Jurnal Pustaka Indonesia (JPI)
2026, Vol. 6, No. 2, 494–514
https://doi.org/10.62159/jpi.v6i2.2634

OPEN ACCESS Research Article

# Perceived Usefulness, Perceived Ease of Use, and Self-Reported Critical Thinking Among Indonesian Vocational Students: The Indirect Role of Self-Regulated Learning

Fani Okta Srinawati[1] and Rose Rahmidani[2*]
[1] Master's Program in Economics Education, Universitas Negeri Padang, Padang, Indonesia
[2] Faculty of Economics and Business, Universitas Negeri Padang, Padang, Indonesia

**Abstract**
The spread of generative artificial intelligence (AI) in vocational classrooms has raised a concern that easy, on-demand answers may displace the reasoning vocational graduates are expected to master. This cross-sectional survey examined whether students' perceptions of AI are associated with their self-reported critical thinking, and whether self-regulated learning accounts for any such association. Drawing on social cognitive theory and the Technology Acceptance Model, questionnaire data were collected from 365 students in nine state vocational schools (SMK) in Pesisir Selatan Regency, West Sumatra, and analysed with PLS-SEM. All four constructs, including critical thinking, were measured by self-report rather than by performance tasks. Perceived usefulness and perceived ease of use were not associated with critical thinking directly (β = -.026, 95% CI [-.163, .111]; β = .064, 95% CI [-.072, .200]) but were weakly associated with self-regulated learning (β = .161; β = .158), which was in turn the strongest correlate of critical thinking (β = .383, 95% CI [.284, .482]). Both indirect paths were small and positive (β = .062 and .061), a configuration indicating indirect-only mediation. Explanatory power was weak, with $R^2$ = .085 for self-regulated learning and .157 for critical thinking, and $f^2$ values for all four technology paths below .02. A clustering sensitivity analysis showed that the technology paths lose significance at an intraclass correlation above about .01, so these associations are provisional. Because all constructs were measured on one occasion, the ordering is a theoretical assumption rather than an observed sequence.

**ARTICLE HISTORY**
Received : 23 March 2026
Revised : 30 July 2026
Accepted : 25 August 2026

**KEYWORDS**
Artificial intelligence; critical thinking; perceived usefulness; self-regulated learning; vocational education



**CORRESPONDING AUTHOR**
*** Rose Rahmidani**, Faculty of Economics and Business, Universitas Negeri Padang, Padang, Indonesia. Email: rose_rahmidani@fe.unp.ac.id

## Introduction

The digitalised landscape of the twenty-first century has reshaped what employers expect from a competent workforce. For graduates of vocational high schools, known in Indonesia as Sekolah Menengah Kejuruan (SMK), the expectation is twofold (Jia & Huang, 2023; Riyanda et al., 2025). They must command the technical skills of their field and, increasingly, demonstrate the higher-order cognition needed to interpret problems, weigh evidence, and decide under uncertainty (Le, 2022; Siregar et al., 2023).

Critical thinking is a broad construct, and the earlier version of this manuscript used it generically. In a business-management vocational programme the capability takes a specific form, and stating that form matters both for the hypotheses and for the interpretation of the measure. The reasoning these students are trained for is applied and commercial: reading a sales or stock record and deciding what it shows about a trend rather than merely reporting the numbers; judging whether a supplier's claim or a promotional figure is supported by the evidence offered; comparing two

pricing, display, or service options against a stated business objective; and inferring the likely consequence of a decision for margin, stock turnover, or customer response. These correspond to the dimensions of interpretation, analysis, evaluation, and inference that the critical-thinking literature identifies (Ninghardjanti et al., 2025; Waliyuddin & Sulisworo, 2022), but their content is domain-specific: the evidence is commercial, the criteria are business criteria, and the conclusions are operational decisions.

This has a direct bearing on the measure used here, and the limitation is stated at the outset rather than deferred. Critical thinking was assessed with a self-report questionnaire in which students rated how far statements about their own reasoning described them, not with a task in which they analysed a commercial case and were scored against a rubric. A self-report scale of this kind indexes a student's perception of their own reasoning, which is related to but not the same as demonstrated reasoning, and it is generic rather than domain-specific: it does not present business evidence and cannot show whether a student can reason with it. The construct examined in this study is therefore more precisely described as self-reported critical-thinking disposition, and every result below is stated in those terms. Whether these associations would hold for performance-based critical thinking in a business context is an open question that this design cannot answer.

For vocational graduates the stakes attached to this capacity are high, since their training is meant to lead directly into employment where the ability to reason through unfamiliar problems separates a competent practitioner from a merely procedural one (Harianto et al., 2025; J.E. Sutanto et al., 2021; Le, 2022). Evidence on the extent of the problem in Indonesian vocational education should be stated more carefully than the earlier version did. Instrument-development and intervention studies conducted in SMK settings, which assess reasoning through tests and rubrics rather than through students' own ratings, have been motivated by consistently weak performance on the higher-order dimensions relative to recall (Sukmawati et al., 2024; Waliyuddin & Sulisworo, 2022; Yuniarti et al., 2024), and the broader evidence on teaching for these skills in vocational contexts points in the same direction (Akmal et al., 2025; Siregar et al., 2023). Several sources cited at this point in the earlier version reported perceptions, general learning outcomes, or non-vocational samples, and have been removed, because evidence about perceived ability or about achievement in general does not establish a deficit in demonstrated reasoning.

Social cognitive theory offers a lens for understanding how such cognition develops. It treats the learner not as a passive recipient of information but as an agent whose functioning arises from the reciprocal interaction of personal, behavioural, and environmental determinants, and its account of self-regulation specifies the subfunctions of self-observation, judgement against personal standards, and self-reaction through which that agency operates (Entwistle et al., 2025; Kholifah et al., 2023). The earlier version supported these propositions with a citation to a study of career goals among Taiwanese college athletes framed by social cognitive career theory, which cannot establish them; that attribution has been removed.

Generative models have made AI a routine study companion, and tools of this kind can personalise learning, return rapid feedback, and provide scaffolding that accelerates understanding (Martínez-Plumed et al., 2021; Zhou, Zhang, et al., 2024). Treating all such engagement as one thing, however, obscures a distinction that matters for the argument. At least four purposes can be separated. A student may ask AI to explain a concept, which engages comprehension; to generate an answer or a finished piece of work, which permits the reasoning to be bypassed entirely; to give

feedback on work the student has already produced, which requires the student to have reasoned first; or to verify a claim or calculation, which is itself an evaluative act. Only the second of these substitutes for thinking; the third and fourth arguably constitute it. Intensity and frequency vary independently of purpose, and quality of use, meaning whether the student interrogates the output or accepts it, varies independently of both.

The present study measures none of these. Its two technology constructs are perceptions, not behaviours: how useful a student finds AI and how easy they find it to operate. A student who uses AI to generate finished answers and one who uses it to check their own reasoning may report identical usefulness. This is a substantial limitation on what the model can show, and it is one reason the construct formerly labelled AI utilisation has been renamed perceived usefulness throughout, as set out below. Distinguishing purpose, intensity, and quality of use requires behavioural or log data and is proposed as the next step rather than claimed as a feature of this one.

The second technological factor is the perceived ease of using AI. When students judge a tool easy to learn and operate, the technical friction of adoption falls and attention can be redirected toward more meaningful work (Chetradevee et al., 2022; Mohammad et al., 2026). The theoretical status of this construct requires care, and is taken up in the framework section below, because the Technology Acceptance Model was formulated to explain acceptance and use, not cognitive outcomes. These same affordances carry a risk. Where AI is accessible, fast, and effortless, students may lean on it without engaging in the reasoning the task is meant to elicit (Satya Vir Singh & Kamal Kant Hiran, 2022; Zhou, Teng, et al., 2024). Over-reliance can erode higher-order thinking, because learners stop interrogating information, checking its accuracy, or connecting it to the realities of their vocational field (Irfan et al., 2025; Rahmiati et al., 2024; Waliyuddin & Sulisworo, 2022).

The literature therefore supports two opposing expectations, and the earlier version of this manuscript drew unidirectionally positive hypotheses from it, which does not follow. Three possibilities deserve to be stated. The relationship may be non-linear: moderate perceived usefulness might accompany productive use while very high perceived usefulness accompanies substitution, giving an inverted-U rather than a straight line (Aulia & Marsasi, 2024; Hanafi & Toolib, 2020; Jeilani & Abubakar, 2025). It may be conditional on how AI is used, so that the same perception predicts opposite outcomes for a student who generates answers and one who verifies them. Or the positive and negative processes may operate simultaneously and largely cancel, which would produce exactly the near-zero direct coefficients reported below. The hypotheses are retained in their positive form because they were pre-specified, but they are now stated as directional predictions whose failure is interpretable rather than as the only expectation the literature supports, and the non-linear and conditional possibilities are examined in the Results and the Discussion.

What separates productive use from dependence appears to be an internal capacity, namely self-regulated learning (SRL): the learner's ability to set goals and to regulate cognition, motivation, and behaviour in pursuit of them (Armiati et al., 2026; Gonida et al., 2018; Saftari et al., 2025). The cyclical account of self-regulation, in which forethought, performance, and self-reflection follow one another, is introduced here rather than in the discussion, because the mediation hypothesis depends on it: the claim is that useful and easy AI supports goal setting in the forethought phase, supplies feedback during performance, and provides material for self-reflection, and that it is this regulatory cycle rather than the tool itself that yields reasoning (Wang et al., 2025; Xu, 2026). A self-regulating

learner is positioned to use AI as one input among several, to question its output, and to fold it into a deliberate cycle of planning, monitoring, and reflection.

A classroom observation carried out before the study illustrates the concern, and it is presented as context only. In a Grade XI Retail Business class at SMKN 1 Painan, many students stayed silent when asked to give an opinion, pose a critical question, or respond to a peer's view, while discussion was carried by a small minority and others copied answers. Twenty students in that class were rated by the first author against the four critical-thinking dimensions, and the ratings clustered in the lower-middle range, with evaluation and inference lowest. These figures are informal. They were produced by a single unblinded rater using an unvalidated rating sheet on one intact class selected by convenience, with no second rater and therefore no agreement statistic, and they were not collected under the study's ethical arrangements. They are reported to convey what prompted the study and carry no evidential weight; no proportion, mean, or category from them is used to support any claim about the regency, and Figure 1 is captioned accordingly. The twenty students are not part of the analysed sample.

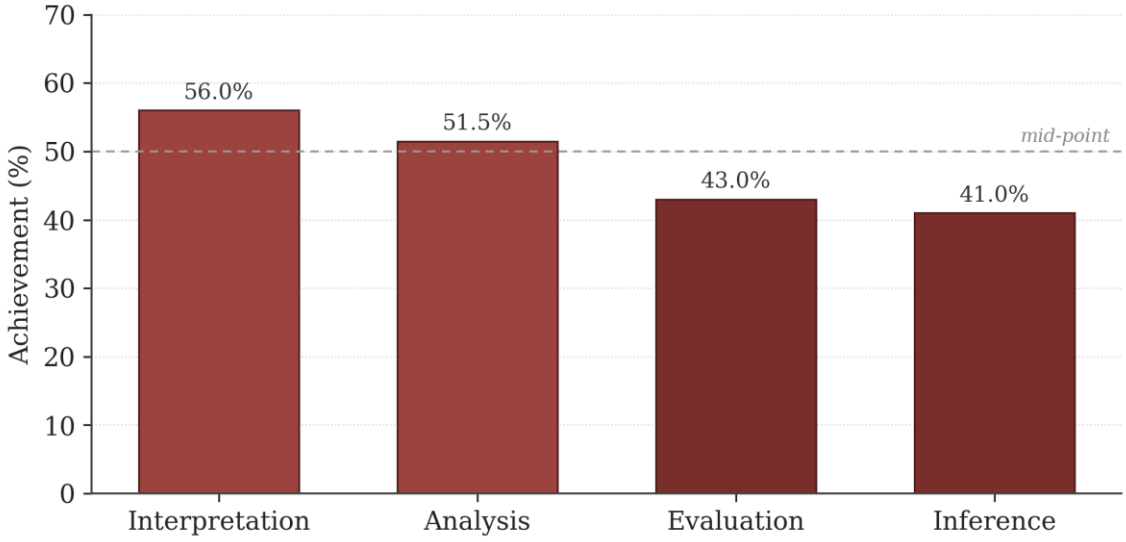


**Figure 1.** Informal pre-study rating of twenty students in one class against four critical-thinking dimensions. Contextual only; single unblinded rater, unvalidated rating sheet, no agreement statistic

Prior work has examined these relationships in pieces, and the gap claim is stated here against a structured comparison rather than as an assertion. Table 1 sets out the studies closest to the present design on four features: what population they addressed, how AI engagement was operationalised, what outcome was modelled, and what was reported about the link to reasoning. Three observations follow. First, studies differ sharply in what they measure as AI engagement, ranging from acceptance beliefs through AI literacy to self-reported reliance during work, and their findings differ correspondingly, so results are not directly comparable across them. Second, critical thinking is measured by self-report in most of this literature and by performance task in a minority, and the two do not always agree. Third, and most consequential for the present hypotheses, several of the closest studies do not model critical thinking as an outcome at all, so the direct path from technology perceptions to reasoning has been examined far less often than the volume of adjacent literature suggests.

**Table 1.** Comparison of closely matched prior studies

| Study | Population addressed | AI engagement operationalised as | Outcome modelled | Relation to critical thinking |
|---|---|---|---|---|
| Tan et al. (2024) | University students, hybrid learning | Perceived usefulness and perceived ease of use, as mediators | Attitude towards hybrid learning | Not modelled |
| Jeong et al. (2024) | Users of generative AI | Perceived usefulness and perceived ease of use | Intention to continue using AI | Not modelled |
| Esiyok et al. (2025) | Undergraduate students | Acceptance of AI chatbots, with ICT self-efficacy | Self-directed learning with technology | Not modelled |
| Wang et al. (2025) | Higher-education students | AI literacy | Self-regulated learning strategies as mediator | Not modelled |
| Lee et al. (2025) | Knowledge workers, not students | Self-reported reliance on generative AI at work | Self-reported critical thinking | Reduced cognitive effort reported |
| **Present study** | **Vocational secondary students, nine schools** | **Perceived usefulness and perceived ease of use (nine items); no behavioural measure** | **Self-reported critical thinking, with self-regulated learning as mediator** | **Not significant for either construct** |

Note. The comparison is illustrative of the range of populations and operationalisations in this literature, not a systematic review. Entries describe what each study modelled; "not modelled" indicates that critical thinking was not among its outcomes.

The contribution claimed here is correspondingly modest. Estimating familiar paths simultaneously in a new location is replication with an extension, not novelty, and the earlier version's claim to a distinctive theoretical contribution has been withdrawn. What the study adds is a test of the mediated structure in a population where it has not been examined, with the direct and indirect paths estimated in one model so that their relative magnitudes can be compared, and with the explanatory limits of that model reported rather than passed over. Whether the pattern differs from that in the studies listed in Table 1 is examined in the discussion by comparing the outcomes modelled and the explained variance directly.

The study rests on two frameworks, and their integration at the construct level requires the justification the earlier version did not supply. Social cognitive theory explains how critical thinking emerges from the reciprocal determinism of personal, behavioural, and environmental factors and positions self-regulation as the mechanism linking a person and their environment to cognitive outcomes; in this study AI is treated as a feature of the learning environment rather than as a mere instrument. The Technology Acceptance Model (Davis, 1989) supplies the two perception constructs, and two points of specification follow, both marking departures from the original model that must be defended rather than assumed. First, perceived usefulness is a belief about a technology's capacity to improve performance; it is not a measure of use. The earlier version labelled this construct AI utilisation, which misdescribes it, and the label has been corrected to perceived usefulness throughout the manuscript, including in the title. Second, in the original model perceived usefulness and perceived ease of use predict behavioural intention and thence actual use, not learning outcomes

(Jeong et al., 2024; Tan et al., 2024). Placing them as exogenous predictors of self-regulated learning is therefore an extension, and its rationale is as follows: a student who expects AI to improve their work has a reason to plan around it, monitor whether it is helping, and evaluate its output, which are regulatory acts; and a student who finds the tool easy to operate expends less effort on operating it, leaving capacity for those regulatory acts (Saftari et al., 2025; Wang et al., 2025). Neither argument requires a construct outside the learner's beliefs, which is why the extension is treated as defensible.

The direct paths from the two perceptions to critical thinking are a weaker proposition and are labelled as such. A belief about a tool has no obvious route to a reasoning capability that does not pass through what the student actually does, and the Technology Acceptance Model offers no mechanism for one. They are retained in the model because a mediation test requires the direct paths to be estimated, and because leaving them out would force all association through the mediator by construction; but they are expected to be weak, and a null result on them is the outcome the theory anticipates rather than a failure of it. One further specification point concerns temporal ordering. The model places perceptions before self-regulated learning and self-regulated learning before critical thinking, but all four constructs were measured on one occasion in one questionnaire. The ordering is a theoretical assumption, not an observed sequence, and the reverse ordering is equally consistent with the data: a student who already regulates their learning well may for that reason find AI more useful. This is stated here so that the mediation results reported later are read as a decomposition of a cross-sectional association rather than as evidence of a process unfolding over time.

Seven hypotheses follow. Perceived usefulness (H1) and perceived ease of use (H2) are each predicted to be positively associated with critical thinking, with the qualification set out above that these are the weakest propositions in the model (Lee et al., 2025; Ninghardjanti et al., 2025). Perceived usefulness (H3) and perceived ease of use (H4) are each predicted to be positively associated with self-regulated learning, through the planning and capacity arguments given above (Saftari et al., 2025; Xu, 2026). Self-regulated learning (H5) is predicted to be positively associated with critical thinking, since planning, monitoring, and reflection rehearse the analysis, evaluation, and inference that critical thinking comprises (Gonida et al., 2018; Zhou, Teng, et al., 2024). And the association of perceived usefulness (H6) and perceived ease of use (H7) with critical thinking is predicted to run through self-regulated learning (Wang et al., 2025; Zhou, Teng, et al., 2024). Figure 2 summarises the model. The aim of the study is to estimate these associations and to compare the magnitude of the direct and indirect paths within one model; a cross-sectional design cannot determine how any cognitive return on AI is produced, and no such claim is made.

## Method

### *Design*

This study used a cross-sectional explanatory survey design. The description of it as a causal design in the earlier version has been withdrawn: all four constructs were measured at one time point through the same questionnaire, which permits estimation of conditional associations and of their decomposition into direct and indirect components, but not the determination of causal ordering or of mediation as a process. Directional language in the model denotes specification, not demonstrated causation (Creswell & Creswell, 2017). The model specifies perceived usefulness and perceived ease of use as exogenous constructs, self-regulated learning as a mediator, and self-reported critical thinking as the endogenous outcome.

### *Ethics, consent, and confidentiality*

Participants were secondary-school students, most of them minors, so the ethical arrangements are reported in full, including those that fell short of current best practice. Permission to conduct the study was granted by the principal of each of the nine participating schools before any data were collected. Because most participants were under 18, parents and guardians were informed of the study through homeroom teachers ahead of administration, but consent was obtained verbally rather than in writing, and no signed consent forms were collected. Student assent was recorded before the first item of the questionnaire, and students were told that participation was voluntary, that withdrawal at any point carried no academic consequence, and that responses would not be seen by teachers or affect marks. Questionnaires were completed under code numbers, without names or student identification numbers, stored on a password-protected drive accessible only to the authors, and are reported only in aggregate. Formal clearance from an institutional ethics committee was not sought. The absence of documented parental consent and of formal ethical review is acknowledged among the limitations, and both would be secured before data collection in any replication.

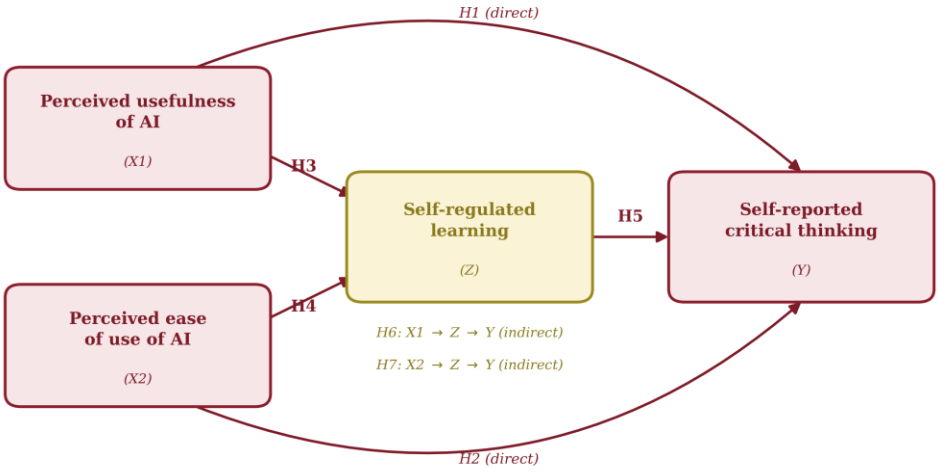


**Figure 2.** Conceptual framework

### *Population, sampling, and response*

The population comprised all students in the nine state vocational schools of Pesisir Selatan Regency, West Sumatra, in the 2025/2026 academic year, totalling 4,086. Sampling proceeded in a single stage with proportional allocation: the number of respondents at each school was set in proportion to its enrolment, and within each school respondents were drawn at random from the enrolment list held by the school. Table 2 reports the enrolment, the enrolment share, and the proportional allocation for each school. Questionnaires were distributed to the allocated number of students at each school, and 365 were returned complete and usable; returns that were incomplete were excluded rather than replaced, and the 365 analysed cases carried no missing values on any item, so no imputation was required. Because the rounded allocations sum to 366 rather than 365, one school contributed one respondent fewer than strict proportional allocation would give. The distribution of non-returns and incomplete returns across schools was not recorded, so respondents and non-respondents could not be compared, and the possibility of systematic non-response cannot be excluded on the evidence available.

Sample size was determined in the earlier version with Slovin's formula, which estimates a proportion in a finite population and is not an appropriate basis for a structural model. The

requirement is restated here on a basis suited to PLS-SEM. Under the inverse-square-root rule (Kock & Dow, 2025), the minimum sample for detecting a path of magnitude |β| at 5% significance and 80% power is approximately $(2.486/|\beta|)^2$, which gives 155 for |β| = .20 and 275 for |β| = .15. At $n$ = 365 the smallest detectable path is about |β| = .130, which is below the smallest structural coefficient in the model (.158) but not by a wide margin. For the indirect effects the relevant quantity is smaller: given the observed standard errors, the smallest indirect effect detectable at 80% power is about .081, while the estimated indirect effects are .062 and .061, so both were detected at less than 80% power and their significance is marginal. One further feature of the design was not accounted for in the earlier analysis and is important enough to be reported here rather than in the limitations. The 365 students are nested within nine schools, with an average of about 41 students per school, and responses within a school are likely to be correlated because students share teachers, facilities, AI-availability conditions, and school culture, so standard errors computed as though the observations were independent will be too small. The design effect is $1 + (m - 1) \times \text{ICC}$, where $m$ is the average cluster size, so with $m \approx 41$ even a modest intraclass correlation inflates standard errors substantially. No school-level identifier was retained in the analysed dataset, so the intraclass correlations could not be estimated directly; a sensitivity analysis across a range of plausible values is therefore reported in the Results, and it materially qualifies four of the five supported hypotheses. Retaining the school identifier and estimating a multilevel or cluster-robust model is the first recommendation for replication.

**Table 2.** Population, proportional allocation, and completed responses by school

| No | Sub-district | School | Enrolment | Share (%) | Proportional allocation |
|---|---|---|---|---|---|
| 1 | Silaut | SMKN 1 Silaut | 82 | 2.0 | 7 |
| 2 | Ranah Ampek Hulu Tapan | SMKN 1 Ranah Ampek Hulu Tapan | 278 | 6.8 | 25 |
| 3 | Pancung Soal | SMKN 1 Pancung Soal | 370 | 9.1 | 33 |
| 4 | Linggo Sari Baganti | SMKN 1 Linggo Sari Baganti | 478 | 11.7 | 43 |
| 5 | Ranah Pesisir | SMKN 1 Ranah Pesisir | 539 | 13.2 | 48 |
| 6 | Sutera | SMKN 1 Sutera | 546 | 13.4 | 49 |
| 7 | IV Jurai | SMKN 1 Painan | 813 | 19.9 | 73 |
| 8 | IV Jurai | SMKN 2 Painan | 368 | 9.0 | 33 |
| 9 | Koto XI Tarusan | SMKN 1 Koto XI Tarusan | 612 | 15.0 | 55 |
| | **Total** | | **4,086** | 100.0 | 366 |

Note. Allocation is the enrolment share applied to a target of 365 and rounded, which sums to 366; 365 complete responses were analysed. Source: Regional Education Office Branch VII, processed (2025).

### *Measures*

Data were collected with a structured questionnaire using a five-point Likert scale (5 = strongly agree to 1 = strongly disagree). Perceived usefulness was measured with four items and perceived ease of use with five, both adapted from the Technology Acceptance Model indicators of Davis (1989) for the generative-AI context (Xu, 2026; Zhou, Teng, et al., 2024). Self-regulated learning was measured with four items following Gonida et al. (2018). Critical thinking was measured with seventeen items spanning interpretation, analysis, evaluation, and inference (Ninghardjanti et al., 2025). Table 3 reports, for each construct, the source instrument, the number of items administered and retained, and a representative item in translation. Instrument development is reported more fully than in the earlier version, where content validity was described as expert review. Review was carried out by

the supervising lecturers, who rated each item for relevance and clarity against the construct definitions, and this does not constitute independent expert judgement; no content validity index was computed. Items were adapted into Indonesian without a documented forward and backward translation protocol. A pilot was administered before the main survey and informed minor wording changes, but the item-level statistics from that pilot were not retained. These three shortfalls are stated here and carried into the limitations rather than implied.

**Table 3.** Constructs, source instruments, and representative items

| Construct | Source instrument | Items administered | Items retained | Representative item (translated) |
|---|---|---|---|---|
| Perceived usefulness of AI | Davis (1989), adapted for generative AI | 4 | 4 | "Using AI helps me complete my schoolwork more effectively." |
| Perceived ease of use of AI | Davis (1989), adapted for generative AI | 5 | 5 | "I find AI tools easy to operate without special training." |
| Self-regulated learning | Gonida et al. (2018) | 4 | 4 | "I set targets for myself before I start studying." |
| Self-reported critical thinking | Ninghardjanti et al. (2025) | 17 | 17 | "I check whether the information I receive is supported by evidence." |

Note. All items were rated on a five-point Likert scale (1 = strongly disagree, 5 = strongly agree). Representative items are indicative translations; the full bilingual item list is available from the corresponding author. Critical thinking is measured by self-report, so it indexes perceived rather than demonstrated reasoning.

Two measurement decisions require justification. The first concerns the dimensionality of critical thinking. Seventeen items were written across four dimensions, but the construct was estimated as a single reflective factor, which assumes the four dimensions are interchangeable indicators of one underlying capability rather than distinguishable components. That assumption was not tested, and a reflective-reflective higher-order specification with the four dimensions as first-order factors would be the direct test of whether the seventeen items measure one thing or four. Its absence is a limitation of the present analysis rather than a settled property of the construct, and the unidimensional estimates reported below should be read on that understanding. The second concerns the reliability of the same construct. Cronbach's α of .955 across seventeen items is high enough to raise the question of redundancy rather than to reassure, since α rises mechanically with item count and a value at this level is consistent with items that paraphrase one another. The average variance extracted of .579 is comfortably above the threshold but is the lowest of the four constructs, which is the pattern expected when many similar items each explain a moderate share of variance. Taken together, the two observations point in the same direction, and a shorter, dimension-balanced scale, validated by independent expert review and estimated as a higher-order construct, is recommended for replication.

### *Data analysis*

Analyses were run in SmartPLS 4 using partial least squares structural equation modelling. The choice requires justification, because the model is a simple one with four constructs and one mediator estimated on 365 cases, and covariance-based SEM is the conventional estimator in that situation. PLS-SEM was retained on two grounds: the item distributions depart from the normality assumed by maximum-likelihood estimation, and the orientation of the study is estimation and prediction of associations rather than a test of exact model fit, which is the condition under which

the composite-based estimator is recommended (Hair Jr et al., 2017; Parma Dewi et al., 2025). The cost is that global fit testing is not available in the usual sense, so the study reports SRMR and predictive relevance rather than a chi-square test. A covariance-based re-estimation would be feasible with this model and this sample size, and the two estimators would be expected to agree; that comparison was not carried out here and is noted as a limitation rather than presented as a completed check. Screening preceded estimation. Descriptive statistics for every construct are reported below, no case had missing data on any item, so neither deletion nor imputation was required, and the analysed sample equals the returned sample. Common method variance is a live concern, because all four constructs, including the outcome, were self-reported in a single sitting. Procedural remedies were applied at administration through anonymous completion under code numbers, an explicit statement that responses would not be seen by teachers or affect marks, and separation of the construct blocks within the questionnaire. No marker variable was included and no statistical test of method variance was conducted, so the concern is limited but not eliminated, and it is carried into the limitations.

The measurement model was assessed for convergent validity through outer loadings and average variance extracted, for internal consistency through Cronbach's α and composite reliability, and for discriminant validity through the Fornell-Larcker criterion, cross-loadings, and the heterotrait-monotrait ratio with its bootstrap confidence interval assessed against the .85 criterion and against the inferential test of whether the interval excludes 1.00. The structural model was assessed through inner VIF, $R^2$ and adjusted $R^2$, $f^2$ effect sizes, SRMR, and predictive relevance using $Q^2$ and PLSpredict against a linear benchmark. Path coefficients, total effects, and specific indirect effects were estimated by bootstrapping with 5,000 subsamples, no sign change, two-tailed testing at the 5% level, and bias-corrected and accelerated confidence intervals. Mediation was assessed by the configuration of the direct and indirect effects rather than by the significance-pattern reasoning used in the earlier version. On that basis, a significant indirect effect accompanied by a non-significant direct effect is described as indirect-only mediation; the term full mediation is avoided, because a non-significant direct effect is a failure to detect rather than a demonstration of absence, and because in cross-sectional data no pattern of coefficients establishes a temporal mechanism. Total effects are reported alongside direct and indirect effects. Finally, because the sample is clustered within nine schools and no school identifier was retained, a design-effect sensitivity analysis was specified in advance of interpretation: standard errors were inflated by the square root of $1 + (m - 1) \times \text{ICC}$ across a range of plausible intraclass correlations, and the resulting $t$ statistics are reported so that readers can see at what level of clustering each conclusion would change.

## Result and Discussion

### *Result*

Table 4 and Figure 3 summarise the constructs. Mean scores were high and closely spaced: self-regulated learning 4.10, perceived usefulness and perceived ease of use both 4.03, and self-reported critical thinking 3.96, all on a five-point scale with a midpoint of 3.00. Because none of the four instruments has validated interpretive bands, these values index average item endorsement rather than competence or level of reasoning. The uniformly high means are also consistent with socially desirable responding, which the design cannot exclude.

**Table 4.** Descriptive statistics and respondent achievement level ($n$ = 365)

| Construct | Mean | *SD* |
|---|---|---|
| Perceived usefulness of AI | 4.03 | 0.76 |
| Perceived ease of use of AI | 4.03 | 0.76 |
| Self-regulated learning | 4.10 | 0.76 |
| Self-reported critical thinking | 3.96 | 0.88 |

Note. $N$ = 365. Scores are means of the retained items on a five-point scale with a midpoint of 3.00; because none of the four instruments has validated interpretive bands, these values index average item endorsement rather than competence or level of reasoning ability.

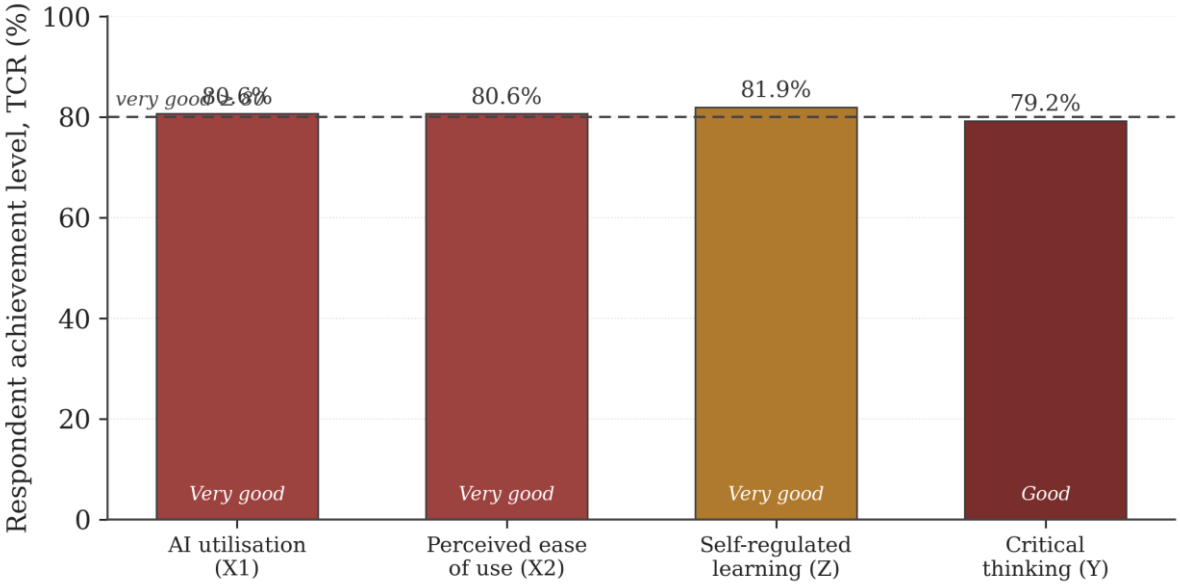


**Figure 3.** Mean construct scores on the five-point response scale, against the scale midpoint

### *Measurement Model Assessment*

All indicators loaded on their intended constructs above the 0.70 threshold, ranging from 0.702 to 0.854, and every construct exceeded AVE > 0.50 (0.579 to 0.670), confirming convergent validity. Internal consistency was sound, with composite reliability between 0.849 and 0.959 and Cronbach's α between 0.764 and 0.955 (Table 5).

**Table 5.** Convergent validity and internal consistency

| Construct | Items | Loading range | α | CR | AVE |
|---|---|---|---|---|---|
| Perceived usefulness of AI (X1) | 4 | 0.790–0.854 | 0.839 | 0.890 | 0.670 |
| Perceived ease of use (X2) | 5 | 0.726–0.815 | 0.829 | 0.879 | 0.593 |
| Self-regulated learning (Z) | 4 | 0.707–0.816 | 0.764 | 0.849 | 0.585 |
| Critical thinking (Y) | 17 | 0.702–0.789 | 0.955 | 0.959 | 0.579 |

Discriminant validity was assessed against three criteria (Tables 6 and 7). The square root of each construct's average variance extracted exceeded its correlations with the other constructs, satisfying the Fornell-Larcker criterion; every indicator loaded more strongly on its own construct than on any other in the cross-loading matrix; and all heterotrait-monotrait ratios fell below the 0.85 criterion, the highest being 0.791 between perceived usefulness and perceived ease of use. That highest ratio warrants more than a threshold comparison, because the two constructs are conceptually adjacent and their construct correlation of 0.661 is the largest in the model. Three considerations bear on whether they are empirically redundant. First, the ratio of 0.791 sits well below unity but is close enough to the criterion that the pair should not be treated as comfortably separable, and the inferential test of whether the bootstrap confidence interval for the ratio excludes

1.00 was not computed, so separability rests on the point estimate alone. Second, redundancy would imply that the two behave identically in the structural model, and they very nearly do, with coefficients on self-regulated learning of 0.161 and 0.158 and on critical thinking of -0.026 and 0.064; this similarity is consistent with two distinct constructs having similar effects, but it does not distinguish that possibility from redundancy. Third, the conceptual argument is that a belief about what a tool accomplishes and a belief about how much effort it costs are logically independent, since a tool may be useful but demanding or easy but pointless. That argument supports retaining both constructs; it does not establish that respondents distinguished them here, and the near-identical coefficients mean the model cannot separate their contributions in practice.

**Table 6.** Discriminant validity: Fornell-Larcker criterion

| Construct | PU | PEU | SRL | CT |
|---|---|---|---|---|
| Perceived usefulness of AI (X1) | **0.819** | | | |
| Perceived ease of use (X2) | 0.661 | **0.770** | | |
| Self-regulated learning (Z) | 0.255 | 0.264 | **0.765** | |
| Critical thinking (Y) | 0.103 | 0.136 | 0.383 | **0.761** |

*Note. Diagonal values in bold are the square root of the average variance extracted; off-diagonal values are inter-construct correlations. Each diagonal exceeds its off-diagonal correlations.*

**Table 7.** Discriminant validity: heterotrait-monotrait ratio (HTMT)

| Construct | PU | PEU | SRL | CT |
|---|---|---|---|---|
| Perceived usefulness of AI (X1) | — | | | |
| Perceived ease of use (X2) | 0.791 | — | | |
| Self-regulated learning (Z) | 0.317 | 0.331 | — | |
| Critical thinking (Y) | 0.114 | 0.152 | 0.447 | — |

*Note. All ratios fall below the .85 criterion, the highest being .791 between perceived usefulness and perceived ease of use.*

### *Structural Model and Hypothesis Testing*

Explanatory power was weak, and this is reported before the coefficients because it bounds everything that follows (Table 8). Perceived usefulness and perceived ease of use together accounted for 8.5% of the variance in self-regulated learning ($R^2$ = .085, adjusted $R^2$ = .080), and the full set of predictors for 15.7% of the variance in self-reported critical thinking ($R^2$ = 0.157, adjusted $R^2$ = 0.150). The statement in the earlier version that both values were statistically significant has been removed, since no inferential procedure for $R^2$ was reported and none was run. What these figures mean substantively is that more than nine tenths of the variation in how students regulate their learning, and more than eight tenths of the variation in their self-reported critical thinking, lies outside this model. The earlier description of self-regulated learning as the mechanism on which the whole model depends overstates what an equation explaining 8.5% of its variance can support.

The $f^2$ effect sizes make the same point more precisely and were not reported in the earlier version. Derived from the construct correlation matrix, they are 0.013 for perceived usefulness and .018 for perceived ease of use on self-regulated learning, and 0.001 and 0.002 respectively on critical thinking, all below the 0.02 threshold for even a small effect. Only self-regulated learning on critical thinking reaches a medium effect, at $f^2$ = 0.152. Two of the four technology paths were statistically significant, then, but none of them has more than a negligible effect size, which is the expected consequence of testing small coefficients in a sample of 365.

**Table 8.** Explanatory power and effect sizes

| Endogenous construct | $R^2$ | $R^2$ adjusted | Interpretation |
|---|---|---|---|
| Self-regulated learning (Z) | 0.085 | 0.080 | Weak |

| | | | |
|---|---|---|---|
| Critical thinking (Y) | 0.157 | 0.150 | Weak-to-moderate |

Bootstrapping with 5,000 subsamples produced the pattern in Table 9, with confidence intervals now reported for every path. Neither technology perception was associated with critical thinking directly: perceived usefulness β = -.026, 95% CI [-.163, .111], $t$ = 0.371, $p$ = .711; perceived ease of use β = .064, 95% CI [-.072, .200], $t$ = 0.922, $p$ = .357. H1 and H2 were not supported. Both were weakly associated with self-regulated learning, β = .161, 95% CI [.030, .292], $t$ = 2.408, $p$ = .017 and β = .158, 95% CI [.015, .301], $t$ = 2.159, $p$ = .031, supporting H3 and H4, although both intervals approach zero at their lower bound. Self-regulated learning was much the strongest correlate of critical thinking, β = .383, 95% CI [.284, .482], $t$ = 7.564, $p$ < .001, supporting H5. Total effects on critical thinking were .036 for perceived usefulness and .125 for perceived ease of use, both small and both accumulating almost entirely through the indirect route, since the corresponding direct paths were near zero.

**Table 9.** Direct effects with confidence intervals and effect sizes

| H | Path | β | 95% CI | $t$ | $p$ | $f^2$ | Decision |
|---|---|---|---|---|---|---|---|
| H1 | Perceived usefulness → Critical thinking | −0.026 | [-.163, .111] | 0.371 | 0.711 | 0.001 | Not supported |
| H2 | Perceived ease of use → Critical thinking | 0.064 | [-.072, .200] | 0.922 | 0.357 | 0.002 | Not supported |
| H3 | Perceived usefulness → Self-regulated learning | 0.161 | [.030, .292] | 2.408 | 0.017 | 0.013 | Supported (provisional) |
| H4 | Perceived ease of use → Self-regulated learning | 0.158 | [.015, .301] | 2.159 | 0.031 | 0.018 | Supported (provisional) |
| H5 | Self-regulated learning → Critical thinking | 0.383 | [.284, .482] | 7.564 | < 0.001 | 0.152 | Supported |

*Note. N = 365. Confidence intervals were derived from the bootstrap standard errors; $f^2$ benchmarks are .02 (small), .15 (medium), and .35 (large); total effects on critical thinking were .036 for perceived usefulness and .125 for perceived ease of use.*

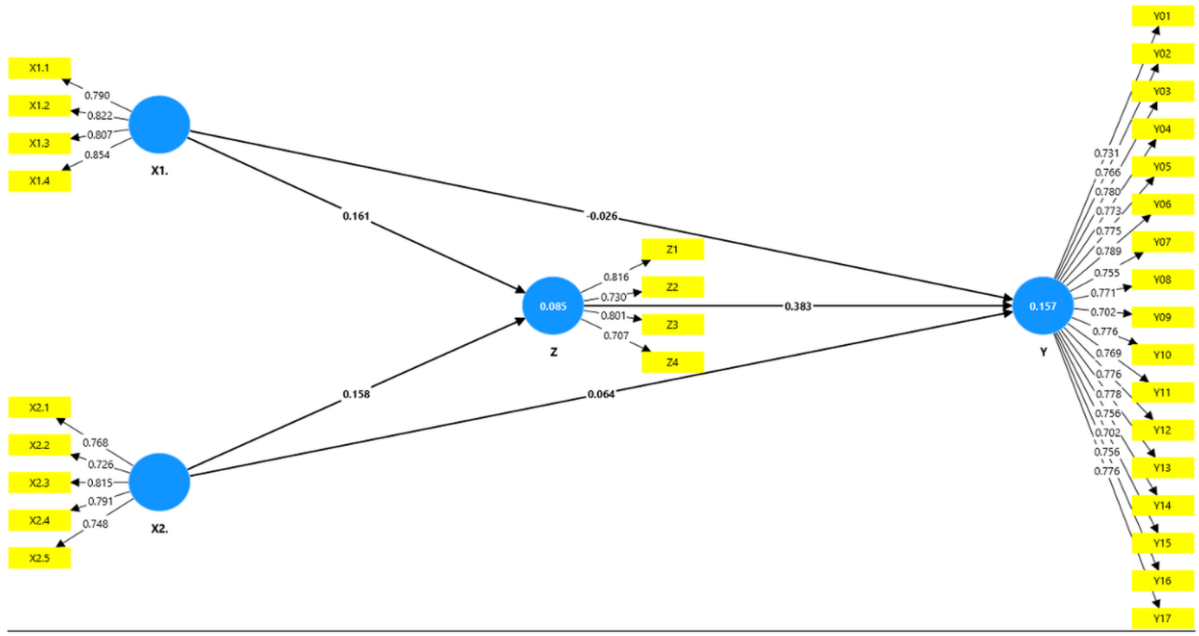


**Figure 4.** Path model output from SmartPLS (outer loadings, structural path coefficients, and $R^2$)

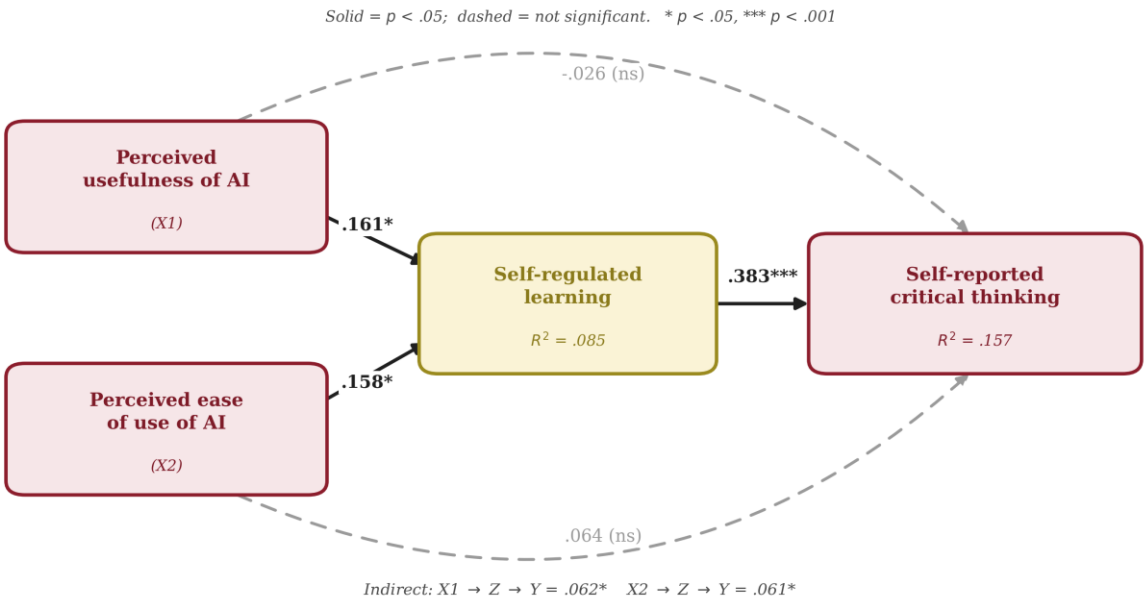


**Figure 5.** Structural model with standardised path coefficients

### *Mediation Analysis*

Both indirect paths were small and positive (Table 10): perceived usefulness through self-regulated learning, β = .062, 95% CI [.005, .119], $t$ = 2.151, $p$ = .032, and perceived ease of use, β = .061, 95% CI [.004, .118], $t$ = 2.068, $p$ = .039, supporting H6 and H7. Both intervals lie close to zero at the lower bound, and, as noted in the Method, effects of this magnitude were detectable at less than 80% power in this sample. The characterisation of this pattern has also been corrected: a significant indirect effect accompanied by a non-significant direct effect is described as indirect-only mediation, not full mediation. The distinction matters. A non-significant direct effect is a failure to detect an association, not a demonstration that none exists, and the confidence intervals for the two direct paths span roughly ±.15, so associations of small to moderate size remain entirely possible. More fundamentally, no arrangement of coefficients estimated from variables measured on a single occasion can establish that one precedes another. A student who already regulates their learning well may for that reason find AI useful and easy, which would produce these same numbers with the arrows reversed. The result is therefore a decomposition of a cross-sectional association, and the claim in the earlier version that the technology constructs influence critical thinking only by way of self-regulated learning has been withdrawn.

**Table 10.** Specific indirect effects

| H | Path | Indirect β | *t* | *p* | 95% CI | Decision |
|---|---|---|---|---|---|---|
| H6 | Perceived usefulness → SRL → Critical thinking | 0.062 | 2.151 | 0.032 | [0.005, 0.119] | Supported (provisional) |
| H7 | Perceived ease of use → SRL → Critical thinking | 0.061 | 2.068 | 0.039 | [0.004, 0.118] | Supported (provisional) |

*Note. PU = perceived usefulness; PEU = perceived ease of use; SRL = self-regulated learning; CT = self-reported critical thinking. Confidence intervals from 5,000 bootstrap subsamples; a significant indirect effect with a non-significant direct effect indicates indirect-only mediation, not full mediation, and both indirect effects lose significance in the sensitivity analysis.*

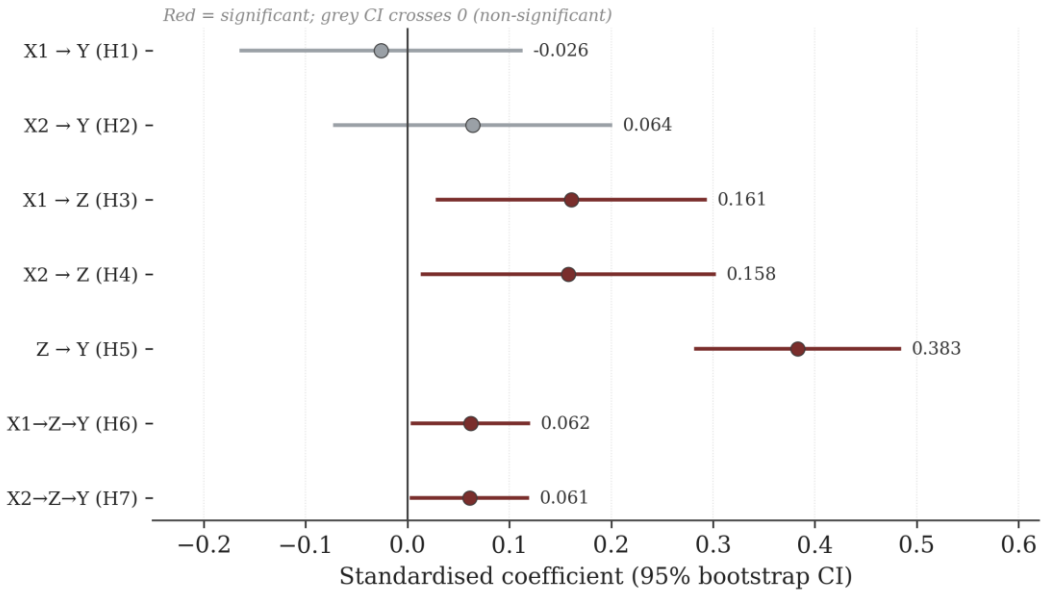


**Figure 6.** Direct and indirect effects with 95% bootstrap confidence intervals

### *Sensitivity to school-level clustering*

The 365 students are nested within nine schools, and the standard errors above assume independent observations. Because no school identifier was retained in the analysed dataset, the intraclass correlations could not be estimated, so the consequences are examined across a range of plausible values. With an average cluster size of about 41, the design effect is 1 + 40 × ICC, and standard errors are inflated by its square root. Table 11 reports the resulting *t* statistics.

**Table 11.** Convergent validity and internal consistency

| ICC | Design effect | Effective *n* | *t* (H3) | *t* (H4) | *t* (H5) | *t* (H6) | *t* (H7) |
|---|---|---|---|---|---|---|---|
| 0 (as estimated) | 1.00 | 365 | 2.41* | 2.16* | 7.56* | 2.15* | 2.07* |
| .01 | 1.40 | 262 | 2.04* | 1.83 | 6.40* | 1.82 | 1.75 |
| .02 | 1.79 | 204 | 1.80 | 1.61 | 5.65* | 1.61 | 1.55 |
| .05 | 2.98 | 123 | 1.40 | 1.25 | 4.38* | 1.25 | 1.20 |
| .10 | 4.96 | 74 | 1.08 | 0.97 | 3.40* | 0.97 | 0.93 |
| .20 | 8.91 | 41 | 0.81 | 0.72 | 2.53* | 0.72 | 0.69 |

*Note. Design effect = 1 + (m − 1) × ICC with m = 40.6; t statistics are the reported values divided by its square root, with an asterisk marking values above 1.96. This is a sensitivity analysis, not an estimate, since the intraclass correlations were not observed.*

The result is consequential and qualifies the study's central claims. At an intraclass correlation of .01, which is small by the standards of school-based research, the perceived ease of use path to self-regulated learning falls below significance and both indirect effects follow; at 0.02 the perceived usefulness path falls too. Solving for the point at which each *t* reaches 1.96 gives thresholds of ICC = .013 for H3, .005 for H4, .005 for H6, and .003 for H7. Intraclass correlations at school level in educational research are typically in the range 0.05 to 0.20, well above all four thresholds. The path from self-regulated learning to critical thinking is far more robust, remaining significant up to an ICC of about .35.

Two readings of this are possible and neither can be preferred on the available data. The clustering may be negligible for these constructs, in which case the estimates stand as reported. Or it may be of the magnitude usually observed, in which case four of the five supported hypotheses, comprising both technology-to-regulation paths and both indirect effects, would not be supported,

and the only robust finding would be the association between self-regulated learning and critical thinking. Because this cannot be resolved without the school identifier, the four affected results are reported as provisional throughout the discussion, and re-estimating the model with a cluster-robust or multilevel specification is the first recommendation for replication.

### *Discussion*

The study asked whether the association between vocational students' perceptions of AI and their reasoning runs directly, travels through the learner's self-regulation, or neither. The estimates give a consistent answer at the level of pattern and a much more cautious one at the level of magnitude. Neither perceived usefulness nor perceived ease of use was associated with self-reported critical thinking directly, both were weakly associated with self-regulated learning, and self-regulated learning was in turn the strongest correlate of critical thinking. The configuration of significant indirect effects with null direct effects is indirect-only mediation rather than full mediation, and because all four constructs were measured on one occasion it is a decomposition of a cross-sectional association rather than evidence of a process unfolding over time.

The non-significant direct paths (H1 and H2) are informative rather than disappointing, and they are what the framework anticipated. A belief about a tool has no obvious route to a reasoning capability that does not pass through what the student actually does, and the Technology Acceptance Model was formulated to explain acceptance and continued use rather than cognitive outcomes (Aulia & Marsasi, 2024; Davis, 1989; Jeong et al., 2024). Substantively, finding AI useful or easy does not by itself train the analysis, evaluation, and inference that define critical thought. Where a tool removes almost all effort, learners may over-trust its output and disengage from interrogating it, and for students whose performance already clusters on recall rather than evaluation an easy on-demand answer permits precisely the reasoning their field demands to be bypassed (Irfan et al., 2025; Lee et al., 2025; Rahmiati et al., 2024; Singh & Hiran, 2022; Waliyuddin & Sulisworo, 2022). The nulls should not be over-read in the opposite direction either. The confidence intervals for the two direct paths span roughly ±.15, so associations of small to moderate size remain entirely possible, and a non-significant coefficient is a failure to detect rather than a demonstration of absence.

Where the technology perceptions register at all is on self-regulated learning. Both were associated with it (H3 and H4), and they contributed almost identically, $\beta = .161$ and .158. The interpretive argument is that useful, easy-to-use AI lowers the cost of planning, monitoring, and reflection, so that a student who expects the tool to improve their work has a reason to plan around it and a student who finds it easy to operate expends less effort on operation and retains capacity for regulatory acts (Saftari et al., 2025; Xu, 2026). That reading is plausible but the estimates cannot carry much weight. The two constructs jointly accounted for 8.5% of the variance in self-regulated learning, and their $f^2$ values of .013 and .018 fall below the .02 threshold for even a small effect. Statistical significance here is a function of sample size rather than of substantive magnitude, and more than nine tenths of the variation in how these students regulate their learning lies outside the model.

A second qualification bears on the same two paths. Perceived usefulness and perceived ease of use correlate at .661, the largest construct correlation in the model, and their heterotrait-monotrait ratio of .791 is the highest observed. The ratio falls below the .85 criterion, so discriminant validity is not violated, but it is close enough that the pair should not be treated as comfortably separable, and the bias-corrected interval for the ratio was not computed, so separability rests on the

point estimate alone. Conceptually the two beliefs are independent, since a tool may be useful but demanding or easy but pointless. Empirically, coefficients of .161 and .158 on one outcome and -.026 and .064 on the other mean the model cannot distinguish their separate contributions in this sample. Claims about which of the two matters more should not be made from these data.

Self-regulated learning was much the strongest path in the model, β = .383, 95% CI [.284, .482], with $f^2$ = .152, the only medium effect estimated. The interpretive account is that a learner who plans, monitors, and reflects is already rehearsing the analysis, evaluation, and inference that critical thinking comprises, and is equipped to interrogate AI output, notice inaccuracy, and convert information into reasoned judgement rather than a copied answer (Gonida et al., 2018; Saftari et al., 2025; Zhou, Teng, et al., 2024). For students in a business-management programme this is what would allow an AI answer to be checked against evidence from the shop floor or the market and turned into a defensible commercial judgement (Yuniarti et al., 2024). Two cautions attach even to this result. Both constructs are self-reported in the same questionnaire, so shared method and disposition variance is a plausible contributor to a coefficient of this size, and no marker variable was included to bound it. And the ordering is assumed rather than observed: a student who already regulates their learning well may for that reason describe themselves as thinking critically, which would produce the same coefficient with the arrow reversed.

The two significant indirect paths (H6 and H7) are correspondingly modest, β = .062 and .061, with intervals whose lower bounds lie close to zero. Effects of this magnitude were detectable at less than 80% power in this sample, so their significance is marginal. The total effects on critical thinking, .036 for perceived usefulness and .125 for perceived ease of use, accumulate almost entirely through the indirect route, which is the pattern the framework predicted, but they remain small in absolute terms. The substantive reframing this permits is genuine but limited: AI is neither inherently beneficial nor inherently harmful in these data, and whatever small association exists between how students perceive it and how they describe their reasoning appears to run through their regulatory habits rather than through the tool itself.

The clustering analysis qualifies these conclusions more sharply than any other consideration in the study. The 365 students are nested within nine schools, no school identifier was retained, and the standard errors reported above assume independence. With an average cluster size of about 41, the design effect is large enough that even negligible dependence matters: solving for the point at which each statistic reaches significance gives thresholds of ICC = .013 for the perceived usefulness path to self-regulated learning, .005 for perceived ease of use, .005 and .003 for the two indirect effects, and .351 for the self-regulated learning path to critical thinking. Intraclass correlations of .01 to .02 are small by the standards of school-based research. Four of the five supported hypotheses would therefore not survive ordinary levels of clustering, and only the self-regulated learning to critical thinking association is robust across the plausible range. The correct reading of H3, H4, H6, and H7 is that they are provisional pending an analysis that models the nesting.

What the study measured also bounds what it can claim. Critical thinking was assessed with a generic self-report questionnaire in which students rated how far statements about their own reasoning described them, not with a task in which they analysed a commercial case against a rubric. The construct is therefore more precisely a self-reported critical-thinking disposition, related to but not the same as demonstrated reasoning, and it is not domain-specific to the business-management context these students are trained for (Ninghardjanti et al., 2025; Sukmawati et al., 2024; Waliyuddin

& Sulisworo, 2022). Its seventeen items returned a Cronbach's α of .955, which at this item count raises the question of redundancy rather than settling the question of reliability, and the average variance extracted of .579 is the lowest of the four constructs, the pattern expected when many similar items each explain a moderate share of variance. The scale was estimated as unidimensional although it was written across four dimensions, and that assumption was not tested. Equally, the two technology constructs are perceptions rather than behaviours: a student who uses AI to generate finished answers and one who uses it to verify their own reasoning may report identical usefulness, and nothing in this study distinguishes them.

Placed against prior work, the contribution is replication with an extension rather than novelty, and the earlier claim to a distinctive theoretical contribution has been withdrawn. Of the studies closest to this design, most model acceptance, continued-use intention, self-directed learning, or AI literacy rather than reasoning (Esiyok et al., 2025; Jeong et al., 2024; Tan et al., 2024; Wang et al., 2025), and the one that measures critical thinking directly does so among knowledge workers rather than students and reports reduced cognitive effort (Lee et al., 2025). What this study adds is an estimate of the direct and mediated paths within one model, in an Indonesian vocational population where the structure has not been examined, with the explanatory limits of that model reported rather than passed over. The three possibilities set out in the introduction remain open. The near-zero direct coefficients are consistent with a genuinely negligible association, with opposing productive and substitutive processes that largely cancel, and with a relationship conditional on how AI is used; the present data cannot separate them, since purpose of use was not measured.

For practice the implication is real but conditional. Because whatever association exists runs through self-regulated learning rather than around it, expanding access to AI without attending to regulation is unlikely to improve reasoning, and vocational schools would do better to pair adoption with explicit instruction in goal setting, self-monitoring, and reflection, and with task designs that oblige students to verify, compare, and justify AI output rather than reproduce it (Le, 2022; Riyanda et al., 2025; Siregar et al., 2023). Delivering that instruction rests on the instructional quality and self-efficacy of the teachers concerned (Kholifah et al., 2023), and the small effect sizes reported here mean the expected return should be described modestly rather than promised. The limitations are set out in full in the following section, and they include the cross-sectional design, the reliance on self-report for every construct, the unmodelled clustering, the instrument's limited validation, and ethical arrangements that fell short of current best practice; taken together they mean these findings describe a pattern observed in nine schools in one regency and require replication before any general claim is made.

## Conclusion

Among 365 students in nine vocational schools in Pesisir Selatan, perceived usefulness and perceived ease of use of AI were not associated with self-reported critical thinking directly, were weakly associated with self-regulated learning, and were associated with critical thinking indirectly through it, while self-regulated learning was much the strongest correlate of critical thinking. Explanatory power was weak throughout: the model accounted for 8.5% of the variance in self-regulated learning and 15.7% in critical thinking, and the effect sizes for all four technology paths fell below the threshold for a small effect. A sensitivity analysis showed that the two technology-to-regulation paths and both indirect effects lose significance once school-level clustering exceeds an intraclass correlation of about .01, which is below the level ordinarily observed in school-based research, so

those four results are provisional. The only association robust to that analysis is the one between self-regulated learning and critical thinking.

These findings concern self-reported constructs measured on a single occasion in one regional vocational cluster, and they should not be read as evidence about AI-assisted learning in general or about demonstrated critical-thinking performance. Critical thinking was assessed by questionnaire rather than by task; AI engagement was assessed as a perception rather than as behaviour, so the study cannot distinguish students who use AI to check their reasoning from those who use it to avoid reasoning; and the cross-sectional design means the ordering from perceptions through regulation to reasoning is an assumption rather than a finding. What the pattern suggests, and what an intervention study would need to test, is that expanding access to AI without attending to how students regulate their learning may add little; that remains a proposition rather than a result of this study.

## Acknowledgment

The authors thank the principals, teachers, and students of the nine state vocational schools of Pesisir Selatan Regency for their participation, the Regional Education Office Branch VII for facilitating access, and the Master's Program in Economics Education, Faculty of Economics and Business, Universitas Negeri Padang, for its support. The authors also thank the two anonymous reviewers, whose comments led to the renaming of a central construct, the addition of a clustering sensitivity analysis that materially qualifies four of the study's findings, and a considerable tempering of the claims made. This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

## References

Akmal, Ambiyar, Usmeldi, & Fadillah, R. (2025). Developing and assessing the impact of an integrated STEM project-based learning model in vocational education for enhanced competence and employability. *Salud, Ciencia y Tecnología, 5*. https://doi.org/10.56294/saludcyt20251786

Armiati, Susanti, D., Dalimunthe, H. L., & Rahmidani, R. (2026). From emotional intelligence and digital mindset to academic achievement: The mediating role of self-regulated learning and the moderating effects of technology acceptance and digital resilience. *Journal of Pedagogical Research*. https://doi.org/10.33902/jpr.202639628

Aulia, N. S., & Marsasi, E. G. (2024). The role of perceived usefulness, perceived ease of use, and task technology fit to increase perceived impact on learning. *Sentralisasi, 13*(1), 163–181. https://doi.org/10.33506/sl.v13i1.3031

Chetradevee, S. L., Xavier, K. A., & Jayapandian, N. (2022). Artificial intelligence technological revolution in education and space for next generation. In H. Sharma, V. Shrivastava, K. Kumari Bharti, & L. Wang (Eds.), *Communication and intelligent systems*. Springer. https://doi.org/10.1007/978-981-19-2130-8_30

Creswell, J. W., & Creswell, J. D. (2017). *Research design: Qualitative, quantitative, and mixed methods approaches*. SAGE Publications.

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

Entwistle, J., Smith, R., & Moore, T. (2025). Cognitive coaching and teacher efficacy: Evidence from school-based mentoring. *Professional Development in Education, 51*(2), 233–248. https://doi.org/10.1080/17521882.2025.2570687

Esiyok, E., Gokcearslan, S., & Kucukergin, K. G. (2025). Acceptance of educational use of AI chatbots in the context of self-directed learning with technology and ICT self-efficacy of undergraduate students. *International Journal of Human–Computer Interaction, 41*(1), 641–650. https://doi.org/10.1080/10447318.2024.2303557

Gonida, E. N., Karabenick, S. A., Stamovlasis, D., Metallidou, P., & Greece, C. T. Y. (2018). Help seeking as a self-regulated learning strategy and achievement goals: The case of academically talented adolescents. *High Ability Studies, 30*(1–2), 147–166. https://doi.org/10.1080/13598139.2018.1535244

Hair, J. F., Jr., Matthews, L. M., Matthews, R. L., & Sarstedt, M. (2017). PLS-SEM or CB-SEM: Updated guidelines on which method to use. *International Journal of Multivariate Data Analysis, 1*(2), 107–123. https://doi.org/10.1504/IJMDA.2017.087624

Hanafi, W. N. W., & Toolib, S. N. (2020). Influences of perceived usefulness, perceived ease of use, and perceived security on intention to use digital payment: A comparative study among Malaysian younger and older adults. *International Journal of Business Management, 3*(1), 15–24.

Harianto, A., Buwani, Faradina, E., & Tuwoso. (2025). Factors affecting work readiness of vocational school graduates: A systematic literature review. *International Journal of Studies in International Education, 2*(2), 90–106. https://doi.org/10.62951/ijsie.v2i2.281

Irfan, D., Watrianthos, R., & Yunus, F. A. N. B. (2025). AI in education: A decade of global research trends and future directions. *International Journal of Modern Education and Computer Science, 17*(2), 135–153. https://doi.org/10.5815/ijmecs.2025.02.07

Jeilani, A., & Abubakar, S. (2025). Perceived institutional support and its effects on student perceptions of AI learning in higher education: The role of mediating perceived learning outcomes and moderating technology self-efficacy. *Frontiers in Education, 10*, Article 1548900. https://doi.org/10.3389/feduc.2025.1548900

Jeong, S., Kim, S., & Lee, S. (2024). Effects of perceived ease of use and perceived usefulness of technology acceptance model on intention to continue using generative AI: Focusing on the mediating effect of satisfaction and moderating effect of innovation resistance. In *Conceptual modeling* (pp. 99–106). Springer. https://doi.org/10.1007/978-3-031-75599-6_7

Jia, W., & Huang, X. (2023). Digital literacy and vocational education: Essential skills for the modern workforce. *International Journal of Academic Research in Business and Social Sciences, 13*(5). https://doi.org/10.6007/ijarbss/v13-i5/17080

Kholifah, N., Kurdi, M. S., Nurtanto, M., Mutohhari, F., Fawaid, M., & Subramaniam, T. S. (2023). The role of teacher self-efficacy on the instructional quality in 21st century: A study on vocational teachers, Indonesia. *International Journal of Evaluation and Research in Education, 12*(2), 998–1006. https://doi.org/10.11591/ijere.v12i2.23949

Kock, N., & Dow, K. E. (2025). *Statistical significance and effect size tests in SEM: Common method bias and strong theorizing*.

Le, S. K. (2022). 21st-century competences and learning that technical and vocational training must develop: Focus 4C skills. *Journal of Engineering Researcher and Lecturer, 1*(1), 1–6.

Lee, H.-P., Sarkar, A., Tankelevitch, L., Drosos, I., Rintel, S., Banks, R., & Wilson, N. (2025). The impact of generative AI on critical thinking: Self-reported reductions in cognitive effort and confidence effects from a survey of knowledge workers. *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, 1–22. https://doi.org/10.1145/3706598.3713778

Martínez-Plumed, F., Gómez, E., & Hernández-Orallo, J. (2021). Futures of artificial intelligence through technology readiness levels. *Telematics and Informatics, 58*, Article 101525. https://doi.org/10.1016/j.tele.2020.101525

Mohammad, F., Ladin, C. A., & Shaharom, M. S. N. (2026). A recent systematic review of technological advancements in art education research. *Journal of Advanced Research in Applied Sciences and Engineering Technology, 58*(1), 145–174. https://doi.org/10.37934/araset.58.1.145174

Ninghardjanti, P., Umam, M. C., Subarno, Winarno, Langgi, N. R., & Widodo, J. (2025). Evaluating the impact of AI on the critical thinking skills among the higher education students by combining the TAM model and critical thinking theory. *Frontiers in Education, 10*, Article 1719625. https://doi.org/10.3389/feduc.2025.1719625

Parma Dewi, I., Aditya Fiandra, Y., Fadillah, R., Marta, R., Rosalina, L., Azima Noordin, N., & Khan, D. (2025). Explaining VR/AR learning in medical education: A comparative PLS-SEM analysis of TAM, SDT, TTF, and flow theory. *Seminars in Medical Writing and Education, 4*, Article 799. https://doi.org/10.56294/mw2025799

Rahmiati, Jalinus, N., Effendi, H., Fadillah, R., & Wulansari, R. E. (2024). Assessing the impact of a STEM learning project model on artificial intelligence education in higher learning institutions. *Data and Metadata, 3*. https://doi.org/10.56294/dm2024.623

Riyanda, A. R., Parma Dewi, I., Jalinus, N., Ahyanuardi, Sagala, M. K., Rinaldi, D., Prasetya, R. A., & Yanti, F. (2025). Digital skills and technology integration challenges in vocational high school teacher learning. *Data and Metadata, 4*. https://doi.org/10.56294/dm2025553

Saftari, R., Mulyadi, H., & Supardi, E. (2025). Tracing the role of artificial intelligence in self-regulated learning: A systematic review using the Winne and Hadwin framework (2007–2025). *JP (Jurnal Pendidikan): Teori dan Praktik, 10*(1), 78–92. https://doi.org/10.26740/jp.v10n1.p78-92

Singh, S. V., & Hiran, K. K. (2022). The impact of AI on teaching and learning in higher education technology. *Journal of Higher Education Theory and Practice, 22*(13). https://doi.org/10.33423/jhetp.v22i13.5514

Siregar, Y. S., Daryanto, E., & Siman, S. (2023). Increasing the competence of vocational education teachers with 4C skills-based training management: Critical thinking, creativity, communication, collaboration. https://doi.org/10.4108/eai.19-9-2023.2340493

Sukmawati, F., Prihatin, R., & Santosa, E. B. (2024). Design and evaluation a mobile augmented reality to enhance critical thinking skills for vocational high schools. *Salud, Ciencia y Tecnología, 4*. https://doi.org/10.56294/saludcyt2024.1000

Sutanto, J. E., Sukardi, D., & Christiani, N. (2021). Competency based training entrepreneurship to improve student's entrepreneur mentality: Case study in East Java, for vocational high school graduates. *International Journal of Economics, Business and Management Research, 5*(10), 28–36.

Tan, P. S. H., Seow, A. N., Choong, Y. O., Tan, C. H., Lam, S. Y., & Choong, C. K. (2024). University students' perceived service quality and attitude towards hybrid learning: Ease of use and usefulness as mediators. *Journal of Applied Research in Higher Education, 16*(5), 1500–1514. https://doi.org/10.1108/JARHE-03-2023-0113

Waliyuddin, D. S., & Sulisworo, D. (2022). High order thinking skills and digital literacy skills instrument test. *Ideguru: Jurnal Karya Ilmiah Guru, 7*(1). https://doi.org/10.51169/ideguru.v7i1.310

Wang, K., Cui, W., & Yuan, X. (2025). Artificial intelligence in higher education: The impact of need satisfaction on artificial intelligence literacy mediated by self-regulated learning strategies. *Behavioral Sciences, 15*(2), Article 165. https://doi.org/10.3390/bs15020165

Xu, J. (2026). How generative AI enhances self-regulated learning in EFL learners: A chain mediation model of "intention to use" and "learning engagement." *Frontiers in Psychology, 17*, Article 1808183. https://doi.org/10.3389/fpsyg.2026.1808183

Yuniarti, N., Rahmawati, Y., Anwar, M., Al Hakim, V. G., Hidayat, H., Hariyanto, D., Husna, A. F., & Wang, J.-H. (2024). Augmented reality-based higher order thinking skills learning media: Enhancing learning performance through self-regulated learning, digital literacy, and critical thinking skills in vocational teacher education. *European Journal of Education, 59*(4). https://doi.org/10.1111/ejed.12725

Zhou, X., Teng, D., & Al-Samarraie, H. (2024). The mediating role of generative AI self-regulation on students' critical thinking and problem-solving. *Education Sciences, 14*(12), Article 1302. https://doi.org/10.3390/educsci14121302

Zhou, X., Zhang, J., & Chan, C. (2024). Unveiling students' experiences and perceptions of artificial intelligence usage in higher education. *Journal of University Teaching and Learning Practice, 21*(6). https://doi.org/10.53761/xzjprb23