IDEAS

Journal on Language Teaching and Learning, Linguistics and Literature

ISSN 2338-4778 (Print)
ISSN 2548-4192 (Online)

Volume 14, Number 2, December 2026
pp. 9279 - 9300

Copyright © 2026 The Author
IDEAS is licensed under CC-BY-SA 4.0 License



Issued by English study program of UIN Palopo

Transforming Reading into Deep Learning: Evaluating an AI-Based Adaptive Reading System for Generation Z Learners

Mulyarti[1], Muhammad Danial[2], Sry Wahyuni[3], Triadi Setiawan[4]

[1, 2 ,4]Universitas Muhammadiyah Barru, Barru, Indonesia
[3]Universitas Negeri Makassar, Makassar, Indonesia

Article Info

Received: 2026-10-28
Revised: 2026-11-04
Accepted: 2026-11-05

Keywords:
Adaptive Learning;
AI-Based Adaptive Reading System;
Deep Learning;
Generation Z;
Reading Activities.

DOI:
10.24256/ideas. v14i2. 13030

Corresponding Author:
Mulyarti
mulyarti14@gmail.com
Universitas Muhammadiyah Barru, Barru, Indonesia

Abstract

This study evaluates the implementation of an Artificial Intelligence (AI)-based adaptive reading system that uses ChatGPT and examines reading comprehension gains and the predictors of self-reported deep learning among Generation Z senior high school students in Indonesia. A key feature of the system is its flexible and interactive discussion assistant, particularly its ability to paraphrase and adapt language use based on user needs. The study used a quantitative approach with a one-group pre-test–post-test design without a control group. The participants were 45 students from three senior high schools in Barru Regency, selected through purposive sampling. Data were collected using a reading comprehension test (pre-test and post-test) and a 30-item questionnaire on the adaptive reading system, reading activities, deep learning, and Generation Z characteristics. Data were analyzed using descriptive statistics, a paired-sample t-test, and multiple linear regression. Students responded positively to the system, with mean scores of 3.92 (adaptive reading system), 3.98 (reading activities), and 4.05 (deep learning). Reading comprehension scores improved significantly following implementation, from 80.93 to 88.02 (mean gain = 7.09, $t(44) = -10.809$, $p < .001$, Cohen's $d = 1.61$). In a separate regression analysis, the adaptive reading system and reading activities jointly accounted for 77.6% of the variance in self-reported deep learning ($R^2 = .776$), and both were significant predictors; this indicates association, not causation. Because there was no control group, the gains are interpreted as improvement following implementation and not as an effect of the system alone. The study offers preliminary evidence for AI-supported reading instruction among Indonesian secondary-school learners and indicates that personalization combined with structured reading activity may create conditions for deeper learning.

ISSN 2338-4778 (Print)
ISSN 2548-4192 (Online)

1. Introduction

Generation Z has grown up in a digital environment and is often referred to as "digital natives." Frequent exposure to video, text, and images on social media, combined with the largely unfiltered nature of online content production, has been linked to learning characteristics that differ from those of previous generations (Nasrudin, Sumarna, & Surahman, 2024; Shirazian, Vatanpour, & Guity, 2025). Grace & Seemiller (2019) found that Generation Z tends to be strong in multitasking and to prefer visual and interactive information and microlearning-based instruction. Consistent with this profile, Čekrlija (2024) reports that Generation Z learners favor content that is concise, fast-paced, and visually stimulating over conventional reading texts and gravitate toward practical exercises that can be applied immediately.

These generalizations should be treated cautiously. Members of the same generation differ considerably in reading ability, digital literacy, and motivation, and digital exposure does not automatically make learners digitally competent or capable of deep learning. The more defensible concern is that heavy reliance on digital media may leave many students with difficulty comprehending complex texts and with underdeveloped critical thinking skills, both of which are central to deep literacy (Delgado, Vargas, Ackerman, & Salmerón, 2018; Gerlich, 2025).

The scale of this challenge in Indonesia is evident in the 2022 Programme for International Student Assessment (PISA) results, which show that 75% of Indonesian students have not yet reached the proficient-reader category, with an average score of 359, far below the OECD average of 476 (OECD, 2023). At the upper end of the distribution, fewer than 7% of Indonesian students reached the highest literacy levels (Levels 5–6), which require comprehending complex texts, evaluating abstract ideas, and distinguishing facts from opinions (Schleicher, 2019).

These figures suggest that many students struggle with the inferential and evaluative reading that deep learning requires, which is the kind of reading that individualized, feedback-rich practice is meant to support. National education policy likewise emphasizes producing AI-based digital reading materials that are adaptive to students' characteristics (D, Khasanah, & Putri, 2022; Haider, Majeed, & Irfan, 2025; Salman, 2025). The present study responds to this need at the classroom level by developing an adaptive reading system for senior high school students and examining it in a limited trial; it does not claim to resolve the national literacy problem.

Three terms that are often used interchangeably need to be distinguished. Adaptive learning refers to instruction in which content, difficulty, and feedback change in response to a learner's performance. AI-assisted learning is the broader use of AI tools to support learning, whether or not the instruction adapts to the individual learner. Generative AI refers to models that produce new text, such as explanations, questions, and feedback, in response to learner input.

In this study, an AI-based adaptive reading system is an adaptive system in which ChatGPT, a generative AI chatbot, is used to generating comprehension question and reading difficulty and feedback are adjusted to each student's performance. Reading comprehension refers to students' performance on a reading comprehension test, whereas deep learning refers to students' self-reported critical thinking and meaningful learning during reading; the two outcomes are analyzed separately.

AI-based adaptive learning systems function by continuously assessing students' strengths, weaknesses, and learning gaps and then adjusting instructional content responsively. Hariyanto, Kristianingsih, & Maharani (2025) demonstrate that AI can adapt learning content to individual learner profiles, thereby increasing engagement and personalizing the learning experience. Intelligent tutoring systems and adaptive learning platforms use machine learning to detect knowledge gaps and provide tailored feedback, which may foster deeper understanding. By converting traditional reading activities into more individualized and responsive experiences, such systems have the potential to encourage engagement, meaningful comprehension, and critical thinking, all central components of deep learning. Yet many educational institutions have not incorporated such technology into teaching and learning, and evidence on its use in secondary-school reading instruction remains limited (Mulyarti, Idham, Danial, & Rauf, 2026).

A number of studies report promising results for AI in learning. AI-based methods have been shown to improve student performance by up to 45% and learning outcomes by up to 32% compared with conventional methods (Patnaik, Baid, Sailaja, Chandar, & Sai, 2025; Sakhipov, Tukenova, & Omirzak, 2025; Sari, Tumanggor, & Efron, 2024). Gains in reading comprehension have been reported particularly for students with low initial reading proficiency who used adaptive AI-based platforms, relative to a control group (Etkin, Etkin, Carter, & Rolle, 2025; Silor & Silor, 2025).

A two-tiered problem-based learning approach that integrates generative AI has similarly been shown to encourage students to connect new information with prior experience, thereby strengthening critical thinking and deep understanding (Gou & Ren, 2025; Scholkmann, Schurer, & Hviid, 2025; Tobondo et al., 2024). These findings suggest that AI can function not merely as a content provider but as a facilitator of exploration, reasoning, and cognitive-skill development, provided it is designed appropriately.

At the same time, other findings point to negative effects of excessive AI use. High-achieving students have shown slightly slower progress when relying too heavily on AI-generated summaries, suggesting that easy access to AI may reduce the deep cognitive effort that critical understanding requires (Özer, Tanberkan, & Perc, 2025; Silor & Silor, 2025). The effectiveness of AI in learning therefore depends heavily on how it is integrated into instruction: not as a substitute for the thinking process, but as a scaffold that encourages inference, synthesis, and critical reflection rather than passive consumption of AI-generated content. Whether Generation Z learners benefit from such systems is an empirical question and should not be assumed, because familiarity with digital media does not guarantee the ability to read deeply with it. The move from reading to deep learning is possible, but it is not automatic and requires thoughtful pedagogical integration.

Four theoretical perspectives inform the design of the adaptive reading system, and Table 1 summarizes how each is operationalized: what the system does, which pedagogical mechanism is expected to operate, and how the resulting change in learner behavior should lead to deeper learning. Constructivist theory and Vygotsky (1978) notion of the Zone of Proximal Development hold that meaningful learning occurs when students actively construct their understanding through interaction with their learning environment, aided by scaffolding tailored to their ability level (Mishra, 2023; Sumarna & Gunawan, 2022).

ISSN 2338-4778 (Print)
ISSN 2548-4192 (Online)

Because human working memory capacity is limited (Sweller, 1988), Cognitive Load Theory holds that effective instructional design must manage extraneous cognitive load so that learners' cognitive resources remain focused on deep, or germane, processing (Liu, 2025); adaptive systems are assumed to reduce extraneous load by presenting material according to students' readiness levels (Barrón Pérez & Morales Piña, 2026; Mahdi, Haseeb, Rafiq, & Usman, 2025).

Self-Regulated Learning theory by Zimmerman (2002) explains how learners monitor, evaluate, and adjust their own strategies, and it also accounts for the overreliance risk noted above: when cognitive processes are fully "delegated" to AI, high-achieving students may have fewer opportunities to practice metacognitive strategies (Özer et al., 2025; Pakili, Yunus, Novian, & Bouty, 2025). Finally, Adaptive Learning Theory describes effective learning as a continuous cycle of assessment, content adaptation, and feedback (Atun, 2020; Erümit & Çetin, 2020; Huu, Tangworakitthaworn, & Gilbert, 2022), which frames the architecture of the system proposed in this study.

Table 1. Alignment between Theoretical Perspectives and the Design of the Adaptive Reading System

Theory	**What the system does**	**Expected pedagogical mechanism**	**Expected change in learner behavior and learning outcome**
Constructivism and Zone of Proximal Development	Adjusts reading difficulty and recommends texts according to each student's performance on the comprehension checks after each text, across [three] difficulty levels (easier, same, more difficult)	Material stays within the learner's zone of proximal development, with scaffolding adjusted to ability	Students work with texts that are challenging but manageable and connect new information with prior knowledge, which supports meaningful learning
Cognitive Load Theory	Presents reading content in units matched to readiness (one text at a time, followed by a short comprehension check of [five] items; two texts per session)	Extraneous load is reduced so that working memory can be devoted to germane processing	Students invest effort in inference and in integrating ideas, not in coping with unsuitable text, which supports critical thinking
Self-Regulated Learning	Gives automatic feedback and records performance and reading behavior in system logs (immediate ChatGPT-generated explanations of each answer; students see their check score and the explanation, while time-on-task, difficulty progression, and check results are logged)	Performance becomes visible, which supports monitoring, evaluation, and strategy adjustment	Students monitor their own comprehension and adjust reading strategies instead of delegating thinking to the system
Adaptive Learning Theory	Runs a cycle of comprehension checks, content adaptation, and feedback (a score of	The learning path changes continuously in response to performance	Each student receives material and feedback that follow from previous performance, which

Theory	What the system does	Expected pedagogical mechanism	Expected change in learner behavior and learning outcome
	[≥80]% moves the student up one level, [60–79]% keeps the same level with a new text, and below [60]% moves the student down one level)		supports sustained engagement

Source: Authors' synthesis.

These perspectives yield the following conceptual framework. The adaptive reading system is expected to support deep learning through content personalization, scaffolding, and feedback, which are hypothesized mediating mechanisms, and the level of AI dependence is hypothesized to moderate this relationship. Structured reading activity is expected to contribute to deep learning alongside the system. Only part of this framework was tested empirically. The study measured the adaptive reading system (X1), reading activities (X2), and self-reported deep learning (Y) and tested X1 and X2 as direct predictors of Y. Personalization, scaffolding, feedback, and AI dependence were not measured as separate variables; mediation and moderation are therefore treated as hypothesized mechanisms for future research and not as tested relationships. Generation Z is treated as the learner context of the study and is described through a separate descriptive scale; it is not modeled as a moderator.

Taken together, this body of research indicates that although AI holds great potential to support deep literacy, its success is not guaranteed and depends heavily on pedagogical design. Most prior research has examined the general effectiveness of AI in learning, and studies of adaptive reading have typically focused on comprehension scores as the main outcome (Etkin et al., 2025; Silor & Silor, 2025), with adaptive systems commonly adjusting text difficulty using indicators such as comprehension accuracy, reading speed, and vocabulary knowledge. At least three aspects remain underexamined: (a) the combination of generative-AI feedback with adaptive adjustment of reading difficulty in a single system; (b) the distinction between reading comprehension gains and deep learning, that is, critical thinking and meaningful learning, as separate outcomes (Etkin et al., 2025; Silor & Silor, 2025); and (c) the joint role of an adaptive system and the reading activities that learners undertake with it, particularly among Indonesian senior high school students (Ariqo, Rizal, & Winoto, 2025; Mulyarti, 2025; Sa'diyah & Savitri, 2024).

This study addresses these aspects by implementing an adaptive reading system that integrates generative AI (ChatGPT) and combines immediate, explanatory feedback on each answer with comprehension-check-based adjustment of text difficulty and logged tracking of students' performance within one system. It also evaluates reading comprehension gains and the predictors of self-reported deep learning among senior high school students.

Therefore, this paper aims to evaluate the implementation of an AI-based adaptive reading system and to examine reading comprehension gains and the extent to which the system and reading activities are associated with deep learning among Generation Z learners. It addresses three research questions: (RQ1) Do students' reading comprehension scores differ significantly before and after implementation of the system?

ISSN 2338-4778 (Print)
ISSN 2548-4192 (Online)

(RQ2) How do students respond to the adaptive reading system, reading activities, deep learning, and Generation Z characteristics? (RQ3) To what extent do the adaptive reading system (X1) and reading activities (X2) predict students' self-reported deep learning (Y)? Corresponding to RQ1 and RQ3, the study tests four hypotheses: H1, students' reading comprehension scores differ significantly before and after using the system; H2, X1 significantly predicts Y; H3, X2 significantly predicts Y; and H4, X1 and X2 together significantly predict Y.

2. Method

This study employed a quantitative approach to evaluate the implementation of an Artificial Intelligence (AI)-based Adaptive Reading System that uses ChatGPT to support reading among Generation Z learners. It examined changes in students' reading comprehension before and after implementation and the relationship between students' questionnaire responses on the system, reading activities, and self-reported deep learning.

Research Design

The study used a pre-experimental, one-group pre-test–post-test design. A single group of students completed a reading comprehension pre-test, used the AI-based adaptive reading system, and then completed a post-test and a questionnaire on their experience. The questionnaire data were analyzed with multiple linear regression to examine whether the adaptive reading system and reading activities predicted self-reported deep learning.

The study had no control or comparison group, and participants were not randomly assigned. The pre-test–post-test comparison is therefore interpreted as change following implementation and not as proof that the system alone caused the change, and the regression results are interpreted as predictive associations. This article reports quantitative data only, that is, test scores and questionnaire responses.

The AI-Based Adaptive Reading System

AI technology. The system uses ChatGPT, a generative AI chatbot developed by OpenAI to generate explanatory feedback on students' answers to the comprehension checks [and the comprehension-check questions, if ChatGPT also produced them]. ChatGPT was configured through a fixed prompt specifying that the users were senior high school EFL students, [the response language], and the instruction to explain why an answer was correct or incorrect.

Adaptation mechanism. Reading difficulty is determined by students' scores on the comprehension checks that follow each text. Students who reached [≥80]% received a more difficult text, those at [60–79]% stayed at the same level with a new text, and those below [60]% received an easier text. The feedback generated by the system is explanation and is given immediately. Students' performance, time-on-task, difficulty progression, and comprehension-check results are recorded in automated system logs, and these records determine the materials and feedback that follow.

Participants

The research was conducted at three senior high schools in Barru Regency: SMAN 1 Barru, SMAN 3 Barru, and SMAN 6 Barru. These schools were chosen because their students vary in English proficiency and the schools have sufficient digital infrastructure. The participants were 45 students (15 from SMAN 1 Barru, 15 from SMAN 3 Barru, and 15 from SMAN 6 Barru), and one English teacher from each school (three teachers in total) acted as practitioners. All 45 students completed the pre-test, the post-test, and the questionnaire. Purposive sampling, a non-probability method that selects participants according to predetermined criteria to ensure the relevance of the data, was used, with the following criteria: (a) active students of SMAN 1, 3, or 6 Barru enrolled in English lessons, (b) able to access the system on a [device] with internet connection, and (c) willing to join the four-week trial and complete the pre-test, post-test, and questionnaire.

Implementation Procedure

The trial lasted 4 weeks and consisted of 2 sessions of 40 minutes each, held outside class hours. In each session, students read 2 texts selected by the system, completed comprehension-check questions after each text and read the ChatGPT-generated explanation of each answer. The reading activities (X2) refer to students' engagement with the digital texts: identifying main ideas, connecting concepts, reflecting on the information obtained, and taking part actively and independently in reading, which are the dimensions measured by the questionnaire. The teacher's role was to orient students before the first session, ensure that devices and internet access were available, and monitor the sessions without giving answers; the teachers were briefed on how the system works. The pre-test was administered [within one week] before the first session and the post-test [within one week] after the last session.

Instruments

The research instruments consisted of a reading comprehension test and a questionnaire. Table 2 summarizes each instrument, the variable it measures, and the analysis in which it was used.

Table 2. Overview of the Research Instruments and Their Role in the Analysis

Instrument	Variable or purpose	Items and scale	Source	Role in the analysis
Reading comprehension test (pre-test and post-test)	Reading comprehension	5 items (it depends on Chat GPT algorithm)	Developed by Chat GPT	Paired-sample t-test (H1)
Questionnaire: Adaptive Reading System (X1)	Students' perception of the system	12 items; 5-point Likert	Developed	Descriptive statistics; predictor in regression (H2, H4)
Questionnaire: Reading Activities (X2)	Students' reading engagement and comprehension activities	7 items; 5-point Likert	Developed	Descriptive statistics; predictor in regression (H3, H4)

ISSN 2338-4778 (Print)
ISSN 2548-4192 (Online)

Instrument	Variable or purpose	Items and scale	Source	Role in the analysis
Questionnaire: Deep Learning (Y)	Self-reported critical thinking and meaningful learning	6 items; 5-point Likert	Developed	Descriptive statistics; outcome in regression (H2–H4)
Questionnaire: Generation Z Characteristics	Digital learning styles and technology preferences	5 items; 5-point Likert	Developed	Descriptive statistics only (learner context)

Source: Authors' compilation.

The questionnaire consists of four focus areas: the AI-based adaptive reading system (X1), reading activities (X2), deep learning (Y), and the characteristics of Generation Z. The first three were analyzed inferentially, with X1 and X2 as independent variables and Y as the dependent variable, whereas the characteristics of Generation Z were analyzed descriptively and describe the learner context.

The first focus area, the AI-Based Adaptive Reading System (X1), has three dimensions: requirements analysis, system flexibility, and artificial intelligence characteristics. The requirements analysis dimension seeks to determine the reading materials that users require in accordance with their learning preferences, interests, skill level, and understanding. The system's capacity to modify reading difficulty levels, make recommendations automatically, and customize information according to user behavior and engagement is measured by the system adaptability dimension. Meanwhile, the AI features dimension is intended to ascertain users' requirements for automatic feedback, reading behavior analysis, and personalized reading.

The second focus is Reading Activities (X2), which consists of the dimensions of reading comprehension and reading engagement. The reading comprehension dimension identifies the extent to which reading materials can help users understand main ideas, connect various concepts, and reflect on the information obtained. The reading engagement dimension, meanwhile, aims to determine the level of interest, independence, and active participation of users in interacting with digital reading materials.

The third focus is Deep Learning (Y), which encompasses the dimensions of critical thinking and meaningful learning. The critical thinking dimension is designed to measure users' ability to analyze and evaluate the information contained in a text. Meanwhile, the meaningful learning dimension identifies users' ability to connect reading material with prior experiences or knowledge and to build understanding that endures over a longer period. Deep learning is therefore measured through students' self-reported responses and not through a performance test.

The final focus is on the characteristics of Generation Z, which are measured descriptively and include the dimensions of digital learning styles and technology preferences. The digital learning styles dimension is used to determine users' preferences for reading materials that are concise, visual, interactive, and responsive. Meanwhile, the technology preferences dimension aims to identify users' comfort levels with AI-based learning systems as well as their perceptions of the benefits of reading recommendations automatically provided by the system.

Based on these focus areas, two instruments were used: a reading comprehension test, administered as a pre-test and a post-test by 5 items but the test depends from chat GPT algorithm by students, and a 30-item questionnaire with a five-point Likert scale (1–5) measuring X1 (12 items), X2 (7 items), Y (6 items), and Generation Z characteristics (5 items). The instruments were tested for validity and reliability using SPSS prior to the analysis. Item validity and reliability were tested on the responses of the 45 participants using corrected item-total correlations and Cronbach's alpha for all four questionnaire scales, including the Generation Z characteristics scale.

Data Collection Technique

Quantitative data were collected in three steps. Before the intervention, all 45 participants completed the reading comprehension pre-test to establish baseline scores. During the intervention, students' use of the AI-based adaptive reading system was tracked through automated system logs, capturing time-on-task, difficulty progression, and comprehension-check results. After the intervention, the same participants completed a parallel form of the reading comprehension test as a post-test and the questionnaire on X1, X2, Y, and Generation Z characteristics, after that the questionnaire was administered. No control or comparison group was used, so all 45 participants received the same treatment.

Data Analysis Technique

Data were analyzed quantitatively using SPSS 26.

Questionnaire responses were analyzed descriptively using mean scores and percentages, interpreted with the criteria in Table 3, to describe students' responses on X1, X2, Y, and Generation Z characteristics. Item validity was tested using corrected item-total correlations against the critical r-value at the 5% significance level, and reliability using Cronbach's alpha ($\alpha > 0.60$). Normality was tested using the Kolmogorov–Smirnov and Shapiro–Wilk tests. Reading comprehension scores before and after implementation were compared using a paired-sample t-test, and the effect size was expressed as Cohen's d for paired samples. The contributions of X1 and X2 to self-reported deep learning (Y) were examined using multiple linear regression, after checking the normality of the residuals and multicollinearity (tolerance and VIF). The significance level was set at .05. Regression results are interpreted as predictive associations and not as causal effects.

3. Result

The results are presented in the following order: students' descriptive responses to the questionnaire, instrument quality and assumption tests, reading comprehension before and after implementation, the regression analysis of self-reported deep learning, and hypothesis testing. Likert-scale means were interpreted using the categories in Table 3.

ISSN 2338-4778 (Print)
ISSN 2548-4192 (Online)

Table 3. Criteria for Interpreting the Average Scores of Indicators

Interval	Category
1.00 – 1.80	Very Low
1.81 – 2.60	Low
2.61 – 3.40	Middle
3.41 – 4.20	High
4.21 – 5.00	Very High

Source: Scale Interval of Class Likert Scale 1-5

Descriptive Learner Responses

Before conducting inferential statistical tests, students' responses to the questionnaire were described. The questionnaire consisted of 30 items using a 1–5 Likert scale to measure four variables: Adaptive Reading System (X1; 12 items), Reading Activities (X2; 7 items), Deep Learning (Y; 6 items), and Generation Z Characteristics in AI learning (5 items). A total of 45 students from three high schools in Barru Regency (SMA Negeri 1 Barru, SMA Negeri 3 Barru, and SMA Negeri 6 Barru) completed the questionnaire, and the same 45 students completed the reading comprehension pre-test and post-test. Mean scores were interpreted using the categories in Table 3. Deep learning was measured by self-report.

Table 4. Summary of Average Scores by Variable

Variable	Total Item	Mean	Category
X1 - Adaptive Reading System	12	3.92	High
X2 – Reading Activity	7	3.98	High
Y – Deep Learning	6	4.05	High
Generation Z Characteristics	5	3.92	High

Source: Data Analysis Results, SPSS

Based on the results of the above analysis, the AI-Based Adaptive Reading variable (X1) had a mean of 3.92, falling into the high category. The Reading Activities variable (X2) had a mean of 3.98, also falling into the high category. On the other hand, the self-reported Deep Learning variable (Y) had a mean of 4.05, also falling into the high category. The Generation Z Characteristics variable had a mean of 3.92 in the high category, suggesting students' readiness and comfort with AI-based learning, although this indicator is descriptive-contextual in nature and was not tested inferentially.

Table 5. Frequency Distribution of Respondents' Total Scores by Category for Each Variable

Variable	Very Low	Low	Middle	High	Very High
X1 - Adaptive Reading System	0 (0.0%)	1 (2.2%)	8 (17.8%)	20 (44.4%)	16 (35.6%)
X2 – Reading Activity	0 (0.0%)	0 (0.0%)	7 (15.6%)	20 (44.4%)	18 (40.0%)
Y - Deep Learning	0 (0.0%)	1 (2.2%)	7 (15.6%)	20 (44.4%)	17 (37.8%)
Generation Z Characteristics	0 (0.0%)	1 (2.2%)	14 (31.1%)	16 (35.6%)	14 (31.1%)

Source: Data Analysis Results, SPSS

Based on the results above, the majority of respondents (approximately 67–84%) fell into the High to Very High categories for all four variables, while the proportion of respondents in the Low and Very Low categories was very small (0–2.2%). This indicates that students' responses were generally positive toward the AI-based adaptive reading system, the reading activities, their own deep learning, and their readiness as Generation Z learners to adopt AI-based learning; the proportion was lowest for Generation Z characteristics (66.7%). These are perception data from a single group of 45 students. They describe students' responses and do not by themselves demonstrate the effectiveness of the system.

Instrument Validity and Reliability

Table 6 presents the validity and reliability results for the questionnaire scales.

Table 6. Results of the Validity and Reliability Tests of the Instruments

Variable	Calculated r	Critical r	Cronbach's Alpha	Description
X1 - Adaptive Reading System	0.402 - 0.730	0.294	0.905	Valid and very reliable
X2 – Reading Activity	0.539 - 0.719	0.294	0.863	Valid and reliable
Y - Deep Learning	0.665 - 0.773	0.294	0.895	Valid and reliable
Generation Z Characteristics	0.611 - 0.778	0.294	0.855	Valid and reliable

Source: Data Analysis Results, SPSS

The validity test used the corrected item-total correlation compared with the critical r-value at the 5% significance level and n = 45, which was 0.294. An item is considered valid if the calculated r-value is greater than the critical r-value. Reliability was assessed using Cronbach's alpha, with a scale considered reliable if alpha was greater than 0.60. All items of the X1, X2, Y, and Generation Z characteristics scales were valid, and the scales were reliable (α = .905, .863, .895, and .855, respectively). The scales were therefore considered suitable for collecting the research data.

Results of the Data Normality Test

A normality test was conducted to determine whether the research data were normally distributed, as a prerequisite for performing parametric tests (paired-sample t-test and multiple linear regression). The normality test used the Kolmogorov-Smirnov and Shapiro-Wilk tests, with the criterion that the data are normally distributed if the significance value is > 0.05. Based on the results of the normality tests, the analysis results are presented as follows:

ISSN 2338-4778 (Print)
ISSN 2548-4192 (Online)

Table 7. Results of the Data Normality Test

Data	K-S Statistic	K-S Sig.	S-W Statistic	S-W Sig.	Description
Pre-Test Score	0.114	0.171	0.972	0.343	Normal
Post-Test Score	0.104	0.200	0.961	0.138	Normal
Regression Model Residual	0.088	0.200	0.980	0.616	Normal

Source: Data Analysis Results, SPSS

Based on the results of the normality test, it can be concluded that the Shapiro-Wilk p-value is greater than 0.05; therefore, the pre-test, post-test, and regression model residual data are considered to be normally distributed. Since the assumption of normality is met, the paired-sample t-test and multiple linear regression analysis can proceed.

Reading Comprehension: Pre-Test and Post-Test Results

A paired-sample t-test was conducted to determine whether students' reading comprehension scores differed significantly before and after the AI-based adaptive reading system was implemented. Table 8 presents the descriptive statistics of the pre-test and post-test scores, and Table 9 presents the correlation and the paired-sample t-test results.

Table 8. Descriptive Statistics of the Reading Comprehension Pre-Test and Post-Test Scores

Variable	N	Minimum	Maximum	Mean	Std. Deviation
Pre-Test	45	55.00	100.00	80.93	9.645
Post-Test	45	70.00	100.00	88.02	7.331

Source: Data Analysis Results, SPSS

Table 9. Paired-Sample Correlation and T-Test Results

Pair	N	Correlation	Sig. Correlation	Mean Difference	t	df	Sig. (2-tailed)
Pre-Test - Post-Test	45	0.901	0.000	-7.089	-10.809	44	0.000

Source: Data Analysis Results, SPSS

The mean reading comprehension score increased from 80.93 (pre-test) to 88.02 (post-test), and the standard deviation decreased from 9.645 to 7.331, indicating that scores became more homogeneous after implementation. The pre-test and post-test scores were strongly correlated ($r = .901$, $p < .001$), which shows that students largely kept their relative rank across the two measurements; this describes the stability of the ordering and does not itself indicate that the intervention was effective. The paired-sample t-test showed a statistically significant difference, $t(44) = -10.809$, $p < .001$, with a mean gain of 7.09 points (95% CI [5.77, 8.41]); the negative sign reflects the direction of subtraction (pre-test minus post-test). The effect size was large (Cohen's d for paired samples = 1.61). Because the study had no control group, these results show that students'

reading comprehension scores improved significantly following implementation of the system. They do not establish that the system alone caused the improvement, since practice effects, familiarity with the test, and other classroom activities cannot be ruled out.

Deep Learning: Multiple Linear Regression Analysis

The regression analysis examined whether the adaptive reading system (X1) and reading activities (X2) predicted students' self-reported deep learning (Y). X1, X2, and Y are the summed scores of the corresponding questionnaire scales. Table 10 presents their descriptive statistics, and Table 11 summarizes the regression model.

Table 10. Descriptive Statistics of the Summed Scores Entered in the Regression Model

Variable	N	Minimum	Maximum	Mean	Std. Deviation
TOTAL_X1 (Adaptive Reading System)	45	27.00	60.00	47.07	7.635
TOTAL_X2 (Reading Activity)	45	19.00	35.00	27.84	4.205
TOTAL_Y (Deep Learning)	45	14.00	30.00	24.31	3.842

Source: Data Analysis Results, SPSS

Table 11. Model Summary of the Multiple Regression Analysis

Model	R	R Square	Adjusted R Square	Std. Error
1	0.881	0.776	0.765	1.862

Source: Data Analysis Results, SPSS

The multiple correlation coefficient was R = .881, and R^2 = .776 (adjusted R^2 = .765; standard error of the estimate = 1.862). This means that X1 and X2 together accounted for 77.6% of the variance in self-reported deep learning (Y) in this sample. This is a statement about explained variance and does not imply that X1 and X2 caused the variation in Y.

Model Significance, Regression Coefficients, and Multicollinearity

Table 12 presents the F-test of the regression model.

Table 12. ANOVA (F-Test) Results

Model	Sum of Squares	df	Mean Square	F	Sig.
Regression	504.024	2	252.012	72.686	0.000
Residual	145.620	42	3.467		
Total	649.644	44			

Source: Data Analysis Results, SPSS

ISSN 2338-4778 (Print)
ISSN 2548-4192 (Online)

The F-test was significant, F(2, 42) = 72.686, p < .001, indicating that X1 and X2 together significantly predicted students' self-reported deep learning (Y).

Table 13. Regression Coefficients

Model	B	Std. Error	Beta	T	Sig.
(Constant)	1.782	1.900	-	0.938	0.354
TOTAL_X1	0.217	0.070	0.432	3.096	0.003
TOTAL_X2	0.442	0.128	0.483	3.463	0.001

Source: Data Analysis Results, SPSS

The regression equation was Y = 1.782 + 0.217 X1 + 0.442 X2. The constant (1.782) was not significant (p = .354) and is the predicted value of Y when X1 and X2 are both zero, which lies outside the range of the observed scores. Holding X2 constant, each one-point increase in X1 was associated with a 0.217-point higher predicted Y (t = 3.096, p = .003). Holding X1 constant, each one-point increase in X2 was associated with a 0.442-point higher predicted Y (t = 3.463, p = .001). Both the adaptive reading system and reading activities were therefore significant positive predictors of deep learning. The standardized coefficients were similar in size (X1: β = .432; X2: β = .483), and because the two predictors are strongly correlated, as shown below, the data do not allow a firm conclusion about which of them is the stronger predictor.

Multicollinearity was examined next (Table 14). The tolerance value for both predictors was 0.274 (> 0.10) and the VIF was 3.648 (< 10), indicating that multicollinearity was within conventional limits and that the model could be interpreted. However, a tolerance of .274 in a two-predictor model implies a correlation of about .85 between X1 and X2, which means that their unique contributions cannot be separated cleanly and that individual coefficients may be less stable. The regression residuals were normally distributed (Table 7). A Breusch–Pagan test, however, indicated some heteroscedasticity (p = .026); with heteroscedasticity-robust (HC3) standard errors both predictors remained significant (X1: t = 2.480, p = .013; X2: t = 3.176, p = .002).

Table 14. Multicollinearity Test Results

Variable	Tolerance	VIF
TOTAL_X1	0.274	3.648
TOTAL_X2	0.274	3.648

Source: Data Analysis Results, SPSS

Hypothesis Testing

Table 15. Summary of Hypothesis Testing

Hypothesis	Test Statistic	Decision
H1: There is a significant difference in students' reading comprehension scores before and after using the AI-based adaptive reading system (pre-test vs. post-test)	t(44) = −10.809; p < .001	Supported (H0 rejected)

H2: The adaptive reading system (X1) significantly predicts deep learning (Y)	t(42) = 3.096; p = .003	Supported (H0 rejected)
H3: Reading activities (X2) significantly predict deep learning (Y)	t(42) = 3.463; p = .001	Supported (H0 rejected)
H4: The adaptive reading system (X1) and reading activities (X2) together significantly predict deep learning (Y)	F(2, 42) = 72.686; p < .001	Supported (H0 rejected)

Source: Data Analysis Results, SPSS

Overall, the results provided empirical support for all four hypotheses (Table 15). Students' reading comprehension scores were significantly higher after the adaptive reading system was implemented (H1), and the adaptive reading system (H2) and reading activities (H3) each significantly predicted self-reported deep learning, as did the two combined (H4), among students at SMAN 1 Barru, SMAN 3 Barru, and SMAN 6 Barru. Given the one-group design and the correlational regression analysis, H1 reflects improvement following implementation, and H2–H4 reflect predictive associations; neither reflects a demonstrated causal effect of the system.

4. Discussion

This study developed an AI-based adaptive reading system and examined two sets of results: students' reading comprehension scores before and after implementation, and the extent to which the adaptive system and reading activities predicted self-reported deep learning. The discussion first interprets the pre-test–post-test gain and the claims it can support, then considers how the adaptive system and reading activity relate to deep learning, and finally discusses Generation Z as the learner context, practical implications, and limitations.

Reading Comprehension Gains Following Implementation

Students' reading comprehension scores increased significantly after the system was implemented, from 80.93 to 88.02 (mean gain = 7.09 points, t(44) = −10.809, p < .001, Cohen's d = 1.61). The safest interpretation is that students improved following implementation of the system. The one-group design cannot separate the contribution of the system from practice effects, familiarity with the test, regular instruction, or the novelty of a new tool.

The gain is nevertheless consistent with prior evidence that adaptive AI-supported platforms can produce measurable comprehension gains (Hariyanto et al., 2025), and with the logic of Adaptive Learning Theory, in which learning is supported by a continuous cycle of assessment, content adaptation, and feedback. The smaller standard deviation at post-test (9.645 to 7.331) may indicate that lower-scoring students gained more, although regression toward the mean and the ceiling of the test (maximum score 100) offer alternative explanations, and individual gain scores were not analyzed. The comprehension gain also should not be equated with deep learning: reading comprehension was measured by a test, whereas deep learning was measured by a self-

ISSN 2338-4778 (Print)
ISSN 2548-4192 (Online)

report questionnaire administered at one point in time, so the study cannot say whether the gains reflect deeper critical engagement or better comprehension of the texts alone.

Two features of the AI-based Adaptive Reading System (X1) are plausible mechanisms for the gain, offered here as interpretations and not as tested explanations. First, the system's capacity to adjust reading difficulty and recommend content based on individual performance corresponds to what Vygotsky (1978) described as the learner's zone of proximal development: material is kept challenging enough to demand effort, but not so far beyond current ability that it produces disengagement. Second, the system's automatic feedback and behavior-tracking features may reduce the burden of self-monitoring that would otherwise fall entirely on the student, freeing cognitive resources for comprehension rather than for judging one's own progress, an effect consistent with Cognitive Load Theory (Sweller, 1988).

In both cases the theoretical label is meaningful only to the extent that the system actually implements the mechanism (see Table 1): the zone-of-proximal-development argument holds because difficulty adjustments followed students' scores on the comprehension checks after each text (moving up, staying, or moving down one level according to the score thresholds), and the cognitive-load argument holds because the feedback was immediate and explanatory, tied to each answer, so that students did not have to judge their own progress unaided.

The Adaptive System, Reading Activity, and Deep Learning

Adaptive Reading System (X1) and Reading Activity (X2) jointly accounted for 77.6% of the variance in self-reported deep learning (R^2 = .776; adjusted R^2 = .765), and both were significant predictors (X1: β = .432, p = .003; X2: β = .483, p = .001). These are associations within a single sample and do not show that the system or the reading activity caused deeper learning. The two standardized coefficients are similar, and because X1 and X2 are strongly correlated (tolerance = .274, VIF = 3.648, implying r ≈ .85), the unique contribution of each predictor cannot be separated cleanly; the data therefore do not show whether the structure and personalization of the reading or the amount of reading activity matters more. The pattern is consistent with Self-Regulated Learning theory (Zimmerman, 2002), which holds that learning gains depend less on time-on-task alone than on feedback loops that let learners monitor, evaluate, and adjust their own strategies, a function the adaptive system was designed to provide, although self-regulation itself was not measured in this study.

The finding that both predictors were significant suggests that the technology may matter through the quality of the reading activity it organizes, and not simply through the presence of AI. The adaptive system is therefore best understood as a complement to sustained reading practice and not as a substitute for it. This points to a double-edged role for Self-Regulated Learning in this design. On one hand, the system's feedback loops may strengthen self-regulation by making performance visible and adjustable in real time. On the other hand, the concern raised in the Introduction, grounded in the Self-Regulated Learning theory of Zimmerman (2002), that AI can weaken self-regulation when cognitive work is fully delegated to the system cannot be ruled out by these results; the present data show an association and cannot confirm whether the system strengthens or substitutes for students' own regulatory effort. Distinguishing these two possibilities is an important direction for future work.

Generation Z as the Learner Context

Generation Z was described in this study through a descriptive scale (mean = 3.92, high category) and was not tested as a predictor or moderator. The study therefore provides no empirical evidence that the system worked better because the participants belong to Generation Z, and the discussion below is contextual interpretation and not a statistically confirmed link. The literature describes Generation Z students as shaped by their upbringing as digital natives (Nasrudin et al., 2024; Shirazian et al., 2025), with strong multitasking tendencies, a preference for visual and interactive information, and a preference for microlearning (Grace & Seemiller, 2019), and as favoring concise, fast, visually stimulating content over conventional text (Čekrlija, 2024). Features of the system such as short, individually calibrated reading units and immediate feedback are compatible with these reported preferences, and students' high ratings of the system may reflect that compatibility.

These generalizations should be applied with caution. Students within the same generation differ in reading ability, digital literacy, motivation, and learning preferences, and familiarity with digital media does not imply readiness for deep reading. The study therefore does not claim that digital immersion is a resource for deep reading; it suggests only that technology designed with learners' preferences and abilities in mind may be received positively. Whether similar results would occur among learners from other generations, or whether Generation Z characteristics moderate the effect of such systems, remains a question for future research.

Potential Implications for Literacy Instruction in Indonesia

The PISA 2022 results show that 75% of Indonesian students fell short of proficient reading, with a national average score of 359 against an OECD average of 476 (OECD, 2023). The present study involved 45 students in three schools in one regency and had no control group, so it cannot provide evidence that adaptive reading systems can address this national problem or be scaled up. The findings are better read as a preliminary, school-level indication that AI-based adaptive reading systems combined with structured reading activities merit testing in larger and controlled studies, in line with national priorities for adaptive, technology-based learning (D et al., 2022; Haider et al., 2025; Salman, 2025).

Implications for Teaching Practice

Several practical implications follow. First, adaptive reading can be incorporated into classroom reading cycles: for example, a pre-reading phase in which the teacher activates prior knowledge, a system-supported reading phase in which students work at an appropriate level, and a post-reading phase of discussion in which students justify their inferences. Second, teachers should monitor the AI-generated recommendations and feedback, for instance by reviewing system logs of difficulty progression and comprehension-check results to identify students who are advanced too quickly or who stagnate, and by correcting feedback that is inaccurate or unhelpful. Third, teachers should encourage students to engage in inference, synthesis, and critical reflection and not in passive consumption, for example by asking students to explain or critique the system's feedback before moving on and by reserving part of each session for unaided reading and discussion.

ISSN 2338-4778 (Print)
ISSN 2548-4192 (Online)

Finally, the value of adaptive reading technology may lie less in its novelty than in its capacity to individualize feedback at a scale that no single teacher could provide unaided, which is why the system should support, and not replace, teachers' professional judgement.

Limitations and a Note of Caution

These conclusions should be read within the boundaries of the study's design. The sample was drawn from three schools in a single regency using purposive sampling of 45 students, which limits the generalizability of the findings. The design relied on a single group with one pre-test–post-test cycle and no non-adaptive control or comparison group, so the gains reported cannot be separated from practice effects, familiarity with the test, maturation, other classroom learning, or the novelty of a new learning tool; for this reason the findings are described as improvement following implementation and not as an effect of the system.

In addition, deep learning was measured by a self-report questionnaire at a single point in time, and reading comprehension scores were close to the ceiling of the test (the maximum score of 100 was reached at both measurements), which may have limited the room for improvement among higher-scoring students. The instruments were validated on the same 45 participants who provided the study data, and students within the Generation Z cohort were treated as one group even though they differ in ability and digital literacy. The regression residuals also showed some heteroscedasticity (Breusch–Pagan $p = .026$), although the predictors remained significant with robust standard errors. Finally, this article reports only quantitative data. Qualitative evidence on how students experienced the system and expert and teacher assessments of its validity and practicality are not part of the present analysis, so the mechanisms proposed above remain interpretive.

It is also worth noting that the conceptual framework proposed content personalization, scaffolding, and feedback as mediating mechanisms, and the level of AI dependence as a moderating variable, between the adaptive system and deep-literacy outcomes. The analysis reported here tested X1 and X2 as direct, parallel predictors of Y and did not formally test this mediation–moderation structure. The mechanisms discussed above are therefore hypothesized mechanisms and not demonstrated ones. The strong direct associations observed are consistent with the conceptual model but do not confirm it; a formal mediation or moderation analysis, ideally with a measured AI-dependence variable, would be needed to test the mechanisms and not infer them after the fact.

Finally, the literature reviewed in the Introduction cautions that over-reliance on AI-generated summaries can slow the progress of already high-performing readers, even as it accelerates struggling readers. Because this study measured aggregate reading comprehension gains and overall deep-learning scores rather than differential effects by baseline proficiency, it cannot confirm or rule out this risk for the present sample. Future implementations should pair adaptive scaffolding with tasks that explicitly require inference, synthesis, and critical reflection, so that the system continues to support deep learning across the full range of student ability rather than only for those who begin with the greatest need.

5. Conclusion

This study evaluated the implementation of an AI-based adaptive reading system in a trial with 45 senior high school students in three schools. Students' reading comprehension scores improved significantly following implementation, rising from a mean of 80.93 in the pre-test to 88.02 in the post-test ($t(44) = -10.809$, $p < .001$; Cohen's d for paired samples = 1.61). Because the study had no control group, this improvement cannot be attributed to the system alone. In a separate analysis, multiple regression showed that the adaptive reading system and reading activities each significantly predicted students' self-reported deep learning and together accounted for 77.6% of its variance ($R^2 = .776$), with similar standardized coefficients for the adaptive reading system ($\beta = .432$, $p = .003$) and for reading activities ($\beta = .483$, $p = .001$). These are predictive associations and not evidence of causal effects.

These findings suggest that adaptive AI technology can provide a more personalized and responsive reading environment and may help create conditions for deeper learning when combined with sustained, structured reading activities. This is the main contribution of the study: preliminary evidence from Indonesian secondary-school learners that adaptive personalization and structured reading activity may work together, and not simply that AI improves reading scores. For teachers, the findings point to using the system within classroom reading cycles, monitoring AI-generated recommendations, and prompting students to infer, synthesize, and reflect critically.

The findings should be interpreted within the limitations of the study, particularly the purposive sample of 45 students from three schools in one regency, the single-group design without a non-adaptive control group, and the reliance on self-report for deep learning. Future research should use controlled or quasi-experimental designs with comparison groups, larger and more diverse samples, longer interventions, and measures of deep learning that go beyond self-report. It should also add qualitative evidence (for example, student interviews and classroom observation) and expert and teacher assessments of the system's validity and practicality to explain how and why the system works. It should also measure AI dependence directly and test whether personalization, scaffolding, and feedback function as mechanisms linking adaptive AI systems to deep learning.

6. Acknowledgement

The authors would like to express their sincere gratitude to the principals, English teachers, and students of SMAN 1 Barru, SMAN 3 Barru, and SMAN 6 Barru for their participation, cooperation, and support throughout the research process. The authors also acknowledge the contributions of the experts who provided valuable feedback during the validation and development of the AI-based Adaptive Reading System. Finally, the authors would like to thank Universitas Muhammadiyah Barru for providing institutional support that facilitated the completion of this study.

ISSN 2338-4778 (Print)
ISSN 2548-4192 (Online)

7. References

Ariqo, W., Rizal, E., & Winoto, Y. (2025). Analisis Penggunaan Kecerdasan Buatan ChatGPT dalam Penulisan Karya Publikasi Ilmiah: Manfaat dan Tantangan. Literatify: Trends in Library Developments, 6(1), 2723–0953. https://doi.org/10.24252/literat

Atun, H. (2020). Intelligent Tutoring Systems (ITS) to Improve Reading Comprehension: A Systematic Review. In Journal of Teacher Education and Lifelong Learning.

Barrón Pérez, S., & Morales Piña, D. (2026). Procesos Cognitivos y Pensamiento Crítico en Entornos Educativos Mediados por Inteligencia Artificial Adaptativa. Imperium Académico Multidisciplinary Journal, 3(1). https://doi.org/10.63969/njwwrh50

Čekrlija, S. (2024). Challenges of Teaching Methods for Generation Z at Universities, Case Study China. https://doi.org/10.52244/c.2024.11.16

D, D., Khasanah, M., & Putri, A. M. (2022). Penguatan Literasi, Numerasi, dan Adaptasi Teknologi pada Pembelajaran di Sekolah. Eksponen, 11(2). https://doi.org/10.47637/eksponen.v11i2.381

Delgado, P., Vargas, C., Ackerman, R., & Salmerón, L. (2018). Don't throw away your printed books: A meta-analysis on the effects of reading media on reading comprehension. Educational Research Review, Vol. 25. https://doi.org/10.1016/j.edurev.2018.09.003

Erümit, A. K., & Çetin, İ. (2020). Design Framework of Adaptive Intelligent Tutoring Systems. Education and Information Technologies, 25(5). https://doi.org/10.1007/s10639-020-10182-8

Etkin, H. K., Etkin, K. J., Carter, R. J., & Rolle, C. E. (2025). Differential effects of GPT-based tools on comprehension of standardized passages. Frontiers in Education, 10. https://doi.org/10.3389/feduc.2025.1506752

Gerlich, M. (2025). AI Tools in Society: Impacts on Cognitive Offloading and the Future of Critical Thinking. Societies, 15(1). https://doi.org/10.3390/soc15010006

Gou, P., & Ren, X. (2025). Developing Students' Computational Thinking through Project-Based Learning Empowered by Generative AI: A Practical Study. Proceedings of 2025 6th International Conference on Education, Knowledge and Information Management, ICEKIM 2025. https://doi.org/10.1145/3756580.3756611

Grace, M., & Seemiller, C. (2019). Generation Z: A Century in the Making tle. In Journal of Chemical Information and Modeling (Vol. 53).

Haider, M. H., Majeed, M. A., & Irfan, M. (2025). Needs of Digital Literacy in the Era of Artificial Intelligence: A Comprehensive Review. Inverge Journal of Social Sciences, 4(4), 396–409.

Hariyanto, Kristianingsih, F. X. D., & Maharani, R. (2025). Artificial Intelligence in Adaptive Education: A Systematic Review of Techniques for Personalized Learning. Discover Education, Vol. 4. https://doi.org/10.1007/s44217-025-00908-6

Huu, P. N., Tangworakitthaworn, P., & Gilbert, L. (2022). The Design and Development of an Adaptive Intelligent Tutoring System Based on Constructive Alignment and Cognitive Theories. 2022 19th International Joint Conference on Computer

Science and Software Engineering, JCSSE 2022. https://doi.org/10.1109/JCSSE54890.2022.9836290

Liu, Y. (2025). Cognitive Load Effects of AI Tutoring Systems Compared to Traditional Instructional Methods. International Journal of Social Sciences and English Literature, 9(11). https://doi.org/10.55220/2576-683x.v9.633

Mahdi, A. M. M., Haseeb, S., Rafiq, M. H., & Usman, M. (2025). Cognitive Load Management through Adaptive AI System an Educational Psychology Perspective. The Critical Review of Social Sciences Studies, 3(3). https://doi.org/10.59075/rycs7x03

Mishra, N. R. (2023). Constructivist Approach to Learning: An Analysis of Pedagogical Models of Social Constructivist Learning Theory. Journal of Research and Development, 6(01). https://doi.org/10.3126/jrdn.v6i01.55227

Mulyarti. (2025). Enhancing Reading Comprehension through Inspiring Novels: A Pre-Experimental Study in an EFL Context. Jurnal Edukasi Saintifik. https://doi.org/https://doi.org/10.56185/jes.v5i2.928

Mulyarti, Idham, A. Z., Danial, M., & Rauf, W. (2026). Transforming Education in the Digital Era: Trends, Innovations, and Future Directions. Jurnal Edukasi Saintifik, 6(2), 51–58. https://doi.org/https://doi.org/10.56185/jes.v6i2.1560

Nasrudin, E., Sumarna, E., & Surahman, C. (2024). Examining the Characteristics of Generation Z and Their Implications for Students' Character Education. Jurnal Iman Dan Spiritualitas, 4(4). https://doi.org/10.15575/jis.v4i4.40971

OECD. (2023). PISA 2022 Results (Volume I): The State of Learning and Equity in Education.

Özer, M., Tanberkan, H., & Perc, M. (2025). Artificial Intelligence Threatens Critical Thinking in Education Systems. Journal of Higher Education and Science, 15(2). https://doi.org/10.5961/higheredusci.1747885

Pakili, N., Yunus, T. M., Novian, D., & Bouty, Abd. A. (2025). Potensi Self-Regulated Learning dalam Pembelajaran Berbasis AI: Kajian Literatur terhadap Adaptabilitas Metode Socratic di Era Digital. Inteligensi : Jurnal Ilmu Pendidikan, 8(2). https://doi.org/10.33366/ilg.v8i2.8124

Patnaik, Dr. R. P. K., Baid, S., Sailaja, A. L. V. N., Chandar, Y. G. P., & Sai, P. R. (2025). Exploring the Role of AI in Enhancing Personalized Learning among College Students. International Journal of Research and Innovation in Applied Science, 10(10). https://doi.org/10.51584/ijrias.2025.101000001

Sa'diyah, I., & Savitri, A. (2024). The Effectiveness of ChatGPT Technology as a Teaching Material for Students' Reading-Writing Literacy in The Digital Era. MEDIASI Jurnal Kajian Dan Terapan Media, Bahasa, Komunikasi, 5(2), 97–109. https://doi.org/10.46961/mediasi.v5i2.818

Sakhipov, A., Tukenova, N., & Omirzak, I. (2025). AI for Personalized Learning: A Comparative Analysis of Traditional and AI-Based Methods. Trudy Universiteta, 4(101). https://doi.org/10.52209/1609-1825_2025_4_263

Salman, A. (2025). Deep Learning Based Educational Transformation as a Roadmap to Indonesia Emas 2045. At Turots: Jurnal Pendidikan Islam, 7(1).

Sari, H. E., Tumanggor, B., & Efron, D. (2024). Improving Educational Outcomes Through Adaptive Learning Systems using AI. International Transactions on Artificial Intelligence (ITALIC), 3(1). https://doi.org/10.33050/italic.v3i1.647

ISSN 2338-4778 (Print)
ISSN 2548-4192 (Online)

Schleicher, A. (2019). PISA 2018: Insights and Interpretations. In Oecd.

Scholkmann, A., Schurer, M. H., & Hviid, C. (2025). Towards a Theory of Generative AI in Problem-based Learning. Journal of Problem Based Learning in Higher Education. https://doi.org/10.54337/ojs.jpblhe.v13i1.10033

Shirazian, S., Vatanpour, M., & Guity, P. (2025). Generation Z: Learners' Profiles and Teaching Challenges. Frontiers in Dentistry, Vol. 22. https://doi.org/10.18502/fid.v22i40.19906

Silor, A. C., & Silor, F. S. C. (2025). Boosting Reading Comprehension Through AI-Based Learning Tools. International Journal of Learning, Teaching and Educational Research, 24(9). https://doi.org/10.26803/ijlter.24.9.4

Sumarna, C., & Gunawan, H. (2022). Foundations of Constructivism Philosophy in Classroom Learning. International Journal of Science and Society, 4(3). https://doi.org/10.54783/ijsoc.v4i3.499

Sweller, J. (1988). Cognitive Load During Problem Solving: Effects on Learning. Cognitive Science, 12(2). https://doi.org/10.1016/0364-0213(88)90023-7

Tobondo, Y. A., Alfian, M., Widnyana, I. G. A. G., Uagadi, A. E. S., Tarindje, S. D., & Melapa, L. (2024). The Effectiveness of the Quasi-Class Action Approach: Integration of Problem-Based Learning, Flipped Classrooms, and Generative Learning Models. Interdisciplinary Journal (IDe), 2(2). https://doi.org/10.61254/idejournal.v2i2.56

Vygotsky, L. S. (1978). Mind of Society:The Development of Higher Psychological Processes. In Mind of Society.

Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. Theory into Practice, Vol. 41, pp. 64–70. Ohio State University Press. https://doi.org/10.1207/s15430421tip4102_2