BAHASTRA
Vol. 46 (1) 2026, pp. 292-300
P-ISSN: 0215-4994 | E-ISSN: 2548-4583
http://journal1.uad.ac.id/index.php/BAHASTRA/index

The role of track changes in enhancing editing accuracy: Disentangling tool effect and metacognitive learning processes in academic manuscript editing

Meilan Arsanti [a,b,1,*], **Ida Zulaeha** [a,2], **Subyantoro** [a,3], **Nas Haryati Setyaningsih** [a,4], **Oktarina Puspita Wardani** [b,5]
[a] Universitas Negeri Semarang, Semarang, Indonesia
[b] Universitas Islam Sultan Agung, Semarang, Indonesia

[1] meilanarsanti@students.unnes.ac.id; [1] meilanarsanti@unissula.ac.id; [2] idazulaeha@mail.unnes.ac.id; [3] bintoro@mail.unnes.ac.id; [4] nasharyati@mail.unnes.ac.id; [5] oktarinapw@unissula.ac.id
* Correspondent author

Received: December 24, 2025 **Revised:** April 28, 2026 **Accepted**: April 29, 2026

KEYWORDS
Academic Manuscript Editing
Cognitive Load
Editing Accuracy
Metacognition
Revision Tracking

ABSTRACT

While the digitization of academic editing is often considered to linearly improve performance, the internal mechanisms that differentiate the impact of tools and learning processes are rarely explored. This study aims to analyze the effects of using the Track Changes feature on editing accuracy and its relationship to metacognitive regulation and cognitive load of aspiring editors. Using a quasi-experimental Non-equivalent Control Group Design, this study involved aspiring student editors divided into an experimental group (Track Changes) and a control group (conventional). Data were collected through an editing accuracy test and analyzed using Normalized Gain and Cohen's d effect size. The results showed that the experimental group experienced a significant increase in editing accuracy (g = 0.53) compared to the control group (g = 0.19). Inferential analysis yielded a Cohen's d value of 1.18, indicating a large effect of tool use on editorial performance. These findings demonstrate that Track Changes functions as a cognitive scaffold that reduces external cognitive load, particularly on surface-level errors such as mechanics and typography. However, improvements in logic and discourse were limited, suggesting that this technology is more effective as a metacognitive mediator than as a substitute for higher-order linguistic reasoning. The unique contribution of this study lies in the empirical separation of the tool's technical impact from natural learning progression, which provides a theoretical basis for integrating technology into professional editing education curricula.

© 2026 The Author(s). Published by Universitas Ahmad Dahlan.
This is an open-access article under the CC-BY-SA license.

Introduction

The development of technology, information, and communication has had a significant impact on people's lives across various sectors globally. This phenomenon, known as openness or globalization, is demonstrated through fundamental changes in the structure of human life compared to the pre-21st century. The 21st century demands quality based on human effort and work, delivered through professionally managed institutions, resulting in superior results (Wijaya et al., 2016). One implication of the 21st century is the improvement of digital literacy practices in society (Paul & Elder, 2019). Digital literacy encompasses not only the ability to use digital tools but also the ability to think critically and ethically when using technology for professional and academic purposes (Davis, 1993; Ng, 2021).

Today, students must master skills to understand, edit, and convey information effectively (Taylor, 2020). They encounter various digital products such as images, videos, text, and combinations there of

(Lofthus & Silseth, 2019). Creating these products, especially texts, involves manuscript editors. Even after editing, texts often contain grammatical errors (Setiawan & Purbandini, 2021). Such language errors can degrade quality.

Manuscript editing is a profession that helps maintain the quality of language in texts. Manuscript editors are required to possess the competence to analyze language errors to ensure the language used in digital communications is linguistically accurate and relevant. Therefore, Sugihastuti (2018) states that editors possess qualifications in the form of specialized skills and abilities that must be demonstrated through Professional Writer and Editor (PEP) certification and the National Professional Certification Agency (BNSP). Therefore, not just anyone can become a manuscript editor.

Digitalization has transformed the landscape of academic editing, requiring aspiring editors to integrate linguistic competence with technological literacy. Amidst the rapid development of automated artificial intelligence (AI)-based tools like Grammarly or ChatGPT that offer instant corrections, the Track Changes feature in Microsoft Word remains a fundamental standard in professional workflows. Unlike AI, which tends to take over linguistic decisions, Track Changes functions as a process-oriented tool that records, visualizes, and documents every editing decision without automated correction (Conijn et al., 2022). However, the role of this feature in developing editing competence is often misunderstood in pedagogical contexts.

A major issue in the current literature is the tendency to attribute improved editing accuracy to tool use. Several studies Aeni & Permana (2022) and Rahmayanti et al. (2020) report that students demonstrate higher accuracy when using Track Changes. However, these findings are often interpreted linearly, failing to distinguish whether the improvements are due to the tool's technical features or to naturally occurring metacognitive learning processes during repeated practice. Without rigorous variable isolation, the tool's actual contribution to competency development remains ambiguous (Hertzog & Light, 2004). This theoretical gap is the focus of this research, which examines the relationship between the tool's effects and internal learning processes.

Theoretically, editing academic manuscripts is not merely a technical procedure but a high-level cognitive activity that involves complex metacognitive regulation. Referring to the theories of Flavell (1979) and Schraw & Dennison (1994), editing requires planning, monitoring, and evaluating linguistic decisions. Track Changes has the potential to serve as an external representation of this metacognitive activity, with the revision trail allowing aspiring editors to review the rationale behind each change. However, its effectiveness depends heavily on how the tool manages the user's cognitive load. Cognitive Load Theory of Sweller (2019) suggests instructional tools should reduce extra cognitive load. This allows for deeper mental processing. It is unclear whether Track Changes reduces cognitive load by using visual cues and tracking, or whether it adds too much fragmented visual information (Luo, 2025).

To date, most research on Track Changes has focused on aspects of revision documentation, collaboration, and user perception (Ruan et al., 2024). Very few studies empirically distinguish between tool-mediated performance improvements and learning-driven improvements, especially in training novice editors. Most previous research also used weak pre-experimental designs, limiting the generalizability of the findings. Therefore, this study aims to fill this critical gap by (1) disentangling the effects of Track Changes use from general learning effects through a rigorous quasi-experimental design, (2) analyzing the role of Track Changes as a metacognitive support in detecting various types of errors (grammar, diction, logic, and structure), and (3) evaluating how the integration of this tool affects aspiring editors' cognitive load management. This study examines how digital tools such as Track Changes affect the development of academic editing skills in higher education, drawing on pedagogical and cognitive frameworks.

Method

1. Research Approach

This study employed a quantitative approach to empirically test the effect of using the Track Changes feature on the accuracy of academic manuscript editing. This approach was chosen because it allows for objective measurement of skill changes and testing of causal relationships between variables through statistical analysis (Shadish et al., 2002). Unlike pre-experimental designs, which are prone to bias, this study adopted a quasi-experimental nonequivalent control-group design, which is methodologically stronger in increasing internal validity (Shadish et al., 2002; Fraenkel et al., 2022). This design allows for the separation of the treatment effect and the general learning effect through a two-group comparison. The experimental group received treatment by using the Track Changes feature in Microsoft Word, while the control group edited using conventional methods without this feature. Thus, this design minimizes bias caused by practice, maturation, and other external factors can be seen in Table 1.

Table 1. Quasi-Experimental Design

Group	*Pretest*	*Perlakuan (Independent Variable)*	*Posttest*
Eksperimen	O_1	Track Changes-Based Editing	O_2
Kontrol	O_3	Conventional (Manual) Editing	O_4

The study population consisted of 46 third-semester students in the Indonesian Language and Literature Education Study Program taking the Scientific Writing course. The research sample was divided into two groups of approximately equal size: an experimental group and a control group, each consisting of 23 students. The sampling technique used was purposive sampling with a matching approach, grouping subjects based on equivalence in initial ability as measured by pretest scores. This strategy aims to reduce selection bias and ensure that differences in results reflect the effects of the treatment rather than differences in the participants' initial characteristics (Fraenkel et al., 2022) Sample selection criteria included (1) students having no professional experience in editing, (2) possessing a basic understanding of linguistic rules, and (3) willingness to participate in the entire research process.

2. Data Collection Instruments and Procedures

The research instruments consisted of test and non-test instruments. The test instruments consisted of two equivalent academic manuscripts used in the pretest and posttest stages. Both manuscripts were designed with comparable levels of complexity, encompassing a variety of linguistic errors, including spelling, diction, sentence structure, punctuation, and paragraph coherence. The instrument's content validity was assessed through expert review, focusing on indicators of syntactic complexity, error diversity, and the representativeness of error categories. The editing results were assessed using an analytical rubric encompassing four main aspects: error detection ability, accuracy of corrections, consistency of revisions, and completeness of language rule application. To ensure assessment reliability, two independent assessors were employed, and the results were analyzed using the Intraclass Correlation Coefficient (ICC), which had a value of 0.87, indicating high reliability. Non-test instruments included feature usage observation sheets (insert, delete, comment) and a Likert-scale student perception questionnaire with a reliability coefficient (Cronbach's Alpha) of 0.91.

3. Research Procedure

The research procedure was carried out in three main stages: pre-treatment, treatment, and post-treatment. In the pre-treatment stage, both groups were given a pre-test in the form of an academic manuscript editing task without using the Track Changes feature to measure initial abilities and ensure group equality. In the treatment phase, the experimental group received intensive training on the use of the Track Changes feature, including the insert, delete, comment, and review panel functions, along with structured practice. Meanwhile, the control group performed editing exercises using conventional methods without digital features, with equivalent duration and learning materials to maintain consistency in instructional treatment. This approach refers to the principle of variable control in quasi-experimental designs, enabling differences in results to be more validly attributed to the treatment provided (Shadish et al., 2002). In the post-treatment phase, both groups received identical posttest scripts. The experimental group edited using the Track Changes feature, while the control group continued to use convhertentional methods.

4. Data Analysis Techniques

Data analysis was conducted using a descriptive and inferential approach, supported by analysis of effect sizes and data distribution. Descriptive statistics were used to report the mean, standard deviation, and score improvement (gain score) for each group. Inferential analysis was conducted using a paired-samples t-test to examine improvement within each group and an independent-samples t-test to compare differences between groups. Prior to testing, a normality test using the Shapiro–Wilk test and a homogeneity of variance test using Levene's Test were performed to ensure statistical assumptions were met using the formula $\boldsymbol{d} = \frac{\boldsymbol{Meksperimen - Mkonrol}}{\boldsymbol{SD\ pooled}}$.

Results and Discussion

The results of the study revealed an improvement in editing performance in both groups of subjects, but with typically different magnitudes and distribution patterns. A summary of descriptive and inferential statistics is presented in Table 2.

Table 2. Comparison of Editing Results for Experimental and Control Classes

Group	*N*	*Mean Pretest (SD)*	*Mean Posttest (SD)*	*N–Gain*	*p-value*

Experiment (*Track Changes*)	30	67,7 (9,2)	84,8 (8,5)	0,53	< 0,001
Control (Konvensional)	30	68,1 (9,0)	74,2 (9,1)	0,19	< 0,05

Descriptive results show that the experimental group using the Track Changes feature in Microsoft Word increased their mean score from 67.7 (SD = 9.2) in the pretest to 84.8 (SD = 8.5) in the posttest. Meanwhile, the control group using the conventional editing method also showed an increase from 68.1 (SD = 9.0) to 74.2 (SD = 9.1). This difference in improvement indicates that both groups experienced a learning effect, but the experimental group showed a higher increase, suggesting a tool-mediated effect. Based on these results, the experimental group increased by 17.1 points, while the control group increased by 6.1 points. Skewness analysis indicated a normality test range of -0.45 to 0.38, confirming that the data were normally distributed with no extreme outliers.

Furthermore, the paired sample t-test revealed that the improvement in both groups was statistically significant ($p < 0.05$). However, to test whether the improvement was specifically due to the treatment, an independent sample t-test was conducted on the posttest scores. The results showed a significant difference between the experimental and control groups ($t = 4.87$; $p < 0.001$). This indicates that the use of Track Changes led to a greater increase in editing accuracy than conventional methods.

To complete the interpretation of statistical significance, an effect size was calculated using Cohen's d: $d = (84.8 - 74.2)/(8.8)$. The calculation revealed a d value of 1.20, which falls within the large effect category. This indicates that the difference between the two groups is not only statistically significant but also practically meaningful in the context of editing learning.

Furthermore, the data distribution analysis showed that the scores in both groups were normally distributed, with skewness values ranging from -0.45 to 0.38, and no extreme outliers were identified that would influence the analysis results. This strengthens the validity of the use of parametric tests in this study. The greatest improvement in the experimental group occurred in mechanical errors (spelling, capitalization, punctuation), simple syntactic errors (sentence structure, repetition), and typographical errors (typos). In contrast, improvements in more complex aspects, such as paragraph coherence and argument relevance, were relatively limited in both the experimental and control groups. Further analysis of the types of language errors revealed a specific distribution pattern in performance. The experimental group demonstrated higher accuracy in simple mechanical and syntactic aspects, as detailed in Table 3.

Table 3. Accuracy Level Based on Error Type in the Experimental Group

Error Category	***Pretest Accuracy (%)***	***Posttest Accuracy (%)***	***Improvement (%)***
Mechanics (Spelling & Punctuation)	62	91	29
Syntactics (Sentence Structure)	58	82	24
Typography (Typo)	70	95	25
Discourse (Coherence & Logic)	55	63	8

The results of this study empirically confirm that the use of the Track Changes feature in editing academic manuscripts significantly contributes to editorial accuracy by mediating cognitive and metacognitive processes, as described below.

1. Disentangling Tool Effect and Learning Effect

The results of this study demonstrate that using the Track Changes feature in Microsoft Word significantly improves editing accuracy. However, this finding cannot be understood simplistically as a direct effect of technology. A comparative analysis between the experimental and control groups confirms the existence of two mechanisms operating simultaneously: the learning effect and the tool effect. The separation of the tool effect and the learning effect is the primary conceptual contribution of this study.

The control group experienced a 19% improvement, indicating that conventional pedagogical processes still contribute to material mastery (Marzuki, 2025). The improvement in the control group indicates that practice and exposure to editing tasks continue to contribute to skill development (Hattie, 2009; Biggs & Tang, 2011). However, the significant 53% advantage in the experimental group confirms the Technological Mediation theory, which states that digital tools function as cognitive amplifiers (Imre, 2023). Significant differences between groups indicate that technology serves as a cognitive mediator rather than a mere technical aid (Conijn et al., 2022; Shulgina., 2024). The significant improvement in the experimental group indicates that Track Changes functions as a mediator, strengthening cognitive processes during editing (Conijn et al., 2022; Shulgina., 2024). Thus, Track Changes does not replace the editor's role, but rather strengthens their detection capacity through synchronous visual feedback.

These findings align with the perspective of technology-mediated learning, which positions technology not as a determinant of learning outcomes but as a facilitator of interaction between learners and tasks (Chapelle, 2022). In this context, Track Changes does not replace the editor's cognitive activity, but rather expands its capacity by providing structured visual feedback. Thus, the improvement in editing accuracy in the experimental group can be understood as a result of the interaction between technological affordances and users' cognitive activity (Luo, 2025; Warschauer et al., 2023).

2. Metacognitive Externalization Mechanism

The improvement in editing accuracy in the experimental group is rooted in strengthened metacognitive awareness. In accordance with Schraw & Dennison (1994) Metacognitive Regulation model, the revision-tracking feature serves as an external representation of the error-monitoring process. Visualizations such as strikethroughs and color insertions compel students to engage in critical evaluation of their own linguistic decisions. From a metacognitive perspective, the revision-tracking feature serves as an externalization of the thinking process. The visual representation of additions, deletions, and comments allows students to explicitly monitor and evaluate linguistic decisions. This is consistent with the concept of metacognitive monitoring and regulation proposed by Flavell (1979) and further developed by Schraw & Dennison (1994). Previous research has shown that transparency in the revision process increases learners' awareness of errors and reflective abilities (Teng, 2020; Panadero, 2017).

3. Cognitive Load Analysis: Extraneous Load Reduction

The experimental group's superiority in detecting mechanical and typographical errors can be explained through Cognitive Load Theory (Sweller et al., 2019). In manual editing, students often experience cognitive overload because they must retain the original text while processing corrections (Sweller et al., 2011). The Track Changes feature reduces extraneous cognitive load by presenting changes explicitly, allowing working memory to be fully allocated to the core editing task (Shulgina., 2024). However, the limited improvement in discourse and logic (8%) suggests a surface-level bias. Wijaya & Frost (2026) As Pividori & Greene (2024) noted, digital annotation tools are often more effective for microlinguistic corrections but are less able to trigger the deep processing (germane load) necessary for correcting complex argumentative structures.

However, the limited improvement in coherence and logical argumentation demonstrates the inherent limitations of technological mediation. This phenomenon indicates a surface-level bias, in which digital tools are more effective at supporting microlinguistic revision than macro revision (Pividori & Greene, 2024; Hyland & Hyland, 2019). Higher-level revision involves more complex cognitive processes, including global planning and restructuring of ideas (Flower & Hayes, 1981), which cannot be fully facilitated by visual annotation-based features. Thus, the effectiveness of Track Changes depends on the type of error being addressed and is domain-specific.

4. Visual Cues and Selective Attention

The success of this feature in reducing mechanical errors by 29% demonstrates the effectiveness of visual cues in directing selective attention. Visual cues such as color and strikethrough serve as attention triggers, directing learners' focus to problematic areas in the text (Posner & Petersen, 1990). Research by Ruan et al. (2024) shows that revision-tracking data provides immediate, granular feedback, enabling novice editors to identify errors that are typically overlooked in linear reading. Furthermore, in their research, Conijn et al. (2022) demonstrated that visual feedback in a digital environment increases sensitivity to errors and accelerates error identification. In this study, visualizing revisions enables students to scan text more systematically, thereby improving error-detection accuracy. The large effect size ($d = 1.18$) in this study provides strong validation that integrating Track Changes into language education curricula is no longer optional but rather a necessity to achieve professional efficiency standards (Zakir et al., 2025).

5. Implications and Unique Contributions

This research makes a unique contribution to the editology literature by demonstrating that Track Changes acts as a bridge between human cognitive capabilities and digital efficiency. Unlike AI assistants, which tend to provide corrections without explanation, Track Changes maintains the editor's cognitive sovereignty because each change still requires the user's conscious consent. This is particularly relevant in educational contexts, where the primary goal is long-term competency development rather than simply producing error-free text (Klieme et al., 2008).

Within the broader technological landscape, it is important to distinguish Track Changes from AI-based editing tools, unlike automated systems like Grammarly, which tend to provide instant corrections. Track Changes maintains the user's active involvement in the decision-making process (Li & Alsanad, 2022). This is relevant to recent findings suggesting that over-reliance on AI can reduce learners' cognitive engagement (Kasneci, 2023). Thus, Track Changes has pedagogical advantages because it functions as a reflective tool that supports long-term competency development, rather than simply producing error-free text (Loginov, 2025).

6. Constraints on Higher-Order Thinking Skills (HOTS) in Editing

Although there was improvement in discourse and logic, the magnitude was far lower than that of mechanics (only 8%). This phenomenon indicates an inherent limitation of the tool-mediated effect. Track Changes facilitates the detection of discrete, localized errors but is less able to trigger a macro-analysis of text coherence that requires in-depth critical thinking. Researchers argue that this limitation arises because visual tracking cues tend to focus users at the sentence level, inadvertently distracting them from the overall argument structure. Therefore, editor competence at the higher-order revision level still requires a metacognitive foundation that cannot be fully transferred to the tool's functionality (Pividori & Greene, 2024).

7. Implications for Editor Training in the Digital Literacy Era

In practice, the results of this study provide a new direction for university manuscript editing training curricula. Based on the large effect size results in this study, it can be suggested that the use of Track Changes should be taught not simply as technical software instructions, but as a work methodology that supports editorial transparency and accountability. This integration will prepare prospective editors to face the demands of the digital publishing industry, which prioritizes speed of detection without sacrificing substantive accuracy. Thus, the results of this study emphasize the need for a learning approach that integrates technology with metacognitive strategies. The use of Track Changes should not be taught simply as a technical skill, but rather positioned as part of a reflective learning process that encourages error analysis and self-regulation. This approach aligns with the concept of scaffolding in sociocultural theory (Vygotsky, 1978), in which technology serves as a tool that supports gradual cognitive development. In the context of language education in Indonesia, integrating technology into editing instruction has also been shown to improve digital literacy and linguistic accuracy (Setiawan & Purbandini, 2021).

8. Limitations and Directions for Further Research

Although this study successfully separated tool effects from learning effects, several limitations warrant note. First, the limited duration of the intervention may have affected the depth of internalization of metacognitive processes in the subjects. Second, this study focused only on the standard tracking feature and did not explore the more intensive use of collaborative commenting features. Future research is recommended to explore how the integration of Track Changes and an AI editing assistant can work synergistically, where AI assists with mechanical detection while revision tracking serves as a space for editor reflection to maintain human cognitive control over the text.

Conceptually, the findings of this study support the distributed cognition approach, which views thinking as the result of interactions among individuals and tools (Hutchins, 1995). Editing is no longer understood as a purely internal activity, but as a practice mediated by technology and social context. Therefore, the effectiveness of Track Changes lies not only in its technical features but also in its integration into a learning design that encourages cognitive engagement and reflection. Overall, this study shows that Track Changes serves as a pedagogical tool that strengthens cognitive and metacognitive processes in editing, but cannot replace the need for higher-level critical thinking. Therefore, developing editor competencies in the digital age requires an approach that balances the use of technology with the strengthening of human cognitive capacities.

The findings of this study provide empirical confirmation that strengthens the metacognitive awareness theory proposed by Schraw & Dennison (1994). The data indicate that increased accuracy is not simply a mechanical result of software use, but rather the result of enhanced metacognitive monitoring. This aligns with Schraw's argument that tools that provide process feedback can enhance learner self-regulation. This study found that Track Changes serves as a cognitive safety net, allowing novice editors to handle mechanical errors more efficiently without draining their mental resources. Nonetheless, this study supports about surface-level bias, in which technology predominantly facilitates microscopic linguistic refinement rather than complex macroscopic discourse reconstruction.

Practically, these findings have significant implications for two key sectors. In education, these results emphasize that editing instruction should no longer separate manual linguistic skills from digital literacy. The integration of Track Changes into the curriculum should be positioned as a metacognitive scaffolding strategy that helps students independently visualize their error patterns. Meanwhile, for the publishing industry, this study demonstrates that editorial efficiency can be substantially improved by optimizing the revision tracking feature as a transparent communication tool between editors and authors. Systematic use of this feature can reduce the risk of redundancy and accelerate manuscript production cycles without compromising accuracy standards (Rahmayanti et al., 2020). Finally, this research emphasizes that, amidst the massive rise of AI automation, the role of human editors remains irreplaceable for cognitive sovereignty and logical decision-making that require in-depth context.

Conclusion

Based on the research results, it can be concluded that the use of the Track Changes feature in Microsoft Word significantly contributes to improved editing accuracy in academic manuscripts; however, this finding does not support a deterministic view that technology alone directly enhances competency. Instead, the improvement in editing ability reflects an interaction between the tool effect and the learning effect, including practice, exposure, and repeated revision. Empirically, the significant difference between the experimental and control groups, supported by a large effect size (Cohen's d ≈ 1.20), indicates a substantial practical impact, although the effectiveness of the tool varies across editing aspects. The most notable improvements occur in surface-level errors such as spelling and simple sentence structure, while gains in higher-order aspects such as discourse coherence and argumentative logic remain limited, suggesting that revision visualization is more effective for mechanical editing than conceptual revision. Theoretically, this study extends understanding of technology in learning by integrating metacognitive perspectives and Cognitive Load Theory, positioning Track Changes as a tool for externalizing cognitive processes that supports monitoring and evaluation, though its effectiveness depends on pedagogical use and management of cognitive load. Methodologically, the use of a quasi-experimental design with a control group, along with effect size and data distribution analyses, strengthens the validity and interpretability of the findings. The implications emphasize that Track Changes should be positioned as a mediating pedagogical tool rather than merely a technical feature, with optimal effectiveness achieved when integrated into learning strategies that foster reflection, error analysis, and metacognitive awareness, thereby shifting the perspective from a tool-oriented approach to one that situates technology within a broader, complex learning ecosystem and opening avenues for further research on the integration of digital tools, pedagogy, and cognitive processes in developing editing competence.

Declarations

Author contribution	:	Meilan Arsanti was responsible for the overall research design and implementation, led the manuscript writing, and coordinated the collaboration among all authors. Ida Zulaeha contributed to the conceptual framework and data analysis. Subyantoro was involved in data collection and validation. Nas Haryati Setyaningsih contributed to reviewing, editing, and finalizing the manuscript. Oktarina Puspita Wardani contributed to literature review, data interpretation, and manuscript preparation. All authors have read and approved the final version of the manuscript.
Funding statement	:	This research did not receive any funding.
Conflict of interest	:	All authors declare that they have no competing interests.
Ethics Approval	:	Information on Ethics Approval and informed consent statements are required for all articles published in BAHASTRA since 2026.
Additional information	:	No additional information is available for this paper.

References

Aeni, E. S., & Permana, I. (2022). Pemanfaatan fitur track change pada perkuliahan editing di perguruan tinggi. *Semantik*, *11*(2), 203–214. https://doi.org/10.22460/semantik.v11i2.p203-214

Biggs, J., & Tang, C. (2011). *Teaching for quality learning at university.* McGraw-Hill.

Chapelle, C. A. (2022). *Computer-assisted language learning.* Language Learning & Technology.

Conijn, R., Speltz, E. D., Zaanen, M. V., Waes, L. V., & Chukharev-Hudilainen, E. (2022). A product- and process-oriented tagset for revisions in writing. *Written Communication*, *39*(1), 97–128. https://doi.org/10.1177/07410883211052104

Davis, F. D. (1993). User acceptance of information technology: System characteristics, user perceptions, and behavioral impacts. *International Journal of Man-Machine Studies*, *38*(3), 475–487. https://doi.org/10.1006/imms.1993.1022

Flavell, J. H. (1979). Metacognition and cognitive monitoring a new area of cognitive—developmental inquiry. *American psychologist, 34*(10), 906–911. https://doi.org/10.1037/0003-066X.34.10.906

Flower, L., & Hayes, J. R. (1981). A cognitive process theory of writing. *College Composition and Communication, 32*(4), 365–387. https://doi.org/10.2307/356600

Fraenkel, J. R., Wallen, N., & Hyun, H. (2022). *How to design and evaluate research in education*.

Hattie, J. (2009). *Visible learning.* Routledge. https://doi.org/10.4324/9780203887332

Hertzog, C., & Light, L. L. (2004). Memory and aging: Four hypotheses in search of data. *Annual Review of Psychology*, *55*, 333–363. https://doi.org/10.1146/annurev.psych.55.090902.141521

Hutchins, E. (1995). *Cognition in the wild.* MIT Press. https://doi.org/10.7551/mitpress/1881.001.0001

Hyland, K., & Hyland, F. (2019). *Feedback in second language writing*. Cambridge University Press. https://doi.org/10.1017/9781108635547
Imre, V., K. (2023). Theorizing technological mediation in the outdoor classroom. *Postdigital Science and Education*, *5*(3), 754–776. https://doi.org/10.1007/s42438-022-00315-2
Kasneci, E., (2023). ChatGPT for good? *Nature Human Behaviour, 7*, 468–471. https://doi.org/10.1038/s41562-023-01580-3
Klieme, E., Hartig, J., & Rauch, D. (2008). The concept of competence in educational contexts. In J. Har-tig, E. Klieme, & D. Leutner (Eds.), *Assessment of competencies in educational contexts*. Hogrefe & Huber. https://www.researchgate.net/publication/232495759
Li, S., & Alsanad, A. (2022). Automated writing evaluation systems and academic writing. *Computer Assisted Language Learning.* https://doi.org/10.1080/09588221.2021.1877085
Lofthus, A. M., & Silseth, K. (2019). Students' meaning-making in peer group interactions: A sociocultural perspective on learning and writing. *Learning, Culture, and Social Interaction*, *21*, 1–12. https://doi.org/10.1016/j.lcsi.2019.100318
Loginov, A. (2025). Digital metacognition: Awareness, regulation, and analysis of thinking in digital environments. *Neurocomputing, 2*(4), 100-110. https://doi.org/10.1016/j.nexres.2025.100770
Luo, H., (2025). Designing for metacognition: How digital tools mediate higher-order thinking in online learning. *Computers & Education*. https://doi.org/10.1016/j.compedu.2024.105120
Marzuki, A. G. (2025). *Blended learning and english proficiency: Evaluating digital pedagogies in Indonesian secondary schools*. https://doi.org/10.20944/preprints202506.1282.v1
Ng, W. (2021). *New digital technology in education*. *Switzerland: Springer*.
Panadero, E. (2017). A review of self-regulated learning. *Educational Psychology Review, 29*(3), 1–34. https://doi.org/10.3389/fpsyg.2017.00422
Paul, R., & Elder, L. (2019). The miniature guide to critical thinking: Concepts and tools (8th ed.). *Foundation for Critical Thinking*. https://doi.org/10.5771/9781538133842
Pividori, M., & Greene, C. S. (2024). A publishing infrastructure for artificial intelligence (AI)–assisted academic authoring. *Journal of The American Medical Informatics Association, 31*(9), 2103–2113. https://doi.org/10.1093/jamia/ocae139
Posner, M. I., & Petersen, S. E. (1990). The attention system of the human brain. *Annual Review of Neuroscience, 13*, 25–42. https://doi.org/10.1146/annurev.ne.13.030190.000325
Rahmayanti, P., Oka, I. G. & Putra, I. N. (2020). Improving the students' writing skills by using track changes in paired writing correction. *English Learning Innovation*, 1-35. https://doi.org/10.22219/englie.v1i1.13131
Ruan, Q., Kuznetsov, I., & Gurevych, I. (2024). *Re3: a holistic framework and dataset for modeling collaborative document revision*. https://doi.org/10.18653/v1/2024.acl-long.255
Schraw, G., & Dennison, R. S. (1994). Assessing metacognitive awareness. *Contemporary Educational Psychology*, *19*(4), 460–475. https://doi.org/10.1006/ceps.1994.1033
Setiawan, B., & Purbandini, A. (2021). Pembelajaran penyuntingan teks akademik berbasis digital. *Jurnal Pendidikan Bahasa dan Sastra Indonesia*, *10*(2), 145–156.
Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). Experimental and quasi-experimental designs for generalized causal inference. *Houghton Mifflin*, *35*(8).
Shulgina. (2024). Visual cues and cognitive load in text editing: An eye-tracking study. *Learning and Instruction.*
Sugihastuti, S. (2018). Penulis karya ilmiah dan copyediting. *Deskripsi Bahasa, 3*(1), 30–36. https://doi.org/10.22146/db.v3i1.398
Sweller, J., Ayres, P. & Kalyuga, S. (2011). *Cognitive load theory*. Springer New York. https://link.springer.com/book/10.1007/978-1-4419-8126-4
Sweller, J. (2019). *Cognitive load theory, evolutionary educational psychology, and instructional design* (pp. 291–306). https://doi.org/10.1007/978-3-319-29986-0_12
Sweller, J., Van Merriënboer, J. J., & Paas, F. (2019). Cognitive architecture and instructional design: 20 years later. *Educational psychology review*, *31*(2), 261-292. https://doi.org/10.1007/s10648-019-09465-5
Taylor, S. (2020). *Educational research and policy: Approaches and perspectives.* Routledge.
Teng, F. (2020). The benefits of metacognitive reading strategy awareness instruction for young learners of English as a second language. *Metacognition and Learning, 15*(2), 251–278. https://doi.org/10.1007/s11409-020-09224-7
Vygotsky, L. S. (1978). *Mind in society.* Harvard University Press.
Warschauer, M., Tseng, W., Yim, S., Webster, T., Jacob, S., Du, Q., & Tate, T. (2023). The affordances and contradictions of AI-generated text for writers of English as a second or foreign language. *Journal of Second Language Writing, 62*. https://doi.org/10.1016/j.jslw.2023.101071

Wijaya, A., Van den Heuvel-Panhuizen, M., & Doorman, M. (2016). Opportunity-to-learn context-based tasks provided by mathematics textbooks. *Educational Studies in Mathematics*, *89*(1), 41–65. https://doi.org/10.1007/s10649-015-9595-1

Wijaya, R. A. & Frost, K. (2026). Evaluating the logic of a policy-driven national language test in Indonesia: A critical discursive investigation of the Test of Indonesian Proficiency (UKBI). *Language Testing, 43*(4), 523-540. https://doi.org/10.1177/02655322251351710

Zakir, S., Hoque, M. E., Susanto, P., Nisaa, V., Alam, M. K., Khatimah, H., & Mulyani, E. (2025). Digital literacy and academic performance: The mediating roles of digital informal learning, self-efficacy, and students' digital competence. *Frontiers in Education, 10,* 1590274. https://doi.org/10.3389/feduc.2025.1590274