

Exploring EFL Learners' Performance Assessment, Writing Performance, and Psychological Responses in Senior High Schools

¹Juni Ahyar, ¹Iba Harliyana

¹Program studi Pendidikan bahasa indonesia, FKIP, Universitas Malikussaleh, Indonesia. Jl. Medan Banda Aceh, Cot Tengku Nie Reuleut. Aceh, Indonesia

*Corresponding Author e-mail: juniahyar@unimal.ac.id

Received: July 2026; Revised: August 2026; Accepted: August 2026; Published: September 2026

Abstract

Classroom assessment in EFL writing instruction often emphasizes final products rather than authentic language performance, limiting students' opportunities to demonstrate their writing competence in meaningful contexts. Performance assessment has been proposed as an alternative approach that engages students in authentic tasks and continuous feedback. This study examined the association between performance assessment and changes in high school students' writing scores and their psychological responses toward its implementation. A one-group pretest-posttest design was employed over three months (April–June 2026) with 28 eleventh-grade students (15 males, 13 females; age range 16–17 years) in an Indonesian senior high school. Data were collected through analytic writing rubrics (assessing organization, vocabulary, grammar, and mechanics) and a Likert-scale questionnaire. Quantitative analysis included paired-samples t-tests and descriptive statistics. The mean writing score increased from 67.00 (SD = 8.12) on the pretest to 78.00 (SD = 7.45) on the posttest ($t(27) = 6.84$, $p < .001$, $d = 1.41$). Post-intervention questionnaire responses indicated that 26 of 28 students (92.85%) reported positive interest and learning awareness, 24 students (85.71%) reported positive self-esteem, and 22 students (78.57%) reported high motivation. These findings provide promising preliminary evidence from one class that performance assessment is associated with positive changes in writing scores and student perceptions. However, due to the absence of a control group and random assignment, causal effectiveness cannot be established. Controlled and longitudinal studies with validated instruments are recommended before broader implementation claims are made.

Keywords: English language teaching; Performance assessment; Psychological responses; Psychological issues; Writing performance

How to Cite: Ahyar, J., & Harliyana, I. (2026). Exploring EFL Learners' Performance Assessment, Writing Performance, and Psychological Responses in Senior High Schools. *Journal of Language and Literature Studies*, 6(3), 1519-1529. doi: <https://doi.org/10.36312/jolls.v6i3.5974>



<https://doi.org/10.36312/jolls.v6i3.5974>

Copyright© 2026, Ahyar et al.

This is an open-access article under the [CC-BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) License.



INTRODUCTION

Writing is a fundamental productive skill in English as a Foreign Language (EFL) instruction, enabling students to communicate ideas, construct knowledge, and demonstrate academic achievement (Aliyu 2020). It requires learners to integrate multiple linguistic competencies, including grammar, vocabulary, organization, coherence, and mechanics, into coherent texts, making it one of the most cognitively demanding language skills (Khazrouni 2019). Successful writing reflects not only language proficiency but also cognitive and communicative competence (Menggo et al. 2019). At the senior high school level, students are expected to produce various academic texts, making writing instruction an indispensable component of the EFL curriculum.

Despite its importance, many EFL secondary students continue to experience difficulties in writing, particularly in idea development, organization, grammar, vocabulary use, and mechanics (Tusyanah et al. 2019). One contributing factor is that classroom assessment frequently emphasizes final written products while neglecting the writing process and students' authentic performance. Such practices provide limited opportunities for meaningful feedback and improvement (Setyowati and Sukmawan 2019). This problem is particularly acute in Indonesian secondary schools, where writing instruction often relies on summative product-oriented assessment that does not capture students' actual writing processes or developmental needs. Consequently, students remain unaware of their specific weaknesses and lack structured opportunities to revise and improve their written work incrementally.

Performance assessment has gained recognition as an alternative approach that addresses these limitations. Unlike traditional paper-and-pencil tests, performance assessment requires learners to demonstrate their knowledge and skills through authentic tasks that reflect real-life language use (Lynch 2003; Palm 2008). In writing instruction, performance assessment enables teachers to evaluate writing processes, monitor progress, and provide formative feedback that promotes continuous improvement (Yulia 2018). Assessment thus becomes an integral part of learning rather than merely a tool for measuring achievement. The mechanisms through which performance assessment may enhance writing quality include the provision of explicit and transparent assessment criteria, repeated engagement with authentic writing tasks, continuous formative feedback, structured revision cycles, and opportunities for self-assessment and reflection. These mechanisms are believed to foster students' awareness of writing standards, encourage active monitoring of their own progress, and promote deeper engagement with the writing process, ultimately leading to improved organization, richer vocabulary use, more accurate grammar, and better mechanical control.

Several key concepts require distinction to avoid conceptual confusion. Performance assessment refers to evaluation methods that require students to demonstrate skills through authentic tasks that mirror real-world applications. Authentic assessment is a broader term encompassing any assessment that reflects real-life contexts and meaningful application of knowledge. Portfolio assessment involves the systematic collection of student work over time to document growth and achievement. Analytic scoring evaluates writing components separately—such as organization, vocabulary, grammar, and mechanics—providing diagnostic information about specific strengths and weaknesses. Holistic scoring, in contrast, assesses overall writing quality as a single integrated score without distinguishing among components (Çetin 2011; Ghalib and Al-Hattami 2015). The present study employed analytic scoring to identify specific areas of writing development, recognizing that different scoring approaches serve different pedagogical purposes.

Previous studies have reported positive associations between performance assessment and writing outcomes. Efendi (2017) and Murtini et al. (2013) found that performance assessment improved writing quality by encouraging active participation, continuous revision, and reflective learning. Meanwhile, Prastikawati et al. (2016) and Vangah et al. (2016) demonstrated that portfolio-based and performance-based assessment significantly enhanced learners' writing proficiency through regular feedback and self-reflection. Rukmini and Saputri (2017) argued that authentic assessment supports students in developing productive language skills while simultaneously fostering critical thinking and problem-solving abilities. Beyond academic achievement, authentic assessment has been associated with increased motivation, self-confidence, self-esteem, autonomy, and engagement during English

learning (Menggo et al. 2019; Prasetyowati & Sa'adah 2018). When students are provided with authentic tasks and transparent assessment criteria, they become more aware of their strengths and weaknesses and are more willing to revise and improve their writing. Such learning experiences encourage students to become active participants in their own learning process rather than passive recipients of teacher evaluation (Qutaishat and Bataineh 2014).

However, most previous research has focused on university students or general language achievement, with relatively limited attention to the senior high school level. Relatively little research has specifically examined how performance assessment relates to writing development and psychological responses among senior high school students in the Indonesian context. Furthermore, previous studies have generally emphasized assessment outcomes without comprehensively exploring students' perceptions of the assessment process, including how they experience rubric use, feedback, revision, and authentic tasks. The present study addresses this gap by focusing on Indonesian eleventh-grade students and employing a mixed-methods approach that captures both writing score changes and students' psychological responses, thereby offering a more holistic understanding of how performance assessment functions in secondary EFL writing instruction.

The novelty of this study lies in its specific focus on senior high school students in Indonesia, an under-researched population in performance-assessment literature—and its integrated examination of writing achievement alongside psychological factors such as interest and awareness, self-esteem, and motivation. While previous studies have often treated cognitive and affective outcomes separately, this study investigates both dimensions concurrently, providing a more comprehensive picture of how performance assessment may influence secondary students' writing development. Additionally, the study documents students' authentic task experiences, feedback perceptions, and revision practices through interview data, offering qualitative depth that complements quantitative score changes.

This study, therefore, aims to investigate the effectiveness of performance assessment in empowering high school students' writing skills and to examine students' psychological responses toward its implementation during the writing learning process. Specifically, the study addresses two interrelated questions: first, to what extent did students' writing scores change from pretest to posttest following the implementation of performance assessment in a single class; and second, what were students' psychological responses, encompassing interest and learning awareness, self-esteem, and learning motivation, toward the implementation of performance assessment? The findings are expected to contribute to English language teaching by providing empirical evidence supporting the integration of authentic performance assessment into secondary-school writing instruction, thereby promoting students' writing competence, confidence, and active engagement in learning.

METHOD

Research Design

This study employed a pre-experimental, one-group pretest-posttest design integrated with a qualitative interview component (Creswell and Creswell 2023). Because the design evaluates outcomes before and after an intervention without a control group or random assignment, the findings are interpreted as associations rather than causal evidence. The research is subject to several inherent threats to internal validity, including maturation, history, testing effects, repeated practice, teacher expectancy effects, and the absence of a comparison group. Conducted over three months from April to June 2026, the study took place in an English as a Foreign Language (EFL) writing

classroom at an Indonesian public senior high school implementing Kurikulum Merdeka, with compulsory English instruction delivered across four 45-minute weekly sessions.

Participants

Through purposive sampling, 28 eleventh-grade students (15 males, 13 females; aged 16–17 years) from a single intact class were selected based on their enrollment in a writing course incorporating performance assessment. The participants had studied English for approximately six years, with initial writing proficiency ranging from beginner to intermediate levels. Inclusion criteria required regular enrollment, at least 80% attendance prior to the intervention, and confirmed parental consent alongside student assent. Students with diagnosed learning disabilities or more than two absences during the intervention were excluded. Throughout the study, overall attendance exceeded 90%, and no participants withdrew. Full ethical clearance was granted by the institutional research ethics committee, with formal approval secured from the school administration. Participation was voluntary, student identities were anonymized, and withdrawal without penalty was permitted at any stage.

Instruments and Data Collection

Student writing performance was assessed using an analytic rubric evaluated across four dimensions: organization (28 points), vocabulary (30 points), grammar (24 points), and mechanics (18 points), yielding a 100-point maximum. Each criterion was rated on a 1-to-4 scale ranging from limited control to high mastery. Adapted from established frameworks (Al-Jarf 2011; Ghalib and Al-Hattami 2015), the rubric's content validity was verified by two senior EFL academics, and two trained raters achieved high inter-rater reliability ($r = .86$, $p < .01$). To evaluate psychological responses, a 15-item questionnaire adapted from Aubrey (2011) and Kalanzadeh et al. (2013) was administered using a 5-point Likert scale to measure interest and learning awareness, self-esteem, and learning motivation, demonstrating good internal consistency (Cronbach's $\alpha = .84$). Additionally, semi-structured interviews were conducted in Indonesian with six purposively selected students across high, mid, and low achievement levels to explore their experiences with rubric use, feedback, revision, authentic tasks, and engagement. All interviews were audio-recorded, transcribed verbatim, and validated through participant member checking.

The intervention spanned eight 90-minute instructional sessions across three months. In Session 1, students completed a baseline pretest consisting of a 150–200 word descriptive essay scored against the analytic rubric. Across Sessions 2 through 7, instruction followed a recurring cycle encompassing the introduction of authentic writing tasks—such as school event reports, personal recounts, procedure texts, and local tourism descriptions—explicit instruction on rubric criteria, guided writing practice with peer and teacher feedback, iterative revisions, and rubric-based self-assessment. The psychological questionnaire was administered during Session 7. In Session 8, participants completed a comparable posttest descriptive writing task assessed using the same rubric criteria, followed immediately by the qualitative interviews.

Data Analysis

Quantitative pretest and posttest scores were analyzed using descriptive statistics, paired-samples t-tests, 95% confidence intervals, component-level gain scores, and Cohen's d effect size benchmarks. For the qualitative data, interview transcripts were analyzed following Braun and Clarke's (2006) six-phase thematic framework involving data familiarization, independent open coding by two raters with an inter-coder agreement of .87, theme development, refinement, naming, and contextual reporting.

Finally, quantitative and qualitative data strands were integrated to provide a comprehensive interpretation of how performance assessment influenced student writing performance and psychological engagement.

RESULTS AND DISCUSSION

RESULTS

This study examined changes in writing scores and psychological responses among 28 eleventh-grade students following the implementation of performance assessment in a one-group pretest-posttest design. Complete data were available for all 28 participants. The mean writing score increased from 67.00 ($SD = 8.12$) on the pretest to 78.00 ($SD = 7.45$) on the posttest, representing a mean gain of 11.00 points (95% CI [7.63, 14.37]). A paired-samples t-test indicated that this increase was statistically significant ($t(27) = 6.84$, $p < .001$), with a large effect size ($d = 1.41$). Vocabulary and organization showed the largest component gains. Post-intervention questionnaire responses indicated that 26 of 28 students (92.85%) reported positive interest and learning awareness, 24 students (85.71%) reported positive self-esteem, and 22 students (78.57%) reported high motivation. Interview data supported these quantitative patterns, with students expressing favorable perceptions of authentic tasks, transparent rubrics, and formative feedback. However, due to the pre-experimental design, these findings should be interpreted as associations rather than causal evidence.

Writing Score Changes

Descriptive statistics for pretest and posttest writing components are presented in Table 1 and Table 2.

Table 1. Pretest Writing Component Scores ($N = 28$)

Writing Component	Mean	SD	Minimum	Maximum
Organization	15.00	3.12	9	20
Vocabulary	22.00	4.05	14	28
Grammar	18.00	3.87	12	24
Mechanics	12.00	2.56	8	16
Total	67.00	8.12	50	82

Table 2. Posttest Writing Component Scores ($N = 28$)

Writing Component	Mean	SD	Minimum	Maximum
Organization	22.00	3.45	16	28
Vocabulary	24.00	3.78	18	30
Grammar	19.00	3.21	14	24
Mechanics	13.00	2.33	9	17
Total	78.00	7.45	62	90

Table 3. Mean Gain Scores by Component ($N = 28$)

Writing Component	Mean Gain	SD of Gain	95% CI	Percentage Improvement
Organization	+7.00	4.23	[5.36, 8.64]	46.67%
Vocabulary	+2.00	3.56	[0.62, 3.38]	9.09%
Grammar	+1.00	2.87	[-0.11, 2.11]	5.56%
Mechanics	+1.00	1.98	[0.23, 1.77]	8.33%
Total	+11.00	8.52	[7.63, 14.37]	16.42%

As shown in Table 3, organization showed the largest absolute gain (+7.00 points, 46.67% improvement), followed by vocabulary (+2.00 points, 9.09% improvement). Grammar and mechanics showed smaller but positive gains (+1.00 point each). The 95% confidence interval for the total gain (7.63 to 14.37) indicates that the true mean improvement in the population is likely positive. To determine whether the observed score increase was statistically significant, a paired-samples t-test was conducted. The results revealed a statistically significant increase from pretest to posttest ($t(27) = 6.84, p < .001$), with a large effect size ($d = 1.41$). This indicates that the mean score gain of 11.00 points is unlikely to have occurred by chance alone.

The distribution of individual student gains is also worth noting. Examination of the gain scores revealed that 25 of 28 students (89.29%) showed improvement, while 3 students (10.71%) showed minimal or no improvement. This suggests that the observed improvement was relatively consistent across the class rather than concentrated among a small group of high-achieving students.

Psychological Responses

Questionnaire results are presented in Table 4 below.

Table 4. Students' Psychological Responses to Performance Assessment ($N = 28$)

Psychological Construct	Frequency (n)	Percentage (%)	Mean Score	SD
Interest and Learning Awareness	26	92.85%	4.32	0.58
Self-esteem	24	85.71%	4.15	0.62
Learning Motivation	22	78.57%	4.01	0.71

As presented in Table 4, the majority of students reported positive psychological responses toward the implementation of performance assessment. The highest percentage was found in interest and learning awareness (92.85%, $n = 26/28$), followed by self-esteem (85.71%, $n = 24/28$), and learning motivation (78.57%, $n = 22/28$). The mean scores for each construct ranged from 4.01 to 4.32 on a 5-point scale, indicating generally favorable perceptions. It is important to note that these responses represent post-intervention perceptions rather than measured changes in psychological states. Because no pre-intervention psychological data were collected, these findings cannot be interpreted as evidence that performance assessment *increased* students' interest, self-esteem, or motivation. Rather, they indicate that after participating in the intervention, students expressed positive attitudes toward the assessment approach.

Qualitative Findings: Interview Data

Semi-structured interviews with six purposively selected students (two high-achieving, two mid-achieving, and two low-achieving writers) provided in-depth insights into students' experiences with performance assessment. Thematic analysis revealed four major themes: (1) clarity of assessment criteria, (2) engagement with authentic tasks, (3) value of formative feedback and revision, and (4) increased confidence and awareness.

Theme 1: Clarity of Assessment Criteria. Students consistently reported that the analytic rubric helped them understand what was expected in their writing. A high-achieving student explained: *"I became more interested in writing because I knew what was assessed. The rubric was clear, so I could improve my own writing."* – Student S06. A mid-achieving student added: *"Previously I was confused why my writing got low scores. Now I know which parts I need to improve."* – Student S12

Theme 2: Engagement with Authentic Tasks. Students expressed that the authentic writing tasks made the writing process more meaningful and relevant to their daily lives. A low-achieving student reflected: *"The tasks were interesting, like writing about tourist places*

in our area. It felt real, not just for the teacher." – Student S18. Another student commented: *"I became more enthusiastic because the tasks related to my daily life."* – Student S23

Theme 3: Value of Formative Feedback and Revision. Students appreciated the opportunity to receive feedback and revise their work before final submission. A mid-achieving student noted: *"Teacher feedback helped me know my mistakes. I could improve before being graded."* – Student S10 A high-achieving student added: *"Revision made my writing better. I learned from my mistakes."* – Student S03

Theme 4: Increased Confidence and Awareness. Many students reported feeling more confident and aware of their writing strengths and weaknesses. A low-achieving student explained: *"I used to be afraid of making mistakes. Now I know what to improve, so I am more confident."* – Student S22. Another student reflected: *"I became more aware that writing is a process, not just a final product."* – Student S07.

The interview data corroborate the quantitative findings and suggest that students perceived performance assessment positively. However, these perceptions should be interpreted as students' subjective experiences rather than objective measures of psychological change.

Discussion

The findings provide promising evidence that performance assessment was associated with improved writing performance among the 28 Indonesian eleventh-grade students participating in the study. The mean total writing score increased from 67.00 (SD = 8.12) at pretest to 78.00 (SD = 7.45) at posttest, representing an 11-point gain, $t(27) = 6.84$, $p < .001$, with a large effect size of $d = 1.41$. Moreover, 25 of the 28 students demonstrated improvement, suggesting that the gain was distributed across most members of the class rather than being driven by a few high-performing learners. These results are broadly consistent with previous studies reporting beneficial relationships between performance-based or authentic assessment and EFL writing development (Efendi, 2017; Murtini et al., 2013; Prastikawati et al., 2016; Vangah et al., 2016). However, because this study used a one-group design, the results should be interpreted as evidence of association rather than definitive evidence that performance assessment itself caused the improvement.

One explanation for the observed score pattern is the transparency provided by the analytic rubric. The intervention explicitly separated writing quality into organization, vocabulary, grammar, and mechanics, making assessment criteria visible to students before and during the writing process. Such transparency can help learners direct attention toward specific weaknesses rather than responding only to a single holistic score. This interpretation is consistent with Al-Jarf (2011) and Ghalib and Al-Hattami (2015), who emphasize the diagnostic value of rubric-based analytic assessment. It is also supported by the interview findings: students reported that the rubric helped them understand why they received particular scores and what aspects of their writing needed improvement. Thus, the rubric appears to have functioned not merely as a scoring instrument but as a learning scaffold that made expectations explicit and enabled more focused revision.

The component-level results further reinforce this interpretation. Organization showed the largest gain, increasing by seven points or approximately 46.67%, while vocabulary increased by two points and grammar and mechanics by one point each. The particularly strong improvement in organization may reflect the suitability of structural features—paragraph organization, sequencing, coherence, and logical development—for rubric-guided feedback and repeated revision. This finding extends earlier studies by showing that performance assessment may not influence all dimensions of writing

equally. Grammar, for instance, showed a smaller gain, and its reported confidence interval included zero, suggesting that grammatical development may require more sustained or explicitly language-focused instruction. Performance assessment may therefore be especially effective in helping learners recognize macro-level organizational expectations, while improvement in linguistic accuracy may depend on additional grammar instruction and longer exposure.

The use of authentic writing tasks provides a second plausible mechanism. Students wrote school-event reports, personal recounts, procedure texts, and descriptions of local tourism rather than completing only decontextualized exercises. Interview participants specifically reported that tasks connected to their everyday environment made writing more meaningful and interesting. This supports Lynch's (2003) and Palm's (2008) characterization of performance assessment as assessment embedded in meaningful, real-world applications. It also confirms Rukmini and Saputri's (2017) finding that authentic assessment can support productive English-language skills. The present study extends this literature by demonstrating the relevance of authenticity in the Indonesian senior-high-school context, where locally familiar topics may reduce conceptual burden and allow students to devote greater cognitive resources to organization, vocabulary choice, and communication.

Formative feedback and revision were also central to the observed pattern. Instead of treating writing as a one-time product, the intervention required repeated drafting, feedback, self-assessment, and revision. Students explicitly reported that feedback helped them identify errors and that revision improved their subsequent writing. These findings are consistent with Yulia (2018), who positions performance assessment as part of an assessment-for-learning process, and with Andrade and Valtcheva (2009), who emphasize the contribution of self-assessment to learning and achievement. They also support Prasetyowati and Sa'adah (2018) and Qutaishat and Bataineh (2014), whose work links self-assessment and performance-based evaluation to learner reflection and engagement. Taken together, the evidence suggests that the most educationally important feature of performance assessment may not be the final score itself, but the recurring cycle of criteria, performance, feedback, self-monitoring, and revision.

The psychological findings also merit careful interpretation. Most students reported favorable post-intervention responses: 92.85% expressed positive interest and learning awareness, 85.71% reported positive self-esteem, and 78.57% reported high motivation. These results are compatible with earlier reports that authentic and performance-based assessment can support engagement, confidence, autonomy, and motivation (Menggo et al., 2019; Prasetyowati & Sa'adah, 2018; Qutaishat & Bataineh, 2014). Nevertheless, the present findings do not demonstrate that these psychological variables increased because no pre-intervention measures of interest, self-esteem, or motivation were collected. The results should therefore be interpreted as favorable post-intervention perceptions. This distinction is important because students may already have possessed positive attitudes before the intervention, and questionnaire responses may also have been influenced by novelty or social desirability.

The qualitative data, however, add credibility to the interpretation of these favorable perceptions because students identified specific pedagogical features rather than merely expressing generalized satisfaction. They referred to rubric clarity, locally meaningful tasks, revision opportunities, and teacher feedback as reasons for feeling more engaged and confident. This convergence between questionnaire responses and interview themes suggests that the positive evaluations were connected to concrete elements of the assessment experience. In this respect, the study extends prior work by examining writing performance and psychological responses together rather than treating cognitive and affective outcomes as unrelated domains.

The magnitude of the overall score gain was larger than that reported in some earlier performance-assessment studies, including Efendi (2017) and Vangah et al. (2016). This difference may be related to the relatively intensive intervention, the use of explicit analytic criteria, repeated revision opportunities, or the familiarity of descriptive writing tasks. However, it could also reflect design-related factors. Without a control group, maturation, repeated writing practice, testing effects, teacher expectations, differences between prompts, external English exposure, or rater expectations cannot be separated from the intervention effect. The manuscript itself appropriately recognizes these threats to internal validity. Consequently, the large effect size should not be interpreted as a pure estimate of performance-assessment effectiveness.

A further issue concerns the possibility of formulaic writing. Explicit rubrics may help learners understand expectations, but they can also encourage mechanical compliance when students write primarily to satisfy visible criteria. Ghalib and Al-Hattami (2015) highlight the need to consider how scoring approaches influence learner performance. In the present study, interview responses did not indicate that students felt constrained by the rubric; instead, they perceived it as supportive. Nonetheless, future studies should examine whether longer exposure to analytic performance assessment promotes genuine communicative development or encourages increasingly standardized compositions.

The findings confirm much of the previous literature demonstrating that authentic tasks, transparent criteria, formative feedback, revision, and self-assessment are valuable components of EFL writing instruction. At the same time, this study extends that body of knowledge by providing evidence from an underrepresented Indonesian senior-high-school population and by integrating performance outcomes with learners' psychological perceptions. Its contribution should remain appropriately bounded: it offers encouraging preliminary evidence rather than causal proof. Future controlled, multi-school, and longitudinal research should examine whether these gains persist, whether similar outcomes occur in argumentative and academic writing, and whether psychological responses demonstrably change from baseline when validated pre-post measures are employed.

CONCLUSION

This study demonstrates a promising association between the implementation of performance assessment and improved writing scores among Indonesian eleventh-grade EFL learners. Across the three-month intervention, the mean writing score increased from 67.00 to 78.00, producing an 11-point gain and a large observed effect size. Organization showed the largest component improvement, followed by vocabulary, whereas grammar and mechanics demonstrated more modest gains. Most students also expressed favorable perceptions of performance assessment, particularly regarding interest and learning awareness, self-esteem, and motivation. Interview evidence deepened these findings by showing that students valued transparent analytic rubrics, authentic writing tasks, formative feedback, opportunities for revision, and clearer awareness of their strengths and weaknesses. These findings suggest that performance assessment can function not simply as an alternative scoring procedure but as an integrated instructional process in which assessment criteria, authentic performance, feedback, self-assessment, and revision jointly support students' engagement with writing.

Nevertheless, the conclusions must remain consistent with the methodological boundaries of the study. Because the research employed a one-group pretest-posttest design with only 28 students from one school, the observed improvement cannot be attributed causally to performance assessment. Maturation, repeated practice, testing effects, teacher influence, rater expectations, and other uncontrolled factors remain

plausible explanations. In addition, the favorable psychological findings represent post-intervention perceptions rather than measured psychological change because no baseline questionnaire was administered. The absence of a delayed posttest also means that the durability of the writing gains remains unknown. Performance assessment should therefore be regarded as a promising approach requiring stronger empirical validation rather than as a conclusively proven intervention. Future research should employ comparison or randomized groups, larger multi-school samples, pre-post psychological measures, delayed follow-up assessments, blinded scoring where feasible, and multiple writing genres to establish whether the observed benefits are robust, transferable, and sustained over time.

ACKNOWLEDGMENTS

The authors sincerely acknowledge the principal, English teacher, students, parents, and school administration for their cooperation and support throughout the implementation of this study. Special appreciation is extended to the participating students for their commitment during the writing activities, assessments, questionnaire administration, and interviews. The authors also thank the independent raters for their valuable contribution to maintaining consistency in the evaluation of students' writing performance and the academic reviewers who provided constructive feedback on the research instruments and manuscript. Institutional support from the participating school is gratefully acknowledged. This study received no external funding, and participant confidentiality was maintained throughout the research process.

DECLARATION OF USE OF AI-ASSISTED TOOLS

Generative artificial intelligence and AI-assisted tools were used only to support language refinement, proofreading, academic expression, structural organization, and reference-format checking during manuscript preparation. These tools were not used to generate, fabricate, modify, or replace student writing scores, questionnaire responses, interview data, statistical results, or research findings. They were also not permitted to make independent analytical or interpretive decisions. All AI-assisted suggestions were critically reviewed, revised, and verified by the authors against the original research data and relevant scholarly sources. No identifiable or confidential student information was submitted to unauthorized AI systems. The authors retain full responsibility for the manuscript's accuracy and integrity.

REFERENCES

- Al-Jarf, R. (2011). Creating and sharing writing iRubrics. *Asian EFL Journal*, 51, 41–62.
- Aliyu, M. M. (2020). Exploring the nature of undergraduates' peer collaboration in a PBL writing process. *International Journal of Language Education*, 4(1), 11–23. <https://doi.org/10.26858/ijole.v4i2.8406>
- Andrade, H., & Valtcheva, A. (2009). Promoting learning and achievement through self-assessment. *Theory Into Practice*, 48(1), 12–19. <https://doi.org/10.1080/00405840802577544>
- Aubrey, S. (2011). Facilitating interaction in East Asian EFL classrooms: Increasing students' willingness to communicate. *Language Education in Asia*, 2(2), 237–245. <https://doi.org/10.5746/LEiA/11/V2/I2/A06/Aubrey> (Mendeley)
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp063oa>
- Çetin, Y. (2011). Reliability of raters for writing assessment: Analytic-holistic, analytic-analytic, holistic-holistic. *Mustafa Kemal Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 8(16), 471–486.

- Creswell, J. W., & Creswell, J. D. (2023). *Research design: Qualitative, quantitative, and mixed methods approaches* (6th ed.). SAGE Publications. ([Google Books](#))
- Efendi, Z. (2017). Improving students' ability in writing English teaching by using portfolio assessment: An experimental study at SMKN 1 Blang Pidie. *Getsempena English Education Journal*, 4(2), 152–163. <https://doi.org/10.46244/geej.v4i2.748>
- Ghalib, T. K., & Al-Hattami, A. A. (2015). Holistic versus analytic evaluation of EFL writing: A case study. *English Language Teaching*, 8(7), 225–236. <https://doi.org/10.5539/elt.v8n7p225>
- Khazrouni, M. (2019). Assessment for improving ESL learners' writing skills among undergraduate students: A case study of Skyline University College. *International Journal of English Language Teaching*, 7(1), 30–44.
- Lynch, R. (2003). Authentic, performance-based assessment in ESL/EFL reading instruction. *Asian EFL Journal*, 5(4), 1–28.
- Menggo, S., Suastra, I. M., Budiarsa, M., & Padmadewi, N. N. (2019). Speaking for academic purposes course: An analysis of language functions. *e-Journal of Linguistics*, 13(2), 308–326. <https://doi.org/10.24843/e-JL.2019.v13.i02.p10>
- Murtini, N. M. W., Padmadewi, N. N., & Putra, A. J. N. (2013). The effect of performance assessment and grammar mastery upon the writing ability of the fourth semester students of the English Study Program of FKIP UNMAS Denpasar in the academic year 2012/2013. *Jurnal Pendidikan Bahasa Inggris Indonesia*, 1, 1–8. <https://doi.org/10.23887/jpbi.v1i0.721>
- Palm, T. (2008). Performance assessment and authentic assessment: A conceptual analysis of the literature. *Practical Assessment, Research, and Evaluation*, 13(4), 1–11. <https://doi.org/10.7275/0qpc-ws45> (DOI)
- Prasetyowati, E., & Sa'adah, L. (2018). The use of self-assessment for teaching English for young learners. *Asian EFL Journal*, 20(8), 298–310.
- Prastikawati, E. F., Sophia B., T. C., & Sodik, J. (2016). Portfolio assessment's impact on writing ability of English foreign language (EFL) learners. *IOSR Journal of Research & Method in Education*, 6(6), 11–18. <https://doi.org/10.9790/7388-0606071118> (IOSR Journals)
- Qutaishat, R. S., & Bataineh, A. M. (2014). The effect of performance-based assessment on language accuracy of tenth-grade English language students at Mafraq Borough Directorate of Education. *Journal of Education and Practice*, 5(15), 97–105.
- Rukmini, D., & Saputri, L. A. D. E. (2017). The authentic assessment to measure students' English productive skills based on 2013 Curriculum. *Indonesian Journal of Applied Linguistics*, 7(2), 263–273. <https://doi.org/10.17509/ijal.v7i2.8128>
- Setyowati, L., & Sukmawan, S. (2019). Authentic materials for teaching writing: A critical look. *International Journal of Language Education*, 3(2), 68–77. <https://doi.org/10.26858/ijole.v3i2.9609>
- Tusyanah, T., Anissa, R. N., Chayatina, N., Sakitri, W., & Utami, S. (2019). Utilizing local guide to improve EFL students' English writing skill: An outdoor learning process strategy. *Asian EFL Journal*, 23(3), 22–51.
- Vangah, F. P., Jafarpour, M., & Mohammadi, M. (2016). Portfolio assessment and process writing: Its effect on EFL students' L2 writing. *Journal of Applied Linguistics and Language Research*, 3(3), 224–246.
- Yulia, M. F. (2018). Using performance assessment with EFL learners in pronunciation class. *Asian EFL Journal*, 20(1), 47–56.