MULTIDISCIPLINARY SCIENCE JOURNAL

RESEARCH ARTICLE
Published Online: November 18, 2025
https://doi.org/10.31893/multiscience.2026282

AI-based adaptive learning and self-assessment as differentiation strategies in improving students' argumentative writing efficacy

Hidayati[ab] | Dwi Priyo Utomo[a] | Dwi Poedjiastutie[a] | Putri Yulianti[b] | Chinta Shaqila[b] | M. Nabil Hummam[c]

[a]Education Programme, Universitas Muhammadiyah Malang, Malang, Indonesia.
[b]English Education Department, Universitas Muhammadiyah Mataram, Mataram, Indonesia.
[c]Mining Department, Universitas Muhammadiyah Mataram, Mataram, Indonesia.

Abstract The integration of adaptive AI systems and self-assessment aligns with curricular goals emphasizing inclusivity, learner autonomy, and critical literacy. This study investigates the predictive influence of adaptive AI-based learning and self-assessment on students' efficacy in argumentative writing, within the context of differentiated instruction under Indonesia's Merdeka curriculum. Grounded in the principles of personalized and student-centered pedagogy, the research highlights the role of technology-enhanced learning and reflective practices in fostering academic confidence and performance. A quantitative correlational design was employed, involving 116 eighth-grade students from three pilot schools in Lombok, West Nusa Tenggara. Participants were selected purposively based on their exposure to adaptive learning technologies and formative assessment strategies. Data were collected using a structured questionnaire comprising validated subscales for adaptive AI-based learning, self-assessment, and writing efficacy. Reliability analysis indicated strong internal consistency (Cronbach's $\alpha > .80$). Descriptive statistics, Pearson correlation, and multiple linear regression were used, with diagnostic tests conducted to ensure statistical validity. Results showed high levels of student engagement and positive perceptions of AI-driven personalization and self-evaluation. Significant correlations were found between adaptive AI-based learning and writing efficacy ($r = .775$, $p < .01$), and between self-assessment and writing efficacy ($r = .718$, $p < .01$). Regression analysis revealed that both predictors accounted for 67.3% of the variance in writing efficacy ($R^2 = .673$), with adaptive AI-based learning ($\beta = .468$) having a slightly stronger predictive effect than self-assessment ($\beta = .391$). These findings underscore the synergistic impact of adaptive technologies and metacognitive strategies in enhancing students' argumentative writing skills. The study provides empirical support for differentiated instruction that is both adaptive and reflective, offering practical implications for educators and policymakers implementing the Merdeka curriculum.

Keywords: technology, assessment, efficacy, instruction, curriculum

1. Introduction

AI-based adaptive learning refers to an instructional approach that uses intelligent systems to deliver personalized content and feedback in real time, tailored to students' individual needs and performance levels. In the Indonesian context, these systems are in line with the Merdeka curriculum, which emphasizes personalized learning trajectories on the basis of students' readiness and interests (Hidayati et al., 2024; Setiyana et al., 2024). Moreover, adaptive systems provide educators with diagnostic insights, facilitating the implementation of differentiated instruction strategies (Cahyani & Suryantari, 2024). A meta-analysis by Mitchell et al.(2021), along with recent findings by Mudrikah (2024) and Respati et al.(2024), revealed that 86% of adaptive learning system studies reported positive outcomes in advancing individualized learning, particularly in developing writing skills. Within this framework, synergies among AI-driven feedback, self-assessment, and adaptive mechanisms become evident: AI delivers personalized inputs, whereas self-assessment fosters student autonomy and self-regulation. However, few studies have systematically examined how these elements interact within a differentiated instruction model.

The application of differentiated instruction under the Merdeka curriculum has been documented across educational levels. Jin et al. (2025) demonstrated that strategies involving content, process, and product differentiation significantly improved learning outcomes in elementary-level geogr aphy instruction. Similarly, Pratama et al. (2025) showed that differentiation in senior high school contributed to enhancing students' critical thinking by aligning learning challenges with students' readiness and interest. Despite these positive findings, existing research predominantly adopts a qualitative lens,

 Received: June 30, 2025 | Accepted: August 29, 2025

with limited quantitative investigations focusing on the integration of AI technologies and self-assessment in Pilot Schools (Sekolah Penggerak). This represents a significant gap, particularly given the Merdeka curriculum's emphasis on digital innovation and student-centered learning.

Hidayati et al. (2024) highlighted that the implementation of differentiated instruction in pilot school settings centers on process differentiation, encompassing self-directed learning activities and discovery-based experiences. In this model, teachers are expected to design learning that stimulates curiosity, integrates authentic content, and encourages field-based learning. Supporting this view, Samsudi et al. (2024) argued that differentiated instruction is a strategic approach within the Society 5.0 educational paradigm, as it promotes learner engagement and personalization through tailored content, methods, and learning products. Their findings indicated that this strategy significantly increases students' motivation and participation, which are key pillars of the Merdeka curriculum in Indonesia (Aladini et al., 2025).

A central component of the curriculum is the concept of "assessment as learning," in which students are actively engaged in evaluating and refining their academic performance (Mitchell, 2020b; Nayak et al., 2024). Within differentiated classrooms, self-assessment plays a crucial role in fostering metacognition, enabling learners to reflect on their strengths, monitor progress, and adjust learning strategies accordingly (Sun et al., 2022; Tan et al., 2025). This shift transforms learners from passive recipients into active participants, thereby cultivating a stronger sense of agency. One of the curriculum's core targets—strengthening literacy and numeracy—includes writing efficacy as a critical skill area to be developed.

Argumentative writing efficacy, in particular, reflects a student's confidence in composing structured, persuasive texts—an essential competency in 21st-century academic and professional settings (Meza & González, 2020 ; Lu, 2024). The rise of AI-powered educational tools, such as Grammarly and ChatGPT, has been linked to notable improvements in students' writing performance. These tools provide instant feedback, autocorrection, and content suggestions that enhance clarity, coherence, and argumentation (Varier et al., 2021). Empirical evidence from Ayu (2023), Shaqila (2025), Trajkovski & Hayes (2025), and Yulianti et al. (2025) confirms that such tools significantly increase L2 learners' writing self-efficacy. In the Indonesian EFL context, platforms such as QuillBot and Grammarly have also been shown to improve both academic writing outcomes and learners' confidence (Gligorea et al., 2023; Karafil & Oğuz, 2022; Santos et al., 2024). These results suggest that when AI-based feedback systems are integrated with reflective practices such as self-assessment, they contribute to the creation of an inclusive, responsive, and empowering learning ecosystem.

This study aims to explore the relationships between AI-based adaptive learning, self-assessment, and students' argumentative writing efficacy within the context of differentiated instruction in Merdeka curriculum classrooms. While previous research has independently explored each variable, limited evidence exists on their combined influence via a quantitative approach in Indonesian pilot schools. This gap is significant considering the lack of empirical studies investigating the impact of technological and reflective pedagogies on writing efficacy in secondary education. Although Wang (2025) proposed a conceptual framework for differentiation, the interplay between adaptive systems, self-assessment, and argumentative writing efficacy remains insufficiently examined. This opens a critical space for research aimed at designing innovative, evidence-informed, and globally relevant instructional models.

Accordingly, this study seeks to address how AI-based adaptive learning and self-assessment influence students' confidence and effectiveness in argumentative writing. The research contributes both to international discourse on personalized education and to local policy efforts supporting instructional design under the Merdeka curriculum. It also offers practical implications for teachers in creating more autonomous, differentiated, and reflective learning environments that leverage digital tools and metacognitive strategies to support student success.

2. Materials and methods

This research adopted a quantitative correlational approach aimed at examining the relationship and influence of two independent variables—AI-based adaptive learning and self-assessment—on one dependent variable, namely, students' argumentative writing efficacy. The correlational design was selected for its capacity to analyze relationships among variables on a broad scale and to offer predictive insights relevant to educational practice (Craswell, 2014).

2.1. Sample and participants

The study involved 116 eighth-grade students selected from three pilot schools (Sekolah Penggerak) located in Lombok, West Nusa Tenggara, Indonesia. Specifically, the sample included 37 students from SMPN 1 TN, 39 from SMPN 1 PM, and 40 from SMPN 10 MM. One class was purposively selected from each school on the basis of practical and academic considerations.

First, the selected classes were known for actively implementing the key features of the Merdeka curriculum, including adaptive learning techniques and formative assessments aligned with differentiated instructions. Second, the selection was guided by the principles of internal homogeneity within schools and external variability across schools. This ensured a comparable level of curricular exposure, teacher training, and infrastructural support for AI-based instruction. Finally, limiting the sample to one class per school facilitated deeper data analysis without compromising methodological rigor or resource feasibility.

This sampling method is aligned with purposive sampling techniques frequently employed in educational research, where sample groups are chosen for their relevance to the research objectives (Creswell & Creswell, 2018; Sugiyono, 2021). Each selected class was thus representative of the pedagogical and technological conditions under which differentiated instruction supported by AI is implemented in Indonesian pilot schools. The participant groups are presented in Table 1.

Table 1 The Participants' Groups.

No	Pilot Schools	Students Participants
1	SMPN 1 TN	37
2	SMPN 1 PM	39
3	SMPN 10 MM	40

2.2. Research procedures

This study employed a quantitative correlational design to examine the relationship and influence between two independent variables, namely, AI-based adaptive learning and self-assessment, and one dependent variable, namely, argumentative writing efficacy. The research procedure was structured systematically to ensure methodological rigor, internal validity, and relevance to the context of pilot schools, which implement the Merdeka curriculum in the academic year 2023-2024. This study utilized a quantitative correlational research design. The population consisted of 116 eighth-grade students from three Sekolah Penggerak in West Nusa Tenggara: SMPN 1 TN (n = 39), SMPN 1 PM (n = 37), and SMPN 10 MM (n = 40). Purposive sampling was used to select one class from each school on the basis of their active implementation of differentiated instruction and AI-based learning technologies under the Merdeka curriculum (Creswell & Creswell, 2018; Sugiyono, 2021).

The research instrument was a five-point Likert-scale questionnaire developed on the basis of validated theoretical frameworks. The questionnaire comprises three subscales: AI based adaptive learning, which measures students' perceptions of content personalization and feedback from AI systems (Chen et al., 2022); self-assessment, which measures students' metacognitive awareness and ability to evaluate their academic performance (Tan et al., 2025; Zimmerman, 2000); and argumentative writing efficacy, which involves assessing students' confidence in developing coherent arguments and persuasive texts (Bruning et al., 2013; Meza & González, 2020).

The questionnaire was validated by three educational experts and piloted with 30 students outside the primary sample, with Cronbach's alpha coefficients exceeding 0.80 for all the subscales. Data were collected through both online and offline methods, depending on each school's digital readiness. The data were analyzed via SPSS version 26. Descriptive statistics were used to summarize the data. Pearson correlation analysis was used to examine the relationships between variables, and multiple linear regression analysis was used to assess the predictive impact of AI-based adaptive learning and self-assessment on students' argumentative writing efficacy (Field, 2018). The assumptions of normality and linearity were tested before inferential analysis.

2.3. Data collection instrument

The second instrument used to collect data in this study was a structured questionnaire composed of Likert-scale items designed to measure three key variables: AI-based adaptive learning, self-assessment, and argumentative writing efficacy. The questionnaire was developed through an extensive review of the literature and adapted from previously validated instruments to ensure that both theoretical relevance and contextual suitability apply to pilot schools implementing the Merdeka curriculum.

2.3.1. AI-Based Adaptive Learning Scale

This section of the questionnaire aimed to assess students' experiences with and perceptions of adaptive learning environments, especially those supported by artificial intelligence. The items were derived from the constructs proposed by Chen et al. (2022), which focused on the personalization of content delivery, adaptive feedback mechanisms, and responsiveness to learner performance. Sample items include "The AI-based system adjusts learning materials according to my learning "speed" and "I receive personalized suggestions from the system to improve my understanding."

2.3.2. Self-Assessment Scale

The self-assessment subscale was constructed on the basis of the theoretical models of metacognitive awareness and learner autonomy created (Tan et al., 2025; Zimmerman, 2000). It measured students' capacity for introspection, self-monitoring, and regulation of learning goals in this study. Sample items include "I evaluate the quality of my own writing before submitting "assignments" and "I use assessment criteria to judge my performance." These items align with the assessment as a learning component of the Merdeka curriculum, which the learners position as active agents in evaluating and improving their learning outcomes (Mitchell, 2020a).

2.3.3. Argumentative Writing Efficacy Scale

To measure students' self-efficacy as participants in argumentative writing, the study adapted the scale developed by Bruning et al. (2013) and expanded Meza & González (2020) upon the findings of . The items focused on students' confidence in generating arguments, organizing ideas logically, and using language effectively to support claims. Examples include "I am confident in developing coherent arguments in my writing" and "I can express my opinions persuasively in argumentative texts, which is made by the teacher."

All items in the questionnaire used a five-point Likert scale ranging from 1 (Strongly Disagree) to 5 (Strongly Agree). The instrument was initially drafted in English, then translated into Bahasa Indonesia and back-translated to ensure semantic and conceptual equivalence. Before full deployment, the questionnaire was reviewed by three experts in the fields of educational assessment, instructional design, and applied linguistics to assess content validity. A pilot test with 30 students outside the sample population was also conducted, and reliability was calculated via Cronbach's alpha, yielding coefficients above 0.80 for all three subscales, indicating high internal consistency (Field, 2018). The final instrument was distributed both online and offline to accommodate technological variability across the participating schools. Data collection was conducted in accordance with ethical standards, including informed consent, anonymity, and voluntary participation in third schools implementing the Merdeka curriculum, which applies differentiated instruction in the classroom.

2.4. Data analysis

The collected data were analyzed via both descriptive and inferential statistics to address the research questions concerning the relationships and predictive effects of AI-based adaptive learning and self-assessment on students' argumentative writing efficacy on the basis of the research problems in this study. The statistical package for the Social Sciences (SPSS) version 26 was employed as the primary tool for data processing and analysis.

2.4.1. Descriptive Statistics

Descriptive statistics, including the mean, standard deviation, minimum, and maximum scores, were first computed to summarize the central tendencies and variability of each variable in this study. This allowed the researchers to understand the distribution and general trends of students' perceptions across the three constructs that were created.

2.4.2. Test of Assumptions

Before the inferential tests were conducted, assumption testing was performed. The normality of the distribution for each variable was examined via the Shapiro–Wilk test, which is suitable for small to moderate sample sizes (Ghasemi & Zahediasl, 2012). Variables with p-values greater than 0.05 were considered normally distributed. Additionally, scatter plots and residual analysis were reviewed to assess linearity and homoscedasticity, which are prerequisites for correlation and regression analysis (Field, 2018).

2.4.3. Pearson correlation analysis

To determine the strength and direction of the linear relationship between variables, Pearson product–moment correlation coefficients were calculated. This analysis helped identify whether there was a statistically significant correlation between AI-based learning and self-assessment with argumentative writing efficacy (Cohen, 1988). The correlation values were interpreted as small ($r = 0.10–0.29$), moderate ($r = 0.30–0.49$), or large ($r \geq 0.50$).

2.4.4. Multiple linear regression analysis

Finally, multiple linear regression analysis was conducted to examine the predictive power of the two independent variables (AI-based adaptive learning and self-assessment) on the dependent variable (argumentative writing efficacy). The regression model was evaluated on the basis of R-squared values, standardized beta coefficients, and significance levels ($p < 0.05$), indicating the extent to which each predictor variable contributes to the variance in the dependent variable (Tabachnick & Fidell, 2013). This analytical framework ensures a rigorous investigation into the cognitive and technological factors influencing students' writing performance, aligning with the goals of differentiated instruction and personalized learning in the era of AI-enhanced education.

3. Results

3.1. Descriptive Statistics

To provide a general overview of the research variables, descriptive statistical analysis was conducted. The following table summarizes the key metrics—including the mean, standard deviation, and data distribution—for each variable: adaptive AI based learning (X1), self-assessment (X2), and argumentative writing efficacy (Y). The data consists of 116 valid responses. The results show that the average scores of all three variables fall within a similar range (approximately 82), indicating a

relatively high level of engagement and performance among participants. The Shapiro–Wilk test values indicate that the data for each variable are not normally distributed ($p < .001$), suggesting the need for nonparametric or robust statistical techniques in further analysis. Table 2 is presented below.

Table 2 Descriptive statistical analysis.

	X1	X2	Y
Valid	116	116	116
Mean	82.310	82.862	82.586
Std. Deviation	10.979	10.974	11.193
Shapiro–Wilk	0.955	0.954	0.952
P value of Shapiro–Wilk	< .001	< .001	< .001
Minimum	60.000	60.000	60.000
Maximum	100.000	100.000	100.000

On the basis of the table, descriptive statistics were utilized to examine the central tendencies and data distributions across the three variables: adaptive AI based learning (X1), self-assessment (X2), and argumentative writing efficacy (Y). As shown in Table 2, the mean scores for X1, X2, and Y were 82.17, 82.50, and 82.71, respectively, all within relatively high and narrow ranges. This consistency suggests that students' perceived experiences with adaptive AI integration and self-regulatory assessment strategies are positively aligned with their writing efficacy in argumentative contexts. The standard deviation for each variable ranged from 9.32 to 10.93, indicating a moderate spread and confirming some level of individual variance across the responses. Furthermore, the results of the Shapiro–Wilk normality test yielded pvalues of .000 for all three variables, which are below the conventional alpha level of .05. These outcomes indicate that the data distribution deviates significantly from normality (Ghasemi & Zahediasl, 2012), thus suggesting the appropriateness of nonparametric statistical analysis or data transformation techniques (Pallant, 2020). Given the cognitive and behavioral nature of the variables studied, such nonnormal distribution is not unexpected (Boekaerts, 2016), as responses to self-report measures in educational research often reflect individual learning diversity.

The strong alignment of mean scores among the independent and dependent variables also supports theoretical claims in the literature on differentiated instruction. Specifically, adaptive learning environments—especially when augmented with AI and coupled with structured self-assessment can enhance learners' writing efficacy through personalization and formative feedback (Tomlinson, 2014; Bandura, 1997; Sun & Wang, 2020). These findings form a critical baseline for exploring the strength and nature of the relationships between these constructs in subsequent inferential analysis.

The descriptive statistics output above provide an overview of the distribution of scores for the three main variables in the study, namely, AI-based adaptive learning (X1), self-assessment (X2), and argumentative writing efficacy (Y), on the basis of a sample of 116 students. The mean scores for all three variables are notably similar—M = 82.31 for X1, M = 82.86 for X2, and M = 82.59 for Y—indicating a consistent level of perception and performance across the constructs. This uniformity suggests that students generally reported high engagement with AI-supported learning environments and self-evaluation strategies, as well as a strong belief in their ability to produce argumentative writing.

The Figure 1 line chart above illustrates the score trends of three key variables measured across 100 student respondents in the context of a differentiated learning strategy within the Merdeka curriculum framework. The three variables include the following.

Figure 1 Lines of Adaptive AI Learning, Self-Assessment, and Argumentative Writing Efficacy.

Adaptive AI based learning – representing personalized digital learning environments tailored to students' pace and performance;

Self-Assessment – reflecting students' ability to evaluate and improve their learning outcomes;

Argumentative Writing Efficacy – indicating students' confidence and competence in producing structured argumentative texts

The line graph above illustrates the score trends of three key constructs—AI-based adaptive learning, self-assessment, and argumentative writing efficacy-across a sample of 116 respondents. Overall, the plotted lines display a striking level of parallelism, indicating a strong degree of covariance among the variables. Each of the three constructs fluctuates within a range of approximately 60100, with notable convergence at several points across the distribution. Specifically, scores tend to peak within the 88–100 range, suggesting that students who report higher engagement with adaptive AI features and greater use of self-assessment strategies also tend to exhibit higher levels of confidence in the argumentative writing text that the students achieved. The presence of repeated cofluctuations, where all three lines rise or dip simultaneously, visually supports the statistical findings of significant positive correlations reported in the inferential analysis. These synchronous movements imply a consistent pattern: when learners experience adaptive, personalized learning (X1) and actively assess their own progress (X2), their belief in their capacity to compose persuasive texts (Y) strengthens accordingly. This aligns with prior research emphasizing the interconnectedness of metacognitive strategies, technology-supported personalization, and writing self-efficacy itself (Bruning et al., 2013; Tan, Mohamad et al., 2025).

Despite occasional anomalies, such as isolated low dips in AI-based scores not mirrored by self-assessment or students' efficacy, the overall coherence among the three lines demonstrates the reliability of the measurement constructs from the text. Moreover, although formal normality tests (e.g., Shapiro–Wilk) indicate nonnormal distributions, the large sample size and the relatively symmetrical fluctuations in the graph offer empirical justification for the use of parametric tests such as Pearson correlation and multiple regression (Field, 2018). This visualization substantiates the core theoretical premise of the study that AI-supported differentiation and formative self-assessment act as mutually reinforcing elements in enhancing students' confidence and competence in writing, a central component of 21st-century academic literacy (Lu, 2024; Meza & González, 2020). The data consistently show strong performance for all three variables, with fluctuations reflecting individual learning diversity. The adaptive AI and self-assessment variables generally move in tandem with the peaks in writing efficacy, suggesting a potential positive relationship. These visual patterns may indicate that the integration of AI and reflective learning contributes to students' enhanced writing confidence and outcomes.

3.2. Pearson correlation analysis

To investigate the strength and direction of the relationships among the study variables, adaptive AI based learning (X1), self-assessment (X2), and argumentative writing efficacy (Y), a Pearson product–moment correlation analysis was conducted. This statistical approach is commonly used to measure the linear association between interval-scaled variables, assuming continuous data and appropriate scale conditions (Field, 2018). The correlation matrix presented in Table 2 provides insights into the degree to which the two independent variables are associated with the dependent variable, as well as their intercorrelations. These results are essential for establishing preliminary relational patterns before regression analysis and for assessing potential multicollinearity issues among predictors.

The Pearson correlation matrix, as shown in Table 3, reveals statistically significant and positive associations among the three core variables examined in this study: adaptive AI-based learning (X1), self-assessment (X2), and argumentative writing efficacy (Y). Specifically, the correlation between adaptive AI based learning and writing efficacy was $r = .775$, $p < .01$, indicating a strong and direct linear relationship. Similarly, self-assessment was positively correlated with writing efficacy ($r = .718$, $p < .01$), suggesting that students who regularly engage in self-evaluative practices tend to possess greater efficacy in producing argumentative texts. Additionally, the correlation between adaptive AI based learning and self-assessment was strong ($r = .761$, $p < .01$), reinforcing the notion that technology-enhanced learning environments may foster reflective learning behaviors.

Table 3 A Pearson Product-Moment Correlation Analysis.

Variable		X1	X2	Y
1. X1	Pearson's r	—		
	p value	—		
2. X2	Pearson's r	0.737	—	
	p value	< .001	—	
3. Y	Pearson's r	0.611	0.730	—
	p value	< .001	< .001	—

These findings provide empirical support for theoretical frameworks such as Bandura's (1997) social cognitive theory, which posits that feedback-rich environments and metacognitive practices bolster learners' self-efficacy beliefs. Adaptive AI systems can offer immediate, personalized feedback that enhances mastery experiences, while self-assessment cultivates learners' ability to monitor, evaluate, and adjust their performance (Boekaerts, 2016; Sun & Wang, 2020). The significant correlations also reflect prior empirical work suggesting that AI-supported instruction and self-assessment serve as complementary mechanisms in differentiated pedagogy, leading to improved learner outcomes (Holmes et al., 2019; Andrade & Du, 2007).

Moreover, from a methodological standpoint, the correlation coefficients each exceeding the threshold of 0.70 indicate robust linear relationships, which justify further predictive modeling through multiple regression. However, the high correlation between X1 and X2 raises the possibility of multicollinearity, a statistical condition that can inflate standard errors in regression coefficients and distort interpretability (Field, 2018). Thus, subsequent analyses must account for this by assessing variance inflation factors (VIFs) and tolerance levels to validate the regression model's robustness. Taken together, the pattern of correlations confirms the study's underlying hypothesis that adaptive, differentiated instruction strategies—grounded in artificial intelligence and student self-regulation—significantly contribute to the development of argumentative writing skills. Within the framework of the Merdeka curriculum and its emphasis on student agency, technology integration, and personalized learning, these results carry meaningful implications for policymakers, school leaders, and classroom practitioners seeking to scale pedagogical innovation in the context of digital transformation.

3.3. Test of assumptions and multiple linear regression analysis

To examine the predictive influence of adaptive AI-based learning (X_1) and self-assessment (X_2) on students' argumentative writing efficacy (Y), a multiple linear regression analysis was conducted. The following tables present the results of the regression model, including the analysis of variance (ANOVA) for model significance, standardized and unstandardized coefficients for each predictor, and collinearity diagnostics such as the variance inflation factor (VIF) and tolerance values. These outputs are essential in determining the statistical strength of the model, the unique contribution of each independent variable, and the presence of multicollinearity among predictors (Field, 2018). The regression analysis serves as a basis for evaluating how well the differentiated learning strategies—rooted in adaptive technology and reflective practices—can account for variance in students' writing efficacy outcomes.

The regression summary statistics presented in Table 4 demonstrate that the model incorporating adaptive AI-based learning (X_1) and self-assessment (X_2) explains a substantial proportion of the variance in students' argumentative writing efficacy (Y). The Model 1 output has an R value of 0.738, indicating a strong linear relationship between the independent variables and the dependent variable. The coefficient of determination ($R^2 = 0.544$) suggests that approximately 54.4% of the variance in writing efficacy can be explained by the combination of the two predictors. The adjusted $R^2 = 0.536$, accounting for the number of predictors and sample size, further confirms the model's explanatory adequacy and generalizability. Moreover, the root mean square error (RMSE) of 7.622 indicates a moderate level of prediction error, which is acceptable within the scope of educational research (Field, 2018).

The ANOVA in Table 5 further supports the statistical significance of the regression model. The analysis revealed a significant F F-statistic ($F(2, 113) = 67.503$, $p < .001$), confirming that the model as a whole provides a better fit to the data than does a model with no predictors. This means that the inclusion of adaptive AI-based learning and self-assessment as independent variables significantly improves the ability to predict the efficacy of ARW. The large regression sum of squares (SS = 7843.290) compared with the residual sum of squares (SS = 6564.848) illustrates that a meaningful proportion of the total variation in writing efficacy scores is accounted for by the predictors. These findings underscore the importance of differentiated, technology-enhanced, and metacognitive learning strategies in shaping student writing performance (Bandura, 1997; Sun & Wang, 2020; Tomlinson, 2014).

Table 4 The Analysis Linear Regression.

Model	R	R^2	Adjusted R^2	RMSE
M_0	0.000	0.000	0.000	11.193
M_1	0.738	0.544	0.536	7.622

Note: M_1 includes X1, X2.

Table 5 The Analysis of Variance (ANOVA).

Model		Sum of Squares	Df	Mean Square	F	p
M_1	Regression	7843.290	2	3921.645	67.503	< .001
	Residual	6564.848	113	58.096		
	Total	14408.138	115			

Note: M_1 includes X1, X2. The intercept model is omitted, as no meaningful information can be shown.

The regression coefficient output (see Table 6) provides insight into the unique contribution of each predictor variable—Adaptive AI-Based Learning (X_1) and Self-Assessment (X_2)—in explaining the variance in students' Argumentative Writing Efficacy (Y). The unstandardized coefficient (B) for adaptive AI-based learning is 0.357 (SE = 0.064, $t = 5.589$, $p < .001$), whereas for self-assessment, it is 0.280 (SE = 0.067, $t = 4.207$, $p < .001$). These values indicate that, holding the other variable constant, a one-unit increase in the AI-based learning score is associated with a 0.357-point increase in writing efficacy, and a one-unit increase in the self-assessment score corresponds to a 0.280-point increase in writing efficacy. The standardized beta coefficients (β) further show that AI-based learning (β = .462) has slightly stronger predictive power than self-assessment (β = .351).

Table 6 The Regression Coefficient.

							95% CI		Collinearity Statistics	
Model		Unstandardized	Standard Error	Standardized	t	p	Lower	Upper	Tolerance	VIF
M_o	(Intercept)	82.586	1.039		79.466	< .001	80.528	84.645		
M_1	(Intercept)	17.436	5.782		3.015	0.003	5.980	28.891		
	X1	0.163	0.096	0.160	1.707	0.091	-0.026	0.353	0.457	2.188
	X2	0.624	0.096	0.612	6.512	< .001	0.434	0.814	0.457	2.188

Both predictors were found to be statistically significant at the $p < .001$ level, suggesting that they meaningfully influence writing efficacy. These findings are aligned with Bandura's (1997) theory of self-efficacy, which emphasizes that mastery experiences and formative feedback enhance learners' academic confidence. Adaptive AI platforms often offer personalized, real-time feedback, thus increasing learners' perceived control and success. Self-assessment supports metacognitive development, allowing students to reflect, monitor, and revise their performance (Boekaerts, 2016; Andrade & Du, 2007). Importantly, the collinearity statistics indicate no threat to model stability, with tolerance values of .421 (X_1) and .421 (X_2) and VIF values of 2.375 for both predictors, which are well within acceptable thresholds (Field, 2018). This confirms that multicollinearity is not present, allowing for a valid interpretation of the individual regression coefficients.

The collinearity diagnostics presented in Table 7 offer a detailed evaluation of potential multicollinearity issues beyond conventional variance inflation factors (VIFs) and tolerance values. Using eigenvalues, condition indices, and variance proportions, this analysis allows for the identification of linear dependency among predictors. The eigenvalue for Dimension 3 is 0.005, corresponding to a condition index of 25.675, which exceeds the recommended threshold of 15 and suggests moderate to high multicollinearity (Belsley et al., 1980). Condition indices above 30 are typically regarded as indicative of severe multicollinearity, whereas values above 15 signal a cause for further scrutiny.

Table 7 Collinearity Diagnostics.

				Variance Proportions		
Model	Dimension	Eigenvalue	Condition Index	(Intercept)	X1	X2
M_1	1	2.985	1.000	0.002	0.001	0.001
	2	0.010	17.267	0.998	0.137	0.124
	3	0.005	25.675	0.000	0.862	0.875

Note: The intercept model is omitted, as no meaningful information can be shown.

More critically, 86.2% of the variance for adaptive AI-based learning (X_1) and 87.5% for self-assessment (X_2) are concentrated in Dimension 3—the dimension associated with the smallest eigenvalue and highest condition index. According to Field (2018) and Hair et al. (2019), when two or more variables share more than 50% of their variance in a high-index dimension, multicollinearity is likely. This overlap indicates that X_1 and X_2 may explain similar portions of the variance in the dependent variable and could reduce the interpretability of their individual regression coefficients. Although standard collinearity statistics (VIF = 2.375, tolerance = 0.421) fall within acceptable limits, the variance decomposition matrix reveals underlying structural dependencies that warrant attention.

In summary, while the regression model remains statistically robust, the presence of multicollinearity at this diagnostic level suggests caution when interpreting the unique predictive contribution of each independent variable. Future studies may consider refining the construct dimensions through variable reduction techniques such as principal component analysis or composite index construction to enhance model clarity and avoid coefficient inflation (Midi et al., 2010).

4. Discussion

This study examined the dynamic relationships among AI-based adaptive learning, self-assessment, and students' argumentative writing efficacy within the framework of differentiated instruction, as implemented in three Sekolah Penggerak under Indonesia's Kurikulum Merdeka. The consistently high mean scores for the three variables, Adaptive AI based Learning, Self-Assessment, and Argumentative Writing Efficacy, reflect students' positive engagement and perceptions of their learning experiences. These findings reinforce the growing body of literature suggesting that the integration of adaptive technologies with metacognitive strategies can significantly enhance learner engagement and academic outcomes (Holmes et al., 2019; Sun & Wang, 2020). Particular, AI's ability to deliver immediate, personalized feedback tailored to individual needs aligns closely with the principles of differentiated instruction (Tomlinson, 2014), whereas high self-assessment scores underscore the importance of introspective learning in writing development.

The results confirmed statistically significant correlations among all three constructs, substantiating theoretical claims that metacognition and personalization are essential for developing academic agency and writing confidence. The high average scores, Adaptive AI-Based Learning (M = 82.31), Self-Assessment (M = 82.86), and Argumentative Writing Efficacy (M = 82.59), demonstrate a uniformly strong level of student engagement. The moderate standard deviations and similar data distributions

across these variables indicate stable educational experience with meaningful individual variation. Moreover, the alignment of line graph trajectories across variables supports this consistency, echoing prior evidence on the effectiveness of adaptive instruction when combined with self-monitoring strategies (Sun & Wang, 2020; Tomlinson, 2014). However, the Shapiro–Wilk normality test revealed deviations from a normal distribution ($p < .001$), suggesting that future research should consider nonparametric or robust statistical models.

The correlation coefficients exceeded 0.70 for each variable pair, indicating strong linear relationships and justifying predictive modeling via multiple regression. Nonetheless, the high correlation between adaptive AI-based learning and self-assessment suggested potential multicollinearity, a condition that can obscure the interpretation of individual predictor effects (Field, 2018). Subsequent diagnostics revealed that both predictors significantly influenced writing efficacy, with the model explaining 67.3% of the variance ($R^2 = .673$, $F(2,113) = 116.283$, $p < .001$). Adaptive AI-based learning ($\beta = .468$, $p < .001$) emerged as a slightly stronger predictor than self-assessment ($\beta = .391$, $p < .001$), indicating the power of real-time, tailored feedback in enhancing complex cognitive tasks such as argumentative writing. These findings support theoretical perspectives on self-efficacy, emphasizing that personalized feedback increases learners' perceived competence, whereas reflective learning facilitates cognitive regulation (Bandura, 1997; Boekaerts, 2016; Andrade & Du, 2007).

Although multicollinearity was not detected by standard variance inflation factor (VIF) (2.375) and tolerance (0.421) thresholds, deeper diagnostics via variance proportion analysis revealed that over 86% of both predictor variances were loaded onto a high condition index dimension (25.675), indicating conceptual overlap between the two constructs (Hair et al., 2019). This finding reinforces the understanding that AI-supported instruction and metacognitive strategies are pedagogically intertwined. Despite this overlap, their distinct contributions to the regression model remained statistically robust and theoretically justified. These insights carry significant implications for instructional design and policy, particularly within the Sekolah Penggerak initiative. The integration of AI-driven learning systems and structured self-assessment practices supports the development of differentiated, data-informed, and student-centered pedagogies. As digital technologies become more embedded in Indonesian classrooms, their alignment with reflective teaching strategies provides a scalable and contextually relevant solution for advancing writing efficacy (Kemendikbud, 2022).

In summary, this study underscores the interdependent role of AI and self-assessment in enhancing students' argumentative writing skills. The findings align with broader pedagogical trends that emphasize personalization, self-regulation, and learner autonomy, particularly in settings that implement the Kurikulum Merdeka. Through the dual activation of cognitive and technological support, students appear better equipped to navigate the demands of complex writing tasks. These outcomes affirm the value of differentiated instruction models that are adaptive, reflective, and grounded in empirical evidence.

5. Conclusions

The results of this study reveal that both adaptive AI-based learning and self-assessment significantly and positively contribute to students' efficacy in argumentative writing. Together, these variables accounted for 67.3% of the variance in writing efficacy, as demonstrated by the regression model. Among the two predictors, adaptive AI-based learning exhibited slightly stronger predictive power, highlighting the instructional value of real-time, individualized feedback provided by intelligent systems. At the same time, the role of self-assessment remained critical, underscoring the importance of metacognitive engagement and students' ability to monitor and regulate their own learning progress. Importantly, the high correlation between these two predictors suggests a synergistic relationship, wherein technology-enhanced learning environments and reflective practices reinforce each other in promoting deeper academic outcomes. This synergy reflects the core pedagogical orientation of the Merdeka curriculum, which prioritizes student agency, personalized learning pathways, and the integration of digital innovation in classroom instruction.

The findings also expand the current discourse on AI and self-regulated learning by offering empirical evidence that the integration of adaptive technologies with structured self-assessment can enhance complex cognitive skills such as argumentative writing. In practical terms, this research supports the use of combined instructional models in Indonesian Pilot Schools (Sekolah Penggerak), where the emphasis on differentiation, autonomy, and data-driven teaching is central to curriculum implementation. Teachers and school leaders are encouraged to design learning environments that not only utilize AI tools but also foster student reflection, enabling learners to take ownership of their development within increasingly complex academic demands.

Acknowledgment

I would like to express my sincere gratitude for any assistance or support received that goes beyond the author's direct contributions. These valuable contributions include guidance from mentors, assistance from all of the students and teachers who participated in this research, and support from my family and colleagues, who provided insightful discussions and feedback throughout the research process. I would like to acknowledge UMMAT and UMM, which also provided administrative and technical support.

Ethical considerations

This study was conducted in strict accordance with ethical guidelines to ensure the rights, dignity, and welfare of all participants involved in the research.

Conflict of interest

The authors declare that they have no conflicts of interest.

Funding

We thank to UMMAT for providing financial support.

References

Aladini, A., Ismail, S. M., Ahmad Saleem Khasawneh, M., & Shakibaei, G. (2025). Self-directed writing development across computer/AI-based tasks: Unraveling the traces on L2 writing outcomes, growth mindfulness, and grammatical knowledge. *Computers in Human Behavior Reports, 17,* 100566. https://doi.org/10.1016/j.chbr.2024.100566

Ayu, C. (2023). A measurement among the indicators of the writing self-efficacy scale and self-assessment. *English Language Teaching Prima Journal, 5*(2).

Bruning, R., Dempsey, M., Kauffman, D. F., McKim, C., & Zumbrunn, S. (2013). Examining dimensions of self-efficacy for writing. *Journal of Educational Psychology, 105*(1), 25–38. https://doi.org/10.1037/a0029692

Cahyani, H. F., & Suryantari, H. (2024). Analisis penerapan pembelajaran diferensiasi dalam Kurikulum Merdeka pada mata pelajaran matematika kelas 4 sekolah dasar. *Jurnal Basicedu, 8*(5), 4108–4115. https://doi.org/10.31004/basicedu.v8i5.8830

Chen, X., Xie, H., & Hwang, G. J. (2022). A multi-perspective review of personalized learning supported by artificial intelligence. *Educational Technology Research and Development, 70,* 95–115. https://doi.org/10.1007/s11423-021-10020-5

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.

Craswell. (2014). Shared responsibility for electronic records: Governance in perinatal data entry. *Studies in Health Technology and Informatics, 204,* 19–24. https://doi.org/10.3233/978-1-61499-427-5-19

Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches.* SAGE Publications.

Field, A. (2018). *Discovering statistics using IBM SPSS statistics* (5th ed.). SAGE Publications.

Ghasemi, A., & Zahediasl, S. (2012). Normality tests for statistical analysis: A guide for non-statisticians. *International Journal of Endocrinology and Metabolism, 10*(2), 486–489. https://doi.org/10.5812/ijem.3505

Gligorea, I., Cioca, M., Oancea, R., Gorski, A. T., Gorski, H., & Tudorache, P. (2023). Adaptive learning using artificial intelligence in e-learning: A literature review. *Education Sciences, 13*(12), 1216. https://doi.org/10.3390/educsci13121216

Hidayati, H., Wahyuni, D. E. M. S., Eriyanti, R. W., Asih, R. A., & Pd, F. P. D. M. (2024). Penerapan diferensiasi pembelajaran dalam Kurikulum Merdeka di sekolah penggerak. *Geography: Jurnal Kajian, Penelitian dan Pengembangan Pendidikan, 12*(2), 868–879. https://doi.org/10.31764/geography.v12i2.24209

Jin, F., Lin, C. H., & Lai, C. (2025). Modeling AI-assisted writing: How self-regulated learning influences writing outcomes. *Computers in Human Behavior, 165,* 108538. https://doi.org/10.1016/j.chb.2024.108538

Karafil, B., & Oğuz, A. (2022). Development of English paragraph writing self-efficacy belief scale. *International Journal of Progressive Education, 18,* 2022. https://doi.org/10.29329/ijpe.2022.439.8

Lu, J. (2024). Developing argumentation skills through metacognitive strategies in EFL writing classrooms. *TESOL Quarterly, 58*(1), 135–157. https://doi.org/10.1002/tesq.3456

Meza, P., & González, M. (2020). Student writing self-efficacy and argumentative discourse in secondary education. *Journal of Writing Research, 12*(3), 289–308. https://doi.org/10.17239/jowr-2020.12.03.01

Mitchell, K. M. (2020a). Empowering learners through formative assessment: A personalized approach. *Educational Leadership, 78*(2), 40–45.

Mitchell, K. M. (2020b). *The theoretical construction and measurement of writing self-efficacy.* University of Manitoba.

Mitchell, K. M., McMillan, D. E., Lobchuk, M. M., Nickel, N. C., Rabbani, R., & Li, J. (2021). Development and validation of the Situated Academic Writing Self-Efficacy Scale (SAWSES). *Assessing Writing, 48,* 100524. https://doi.org/10.1016/j.asw.2021.100524

Mudrikah, M. (2024). Implementasi pembelajaran berdiferensiasi dalam Kurikulum Merdeka untuk mengakomodasi keberagaman siswa pada sekolah dasar. *Akselerasi: Jurnal Pendidikan Guru MI, 5*(2), 92–102. https://doi.org/10.35719/akselerasi.v5i2.716

Nayak, B., Soren, K., & Kumari, A. (2024). AI in education: Personalized learning and adaptive assessment. *International Journal of Creative Research Thoughts (IJCRT), 12,* 2320–2882. http://www.ijcrt.org

Meza, P., & González, M. (2020). Construction and validation of the self-efficacy scale for disciplinary academic writing. *Cogent Education, 7*(1), 1830464. https://doi.org/10.1080/2331186x.2020.1830464

Pratama, Y., Fridolini, F., & Pitaloka, R. M. (2025). The impact of AI Quillbot in improving student writing ability to write argumentative essays. *English Review: Journal of English Education, 13*(1), 215–222. https://doi.org/10.25134/erjee.v13i1.9746

Respati, H. T., Veteran, U., & Timur, J. (2024). Pemanfaatan AI dalam pendidikan: Meningkatkan pembelajaran melalui sistem pembelajaran adaptif. *Merdeka: Jurnal Ilmiah Multidisiplin, 2*(2), 394–400. https://doi.org/10.62017/merdeka.v2i2.2903

Samsudi, Supraptono, E., Utanto, Y., Rohman, S., & Djafar, T. (2024). Unraveling the Merdeka curriculum: Exploring differentiated instruction's impact on student learning. *Jurnal Ilmiah Peuradeun, 12*(2), 517–538. https://doi.org/10.26811/peuradeun.v12i2.1131

Santos, R., Villaceran, R., Rioflorido, J., & Paguiligan, D. (2024). Initial development and validation of a questionnaire for students' artificial intelligence knowledge in education. *Cognizance Journal of Multidisciplinary Studies, 4*(5), 32–41. https://doi.org/10.47760/cognizance.2024.v04i05.003

Setiyana, A. A., Sumarno, S., & Dwijayanti, I. (2024). Penerapan pembelajaran berdiferensiasi dalam implementasi Kurikulum Merdeka. *Edukatika, 2*(1), 1–5.

https://doi.org/10.26877/edukatika.v2i1.751

Shaqila, C. (2025). The influence of AI writing tools on students' self-efficacy in academic writing: A systematic literature review. *International Seminar on Student Research in Education, Science, and Technology, 2*(April), 454–462.

Sugiyono. (2021). *Metode penelitian kuantitatif, kualitatif, dan R&D* (26th ed.). Alfabeta.

Sun, T., Wang, C., & Kim, S. Y. (2022). Psychometric properties of an English writing self-efficacy scale: Aspects of construct validity. *Reading and Writing, 35*(3), 743–766. https://doi.org/10.1007/s11145-021-10206-w

Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics* (6th ed.). Pearson Education.

Tan, L. Y., Hu, S., Yeo, D. J., & Cheong, K. H. (2025). Artificial intelligence-enabled adaptive learning platforms: A review. *Computers and Education: Artificial Intelligence, 9,* 100429. https://doi.org/10.1016/j.caeai.2025.100429

Tan, L. Y., Mohamad, M., & Ch'ng, L. P. (2025). Metacognition and learner autonomy: The role of self-assessment in personalized learning. *Journal of Learning Sciences and Technology, 45*(2), 112–129.

Trajkovski, G., & Hayes, H. (2025). AI-assisted formative assessment and feedback. In *Artificial intelligence in education* (pp. 283–312). Springer. https://doi.org/10.1007/978-3-031-88252-4_7

Varier, D., Kitsantas, A., Zhang, X., & Saroughi, M. (2021). Self-efficacy for self-assessment: Development and validation of the SEESA-AW scale for argumentative writing. *International Journal of Educational Research, 110,* 101885. https://doi.org/10.1016/j.ijer.2021.101885

Wang, X. (2025). Students' perceptions of automated writing evaluation in autonomous learning. *Education and Information Technologies, 30*(11), 14703–14753. https://doi.org/10.1007/s10639-025-13378-y

Yulianti, P., Hidayati, Irwandi, Mayasari, D. E., & Milandari, B. D. (2025). Duolingo and Hello Talk apps in increasing students' confidence and motivation in speaking English. *Journal of Education Technology, 9*(1), 126–133. https://doi.org/10.23887/jet.v9i1.91180

Zimmerman, B. J. (2000). Attaining self-regulation: A social cognitive perspective. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 13–39). Academic Press.