Enigma in Education
https://enigma.or.id/index.php/edu

e-ISSN: 3026-5827

ORIGINAL ARTICLE
Enigma in Education | Vol. 4, No. 1 (2026) | pp. 13–18

Effectiveness of an Automated Feedback System on Creative Writing and Scientific Translation Skills Among Undergraduate Students: A Quasi-Experimental Study

Danila Adi Sanjaya[1], Gayatri Putri[2*]

[1]Department of Literature Science, CMHC Research Center, Palembang, Indonesia
[2]Department of Library and Informatics Science, Enigma Institute, Palembang, Indonesia

ARTICLE INFO

Article history
Received: March 19, 2026
Revised: April 18, 2026
Accepted: May 2, 2026
Available online: May 20, 2026
Published: May 30, 2026

Keywords
Automated feedback; computer-assisted language learning; creative writing; scientific translation; writing self-efficacy

Corresponding author
Gayatri Putri
gayatri.putri@enigma.or.id

ORCID
0009-0006-1314-2157

DOI
https://doi.org/10.61996/edu.v4i1.123

ABSTRACT

Background. Evidence on automated feedback for higher-order skills is thin. Creative writing, scientific translation, and writing self-efficacy among non-native English speakers have received little study.

Objective. This study examined how an automated feedback system (AFS) was associated with creative writing skills (CWS), scientific translation skills (STS), and writing self-efficacy (WSE) among undergraduates.

Methods. A quasi-experimental pretest–posttest design with a non-equivalent control group was used. The 136 participants (experimental group: n = 68; control group: n = 68) studied in the Department of English Education at a public university in Palembang, Indonesia. They attended Writing and Translation courses during a 10-week intervention. The experimental group received automated feedback through an integrated computer-assisted language learning platform. The control group received conventional feedback from instructors. Pretest and posttest data were analyzed with analysis of covariance (ANCOVA).

Results. Large effects favored the experimental group. They appeared for creative writing skills (d = 1.84, 95% CI [1.38, 2.30], p < 0.001), scientific translation skills (d = 2.08, 95% CI [1.62, 2.54], p < 0.001), and writing self-efficacy (d = 1.94, 95% CI [1.49, 2.39], p < 0.001).

Conclusion. In this sample, automated feedback was associated with greater skill development and learner confidence in academic writing, mostly in skill-intensive tasks that need iterative practice. Combining it with periodic instructor guidance and peer collaboration is also emphasized.

Copyright: © 2026 The Author(s). Published by Enigma Institute on behalf of Enigma in Education. The authors retain copyright of their work.

Access license: This is an open access article distributed under the terms of the Creative Commons Attribution–NonCommercial–ShareAlike 4.0 International License (CC BY-NC-SA 4.0), which permits non-commercial use, sharing, adaptation and reproduction in any medium, provided the original author and source are credited and any derivative work is distributed under the same license.

1. Introduction

Technology now plays a growing part in language education. Computer-assisted language learning (CALL) systems have become practical aids for teaching writing.[1,2] Undergraduates in every discipline are expected to write well, whether the task is creative expression or composition in a specialized field.[3,4] Feedback delivered by instructors has clear limits. Each student receives little individual attention, comments arrive late, and corrections are applied unevenly.[5,6]

Automated feedback systems were developed to address these limits by giving learners corrective input that is immediate, consistent, and easy to scale.[7,8] Earlier studies found that such feedback helps with grammar and mechanical errors.[9,10] Evidence for higher-order skills, creative writing and scientific translation in particular, is still scarce.[11]

Scientific translation means conveying specialized terminology and concepts accurately from one language to another while keeping to the conventions of the discipline, and it poses a writing problem of its own.[12] Creative writing, by contrast, rewards originality, stylistic growth, and rhetorical skill.[13] Both depend on metacognitive awareness and on repeated cycles of revision.[14,15]

Writing self-efficacy covers a learner's confidence in writing, belief in the ability to overcome difficulties, and persistence through revision. It predicts writing achievement.[16,17] How automated feedback relates to self-efficacy among non-native English speakers has seldom been studied.[18]

This study asked three related questions. (1) To what extent does automated feedback improve creative writing among undergraduate language learners? (2) Does it improve scientific translation competence? (3) How does exposure to it influence the course of writing self-efficacy? We hypothesized that the experimental group, which received automated feedback, would improve significantly more on every outcome than the control group, which received conventional instructor feedback.

2. Methods

2.1 Study design and setting

The study used a quasi-experimental design with a non-equivalent control group. It took place during the 2024 academic year at a public university in Palembang, Indonesia.

2.2 Participants, sampling, and eligibility

The sample consisted of 136 Indonesian undergraduate students (male: n = 62, 45.6%; female: n = 74, 54.4%; M_{age} = 19.8 years, SD = 1.2) in fourth-semester Writing and Translation

courses of the Department of English Education. The experimental condition had 68 students (male: n = 30, 44.1%; female: n = 38, 55.9%), and the control condition had 68 (male: n = 32, 47.1%; female: n = 36, 52.9%). The conditions did not differ significantly in demographics ($\chi^2(1) = 0.12$, $p = 0.73$ for gender; $t(134) = 0.97$, $p = 0.34$ for age). All participants had Intermediate or Advanced English proficiency (TOEFL scores: M = 520, SD = 35, range: 475–607). Table 1 gives the full demographic characteristics and baseline equivalence statistics.

2.3 Variables, instruments, and procedures

Students in the experimental group used an integrated automated feedback platform (hereafter AFS) designed for English language learners. It gave feedback at three levels: surface-level errors (orthography, grammar, punctuation), sentence-level concerns (clarity, concision, variety), and discourse-level issues (organization, coherence, rhetorical effectiveness). Feedback appeared immediately after each writing assignment was submitted, and students could revise as often as they wished during the 10-week intervention.

The control group followed standard instruction. Instructors gave written comments, which were followed by discussion in class. Feedback usually arrived within 3–5 days of submission, and most assignments allowed a single revision. Both groups followed identical curricula in content coverage, assignment types, and overall course objectives, so the main differences lay in how feedback was delivered and how quickly. Weekly writing assignments of approximately 500–750 words were the main learning activity in both conditions, and students completed approximately 8–10 of them during the intervention.

Creative writing skills. An analytic rubric with five dimensions was used to assess creative writing skills: originality and imagination (4 points), narrative structure and coherence (4 points), character development and voice (4 points), stylistic variation and language use (4 points), and technical accuracy (4 points). Total scores ranged from 0 to 20 points. All writing samples were scored by two independent raters who were blind to condition. Inter-rater reliability (intraclass correlation coefficient) was adequate ($ICC_{2,1} = 0.78$ for pretest; $ICC_{2,1} = 0.82$ for posttest).

Scientific translation skills. Participants translated standardized English-to-Indonesian academic passages (approximately 300 words each) taken from peer-reviewed biology and chemistry articles, and these translations measured scientific translation proficiency. A weighted rubric rated terminology accuracy (35%), meaning preservation (30%), fluency and naturalness in Indonesian (20%), and adherence to disciplinary conventions (15%). Total scores ranged from 0 to 100 points. Three bilingual raters, all native Indonesian speakers with advanced English proficiency and translation experience, rated the translations independently. Inter-rater agreement was acceptable ($ICC_{2,3} = 0.81$ for pretest; $ICC_{2,3} = 0.84$ for posttest).

Writing self-efficacy. Writing self-efficacy was measured with a 10-item Writing Self-Efficacy Scale (WSES). The authors adapted it for translation contexts from the writing self-efficacy scale of Zimmerman and Bandura (1994).[19] Items were rated on 100-point scales (0 = "not confident at all" to 100 = "completely confident") and asked about the capability to perform various writing and translation tasks. Internal consistency, assessed with Cronbach's alpha, was adequate at pretest ($\alpha = 0.87$) and at posttest ($\alpha = 0.89$).

2.4 Statistical analysis

Descriptive statistics (means, standard deviations, ranges) were computed for all dependent variables. The assumptions for parametric testing were checked: homogeneity of variance (Levene's test), normality (Shapiro–Wilk test), and homogeneity of regression slopes for ANCOVA. All were satisfied ($p > 0.05$ for Levene's tests). Analysis of covariance (ANCOVA) was run separately for each dependent variable, with pretest scores as covariates to control baseline differences and raise statistical power. Cohen's d estimated the effect sizes, with 95% confidence intervals from bootstrapping (10,000 resamples). Significance was set at $\alpha = 0.05$ (two-tailed). Analyses used IBM SPSS Statistics version 27.0 and R version 4.3.1.

2.5 Ethical considerations

The Ethics Committee of Social Research, Enigma Institute, Indonesia approved the study (approval number: 2025/201), and the research followed the principles of the Declaration of Helsinki. All participants gave written informed consent before data collection began.

3. Results

3.1 Participant characteristics and baseline comparisons

Table 1 gives the demographic characteristics and baseline descriptive statistics. Before the intervention, the experimental and control groups performed equivalently on all dependent variables. Independent samples t-tests found no significant baseline differences in creative writing skills ($t(134) = 0.29$, $p = 0.77$), scientific translation skills ($t(134) = 0.17$, $p = 0.87$), or writing self-efficacy ($t(134) = 0.17$, $p = 0.86$).

Table 1. Demographic characteristics and baseline statistics for the experimental and control groups.

Variable	Experimental	Control	p
n	68	68	—
Age (M ± SD)	19.9 ± 1.1	19.7 ± 1.3	0.34
Gender (% female)	55.9%	52.9%	0.73
TOEFL score (M ± SD)	522 ± 34	519 ± 36	0.62
Pretest CWS (M ± SD)	8.76 ± 2.34	8.64 ± 2.41	0.77
Pretest STS (M ± SD)	62.14 ± 8.92	61.88 ± 9.24	0.87
Pretest WSE (M ± SD)	58.43 ± 10.18	58.12 ± 10.56	0.86

Note. CWS = creative writing skills (0–20); STS = scientific translation skills (0–100); WSE = writing self-efficacy (0–100); TOEFL = Test of English as a Foreign Language. p-values are from independent samples t-tests (continuous variables) or a χ^2 test (gender).

3.2 Primary outcomes

Table 2 presents the ANCOVA results, and Figure 1 shows the corresponding adjusted between-group means. With pretest performance controlled, the experimental group that received automated feedback improved significantly more on all three dependent variables. Figure 2 summarizes the between-group effect sizes.

Creative writing skills. In creative writing, the experimental group ($M_{adj} = 17.82$, $SE = 0.38$) significantly outperformed the control group ($M_{adj} = 14.26$, $SE = 0.38$) at posttest, $F(1, 133) = 68.44$, $p < 0.001$, with a large effect size ($d = 1.84$, 95% CI [1.38, 2.30]). Table 3 reports the raw pretest and posttest means.

Scientific translation skills. The largest treatment effect among the dependent variables was found for scientific

translation skills after the intervention. The experimental group (M_{adj} = 81.34, SE = 1.87) scored well above the control group (M_{adj} = 69.18, SE = 1.87), $F(1, 133) = 89.76$, $p < 0.001$, and the effect size was very large (d = 2.08, 95% CI [1.62, 2.54]).

Writing self-efficacy. Writing self-efficacy rose significantly more in the experimental group (M_{adj} = 81.25, SE = 2.14) than in controls (M_{adj} = 68.94, SE = 2.14), $F(1, 133) = 76.25$, $p < 0.001$, a large effect (d = 1.94, 95% CI [1.49, 2.39]).

Table 2. Analysis of covariance results: posttest performance by group (controlling for pretest scores).

Variable	Experimental	Control	F	p	d [95% CI]
CWS	17.82 (0.38)	14.26 (0.38)	68.44	< 0.001	1.84 [1.38, 2.30]
STS	81.34 (1.87)	69.18 (1.87)	89.76	< 0.001	2.08 [1.62, 2.54]
WSE	81.25 (2.14)	68.94 (2.14)	76.25	< 0.001	1.94 [1.49, 2.39]

Note. Values in parentheses are standard errors of adjusted means. Adjusted means control for pretest performance via ANCOVA; d = Cohen's d with 95% CI from bootstrap resampling (10,000 replicates). All comparisons p < 0.001. CWS, STS, and WSE as defined in Table 1.

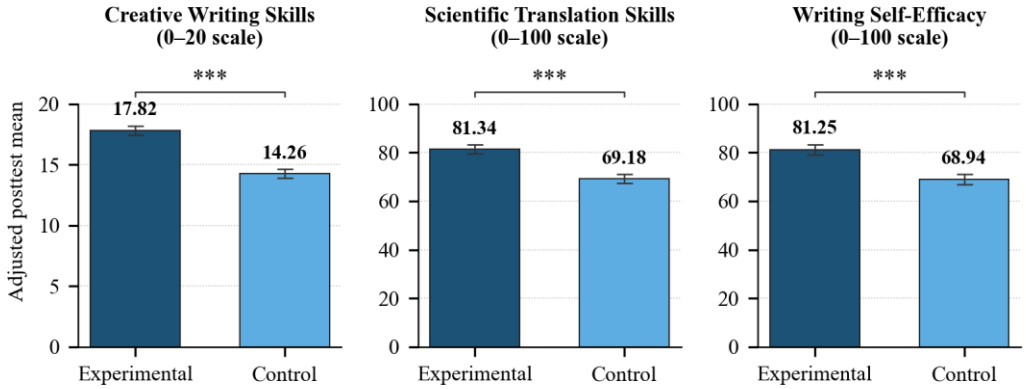


Figure 1. Adjusted posttest means (± SE) by group for the three outcome measures, controlling for pretest scores via ANCOVA. Each panel uses the measure's native scale. The experimental group exceeded the control group on every measure (***p < 0.001). Values correspond to Table 2.

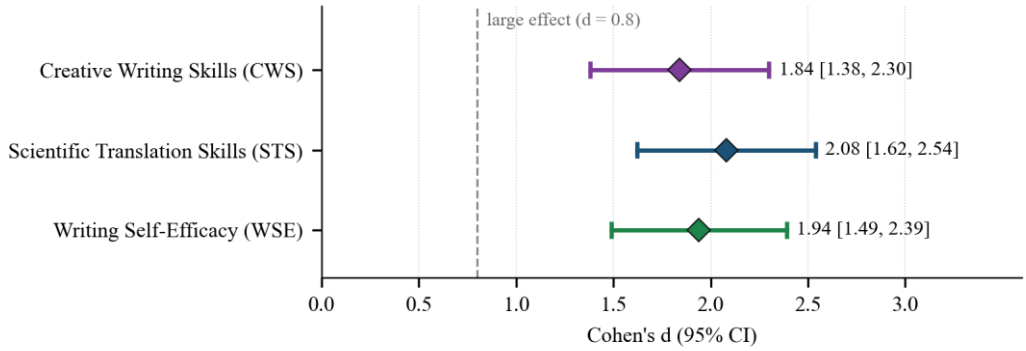


Figure 2. Forest plot of between-group effect sizes (Cohen's d) with 95% confidence intervals for the three outcomes. All estimates exceed 1.8 and none cross zero, indicating very large, statistically significant differences favoring the experimental group. The dashed line marks the large-effect threshold (d = 0.8). Values correspond to Table 2.

3.3 Within-group changes

Table 3 lists the within-group pre-to-post changes, and Figure 3 illustrates these trajectories. Within-group effect sizes are paired-samples d ($d_{within} = t/\sqrt{67}$). In the experimental group, paired samples t-tests showed significant improvement in creative writing skills ($t(67) = 12.84$, $p < 0.001$, $d_{within} = 1.56$), scientific translation skills ($t(67) = 14.29$, $p < 0.001$, $d_{within} = 1.73$), and writing self-efficacy ($t(67) = 13.47$, $p < 0.001$, $d_{within} = 1.64$). The control group also improved significantly, but by much less: creative writing ($t(67) = 4.12$, $p < 0.001$, $d_{within} = 0.50$), scientific translation ($t(67) = 5.31$, $p < 0.001$, $d_{within} = 0.65$), and writing self-efficacy ($t(67) = 4.88$, $p < 0.001$, $d_{within} = 0.59$).

Table 3. Within-group pretest to posttest changes for the experimental and control groups.

Variable	Group	Pretest	Posttest	t	Paired d
CWS	Experimental	8.76 (2.34)	17.89 (1.42)	12.84	1.56
	Control	8.64 (2.41)	12.34 (1.89)	4.12	0.50
STS	Experimental	62.14 (8.92)	82.67 (4.18)	14.29	1.73
	Control	61.88 (9.24)	71.42 (7.56)	5.31	0.65
WSE	Experimental	58.43 (10.18)	82.89 (6.12)	13.47	1.64
	Control	58.12 (10.56)	72.34 (8.94)	4.88	0.59

Note. Values in parentheses are standard deviations; pretest and posttest means are unadjusted raw values. d = paired-samples effect size ($t/\sqrt{67}$). All t values significant at p < 0.001. CWS, STS, and WSE as defined in Table 1.

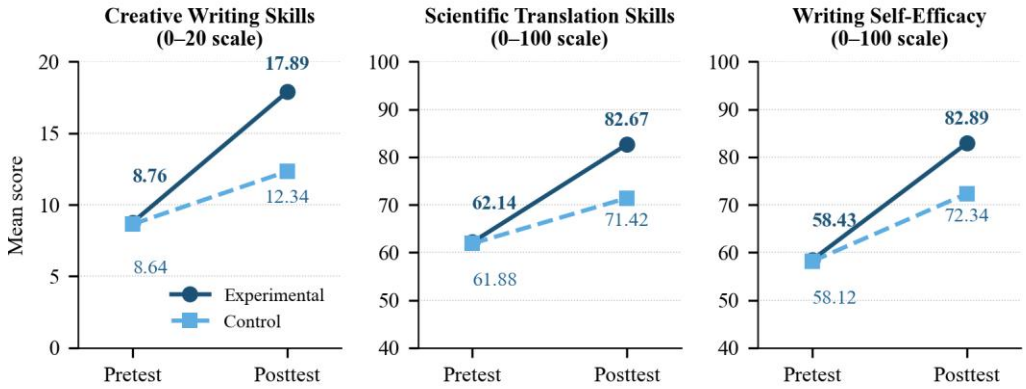


Figure 3. Within-group trajectories from pretest to posttest for the experimental (solid lines) and control (dashed lines) groups across the three outcome measures (CWS, 0–20; STS and WSE, 0–100). Points are unadjusted group means. The steeper experimental-group slopes illustrate the larger gains associated with automated feedback. Values correspond to Table 3.

4. Discussion

4.1 Principal findings

The results suggest that automated feedback is associated with improvement in several dimensions of undergraduate writing competence. Effects were large. Cohen's *d* exceeded 1.80 on every outcome, above the effect sizes usually reported for writing interventions,[20,21] although effects of this size in a single-site, non-randomized study call for replication. The findings add to earlier work on CALL systems and automated feedback. They suggest benefits for demanding writing tasks, not only for the correction of mechanical errors.[8,22]

Creative writing. Creative writing improved considerably in the experimental group (d = 1.84). This fits theories that give immediate feedback a central place in skill acquisition and metacognitive development.[6,23] The system returned corrective information at once. Students may therefore have spotted errors quickly and tested alternative stylistic choices, processes that information-processing models predict.[24,25]

The creative writing rubric weighted technical accuracy alongside higher-order concerns such as originality, character development, and narrative coherence. Gains were seen across these dimensions. This suggests that feedback targeting several linguistic and discourse-level features can support the growth of subtler writing abilities. The system applied the rubric criteria consistently, so learners may have gained a clearer view of quality standards and of what was expected of them.[26]

The within-group effect in creative writing was about 3.1 times larger in the experimental group than in the control group (paired-samples d_{within} = 1.56 vs. 0.50). Instructor feedback is interpretively rich. Even so, this gap suggests that it may not supply input frequent or consistent enough to drive large gains in complex writing domains.[27]

Scientific translation. The largest effect appeared for scientific translation (d = 2.08), which suggests that automated systems may be especially useful in this specialized domain. Translators must attend at once to terminology, meaning, and stylistic convention. Such demands may be met well by systematic feedback on several dimensions.[12]

The system drew on domain-specific terminology databases and translation evaluation algorithms. These may account for its edge over general writing feedback. Feedback on terminology and disciplinary convention could be especially useful to translation learners, since instructors find these dimensions hard to judge consistently in large cohorts.[12]

Translation showed a larger effect than creative writing (2.08 vs. 1.84). Automated systems may therefore suit structured, rule-governed tasks that call for technical precision. Creative writing depends on originality and subjective aesthetic judgment, which algorithms may evaluate less easily; this could explain the somewhat smaller, though still large, effect. The interpretation is tentative and should be tested directly in future comparative studies.

Writing self-efficacy. Writing self-efficacy rose significantly (d = 1.94). This matters theoretically. Self-efficacy is a proximal outcome that shapes persistence, engagement, and later writing efforts.[16] Theories of motivation and self-regulation hold that feedback on competence and experiences of success build confidence.[17,28,29]

Immediate feedback may have produced frequent experiences of success. Students received corrections quickly, revised, resubmitted, and saw their work improve. This cycle supports mastery experiences, which are theorized to be the strongest source of self-efficacy.[28] Unlimited revisions were also allowed in the experimental condition. Students could watch their competence grow step by step, which may have increased efficacy gains relative to the single revision permitted in the control condition.

Self-efficacy also rose in the control group (d_{within} = 0.59), but far less than in the experimental group (d_{within} = 1.64). Traditional instructor feedback thus appears to build some efficacy, whereas the intensity and consistency of automated feedback may have produced much stronger gains in confidence.

4.2 Interpretation and implications

Institutions weighing technology adoption may, on these findings, consider giving priority to automated feedback in writing-intensive courses. The large effects suggest that such investment could improve student writing and learner confidence, although replication in other settings is needed.

Design matters, however. Feedback should address several linguistic and discourse-level dimensions. Systems limited to mechanical errors or generic writing concerns may yield much smaller benefits than those that offer discipline-specific, higher-order feedback.[8] Faculty will also need training. Instructors must understand what a system can and cannot do. Only then can they combine it with their own guidance and with peer collaboration.

Combining automated and instructor feedback may also be worthwhile. Future studies could test hybrid models in which the system supplies rapid, frequent input and instructors concentrate on higher-order commentary and personal guidance, a division of labor that might improve both efficiency and effectiveness.

4.3 Strengths and limitations

The study has several limitations. First, it was conducted at a single institution with a fairly homogeneous sample in terms of English proficiency and academic background. Generalization to learners with lower proficiency or more varied experience should therefore be cautious. Second, the system was

designed specifically for English language learners. The findings may not carry over to off-the-shelf systems without such design. Third, the 10-week intervention gave adequate time for skill development, but longer follow-up would show whether the gains last. Fourth, baseline equivalence was established statistically, yet unmeasured selection factors or differences in motivation may have influenced outcomes. Quasi-experimental designs cannot rule out such threats, although the effect sizes far exceeded those typical of selection artifacts.[29] Fifth, no qualitative data were collected on how automated feedback influenced outcomes. Think-aloud protocols or stimulated-recall interviews in future work would shed light on the mechanisms. Sixth, the two conditions differed not only in the use of automation but also in feedback source, delay, and number of revision opportunities, so the effects cannot be attributed to automation alone, and the procedure by which participants entered each condition is not fully described.

Future research. Future research should test whether the benefits of automated feedback hold across other populations, languages, and educational settings. Longitudinal designs that follow writing development over several academic years would show whether the effects last, and comparisons of system designs would reveal which features matter most. Studies of interactions with individual differences such as cognitive ability, motivation, and writing anxiety could identify subgroups that benefit especially strongly or hardly at all.

5. Conclusion

In this quasi-experimental study, automated feedback was associated with significantly better creative writing, scientific translation, and writing self-efficacy among undergraduates. The effects were large (d = 1.84–2.08). They also stood well above those usually found in research on writing instruction. Automated feedback therefore looks like a promising innovation for language education. Adopting such systems in undergraduate writing curricula is worth considering, above all when the systems address several linguistic and discourse-level dimensions and when they are combined with instructor guidance and peer collaboration, although replication in other settings is needed. Writing and translation matter more and more in higher education. Used with care, technology may improve student learning and give instructors a faster, more efficient way to provide feedback.

Declaration

Artificial Intelligence Use Statement

No generative artificial intelligence tool was used to write or revise the text, analyze the data, create images, or generate any other scholarly content of this manuscript.

Author Contribution Statement

Conceptualization, D.A.S. and G.P.; methodology, D.A.S.; formal analysis, D.A.S.; investigation, D.A.S.; data curation, D.A.S.; writing – original draft, D.A.S.; writing – review and editing, D.A.S. and G.P.; supervision, G.P.; project administration, G.P. All authors read and approved the final manuscript.

Funding Statement

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Conflict of Interest / Competing Interests

The authors declare that they have no competing financial or non-financial interests related to this work.

Ethics Approval and Consent to Participate

The Ethics Committee of Social Research, Enigma Institute, Indonesia approved the study (approval number: 2025/201), and the research followed the principles of the Declaration of Helsinki. Written informed consent was obtained from all participants before data collection.

Consent for Publication

Not applicable. No identifiable personal information is reported.

Data Availability Statement

The data supporting the findings of this study are available from the corresponding author on reasonable request.

Acknowledgments

None.

References

1. Warschauer M, Ware P. Automated writing evaluation: defining the classroom research agenda. Lang Teach Res. 2006;10(2):157-80. doi:10.1191/1362168806lr190oa
2. Chapelle CA. English language learning and technology: lectures on applied linguistics in the age of information and communication technology. Amsterdam: John Benjamins; 2003. doi:10.1075/lllt.7
3. Graham S. Changing how writing is taught. Rev Res Educ. 2019;43(1):277-303. doi:10.3102/0091732x18821125
4. Kellogg RT. Training writing skills: a cognitive developmental perspective. J Writ Res. 2008;1(1):1-26. doi:10.17239/jowr-2008.01.01.1
5. Ellis R. Corrective feedback and teacher development. L2 J. 2009;1(1):3-18. doi:10.5070/l2.v1i1.9054
6. Nicol D. From monologue to dialogue: improving written feedback processes in mass higher education. Assess Eval High Educ. 2010;35(5):501-17. doi:10.1080/02602931003786559
7. Shermis MD, Burstein JC, editors. Automated essay scoring: a cross-disciplinary perspective. Mahwah (NJ): Lawrence Erlbaum Associates; 2003. doi:10.4324/9781410606860
8. Stevenson M, Phakiti A. The effects of computer-generated feedback on the quality of writing. Assess Writ. 2014;19:51-65. doi:10.1016/j.asw.2013.11.007
9. Heift T, Schulze M. Errors and intelligence in computer-assisted language learning: parsers and pedagogues. New York: Routledge; 2007. doi:10.4324/9780203012215
10. Nagata N. Robo-Sensei's NLP-based error detection and feedback generation. CALICO J. 2009;26(3):562-79. doi:10.1558/cj.v26i3.562-579
11. Chapelle CA, Cotos E, Lee J. Validity arguments for diagnostic assessment using automated writing evaluation. Lang Test. 2015;32(3):385-405. doi:10.1177/0265532214565386
12. Bowker L, Corpas Pastor G. Translation technology. In: Mitkov R, editor. The Oxford handbook of computational linguistics. 2nd ed. Oxford: Oxford University Press; 2015. doi:10.1093/oxfordhb/9780199573691.013.007
13. Morley D. The Cambridge introduction to creative writing. Cambridge: Cambridge University Press; 2007. doi:10.1017/cbo9780511803024
14. Flower L, Hayes JR. A cognitive process theory of writing. Coll Compos Commun. 1981;32(4):365-87. doi:10.58680/ccc198115885
15. Bereiter C, Scardamalia M. The psychology of written composition. Hillsdale (NJ): Lawrence Erlbaum Associates; 1987. doi:10.4324/9780203812310
16. Pajares F. Self-efficacy beliefs, motivation, and achievement in writing: a review of the literature. Read Writ Q. 2003;19(2):139-58. doi:10.1080/10573560308222
17. Zimmerman BJ. Becoming a self-regulated learner: an overview. Theory Pract. 2002;41(2):64-70. doi:10.1207/s15430421tip4102_2
18. Sari E, Han T. The impact of automated writing evaluation on English as a foreign language learners' writing self-efficacy, self-regulation, anxiety, and performance. J Comput Assist Learn. 2024;40(5):2065-80. doi:10.1111/jcal.13004
19. Zimmerman BJ, Bandura A. Impact of self-regulatory influences on writing course attainment. Am Educ Res J. 1994;31(4):845-62. doi:10.3102/00028312031004845
20. Lipsey MW, Wilson DB. Practical meta-analysis. Thousand Oaks (CA): SAGE Publications; 2001.

21. Graham S, McKeown D, Kiuhara S, et al. A meta-analysis of writing instruction for students in the elementary grades. J Educ Psychol. 2012;104(4):879-96. doi:10.1037/a0029939
22. Shermis MD. State-of-the-art automated essay scoring: competition, results, and future directions from a United States demonstration. Assess Writ. 2014;20:53-76. doi:10.1016/j.asw.2013.04.001
23. Schunk DH. Self-efficacy for reading and writing: influence of modeling, goal setting, and self-evaluation. Read Writ Q. 2003;19(2):159-72. doi:10.1080/10573560308219
24. Atkinson RC, Shiffrin RM. Human memory: a proposed system and its control processes. Psychol Learn Motiv. 1968;2:89-195. doi:10.1016/s0079-7421(08)60422-3
25. Baddeley A. Working memory: looking back and looking forward. Nat Rev Neurosci. 2003;4(10):829-39. doi:10.1038/nrn1201
26. Panadero E, Jonsson A. The use of scoring rubrics for formative assessment purposes revisited: a review. Educ Res Rev. 2013;9:129-44. doi:10.1016/j.edurev.2013.01.002
27. Ferris DR. Does error feedback help student writers? New evidence on the short- and long-term effects of written error correction. In: Hyland K, Hyland F, editors. Feedback in second language writing: contexts and issues. Cambridge: Cambridge University Press; 2006. p. 81-104. doi:10.1017/cbo9781139524742.007
28. Bandura A. Self-efficacy: the exercise of control. New York: WH Freeman; 1997.
29. Shadish WR, Cook TD, Campbell DT. Experimental and quasi-experimental designs for generalized causal inference. Boston: Houghton Mifflin; 2002.