order to enhance their consistency and reliability levels.

Recommendations

Based from the conclusions, the following recommendations are forwarded:

1. The writing assessment instruments that are tested and found valid can be fully implemented in Grade 7 classrooms since they have been validated as well as being appropriate for evaluating the writing abilities of second language learners. Teachers can include these writing assessment tools within their classroom assessments for learners. Feedback from teachers in classrooms must also be taken into consideration for the development of these tools.

2. The future design and validation of tests need to have better descriptions of what the test is supposed to measure and how it will be measured starting from the very beginning. It would be helpful if orientation or calibration sessions were held before conducting the validation process for all experts involved in the process.

3. To increase reliability, assessment instruments that lack sufficient reliability need to be improved by refining rubric descriptions, making them clearer, precise, and consistent. In order to facilitate ease in interpretation of rubric descriptions by raters, complicated rubric descriptions need to be made simpler. Moreover, pilot tests and rater training are suggested to enhance inter-rater reliability.

For future research, it is advised that there be an increase in validation and reliability tests for the writing assessment tools through increased participation of experts and raters in larger numbers to increase evidence of validity and reliability and thus ensure the generalizability of the results obtained from the tools.

References

- Ajjawi, R., & Bearman, M. (2024). Assessment feedback and evaluation: Improving validity through iterative design. *Assessment & Evaluation in Higher Education*, 49(1), 1–15. <https://doi.org/10.1080/02602938.2024.2301456>
- Andrade, H., & Beekman, K. (2023). Editorial: Classroom assessment as the co-regulation of learning. *Frontiers in Education*, 7. <https://doi.org/10.3389/feduc.2022.1063123>
- Annisa, A., & Gusnawaty, G. (2024). Feedback and revisions in developing writing skills in second language learners: A systematic review. *Kajian Pendidikan, Seni, Budaya, Sosial Dan Lingkungan*, 1(1), 1–8. <https://doi.org/10.58881/kpsbsl.v1i1.7>
- Asli, N. F., Matore, M. E. E. M., & Yunus, M. M. (2024). Construct validity of primary trait writing rubrics based on assessment use argument (AUA) validation framework. *Heliyon*, 10(22), e40053. <https://doi.org/10.1016/j.heliyon.2024.e40053>
- Bitchener, J., & Ferris, D. R. (2022). *Written corrective feedback in second language acquisition and writing*. <https://doi.org/10.4324/9780203832400>
- Bodbod, I. (2024). Exploring writing strategies OF two sophomore students OF a state university in the Philippines: A case study on its implications for pedagogical support AND cognitive awareness. *Balangkas*, 1(1), 10–16. <https://doi.org/10.66317/688571>
- Brookhart, S. M. (2021). *How to create and use rubrics for formative assessment and grading*. ASCD. <https://doi.org/10.4324/9780429356601>
- Brookhart, S. M., & Chen, F. (2024). Improving rubric design through feedback and iterative validation. *Educational Measurement: Issues and Practice*, 43(1), 56–70. <https://doi.org/10.1111/emip.12567>
- Cabudbod, R. O., Osias, N. C., & Comon, J. D. (2024). Predictors of writing skills and learners' writing proficiency in Salay District, Division of Misamis Oriental. *European Modern Studies Journal*, 8(3), 474–503. [https://doi.org/10.59573/emsj.8\(3\).2024.34](https://doi.org/10.59573/emsj.8(3).2024.34)
- Cepillo, J. C. (2024). Minds on Manuscripts: Analyzing the research writing competencies of SHS Students as Catalysts for an innovation. *International Journal of Innovative Science and Research Technology (IJISRT)*, 962–983. <https://doi.org/10.38124/ijisrt/ijisrt24sep745>