English Language Teaching Educational Journal
Vol. 9, No. 1, April 2026, pp. 57-68
ISSN 2621-6485

Automated writing evaluation tools in Business English Writing: An experimental study among Indonesian undergraduates

Binar Winantaka [a1], Anwar Efendi [a2], Nur Hidayanto Pancoro Setyo Putro [a3]

[a] Universitas Negeri Yogyakarta, Jl. Colombo No.1, Karang Malang, Caturtunggal, Yogyakarta, 55281, Indonesia
[1] binar@uny.ac.id , [2] anwar@uny.ac.id, [3] nur_hidayanto@uny.ac.id

*Corresponding author

ARTICLE INFO

Article history
Received 19 September 2025
Revised March 2026
Accepted 02 April 2026
Available Online 30 April 2026

Keywords
Automated writing evaluation
Business English writing
Indonesian EFL students
Self-regulated learning
Writing self-efficacy

ABSTRACT

Business English writing competence is increasingly critical for Indonesian university graduates navigating global professional environments, yet many students continue to experience persistent difficulties with grammar, coherence, and professional register. Automated Writing Evaluation (AWE) tools such as Grammarly and Grammark provide immediate, data-driven feedback that may support writing development beyond surface-level error correction (Barrot, 2021; Calma et al., 2022). This quasi-experimental study examined whether AWE integration in a Business English writing course improved Indonesian undergraduates' perceived usefulness, ease of use, motivation, self-regulated learning (SRL), and writing self-efficacy (WSE), relative to conventional instructor feedback. Sixty undergraduate students were randomly assigned to an AWE group (n = 30) or a control group (n = 30). A validated 38-item Likert questionnaire and open-ended reflective prompts were administered post-intervention. Three key findings emerged: (1) AWE significantly enhanced perceived usefulness (d = 2.92) and ease of use (d = 2.74); (2) motivation, self-regulated learning, and writing self-efficacy increased substantially in the AWE group; and (3) students unanimously recommended a hybrid AWE-teacher feedback model to address AWE's limitations in higher-order writing development (Thi & Nikolov, 2022). These findings suggest that AWE integration holds considerable promise for Indonesian EFL Business English instruction, particularly when systematically complemented by expert human feedback.

How to Cite: Winantaka, B., Efendi, A., & Putro, N. H. P. S. (2026). Automated writing evaluation tools in business English writing: An experimental study among Indonesian undergraduate. *English Language Teaching Educational Journal*, *9*(1), 57-68. https://doi.org/10.12928/eltej.v9i1.15011

1. Introduction

Business English writing is an indispensable skill for contemporary university graduates, particularly those aspiring to careers in global industries, international commerce, and digitally mediated professional environments (Bozdoğan & Kasap, 2019; Nickerson, 2005). In Indonesia, institutions of higher education have increasingly emphasized professional writing competencies in alignment with national economic development agendas and growing engagement in international trade and digital communication. Despite curricular reforms and expanded English language exposure, many Indonesian students continue to face significant difficulties in producing writing that is not only grammatically accurate but also cohesive, rhetorically appropriate, and fit for business purposes (Adi, 2024; Toba et al., 2019).

Empirical research has documented the scope of these challenges. Indonesian EFL students exhibit notable linguistic difficulties—particularly in grammar (45.16%)—alongside affective challenges such as writing anxiety (38.71%) (Hamdani, 2025). Adi (2024) found that common business English writing

errors among Indonesian undergraduates include grammatical inaccuracies (48%), difficulties in conveying business messages (45%), and structural deficiencies in text organization (38%). Ningrum (2023) further documented that inadequate metalinguistic awareness and limited exposure to authentic business genres compound these difficulties even at the postgraduate level.

The integration of AWE tools—such as Grammarly, Grammark, and MI Write—into writing instruction has emerged as a promising pedagogical response (Bryant et al., 2023; Warschauer & Ware, 2006). These tools employ natural language processing and machine learning to deliver immediate, actionable feedback on grammar, mechanics, vocabulary, and stylistic concerns (Al-Inbari & Al-Wasy, 2023; Koltovskaia, 2020; Barrot, 2021). Meta-analytic evidence indicates that AWE exerts large positive effects on writing quality, with particular advantages for EFL learners relative to native speakers (Zhai & Ma, 2023; Ngo et al., 2022; Wilson & Roscoe, 2020). Notably, Calma et al. (2022) demonstrated that Grammarly can serve as a viable instructional intervention in management education specifically, yielding writing improvements on grammar, mechanics, and stylistic clarity—findings of direct relevance to Business English writing courses. However, research consistently demonstrates that AWE performs substantially better on surface-level features than on higher-order concerns such as organizational coherence, argumentation, content development, and genre appropriateness (Ranalli & Link, 2020; Lo et al., 2025; Ware, 2011; Thi & Nikolov, 2022).

Despite the growing body of international AWE research, significant empirical gaps remain. Most existing studies focus primarily on writing quality as an outcome, with fewer examining AWE's influence on motivational, self-regulatory, and self-efficacy variables—constructs central to sustained writing development (Koltovskaia, 2020; Stevenson & Phakiti, 2014; Zhang & Hyland, 2018; Sari, 2024). Studies conducted specifically within Indonesian Business English contexts remain particularly scarce, leaving practitioners without adequate evidence to guide technology-integration decisions.

The novelty of the present study lies in its examination of AWE's multidimensional psychological and behavioral impact within a Business English writing course at an Indonesian undergraduate institution**.** Unlike prior studies that measure writing quality or technology acceptance in isolation, this investigation employs a comprehensive framework simultaneously assessing perceived usefulness, ease of use, motivation, self-regulated learning, and writing self-efficacy—five constructs that together reflect both technology acceptance and learner psychology. To the best of the authors' knowledge, this represents among the first quasi-experimental studies to systematically examine this full spectrum of outcomes in Indonesian Business English writing instruction, thereby contributing original empirical evidence to an underrepresented area of EFL technology research.

The theoretical foundation of this study draws on two complementary frameworks. The Technology Acceptance Model (TAM), developed by Davi's (1989), proposes that technology adoption is primarily driven by perceived usefulness (PU) and perceived ease of use (PEU)—constructs that predict behavioral intention and actual use. TAM has been extensively validated across educational technology contexts and provides a theoretically grounded basis for examining AWE adoption (Asiyah, 2018; Davis, 1989; Chen & Cheng, 2008). Bandura's (1977) social-cognitive theory of self-efficacy posits that individuals' beliefs in their capability to execute specific tasks predict their performance, motivation, and persistence. Writing self-efficacy—confidence in one's ability to produce effective written communication—is an established predictor of writing achievement and engagement (Artino, 2012; Dewanti, 2018; Tsao, 2021; Zimmerman, 2000). Self-regulated learning (SRL), encompassing the metacognitive, motivational, and behavioral processes through which learners independently manage their learning, demonstrates significant positive associations with writing performance and can be cultivated through strategic feedback and tool use (Nugraeni, 2025; Ariyanti, 2018; Pintrich, 2004; Nicol & Macfarlane-Dick, 2006; Rahimi et al., 2025).

The research questions guiding this study are: (1) Does AWE integration significantly improve Indonesian university students' perceived usefulness and ease of use compared to conventional instructor feedback? (2) Does AWE use enhance students' motivation, SRL, and writing self-efficacy? (3) What are students' intentions to continue using AWE beyond the course? (4) What qualitative experiences do students report regarding perceived benefits and limitations of AWE?

2. Method

2.1. Participants

Participants comprised 60 undergraduate students (34 females, 26 males) enrolled in a compulsory Business English writing course at a university in Central Java, Indonesia. All were EFL learners with English proficiency ranging from intermediate to upper-intermediate levels (CEFR B1–C1), as determined by course placement tests and self-reported language background data. Participant ages ranged from 19 to 22 years (M = 19.7, SD = 0.8). Participants were randomly assigned to either an experimental group (n = 30) or a control group (n = 30). Baseline comparability was confirmed across gender distribution, age (experimental: M = 19.8, SD = 0.9; control: M = 19.6, SD = 0.7), and self-reported proficiency levels, with no significant between-group differences on any baseline measure (all $p > .05$). All participants provided written informed consent prior to enrollment. The study received approval from the institutional review board at the participating university. Participation was voluntary, and students were informed of their right to withdraw at any time without penalty.

2.2. Research Design

A quasi-experimental, non-equivalent group comparison design was employed. The experimental group received direct instruction in, and consistent use of, AWE tools—specifically Grammarly and Grammark—throughout all drafting and revision tasks. The control group received written feedback provided exclusively by their instructors, reflecting established pedagogical practice (Warschauer & Ware, 2006). Both groups completed identical writing tasks: business emails, memos, and formal proposals designed to mirror authentic professional communication contexts (Bozdoğan & Kasap, 2019). Assignment prompts, due dates, grading criteria, and instructional content were standardized across conditions to ensure equivalence.

2.3. Procedure

The four-week intervention was embedded within the regular Business English writing course. At baseline (Week 0), all students completed a demographic questionnaire and a pre-intervention survey assessing prior attitudes toward writing technology and digital tool experience. During Week 1, the experimental group participated in a 60-minute structured training session on Grammarly and Grammark. The session included demonstrations of tool interfaces, feature explanations, guidance on interpreting error feedback, and supervised practice with sample business texts. Students were explicitly instructed to engage critically with AWE suggestions and to exercise independent judgment in accepting or rejecting feedback (Koltovskaia, 2020). The control group received no AWE training.

Over Weeks 1–4, both groups completed four writing cycles, each involving drafting, revision, and final submission of a business communication task. Experimental group participants submitted drafts through AWE platforms, reviewed automated feedback, and incorporated selected suggestions into revisions. Instructors for this group remained available for procedural queries but refrained from providing written corrective feedback on student drafts. Control group participants received comprehensive written instructor feedback on grammar, mechanics, organization, and content, consistent with standard pedagogical practice (Warschauer & Ware, 2006). At the end of Week 4, all students completed a post-intervention questionnaire assessing six primary constructs; experimental group participants additionally responded to three open-ended reflective questions. Total data collection time was approximately 30 minutes per participant.

2.4. Instruments

a) Questionnaire development and validation

A 38-item, five-point Likert-type questionnaire was developed drawing on established frameworks: the Technology Acceptance Model (Davis, 1989), self-efficacy theory (Bandura, 1977), and prior AWE research instruments (Palermo, 2020; Ranalli & Link, 2020; Guo et al., 2022). Instrument development followed a four-stage process: (1) review of relevant construct definitions and prior validated scales; (2) item generation guided by construct specifications; (3) expert review by three specialists in EFL writing instruction and educational measurement, who evaluated content validity, item clarity, and cultural-linguistic appropriateness; and (4) item revision based on expert feedback. This process aligns with recommended practices for instrument development in EFL research (Palermo, 2020). The questionnaire measured six constructs as follows.

Perceived Usefulness (PU): Five items assessed students' perceptions of AWE's contribution to error identification, writing accuracy, overall quality, and writing development (e.g., "Using AWE helped me identify errors I would have otherwise missed"). This construct operationalizes TAM's foundational dimension of technology utility (Davis, 1989).

Perceived Ease of Use (PEU): Five items gauged ease of interacting with AWE platforms, including interface navigation, learning curve, and feedback application (e.g., "It was easy to use AWE tools in my writing process"). PEU is a core TAM predictor of technology adoption (Davis, 1989; Chen & Cheng, 2008).

Motivation and Confidence (MC): Five items examined whether AWE increased writing motivation, confidence, and revision engagement (e.g., "Using AWE increased my confidence in my writing") (Artino, 2012; Zimmerman, 2000).

Self-Regulated Learning (SRL): Ten items assessed goal-setting, revision planning, progress monitoring, resource-seeking, self-evaluation, time management, and reflective learning behaviors (e.g., "I set clear writing goals for myself and work toward achieving them") (Nugraeni, 2025; Ariyanti, 2018; Pintrich, 2004; Rahimi et al., 2025).

Writing Self-Efficacy (WSE): Ten items captured confidence in composing diverse business documents—emails, memos, reports, and proposals—with two reverse-coded items included to control for acquiescence bias (e.g., "I am confident I can write a professional business email") (Artino, 2012; Dewanti, 2018; Tsao, 2021).

Future Intention to Use AWE (FI): Three items, administered exclusively to the experimental group, assessed intention to continue AWE use, integrate AWE into regular practice, and recommend AWE to peers (Davis, 1989).

Reliability analyses (Cronbach's α) indicated acceptable internal consistency: PU (.75), PEU (.69), MC (.71), SRL (.68), WSE (.64). The moderate values for SRL and WSE indicate these constructs warrant psychometric refinement in future research; items may not capture their respective constructs with maximum precision in this non-Western EFL context (Palermo, 2020). However, coefficients fall within acceptable ranges for newly developed instruments applied in underrepresented populations. Researchers are encouraged to pilot and factor-analyze revised versions in future applications.

b) Qualitative measures

Experimental group participants responded to three open-ended prompts: (1) "What aspects of AWE tools did you find most helpful in your writing process?"; (2) "What challenges or limitations did you encounter when using AWE?"; and (3) "What recommendations would you make to improve AWE use in Business English writing courses?" Responses averaged 150–300 words per participant. Notably, only 10 of the 30 experimental participants (33%) provided qualitative responses; the remaining 20 either were absent during the post-session or declined participation. This response rate constitutes a notable methodological limitation: qualitative findings should be interpreted as exploratory and illustrative rather than broadly representative (Lo et al., 2025).

2.5 Data Analysis

a) Quantitative analysis

Descriptive statistics (means, standard deviations) were computed for each construct by group. Data normality was assessed using Shapiro-Wilk tests; distributions were approximately normal, supporting the use of parametric procedures. Independent-samples *t*-tests with Welch's correction for unequal variances were conducted for between-group comparisons. The significance level was set at $\alpha = .05$. Effect size was estimated using Cohen's *d*, interpreted as small ($d = 0.20$), medium ($d = 0.50$), and large ($d = 0.80$) (Wilson & Roscoe, 2020).

b) Qualitative analysis

Open-ended responses were analyzed using inductive thematic analysis following the six-phase framework of Braun and Clarke (2006): (1) transcription and data familiarization; (2) initial line-by-line coding; (3) preliminary theme generation; (4) theme refinement and review against the full dataset; (5) theme definition and naming; and (6) selection of representative quotes for reporting. Two independent coders analyzed all responses; disagreements were resolved through discussion and consensus. No

predefined coding schemes were imposed, allowing thematic patterns to emerge organically from participants' responses.

3. Findings and Discussion

3.1 Quantitative Findings

Table 1 presents descriptive statistics and between-group comparisons for each measured construct.

Table 1. Between-Group Comparisons of Learning Constructs (Post-Intervention)

Construct	Experimental (M ± SD)	Control (M ± SD)	t	p	d
Perceived Usefulness	22.13 ± 1.33	17.97 ± 1.52	11.29	<.001	2.92
Perceived Ease of Use	21.10 ± 1.36	17.33 ± 1.51	10.62	<.001	2.74
Motivation & Confidence	20.43 ± 1.27	18.73 ± 1.45	4.80	<.001	1.24
Self-Regulated Learning	39.80 ± 1.67	38.40 ± 1.77	3.15	.003	0.81
Writing Self-Efficacy	36.97 ± 1.63	35.77 ± 1.72	2.78	.007	0.72

All between-group differences reached statistical significance ($p < .01$). Perceived usefulness and perceived ease of use demonstrated very large effect sizes ($d = 2.92$ and 2.74, respectively), indicating substantial practical significance. Motivation and confidence yielded a large effect size ($d = 1.24$). SRL and WSE demonstrated medium to large effects ($d = 0.81$ and 0.72, respectively). These effect sizes are consistent with—and in several cases exceed—those reported in international meta-analyses of AWE research (Zhai & Ma, 2023; Ngo et al., 2022; Wilson & Roscoe, 2020). The gap between groups was particularly pronounced for perceived usefulness (difference of 4.16 points) and ease of use (difference of 3.77 points), suggesting that AWE exposure substantially shaped students' perceptions of the tools' practical value.

3.2 Future Intention to Continue Using AWE

The Future Intention scale, administered exclusively to the experimental group, yielded a composite mean of $M = 12.70$ (of 15; $SD = 1.15$), corresponding to a per-item mean of 4.23 on the five-point scale. This result may indicate strong behavioral intention to continue AWE use beyond the current course—a theoretically important outcome predicting sustained technology adoption within the TAM framework (Davis, 1989; Asiyah, 2018).

3.3 Qualitative Findings

Thematic analysis of open-ended responses (n = 10; 33% of the experimental group) identified three major themes reflecting students' lived experiences with AWE, consistent with patterns documented in prior AWE research (Palermo, 2020; Koltovskaia, 2020; Ranalli & Link, 2020; Guo et al., 2022).

Theme 1: Surface-Level Support and Vocabulary Enhancement. Students consistently reported high value in AWE's rapid detection and correction of surface-level errors—grammar, spelling, and punctuation. Repeated exposure to automated feedback was described as heightening metacognitive awareness of personal error patterns (Barrot, 2021; Guo et al., 2022). Vocabulary enhancement emerged as a secondary benefit, with students appreciating synonym suggestions and enriched word choices appropriate for professional registers (Al-Inbari & Al-Wasy, 2023; Calma et al., 2022). Tone adjustment features were also positively noted for calibrating formality in business communication.

> *"AWE helps me see the mistakes I always make and suggests better words for my business emails."*
>
> *"Grammarly showed me my error patterns. Now I am more aware of problems with tenses and articles."*

Theme 2: Limitations in Higher-Order Feedback. Students frequently identified AWE's limited capacity to support higher-order writing concerns. AWE was reported to offer little guidance on text organization, paragraph structure, logical argumentation, and business genre conventions. Suggestions

were occasionally described as contextually inappropriate in domain-specific business writing contexts (Lo et al., 2025; Thi & Nikolov, 2022). These findings are consistent with meta-analytic evidence indicating that AWE performs poorly on content relevance and organizational coherence (Ranalli & Link, 2020; Wilson & Roscoe, 2020).

> *"Grammarly does not help me organize my ideas or assess whether my arguments are strong. Sometimes the advice does not fit the business context."*

Theme 3: Recommendations for Hybrid Integration. Across nearly all responses, students recommended that AWE function as a complement to—rather than a replacement for—human teacher feedback. Students requested context-specific corrections, business genre templates, and bilingual error explanations. The prevailing recommendation for a hybrid model combining automated surface-level corrections with expert pedagogical guidance is aligned with emerging empirical evidence demonstrating hybrid feedback's superiority over either approach alone (Lo et al., 2025; Thi & Nikolov, 2022; Shang, 2019; Hattie & Timperley, 2007).

> *"It would be best if the teacher explained grammar rules while AWE corrected the text. That way, we not only learn what went wrong, but why."*

4. Discussion

The very large effect sizes for perceived usefulness (d = 2.92) and perceived ease of use (d = 2.74) suggest that students who used AWE tools perceived considerable practical value and found the platforms readily navigable. These findings are consistent with—though they exceed—effect sizes reported in prior international AWE studies (Zhai & Ma, 2023; Geng & Razali, 2022; Ngo et al., 2022), and may partly reflect the novelty of AWE exposure for this sample, as many participants likely had limited prior experience with automated feedback tools. This interpretation warrants caution; future research should examine whether effect magnitudes stabilize with extended familiarity (Koltovskaia, 2020).

4.1 Interpretation of Quantitative Findings

The large effect for motivation and confidence (d = 1.24) is particularly noteworthy. From a social-cognitive perspective, AWE's immediate, concrete feedback may provide learners with tangible mastery experiences—direct evidence of error identification and correction—that may strengthen self-efficacy beliefs and sustain motivational engagement (Bandura, 1977). Bandura (1977) identified mastery experiences as the most potent source of self-efficacy, and the iterative AWE-assisted revision process likely afforded students repeated such experiences across four writing cycles. This aligns with Dewanti's (2018) finding that self-efficacy significantly mediates writing performance among Indonesian EFL students and with Zimmerman's (2000) characterization of self-efficacy as an essential motivator of learning. Sari (2024) similarly found that AWE use significantly enhanced writing self-efficacy and reduced writing anxiety among EFL learners—further corroborating the motivational benefits observed in the present study.

The medium to large effects for SRL (d = 0.81) and WSE (d = 0.72), while smaller than those for TAM constructs, nonetheless indicate substantively meaningful differences. AWE's capacity for immediate feedback may have encouraged greater independent editing engagement, thereby fostering metacognitive awareness and self-regulatory behaviors (Nugraeni, 2025; Ariyanti, 2018; Pintrich, 2004; Rahimi et al., 2025). Rahimi et al. (2025), in an activity-theory-based study, demonstrated that AWCF through Grammarly promoted learner agency and self-directed revisions—an outcome conceptually aligned with the SRL improvements observed in the present study. The high future intention score (M = 12.70 of 15) further suggests sufficient perceived value to motivate sustained technology use—a key TAM outcome (Davis, 1989). However, effect sizes should be interpreted with appropriate caution pending replication with objective writing outcome measures, given the study's exclusive reliance on self-report data.

Tsao (2021) found that EFL learners' writing self-efficacy significantly predicted their engagement with written corrective feedback—suggesting that the WSE improvements observed in the present study may have downstream effects on how students process and utilize AWE feedback. This bidirectional relationship between self-efficacy and feedback engagement warrants further investigation in future AWE research

4.2 Qualitative Findings in Context

The qualitative findings corroborate and extend the quantitative results. Students' appreciation for AWE's surface-level feedback aligns with the high PU scores and with established literature documenting AWE's effectiveness for grammar and mechanics correction (Bryant et al., 2023; Barrot, 2021; Guo et al., 2022; Al-Inbari & Al-Wasy, 2023). Guo et al. (2022) found that EFL students could effectively use AWCF to reduce writing errors in research writing, reporting high rates of error correction and increased learner autonomy in revision—consistent with the surface-level support theme identified in the present study. Simultaneously, the identified limitations in higher-order feedback support are consistent with meta-analytic evidence demonstrating AWE's relative ineffectiveness for content development, argumentative coherence, and genre organization (Ranalli & Link, 2020; Wilson & Roscoe, 2020; Thi & Nikolov, 2022). That students spontaneously articulated these limitations suggests an emerging degree of critical technology literacy—an educationally important outcome in its own right (Koltovskaia, 2020).

The convergent recommendation for a hybrid AWE–teacher feedback model is strongly supported by recent empirical research. Lo et al. (2025) demonstrated that combining AI-generated and teacher feedback produced superior outcomes in terms of motivation, writing quality, and academic performance. Thi and Nikolov (2022), examining teacher and Grammarly feedback in Myanmar EFL students' writing, found that each modality served complementary functions: Grammarly effectively addressed grammatical and mechanical errors, while teacher feedback was indispensable for genre-level and content concerns. Similarly, Shang (2019) demonstrated that combining automated corrective feedback with peer feedback enhanced EFL writing performance beyond what either method achieved alone. Li et al. (2023) further showed that AWE integrated with peer review in large-sized university writing classes yielded meaningful revisions and supported writing development at scale. Hattie and Timperley's (2007) conceptualization of effective feedback as closing the gap between current and desired performance provides additional theoretical grounding for the complementary strengths of automated and human feedback.

4.3 Theoretical Implications

From a TAM perspective, robust PU and PEU scores indicate strong technology acceptance among Indonesian undergraduates, likely attributable to the transparent, objective nature of AWE feedback and the user-friendly interfaces of contemporary AWE systems (Davis, 1989; Asiyah, 2018; Chen & Cheng, 2008). However, TAM's predictive framework may be insufficient for capturing the full complexity of AWE adoption in higher-order writing tasks, where perceived utility and ease alone cannot address genre-specific, rhetorical, and content-related concerns. Future theoretical models may benefit from incorporating writing-specific constructs—such as genre awareness, rhetorical competence, and critical tool literacy—into expanded TAM frameworks for writing technology contexts.

From a self-efficacy perspective, the observed WSE improvements may reflect the cumulative effect of mastery experiences through iterative error identification and revision (Bandura, 1977; Artino, 2012; Tsao, 2021). Sari (2024) similarly identified AWE-mediated self-efficacy gains in EFL writing contexts, suggesting this pattern may generalize across diverse EFL populations. However, the study's brief four-week intervention and reliance on perceptual measures limit conclusions about long-term efficacy development or skill transfer to novel writing contexts. Longitudinal research incorporating objective writing quality assessments would clarify whether AWE-enhanced self-efficacy translates to demonstrable, sustained improvement in authentic business writing performance.

4.4 Theoretical Implications

The findings support implementation of a hybrid instructional model in which AWE tools systematically complement, rather than replace, teacher feedback (Lo et al., 2025; Thi & Nikolov, 2022). Under this model, AWE manages surface-level error correction—grammar, spelling, punctuation, and basic clarity—providing students with rapid, objective feedback at tasks automated systems handle effectively (Bryant et al., 2023; Barrot, 2021). Teachers, relieved of routine error-correction workload, concentrate on higher-order feedback addressing content development, organizational coherence, argumentation quality, and professional genre conventions—domains requiring human expertise and contextual judgment (Warschauer & Ware, 2006; Hattie & Timperley, 2007). In large-enrollment settings, Li et al. (2023) demonstrate that this hybrid approach is practically viable even in classes with limited instructor capacity—a particularly salient consideration for Indonesian universities where class sizes are often large. Nicol and Macfarlane-Dick (2006) further emphasize that effective feedback should

develop students' capacity for self-assessment; integrating AWE with guided teacher feedback may support precisely this goal.

Calma et al. (2022), studying Grammarly's effectiveness in management education specifically, found that the tool improved writing mechanics and increased student engagement with revision—providing direct precedent for its adoption in Business English writing courses. Institutional adoption should be accompanied by structured AWE literacy training guiding students to critically evaluate, selectively apply, and reflectively engage with automated feedback rather than accept suggestions uncritically (Koltovskaia, 2020; Stevenson & Phakiti, 2014). AWE's accessibility outside classroom hours may additionally support independent self-regulatory development (Nugraeni, 2025; Rahimi et al., 2025).

4.5 Theoretical Implications

Several limitations warrant appropriate caution in interpreting and generalizing these findings. Sample and context**:** The sample of 60 students from a single Central Java institution limits generalizability; replication across diverse Indonesian regions and institutional types is necessary (Zhai & Ma, 2023; Ngo et al., 2022). Intervention duration: The four-week period represents brief AWE exposure; longitudinal designs spanning a full semester or academic year would clarify sustainability of effects (Palermo, 2020; Sari, 2024). Outcome measurement: The study's exclusive reliance on self-report measures means objective writing quality improvement cannot be verified; future studies should include pre- and post-intervention writing samples assessed with validated rubrics such as IELTS Writing Descriptors or locally developed scales (Wilson & Roscoe, 2020). Instrument reliability: Moderate Cronbach's α values for SRL (.68) and WSE (.64) indicate construct measurement requires refinement; future research should pilot revised instruments and conduct factor analyses before large-scale deployment (Palermo, 2020). Qualitative sample: Only 10 of 30 experimental participants provided qualitative responses, limiting depth and representativeness of thematic analysis; future studies should consider structured interviews with a purposive subsample (Lo et al., 2025). Individual differences: Despite random assignment, unmeasured variables—including technology comfort, intrinsic motivation, prior writing experience, and writing anxiety—may have influenced outcomes (Sari, 2024).

5. Conclusion

This quasi-experimental study provides evidence that Indonesian undergraduates who integrated AWE tools into their Business English writing processes reported significantly higher perceived usefulness, ease of use, motivation, self-regulated learning, and writing self-efficacy relative to peers receiving conventional instructor feedback. Strong future intention to continue AWE use further underscores the practical value students attributed to these tools. Effect sizes were statistically and practically meaningful, exceeding prior meta-analytic benchmarks for several constructs (Zhai & Ma, 2023; Ngo et al., 2022). The study's primary contribution lies in providing multidimensional empirical evidence of AWE's impact within the Indonesian Business English writing context—an area previously underrepresented in EFL technology literature—extending TAM and self-efficacy frameworks to a non-Western business writing setting and corroborating the growing body of evidence supporting AWE in management and Business English education (Calma et al., 2022; Sari, 2024).

Simultaneously, the findings highlight AWE's inherent limitations in addressing higher-order writing concerns. Students' own identification of these limitations—unprompted—reflects a degree of critical technology literacy, and their near-universal recommendation for a hybrid AWE–teacher feedback model is well supported by current empirical evidence (Lo et al., 2025; Thi & Nikolov, 2022; Shang, 2019). For Indonesian universities and EFL contexts more broadly, the evidence advocates for a balanced, purposefully designed pedagogical approach: leveraging AWE's efficiency in surface-level correction while preserving and enhancing the indispensable role of expert human instruction in developing higher-order writing competence (Li et al., 2023).

Future research should prioritize longitudinal designs, objective writing quality assessments using validated rubrics, and psychometric refinement of research instruments. Expanding geographic and institutional contexts, examining individual learner characteristics as potential moderators, and investigating teacher adoption processes and barriers would further advance both theory and practice in this growing field. The current evidence nonetheless suggests that thoughtfully implemented AWE integration, paired with expert human feedback, holds genuine promise for enhancing Business English writing instruction and student success in Indonesian higher education.

Acknowledgment

The authors express their profound gratitude to all individuals who played a role in the research undertaking, including those involved in the collection and examination of data, as well as those who provided constructive feedback.

Declarations

Author contribution	:	The authors are responsible for the entire research project, from initiating the ideas to revising the manuscript.
Funding statement	:	This research did not receive any funding.
Conflict of interest	:	There is no conflict of interest in this work.
Ethical declaration	:	We, as authors acknowledge that this work has been written based on ethical research that conforms with the regulations of my university and that we have obtained permission from the relevant institute when collecting data. We support ELTEJ in maintaining high standards of personal conduct, practicing honesty in all our professional practices and endeavors.
Additional information	:	No additional information is available for this paper.

REFERENCES

Adi, I. S. M. A., PUTRI, A. M. J., & Sardi, A. (2024). Examining university students' Business English Writing performance: Frequent errors and pitfalls. *Research and Innovation in Applied Linguistics [RIAL]*, *2*(2), 79-94.

Al-Inbari, F. A. Y., & Al-Wasy, B. Q. M. (2023). The impact of automated writing evaluation (AWE) on EFL learners' peer and self-editing. *Education and Information Technologies*, *28*(6), 6645-6665. https://doi.org/10.1007/s10639-022-11432-7

Ariyanti, A., Fitriana, R., & Pane, W. S. (2018). Self-regulated learning in writing of EFL learners. *IJELTAL (Indonesian Journal of English Language Teaching and Applied Linguistics)*, *3*(1), 155-166.

Artino, A. R. (2012). Academic self-efficacy: from educational theory to instructional practice. *Perspectives on medical education*, *1*(2), 76-85. https://doi.org/10.1007/s40037-012-0012-5

Asiyah, S., Budiyanto, C., & Tamrin, A. (2018). Technology acceptance model in the analysis of the influence of e-learning implementation to students' motivation. *Indonesian Journal of Informatics Education*, *2*(1), 51-56.

Bandura, A. (1977). Self-efficacy: toward a unifying theory of behavioral change. *Psychological review*, *84*(2), 191. https://doi.org/10.1037/0033-295X.84.2.191

Barrot, J. S. (2023). Using automated written corrective feedback in the writing classrooms: Effects on L2 writing accuracy. *Computer Assisted Language Learning*, *36*(4), 584-607. https://doi.org/10.1080/09588221.2021.1936071

Bitchener, J., & Knoch, U. (2010). The contribution of written corrective feedback to language development: A ten month investigation. *Applied linguistics*, *31*(2), 193-214. https://doi.org/10.1093/applin/amp016

Bozdoğan, D., & Kasap, B. (2019). Writing skills and communication competence of undergraduate students in an engineering English course. *Journal of Language Teaching and Learning*, *9*(2), 49-65. https://doi.org/10.33335/jltl.171

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative research in psychology*, *3*(2), 77-101. https://doi.org/10.1191/1478088706qp063oa

Bryant, C., Yuan, Z., Qorib, M. R., Cao, H., Ng, H. T., & Briscoe, T. (2023). Grammatical error correction: A survey of the state of the art. *Computational Linguistics*, 643-701. https://doi.org/10.1162/coli_a_00478

Calma, A., Cotronei-Baird, V., & Chia, A. (2022). Grammarly: An instructional intervention for writing enhancement in management education. *The International Journal of Management Education*, *20*(3), 100704. https://doi.org/10.1016/j.ijme.2022.100704

Chen, C. F. E., & Cheng, W. Y. E. (2008). Beyond the design of automated writing evaluation: Pedagogical practices and perceived learning effectiveness in EFL writing classes. Language Learning & Technology, 12(2), 94–112.

Cheng, G. (2017). The impact of online automated feedback on students' reflective journal writing in an EFL course. *The Internet and Higher Education*, *34*, 18-27. https://doi.org/10.1016/j.iheduc.2017.04.002

Damayanti, S., & Azizah, S. (2024). Grammarly's effectiveness in enhancing English writing: A case study of vocational high school students. *PANYONARA: Journal of English Education*, *6*(2), 15-24.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS quarterly*, *13*(3), 319-340. https://doi.org/10.2307/249008

Dewanti, R. (2018). The influence of self-efficacy on students' writing performance. *Journal of English Education*, 7(2), 145–162.

Dewi, U. (2022). Grammarly as automated writing evaluation: Its effectiveness from EFL students' perceptions. *Lingua Cultura*, 16(2), 155–161. https://doi.org/10.21512/lc.v16i2.8315

Dikli, S., & Bleyle, S. (2014). Automated essay scoring feedback for second language writers: How does it compare to instructor feedback?. *Assessing Writing*, 22, 1–17. https://doi.org/10.1016/j.asw.2014.03.006

Ellis, R. (2009). A typology of written corrective feedback types. *ELT Journal,* 63(2), 97–107. https://doi.org/10.1093/elt/ccn023

Ferris, D. R. (2004). The "grammar correction" debate in L2 writing: Where are we, and where do we go from here?. *Journal of Second Language Writing*, 13(1), 49–62. https://doi.org/10.1016/j.jslw.2004.04.005

Geng, J., & Razali, A. B. (2022). Effectiveness of the automated writing evaluation program on improving undergraduates' writing performance. *English Language Teaching*, 15(7), 49–58. https://doi.org/10.5539/elt.v15n7p49

Guo, Q., Feng, R., & Hua, Y. (2022). How effectively can EFL students use automated written corrective feedback (AWCF) in research writing?. *Computer Assisted Language Learning*, 35(9), 2312–2331. https://doi.org/10.1080/09588221.2021.1879161

Hamdani, M. H. (2025). Indonesian EFL students' writing challenges: A meta-analysis of affective, linguistic, and cognitive dimensions. *Global Education Journal*, 8(1), 45–73.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81–112. https://doi.org/10.3102/003465430298487

Hyland, K. (2003). *Second language writing*. Cambridge University Press.

Kellogg, R. T., & Whiteford, A. P. (2009). Training advanced writing skills: The case for deliberate practice. *Educational psychologist*, *44*(4), 250-266. https://doi.org/10.1080/00461520903213600

Koltovskaia, S. (2020). Student engagement with automated written corrective feedback (AWCF) provided by Grammarly: A multiple case study. *Assessing Writing*, 44, 100450. https://doi.org/10.1016/j.asw.2020.100450

Lee, S.-M. (2020). The impact of using machine translation on EFL students' writing. *Computer Assisted Language Learning*, 33(3), 157–175. https://doi.org/10.1080/09588221.2018.1553186

Li, W. Y., Kau, K., & Shiung, Y. J. (2023). Pedagogic exploration into adapting automated writing evaluation and peer review integrated feedback into large-sized university writing classes. *SAGE Open*, 13(4). https://doi.org/10.1177/21582440231209087

Lo, N., Chan, S., & Wong, A. (2025). Evaluating teacher, AI, and hybrid feedback in English language learning: Impact on student motivation, quality, and performance in Hong Kong. *SAGE Open*, 15(3). https://doi.org/10.1177/21582440251352907

Ngo, T. T. N., Chen, H. H. J., & Lai, K. K. W. (2022). The effectiveness of automated writing evaluation in EFL/ESL writing: A three-level meta-analysis. *Interactive Learning Environments*. https://doi.org/10.1080/10494820.2022.2096642

Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education*, 31(2), 199–218. https://doi.org/10.1080/03075070600572090

Nickerson, C. (2005). English as a lingua franca in international business contexts. *English for Specific Purposes*, 24(4), 367–380. https://doi.org/10.1016/j.esp.2005.02.001

Ningrum, R. W. (2023). Academic writing difficulties for Indonesian postgraduate students: A phenomenological study. *Journal of English Education*, 8(4), 365–384.

Nugraeni, R. (2025). The correlation between self-regulated learning strategies and EFL students' writing performance. *Indonesian EFL Journal,* 11(1), 1–15.

Palermo, C. (2020). Implementing automated writing evaluation in different instructional contexts. *Journal of Writing Research*, 12(1), 1–32. https://doi.org/10.17239/jowr-2020.12.01.01

Pintrich, P. R. (2004). A conceptual framework for assessing motivation and self-regulated learning in college students. *Educational Psychology Review*, 16(4), 385–407. https://doi.org/10.1007/s10648-004-0006-x

Rahimi, M., & Fathi, J. (2022). Exploring the impact of wiki-mediated collaborative writing on EFL students' writing performance, writing self-regulation, and writing self-efficacy: A mixed methods study. *Computer Assisted Language Learning*, 35(1–2), 120–163. https://doi.org/10.1080/09588221.2019.1683700

Rahimi, M., Fathi, J., & Zou, D. (2025). Exploring the impact of automated written corrective feedback on the academic writing skills of EFL learners: An activity theory perspective. *Education and Information Technologies*, 30, 2691–2735. https://doi.org/10.1007/s10639-024-12896-5

Ranalli, J., & Link, S. (2020). Student perceptions of automated writing feedback and the role of teacher mediation. *Journal of Writing Research*, 12(1), 141–171. https://doi.org/10.17239/jowr-2020.12.01.05

Saricaoglu, A. (2019). The relationship between automated writing evaluation feedback uptake and text quality. System, 82, 1–10. https://doi.org/10.1016/j.system.2019.02.009

Sari, T. (2024). The impact of automated writing evaluation on EFL learners' writing self-efficacy, self-regulation, anxiety, and performance. *Journal of Computer Assisted Learning*, 40(5), 2065–2080. https://doi.org/10.1111/jcal.13004

Shang, H. F. (2019). Exploring online peer feedback and automated corrective feedback on EFL writing performance. *Interactive Learning Environments*. https://doi.org/10.1080/10494820.2019.1629601

Stevenson, M., & Phakiti, A. (2014). The effects of computer-generated feedback on the quality of writing. *Assessing Writing*, 19, 51–65. https://doi.org/10.1016/j.asw.2013.11.007

Thi, N. K., & Nikolov, M. (2022). How teacher and Grammarly feedback complement one another in Myanmar EFL students' writing. *Asia-Pacific Education Researcher*, 31(6), 767–779. https://doi.org/10.1007/s40299-021-00625-2

Toba, R., Noor, W., & Sanu, L. (2019). The current issues of Indonesian EFL students' writing skills: Ability, problem, and reason in writing comparison and contrast essay. *Dinamika Ilmu*, 19(1), 57–73. https://doi.org/10.21093/di.v19i1.1506

Tsao, J. J. (2021). Effects of EFL learners' L2 writing self-efficacy on engagement with written corrective feedback. *Asia-Pacific Education Researcher*, 30(6), 575–584. https://doi.org/10.1007/s40299-021-00591-9

Ware, P. D. (2011). Computer-generated feedback on student writing. *TESOL Quarterly*, 45(4), 769–774. https://doi.org/10.5054/tq.2011.272525

Warschauer, M., & Ware, P. (2006). Automated writing evaluation: Defining the classroom research agenda. *Language Teaching Research*, 10(2), 157–180. https://doi.org/10.1191/1362168806lr190oa

Wilson, J., & Roscoe, R. (2020). Automated writing evaluation and feedback: Multiple metrics of efficacy. *Journal of Educational Computing Research*, 58(1), 87–125. https://doi.org/10.1177/0735633119830764

Zhai, N., & Ma, X. (2023). The effectiveness of automated writing evaluation on writing quality: A meta-analysis. *Journal of Educational Computing Research*, 61(4), 875–900. https://doi.org/10.1177/07356331221127300

Zhang, Z. V., & Hyland, K. (2018). Student engagement with teacher and automated feedback on L2 writing. *Assessing Writing*, 36, 90–102. https://doi.org/10.1016/j.asw.2018.02.004

Zimmerman, B. J. (2000*). Self-efficacy: An essential motive to learn. Contemporary Educational Psychology*, 25(1), 82–91. https://doi.org/10.1006/ceps.1999.1016