

Generative AI Use, Self-Regulated Learning, and Creativity Among Indonesian Design Students

Yulyanto^{1,*}, Jerry Dounald Raharjaan¹, Sigit Setya Kusuma¹, Ilhamsyah², Yayat Sudaryat², & Munya Saleh Ba Matraf³

¹Visual Communication Design, Universitas Kuningan, Indonesia

²School of Creative Industries, Universitas Telkom, Indonesia

³School of Computing, Universiti Utara Malaysia, Malaysia

*Corresponding email: yulyanto@uniku.ac.id

Received: 15 August 2026

Accepted: 21 September 2026

Published: 07 October 2026

Abstract: Generative artificial intelligence is increasingly embedded in design education, yet whether its use supports or substitutes for the self-directed creative process remains unclear, particularly outside the Chinese samples that dominate prior research. This study examined whether self-regulated learning (SRL) mediates the relationship between generative AI use and creativity among Indonesian design students. Data were collected through an online questionnaire from 120 Indonesian visual communication design students and analyzed using partial least squares structural equation modeling (PLS-SEM) with bootstrapped confidence intervals and four pre-specified sensitivity analyses. SRL was strongly and consistently associated with creativity ($\beta = .674$, $f^2 = .962$), and this association remained stable across all robustness checks. AI usage's relationships with SRL, with creativity directly, and with creativity indirectly through SRL were significant in the full sample but lost significance under two independent sensitivity analyses. However, the sample was adequately powered to detect these effect sizes (post-hoc power $\geq .91$); treat the point estimates as preliminary and confirm them in a larger, independent sample before treating them as established. A supplementary analysis found that frequency of AI use, unlike a richer measure of how AI was used, was unrelated to SRL or creativity. SRL is a robust correlate of creativity in AI-assisted design education, while AI use's own contribution requires replication; design pedagogy may be better served by cultivating self-regulation directly than by monitoring AI usage frequency alone.

Keywords: generative AI, self-regulated learning, creativity, design education, PLS-SEM.

Article's DOI: <https://doi.org/10.23960/jpp.v16i4.pp1975-1992>

■ INTRODUCTION

Generative artificial intelligence has moved, in a few years, from a novelty to a near-default feature of how students produce creative work. A 2026 national survey of UK higher education found that 95% of students now report using generative AI in some form, with 94% using it specifically for assessed coursework (Higher Education Policy Institute, 2026); in the creative industries into which today's design students will graduate, 86% of creators report using AI tools as part of everyday production workflows, and generative image tools are already embedded in

72% of graphic design workflows specifically. For visual communication design (DKV) education, this raises a pressing question: does routine reliance on a tool that can generate finished visual output on demand support independent, creative designers, or does it substitute for the exploration and revision design education is meant to cultivate? A growing empirical literature has begun to test this question, but with a consistent blind spot that motivates the present study.

Empirical work on generative AI and student creativity has expanded rapidly since 2023. A clear pattern has emerged: across

engineering (Zhou et al., 2025), sports science (Gu, 2026), music (Zhuang & Li, 2026), language learning (Xu, 2026), and product design ideation specifically (Lin & Xie, 2026), AI usage is consistently associated with higher creativity or creativity-adjacent outcomes, typically through some psychological mechanism rather than as a raw, unmediated effect. The specific mechanism tested has varied by study: critical thinking and AI self-efficacy in chain-mediation models (Zhou et al., 2025), dependence and flow (Gu, 2026), and self-regulated learning alongside creative self-efficacy in an experimental intervention (Zhuang & Li, 2026); the underlying theoretical logic, however, is shared: AI's effect on creativity is not direct so much as it operates by changing how students think about and manage their own creative process. This logic draws most directly on Zimmerman's (2020) model of self-regulated learning (SRL), in which the capacity to set goals, monitor progress, and evaluate outcomes is treated as a general-purpose cognitive-motivational skill that should, in principle, extend to creative work as readily as to conventional academic tasks, and on componential models of creativity (Guilford, 1967; Torrance, 1974) that decompose creative performance into fluency, flexibility, originality, and elaboration, dimensions that self-regulatory oversight plausibly shapes. Cross-disciplinary survey evidence corroborates this pattern outside experimental settings as well: university students' self-efficacy in AI use predicts how much perceived creative and critical-thinking benefit they report from generative AI (Alghamdi, 2025).

Within design education specifically, the evidence is thinner and more mixed. Hwang and Wu's (2025) study of Chinese design students, the closest population match to the present study, found that AI self-efficacy and anxiety reduction, rather than SRL, mediated the relationship between generative AI use and creative cognition, leaving open whether SRL operates similarly in

a design-studio context or whether the self-efficacy/anxiety pathway documented there is the more relevant mechanism for this population. Qualitative work from Indonesia (Rahiem, 2026; Wiradarmo, 2026) has documented how design students describe their AI use as a conceptual exploration tool constrained by prompting difficulty and concerns about cultural relevance, but has not tested whether these descriptive patterns correspond to measurable differences in self-regulation or creative output. A recurring tension across this literature is that studies exploring how AI use relates to self-regulatory processes and studies formally testing SRL as a structural mediator of the AI-creativity link have rarely been the same studies, and never, to our knowledge, in a design-education population outside China.

Three limitations in this literature motivate the present study. First, a population and context gap: every study testing a comparable structural mediation model used Chinese university students; none has tested whether these relationships hold among Indonesian design students, whose institutional and design-pedagogical context differs in non-cosmetic ways. Second, a mediator-specificity gap: self-efficacy and anxiety have been tested as mediators in a design-student sample (Hwang & Wu, 2025), and SRL has been tested as a mediator in non-design populations, but no study has tested SRL specifically as a mediator of AI use and creativity within design education, despite its theoretical centrality to the iterative, self-directed process design work demands. Third, a transparency gap: the studies reviewed above uniformly report a single significant path estimate without examining whether it is stable under alternative analytic specifications, even though most operate with sample sizes where a handful of cases can materially shift bootstrapped confidence interval.

Indonesia offers a particularly informative setting for closing these gaps. Design education

there operates within its own national accreditation structure and studio-teaching traditions. Its expanding but still uneven higher-education technology infrastructure (Helmiatin et al., 2024) means students' access to, and prior exposure to, generative AI tools plausibly differs from that of students in wealthier, more heavily AI-invested educational markets where nearly all comparable evidence originates. Testing whether an SRL-mediated pathway travels to this setting is therefore not a routine replication but a genuine test of the pathway's generality.

The present study addresses these gaps through a cross-sectional survey of 120 Indonesian visual communication design students, testing a structural model in which self-regulated learning mediates the relationship between generative AI use and creativity. Its contribution is threefold. Empirically, it provides the first test of this mediation pathway within design education

and the first in an Indonesian sample. Theoretically, it evaluates whether SRL functions as a mediator in a design-studio context, offering a point of comparison to the self-efficacy/anxiety pathway identified in the one comparable design-student study. Methodologically, it sets a standard of transparency uncommon in this literature by subjecting its findings to pre-specified robustness checks and explicitly reporting which effects hold stably and which do not.

Formally, the study tests four hypotheses that jointly specify this mediation model, depicted in Figure 1:

- H1. AI Usage is positively associated with SRL.
- H2. SRL is positively associated with Creativity.
- H3. AI Usage is positively associated with Creativity directly.
- H4. AI Usage is positively associated with Creativity indirectly, through SRL.

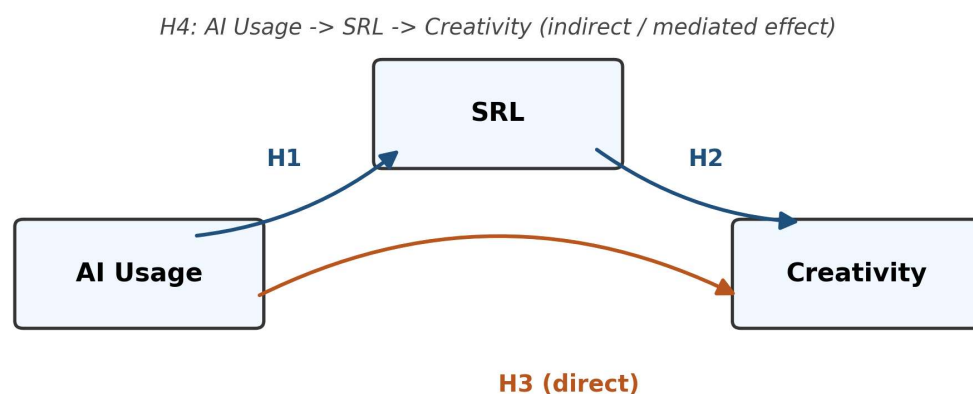


Figure 1. Hypothesized structural model

■ METHOD

Participants

Participants were undergraduate students enrolled in visual communication design programs in Indonesia, recruited through an open online survey link distributed across design-program networks. A total of 125 responses were received; data screening removed four duplicate submissions (matched by email) and one respondent missing more than 10% of core items

(a threshold set prior to analysis), yielding a final sample of $N = 120$. Respondents were drawn predominantly from a single institution (80.8%, $n = 97$), with the remainder from three other institutions (19.2%, $n = 23$). The sample was 58.3% male and 41.7% female, spanning semesters 1 through 9 (46.7% in semesters 5-6), with 69.2% reporting a GPA above 3.50. Regarding AI experience, 37.5% had used generative AI tools for over a year, and frequency

of use ranged from never (3.3%) to almost daily (3.3%), with occasional use (1-2 times/week) most common (44.2%).

Research Design and Procedures

Data were collected via an anonymous online questionnaire (Google Forms) over five days, following electronic informed consent; all retained responses affirmed consent before completing the substantive items. For the small number of respondents missing a single item within a given construct's block (fewer than 10% of that block, the same threshold used for the overall exclusion criterion described under Participants), the missing value was imputed using that respondent's own mean across the completed items of the same construct (within-construct person-mean imputation); this affected three respondents, one missing item each, within the SRL and Creativity blocks that enter the confirmatory structural model. One additional respondent had a single missing item in AI Self-Efficacy, a construct reported descriptively rather than entered into the confirmatory model, imputed instead via the sample column mean for that item. The study did not undergo review by a formal institutional research-ethics committee; given its anonymous, minimal-risk survey design with no collection of directly identifying information, data collection proceeded on electronic informed consent alone, and formal institutional ethical clearance was not obtained beforehand.

Instruments

The questionnaire included demographic items, AI usage history, and five Likert-scale instruments (1 = strongly disagree to 5 = strongly agree): AI Usage (10 items), Self-Regulated Learning (18 items), Creativity (16 items), AI Self-Efficacy (8 items), and AI-related Anxiety (8 items). All instruments were developed for this study rather than adapted wholesale from prior scales, as no validated Indonesian-language instrument existed for the visual-design context

specifically. Item development proceeded in three stages: a review of relevant theoretical and empirical literature to generate an initial item pool for each construct; expert judgment by academic staff with backgrounds in design education and psychometrics, who evaluated item relevance, clarity, and content validity and recommended revisions (expert ratings were used qualitatively to guide item revision rather than aggregated into a quantitative content-validity index such as Aiken's V; Aiken, 1980); and a pilot test with a small group of design students not included in the main sample, used to check item comprehensibility and approximate completion time before the instrument was finalized. Composite scores for each construct were computed as the mean of its constituent items.

Data Analysis

The hypothesized model specified AI Usage as a predictor of SRL and Creativity, with SRL mediating the AI Usage-Creativity relationship. AI Self-Efficacy and AI-related Anxiety were measured and are reported descriptively (Table 1) but were not entered as additional predictors or mediators in this confirmatory model: the pre-registered hypotheses target the SRL-specific mediator-specificity gap identified in the Introduction, since self-efficacy and anxiety had already been tested as mediators in a comparable design-student sample (Hwang & Wu, 2025), and adding them post hoc as further structural paths would convert a confirmatory test of a single pre-specified model into exploratory model-building; an integrated model testing all three candidate mediators jointly is proposed as a future direction (see Result and Discussion). Measurement adequacy was assessed via exploratory factor analysis (Kaiser-Meyer-Olkin sampling adequacy, Kaiser, 1974; Bartlett's test of sphericity; eigenvalue and factor-loading inspection) and internal consistency via Cronbach's alpha, prior to any hypothesis testing.

In response to peer review, outer-model evaluation was extended beyond this preliminary screening to include standardized outer loadings, average variance extracted (AVE), composite reliability (ρ_c , Joreskog's coefficient, and ρ_A , Dijkstra & Henseler, 2015), and discriminant validity via the heterotrait-monotrait ratio of correlations (HTMT, Henseler et al., 2015) for all five instruments, computed with the same custom Python implementation used for the structural model, as dedicated software such as SmartPLS or the *sempr* package in R was not available in the authors' computing environment; full results are reported in the Result and Discussion section. Structural assumptions were checked using standard diagnostics: variance inflation factors for multicollinearity, the Breusch-Pagan test for heteroscedasticity, Shapiro-Wilk tests for univariate normality of composite scores, and Harman's single-factor test to screen for common method bias (Podsakoff et al., 2003). The structural model was estimated using partial least squares structural equation modeling (PLS-SEM), chosen over covariance-based SEM because it does not require multivariate normality, performs well with moderate sample sizes, and suits models that emphasize predicting an outcome (creativity) from multiple predictors. Estimation used a path-weighting scheme with reflective (Mode A) outer estimation, implemented as a custom procedure in Python (NumPy/SciPy) following the estimation logic described by Hair et al. (2022), as dedicated PLS-SEM software was not available in the authors' computing environment; this is noted explicitly as a detail relevant to independent replication. Path coefficients, indirect (mediated) effects, and f^2 effect sizes were estimated, and statistical significance was assessed via bootstrapping (5,000 resamples, fixed random seed for reproducibility) using bias-corrected percentile 95% confidence intervals (Preacher & Hayes, 2008). Because the four hypotheses

tested a single a priori structural model rather than multiple independent comparisons, no correction for multiple testing was applied, consistent with the pre-registered analysis plan. To evaluate the robustness of the primary findings, four pre-specified sensitivity analyses were conducted: re-estimation after excluding respondents identified via item-level response-invariance screening as exhibiting careless-response patterns; re-estimation under an alternative missing-data To address whether the sample provides adequate statistical power to detect the observed structural effect sizes, post-hoc power was additionally computed for each structural path using the noncentral F distribution for the associated R^2 increase (Cohen, 1988), following the same logic implemented in G*Power (Faul et al., 2007), which was not available in the authors' computing environment; results are reported alongside the structural path estimates.

■ RESULT AND DISCUSSION

Of 125 questionnaire responses received, 120 were retained for analysis after removing duplicate submissions and one case with excessive missing data on the core constructs. Results are presented in three parts: measurement adequacy, tests of the four pre-registered hypotheses, and sensitivity analyses evaluating the robustness of the structural findings, followed by a discussion of their implications for theory and practice.

Sample Characteristics

Before turning to the structural results, several sample characteristics are worth noting because they bear on how the findings should be interpreted. Respondents were concentrated at a single institution (80.8%), were slightly more often male (58.3%) than female (41.7%), and were most commonly in their third year of study (semesters 5-6, 46.7%); most reported a strong academic record (69.2% with a GPA above

3.50). AI experience was heterogeneous: over a third of respondents (37.5%) had used generative AI tools for more than a year, while a small minority (3.3%) reported never having used them, giving the sample meaningful variation on the study's key predictor despite its institutional concentration.

Validity and Reliability of Measurement and Structural Model

Exploratory factor analysis supported treating each of the three core scales as a coherent single composite. Sampling adequacy was good to excellent for all three (Kaiser-Meyer-Olkin: AI Usage = .877, SRL = .879, Creativity = .910), and Bartlett's test of sphericity was significant in each case (all $p < .001$). Single-factor loadings were acceptable overall, with only one weak item (loading $< .40$) each for AI Usage and SRL, and none for Creativity. Kaiser's eigenvalue criterion nonetheless indicated some multidimensionality beneath the dominant factor, most pronounced for SRL, where four eigenvalues exceeded 1.0, and the first factor accounted for only 40.4% of item variance. This pattern aligns with SRL's theorized sub-domains (e.g., goal-setting, planning, monitoring, and self-evaluation) and Creativity's four hypothesized facets (fluency, flexibility, originality, elaboration). It provides direction for future higher-order measurement modeling rather than evidence against the composite-scoring approach adopted here, given

that loadings on the dominant factor remained acceptable for nearly every item. Internal consistency was high for all five scales: AI Usage ($\alpha = .851$), SRL ($\alpha = .909$), Creativity ($\alpha = .930$), AI Self-Efficacy ($\alpha = .909$), and AI-related Anxiety ($\alpha = .927$). These reliability estimates compare favorably with previously validated instruments used in related Indonesian higher-education samples, offering indirect support for the newly developed scales. However, formal cross-sample validation remains a task for future research.

Composite scores for AI Usage and SRL did not deviate significantly from normality (Shapiro-Wilk $W = .990$, $p = .493$, and $W = .988$, $p = .348$, respectively); Creativity showed a small but statistically significant deviation ($W = .972$, $p = .014$), not considered consequential given that bootstrapped PLS-SEM does not require multivariate normality. Multicollinearity was negligible (variance inflation factors = 1.10 for both predictors of Creativity). The Breusch-Pagan test indicated homoscedastic residuals for SRL ($LM = 2.29$, $p = .130$) but mild heteroscedasticity for Creativity ($LM = 6.97$, $p = .031$), which bootstrapped inference mitigated. Harman's single-factor test indicated that one unrotated factor accounted for 31.97% of variance across all 44 core items, below the 50% threshold conventionally used to flag common method bias. Table 1 presents descriptive statistics for the five study constructs.

Table 1. Descriptive statistics for study constructs (N = 120)

Construct	k items	M	SD	Skewness	Kurtosis	α
AI Usage	10	3.40	0.61	-0.09	0.32	.851
Self-Regulated Learning	18	3.84	0.50	0.06	-0.02	.909
Creativity	16	3.44	0.56	0.56	0.37	.930
AI Self-Efficacy	8	3.22	0.64	0.52	1.13	.909
AI-related Anxiety	8	3.81	0.80	-0.51	0.26	.927

Note. All items were scored on a 5-point Likert scale (1 = strongly disagree, 5 = strongly agree). AI Self-Efficacy and AI-related Anxiety were assessed descriptively and were not entered into the confirmatory structural model.

Table 2 presents the outer-model evaluation (standardized loading range, average variance extracted, and two composite-reliability coefficients) for all five instruments, extending the preliminary EFA/alpha screening above in response to peer review.

Table 2. Outer-Model evaluation: standardized loadings, AVE, and composite reliability (N = 120)

Construct	k items	Loading range	AVE	CR (ρ_c)	ρ_A
AI Usage	10	0.172 to 0.841	0.457	0.886	0.886
Self-Regulated Learning	18	0.324 to 0.749	0.404	0.922	0.918
Creativity	16	0.636 to 0.815	0.493	0.939	0.932
AI Self-Efficacy	8	0.737 to 0.843	0.616	0.928	0.912
AI-related Anxiety	8	0.750 to 0.867	0.667	0.941	0.929

Note: AVE = average variance extracted; CR = composite reliability (Jöreskog's $\hat{\rho}_c$); ρ_A = Dijkstra and Henseler's (2015) consistent reliability coefficient. AVE and CR/ $\hat{\rho}_A$ were computed from standardized outer loadings using the project's own custom PLS-SEM implementation (see Data analysis). One AI Usage item ("I use AI as a tool, not as a replacement for my own creative process") loaded weakly ($\lambda = 0.17$), consistent with the single weak item already flagged by the EFA above; AVE fell below the conventional .50 threshold for AI Usage (.457) and SRL (.404), mirroring the multidimensionality the EFA identified for SRL and motivating the second-order re-specification reported under Sensitivity and Robustness Analyses below. Creativity, AI Self-Efficacy, and AI-related Anxiety all met the .50 threshold; all five constructs met the conventional .70 threshold for CR and $\hat{\rho}_A$.

Table 3 presents the heterotrait-monotrait ratio of correlations (HTMT; Henseler et al., 2015) among all five constructs to evaluate discriminant validity.

Bivariate correlations were broadly consistent with the hypothesized structural pattern: SRL and Creativity were strongly correlated ($r = .73$), AI Usage correlated moderately with both

Table 3. Heterotrait-Monotrait ratio (HTMT) discriminant-validity matrix (N = 120)

	AI Usage	Self-Regulated Learning	Creativity	AI Self-Efficacy	AI-related Anxiety
AI Usage	N/A	0.385	0.444	0.523	0.172
Self-Regulated Learning	0.385	N/A	0.790	0.478	0.526
Creativity	0.444	0.790	N/A	0.421	0.445
AI Self-Efficacy	0.523	0.478	0.421	N/A	0.194
AI-related Anxiety	0.172	0.526	0.445	0.194	N/A

Note: All HTMT values fall below the conservative .85 threshold (Henseler et al., 2015), supporting discriminant validity among all five constructs; the highest value (SRL-Creativity, HTMT = .790) is consistent with, and expected given, the strong substantive association between these two constructs (H2).

SRL ($r = .29$) and Creativity ($r = .38$), and AI Usage correlated with AI Self-Efficacy ($r = .47$) but was essentially unrelated to AI-related Anxiety ($r = .07$). The near-zero AI Usage-Anxiety correlation is notable against the two deferred constructs' otherwise sensible pattern: AI-related

Anxiety correlated positively with both SRL ($r = .48$) and Creativity ($r = .41$), suggesting anxiety in this sample may reflect a reflective, quality-conscious stance toward one's own creative process rather than simple technology aversion, though these two constructs were not entered into

the confirmatory structural model and this interpretation remains exploratory.

Figure 2 displays the full correlation matrix.

	AI Usage	Self-Regulated Learning	Creativity	AI Self-Efficacy	AI-related Anxiety	AI Use Frequency (ordinal)
AI Usage	1.00	0.29*	0.38*	0.47*	0.07	0.44*
Self-Regulated Learning	0.29*	1.00	0.73*	0.43*	0.48*	-0.05
Creativity	0.38*	0.73*	1.00	0.38*	0.41*	0.02
AI Self-Efficacy	0.47*	0.43*	0.38*	1.00	0.14	0.10
AI-related Anxiety	0.07	0.48*	0.41*	0.14	1.00	-0.07
AI Use Frequency (ordinal)	0.44*	-0.05	0.02	0.10	-0.07	1.00

Figure 2. Correlation matrix among AI usage, SRL, creativity, AI self-efficacy, and AI-related anxiety. Note: Pearson r ; * $p < .05$. Cell shading follows a red (positive) to blue (negative) scale proportional to $|r|$.

A supplementary comparison using participants' self-reported frequency of AI use, rather than the multi-item AI Usage composite, showed a different pattern: frequency correlated with the AI Usage composite ($r = .44$) but not with SRL ($r = -.05$) or Creativity ($r = .02$), and an independent-samples t test comparing high- and low-frequency users found a significant

difference on the AI Usage composite ($t = 4.89$, $p < .001$) but not on SRL ($t = -0.84$, $p = .406$) or Creativity ($t = -0.15$, $p = .885$), as shown in Figure 3. This divergence suggests that how students use AI, rather than simply how often, is the more relevant correlate of self-regulation and creative outcomes, discussed further below.

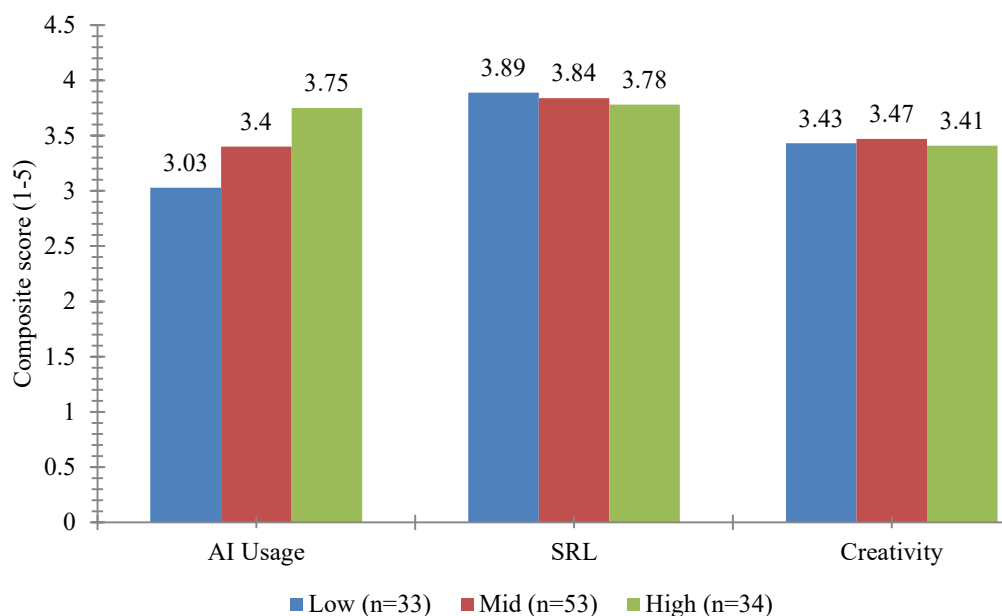


Figure 3. AI usage, SRL, and creativity by frequency-of-use group (low/mid/high). Note: Bars show group means; Table 5 below reports the full five-number summary (min/Q1/median/Q3/max) and n per group and construct, capturing the distributional spread and outlier information a boxplot would otherwise convey

Table 5. Five-Number summary of AI usage, SRL, and creativity by frequency-of-use group (N = 120)

Construct	Group	n	Min	Q1	Median	Q3	Max
AI Usage	Low	33	1.70	2.60	3.00	3.30	5.00
	Mid	53	2.10	3.10	3.40	3.70	4.60
	High	34	2.80	3.60	3.75	4.00	5.00
Self-Regulated Learning	Low	33	2.83	3.61	3.83	4.22	4.89
	Mid	53	2.83	3.44	3.78	4.28	4.94
	High	34	2.33	3.47	3.78	4.06	5.00
Creativity	Low	33	2.13	3.00	3.25	3.81	4.94
	Mid	53	2.50	3.25	3.44	3.62	4.81
	High	34	2.38	3.08	3.34	3.69	5.00

Hypothesis Testing: Structural Model Results

The structural model was estimated using PLS-SEM with bootstrapped bias-corrected 95% confidence intervals (5,000 resamples). Table 4 presents the path estimates; all four pre-registered hypotheses were supported at

conventional significance levels in the primary sample. Because the hypotheses jointly specify a single mediation model rather than four independent comparisons, no correction for multiple testing was applied, consistent with the pre-registered analysis plan described in the Method section.

Table 4. Structural path estimates for the primary model (N = 120)

Hypothesis	Path	β	95% CI (BC bootstrap)	p	f^2
H1	AI Usage $\rightarrow$ SRL	.345	[.021, .505]	.014	.135
H2	SRL $\rightarrow$ Creativity	.674	[.560, .757]	< .001	.962
H3	AI Usage $\rightarrow$ Creativity (direct)	.210	[.016, .335]	.017	.094
H4	AI Usage $\rightarrow$ SRL $\rightarrow$ Creativity (indirect)	.233	[.018, .357]	.014	N/A

Note. CI = confidence interval; BC = bias-corrected. $R^2 = .119$ for SRL, $.596$ for Creativity

H1. AI Usage was positively associated with SRL, supporting H1 ($\beta = .345$, 95% CI [.021, .505], $p = .014$, $f^2 = .135$, post-hoc power = .979). Post-hoc power indicates the primary sample was more than adequately powered to detect an effect of this size; nonetheless, as detailed under Sensitivity and Robustness Analyses below, this path's confidence interval widened to include zero once two careless-response cases were excluded, indicating the effect is not robust to a small number of specific cases even though the design itself was well powered to detect it.

H2. SRL was positively associated with Creativity, supporting H2 and representing the strongest effect in the model ($\beta = .674$, 95% CI [.560, .757], $p < .001$, $f^2 = .962$, post-hoc power > .999). This path was the only one to remain fully robust across every sensitivity check (see below), and its very large effect size indicates SRL alone accounts for most of the explained variance in Creativity ($R^2 = .596$ for the full model).

H3. AI Usage retained a significant direct effect on Creativity after accounting for SRL, supporting H3 ($\beta = .210$, 95% CI [.016,

.335], $p = .017$, $f^2 = .094$, post-hoc power = .913). Although $f^2 = .094$ is small by conventional benchmarks, post-hoc power analysis indicates the sample was adequately powered (power = .91; the minimum effect size detectable at 80% power for this test is $f^2 = .067$, smaller than the effect observed) to detect an effect of this magnitude. Hence, the instability of this path under sensitivity analysis (see below) reflects genuine case-level sensitivity rather than an underpowered design.

H4. The indirect effect of AI Usage on Creativity through SRL was significant, supporting H4 ($\beta = .233$, 95% CI [.018, .357], $p = .014$). Because the direct effect (H3) remained significant alongside this indirect effect, this indicates partial rather than full mediation (Baron & Kenny, 1986).

Theoretically, the sheer disparity between SRL's effect size and AI Usage's is informative in its own right. SRL is mechanistically proximate to Creativity: the planning, monitoring, and evaluative processes it comprises are largely the same processes through which a creative idea is generated, tested, and revised, so a measure of those processes should track creative output closely. AI Usage, by contrast, is a tool-use behavior several causal steps removed from the outcome; it can shape Creativity only by first changing how a student thinks about and manages their own work, which is precisely what SRL is theorized to capture. A mediator this mechanistically close to its outcome would be expected to absorb a large share of explained variance almost independent of how strongly the more distal predictor (AI Usage) is measured, which is consistent with the pattern observed here: a small, moderately reliable AI Usage-SRL link (H1) coexisting with a very large SRL-Creativity link (H2).

Sensitivity and Robustness Analyses

Because a small number of data points can disproportionately influence bootstrapped

estimates in a sample of this size, we evaluated the stability of the four path estimates across independent perturbations of the analysis (Liu et al., 2025).

First, item-level response-invariance screening identified two respondents who gave an identical response to all 60 Likert items spanning constructs of opposite conceptual valence, a pattern consistent with careless or inattentive responding (Ward & Meade, 2023). Re-estimating the model excluding these two cases ($n = 118$) reproduced the same pattern of effects in direction. However, the bias-corrected confidence intervals for H1, H3, and H4 widened to include zero (AI Usage $\rightarrow$ SRL: $\beta = .311$, 95% CI [“.071, .483]; AI Usage $\rightarrow$ Creativity: $\beta = .206$, 95% CI [-.051, .324]; indirect effect: $\beta = .208$, 95% CI [-.051, .335]), while H2 remained fully robust ($\beta = .671$, 95% CI [.547, .755]). A simple (non-bias-corrected) percentile bootstrap CI and the raw bootstrap-based p-value both remained below .05 for all paths in this reduced sample, indicating that the discrepancy reflects the conservatism of the bias-corrected method under a smaller, less symmetric bootstrap distribution rather than a qualitative reversal of the findings.

Second, we re-estimated the model under an alternative missing-data specification, using listwise deletion of the three cases whose single missing core item had originally been imputed with a within-construct person mean ($n = 117$; Amusa & Hossana, 2024). This independent perturbation reproduced the identical pattern: H2 remained robust ($\beta = .671$, 95% CI [.570, .761]), while the confidence intervals for H1 and the indirect effect again included zero, with H3 remaining just significant ($\beta = .215$, 95% CI [.015, .335]).

Third, Mahalanobis distance screening on the three composite scores identified no multivariate outliers at $p < .001$ (critical $\chi^2 = 16.27$, $df = 3$), including the two careless-response cases identified above; those cases were evidently not extreme in average response level,

only in item-to-item invariability. This distinction matters methodologically: a screen based on multivariate distance alone would not have flagged these two cases, underscoring the value of combining several complementary diagnostic approaches rather than relying on any single outlier-detection method.

Fourth, a five-fold stability check, re-estimating the structural model on five independent 96-case subsamples, found that all three direct path coefficients retained a consistent positive sign in every fold (AI Usage $\rightarrow$ SRL ranged from .294 to .383; AI Usage $\rightarrow$ Creativity from .175 to .263; SRL $\rightarrow$ Creativity from .643 to .700), with SRL $\rightarrow$ Creativity showing markedly lower variability than the other two paths across folds.

Fifth, in response to the exploratory factor analysis finding reported below that SRL's first unrotated factor accounted for only 40.4% of item variance (four eigenvalues > 1.0), we re-specified SRL as a second-order (higher-order) reflective construct rather than a flat unidimensional composite, following the two-stage approach described by Sarstedt et al. (2019). The 18 SRL items were regrouped into three first-order sub-dimensions corresponding to Zimmerman's (2020) forethought, performance, and self-reflection phases, consistent with recent evidence that these three phases are the natural grain at which generative-AI-supported self-regulation is modeled (Xia et al., 2026; Xu et al., 2025); each phase's item-level latent-variable score was then used as a reflective indicator of the second-order SRL construct in an otherwise identical structural model. This re-specification reproduced the primary results essentially unchanged: AI Usage $\rightarrow$ SRL, $\beta = .355$, 95% CI [.046, .506]; SRL $\rightarrow$ Creativity, $\beta = .662$, 95% CI [.552, .764]; AI Usage $\rightarrow$ Creativity (direct), $\beta = .209$, 95% CI [.008, .332]; indirect effect, $\beta = .235$, 95% CI [.049, .355]; and, under the same two-case exclusion used in the first sensitivity check, the

same pattern of H2 remaining fully robust ($\hat{\alpha} = .657$, 95% CI [.538, .760]) while H1, H3, and the indirect effect again widened to include zero. The substantive conclusions of this study therefore do not depend on treating SRL as unidimensional.

Taken together, the five sensitivity analyses (four pre-registered, plus the post-review second-order re-specification above) converge on a consistent picture: the SRL-Creativity relationship (H2) is robust across every specification tested, whereas the relationships involving AI Usage as a predictor (H1, H3) and the mediated pathway (H4) are directionally consistent, always positive and never reversing sign, but statistically marginal, in the sense that their confidence intervals depend on the inclusion of a small number of specific cases rather than reflecting an effect so large it would survive almost any reasonable data-cleaning decision.

The clearest result is that SRL is strongly associated with creativity. This association held under every robustness check applied and accounted for a substantial share of variance in creative outcomes. Evidence bearing more directly on the role of AI itself was directionally consistent with our hypotheses but did not hold up as reliably once a small number of atypical cases were removed; we treat these latter results as preliminary rather than settled, and return throughout this discussion to the distinction between what the data support with confidence and what they merely suggest.

The strength of the SRL-creativity association ($\beta = .674$, $f^2 = .962$) is broadly consistent with theoretical accounts treating self-regulation as constitutive of, rather than merely conducive to, creative performance: the planning, monitoring, and evaluative processes SRL comprises are largely the same processes through which a creative idea is generated, tested, and refined. Empirically, this pattern echoes findings from Zhuang and Li's (2026) mixed experimental-correlational study of Chinese music students, in

which generative-AI-supported SRL was among the strongest predictors of perceived creative competence, and is consistent with earlier evidence that SRL processes, particularly self-evaluation and strategic monitoring, predict creative problem-solving independent of technology use. Our contribution is contextual and constructual: this relationship had not previously been documented in visual design education, nor in an Indonesian sample, nor with SRL modeled as a single composite alongside AI-specific predictors rather than as a general disposition. That SRL only partially, rather than fully, mediated the relationship between AI use and creativity is informative in itself.

A near-identical structural pattern, in which digital skills predicted a creative outcome both directly and through self-directed learning, was recently reported by Jabeen et al. (2025) in a substantially larger sample ($N = 400$) of Pakistani university students, where the mediating pathway was likewise partial. The convergence of two independent samples, in different countries and different creative or academic outcomes, on the same qualitative conclusion suggests this may be a stable feature of the underlying phenomenon rather than an artifact of either study's sample; a fully mediated model, in which SRL absorbed the entirety of AI usage's association with creativity, would arguably have been the less plausible finding, since it would imply that AI usage carries no association with creative outcomes beyond whatever self-regulatory processes it happens to activate. The partial-mediation pattern is more consistent with a view of generative AI as one of several tools and conditions that shape creative output, operating both through and independently of learners' self-regulatory capacity, rather than as a technology whose entire creative benefit is filtered through a single psychological gatekeeper.

The relationships in which AI usage predicted creativity were more fragile than the SRL-creativity link, with confidence intervals that

included zero under two independent robustness perturbations even though point estimates remained positive throughout. One plausible reading draws on a recent meta-analysis of AI-supported SRL interventions, which found AI to be reliably more effective during the performance phase of self-regulated learning than during the preparatory forethought phase (Xu et al., 2026). If AI usage in our sample operates similarly, shaping how students execute and evaluate design work more than how they plan before beginning it, a modest and situationally variable association with an SRL measure spanning both phases would be expected. This is a post hoc interpretation, offered as a hypothesis for future work rather than a conclusion. This preliminary status also contrasts with the one directly comparable study in our review: Hwang and Wu's (2025) design-student sample, where AI self-efficacy and anxiety reduction were robust, well-powered mediators. Our study targeted a conceptually related but distinct mediator (SRL) and found a qualitatively similar but statistically weaker pattern; whether this reflects a genuinely smaller effect, an underpowered test, or measurement limitations in a newly developed scale cannot be adjudicated with the present data. It is also worth emphasizing what the fragility finding does not show: at no point across any of the four sensitivity analyses did a path coefficient change sign or approach a materially different magnitude; what varied was solely whether the bias-corrected interval crossed zero, a threshold that is, by construction, sensitive to a handful of cases in a sample of this size. Readers should therefore weigh the practical difference between an effect that reverses direction under scrutiny, which would undermine the hypothesis outright, and an effect that remains positive but statistically marginal, which more plausibly reflects a genuinely modest effect size still awaiting a well-powered test.

An unanticipated but theoretically informative pattern emerged from the

supplementary analysis: students' self-reported frequency of AI use was essentially unrelated to both SRL and creativity, even though the multi-item AI Usage scale, which captures how AI was incorporated into the design process rather than how often it was opened, related to both. This dissociation is consistent with an emerging methodological argument that frequency counts are an increasingly inadequate proxy for engagement quality. It also has a substantive reading specific to this population: two students who both report opening a generative AI tool several times a week may be doing very different things with it, one treating it as a quick shortcut to a finished image, the other using it as one step in a longer cycle of ideation, critique, and revision, and only the multi-item measure, which asks about the manner of integration rather than the act of opening the application, appears sensitive to that difference. Practically, it implies that policies aimed at simply increasing or restricting how often students use AI tools may miss the more consequential question of how those tools are used, and that future data collection in this literature should resist treating a single frequency item as an adequate stand-in for a richer usage construct.

The Indonesian setting of this study is not incidental to its contribution. Nearly all prior structural tests of AI-creativity mediation identified in our review were conducted with Chinese university samples, in higher-education systems with different norms around instructor authority, assessment, and technology procurement; whether psychological mechanisms established in that context transfer directly to other national systems is an empirical question, not a safe assumption. Indonesian design education operates within its own accreditation structure and studio-teaching traditions, and Indonesian students' access to, and prior exposure to, generative AI tools may differ systematically from that of students in wealthier, more heavily AI-

invested educational markets. That the SRL-creativity link replicated in this setting is therefore a meaningful, non-trivial extension rather than a routine confirmation; that the AI-specific pathways were comparatively fragile here invites the further, open question of whether contextual factors specific to Indonesian design education, such as resource access, curricular structure, or cultural norms around technology-assisted work, systematically attenuate AI's incremental contribution relative to what has been observed elsewhere, a question this cross-sectional single-country study cannot itself resolve but is well positioned to motivate.

This study extends self-regulated learning theory into a domain, AI-assisted visual design, where SRL's relationship to creative performance had not been empirically established, using a mediator conceptually distinct from the self-efficacy and anxiety constructs that have dominated recent AI-creativity research in adjacent design contexts. Because Hwang and Wu (2025) found self-efficacy and anxiety reduction to be the operative mechanisms in a similarly composed design-student sample, while the present study found SRL to be theoretically plausible but empirically more fragile, a natural next step for theory-building is to specify how these mechanisms relate to one another rather than treating them as competing explanations. Future work can test whether SRL, self-efficacy, and anxiety operate as parallel, competing, or sequential mediators within a single integrated model; for instance, whether SRL processes are themselves a downstream consequence of AI self-efficacy, which would imply a chain mediation pathway rather than three independent explanations of the same outcome. Methodologically, the divergence between our frequency-based and composite-based AI usage measures contributes an empirical caution to a field that has often relied on single-item frequency measures as a proxy for technology engagement,

and suggests that instrument validation studies in this space should report the two types of measures separately rather than assuming convergence.

For design educators, the most actionable implication follows from the one robust finding: because SRL is strongly associated with creative performance regardless of the uncertainty surrounding AI's specific role, pedagogical investment in students' planning, self-monitoring, and self-evaluation skills is likely to support creative development whether or not it is explicitly framed around AI. For institutions developing AI-use policies, the frequency-quality dissociation suggests that guidelines framed purely around limiting or mandating usage frequency are unlikely to be the most effective lever; guidance oriented toward how students integrate AI into their workflow is more consistent with what the data relate to creative outcomes. In practice, this could translate into studio assignments that explicitly require students to document their planning, revision, and self-evaluation decisions when generative AI is part of the workflow, rather than assessing only the finished visual output; for example, requiring a short process log alongside a final submission, in which students note which stages of the work involved AI assistance and what they changed or rejected from the AI-generated material and why. Two boundary conditions apply to these recommendations: the sample was drawn overwhelmingly from a single institution (80.8% of respondents), and the preliminary status of the AI-specific pathways means recommendations tied to AI's incremental contribution specifically should be treated as provisional pending replication, even as the recommendation to invest in self-regulation itself rests on firmer ground.

Several limitations bear directly on how these findings should be read. First, respondents were drawn overwhelmingly from one institution, limiting confidence that the observed relationships,

particularly their magnitude, generalize to Indonesian design students as a whole. Second, the analytic sample ($N = 120$), while adequate under conventional benchmarks for the structural complexity of the model tested, fell short of a more conservative sample-size heuristic tied to the largest measurement block (Kock & Hadaya, 2018); this is almost certainly implicated in the fragility of the AI-usage pathways documented above. Third, all three core instruments were newly developed rather than adapted from previously validated scales, and their content validity rested on qualitative expert judgment rather than a quantitative index such as Aiken's V (Aiken, 1980); the absence of both a quantitative content-validity index and confirmatory validation in an independent sample means measurement error remains a plausible partial explanation for the marginal pathways; consistent with this, the outer-model evaluation reported above found AVE below the conventional .50 threshold for AI Usage and SRL, driven substantially by one weakly loading AI Usage item.

Relatedly, post-hoc power analysis indicates the fragility of the AI-usage pathways under sensitivity analysis is not attributable to inadequate statistical power (power e'' .91 for both H1 and H3 at the observed effect sizes); readers should therefore interpret this fragility as genuine case-level sensitivity in a sample of this size, not as a simple power problem that a larger N alone would necessarily resolve, although replication in a larger, multi-institutional sample remains the most direct way to establish whether the observed point estimates themselves hold. Fourth, the cross-sectional design cannot establish causal direction, and at least one recent study has reported evidence consistent with a reversed pathway, in which self-regulation shapes AI reliance rather than the reverse, that our data cannot rule out. Finally, the qualitative phase originally planned to accompany this quantitative

model, intended to illuminate the mechanisms behind exactly the pathways that proved most fragile here, had not been conducted at the time of writing.

Four directions follow most directly from these results. First, replicating the structural model in a larger, multi-institution sample would clarify whether the AI-usage pathways are genuine small effects or artifacts of the present sample's size and institutional composition. Second, the qualitative interviews and design-artifact analysis originally planned for this project remain, in our view, the most direct way to explain the mechanisms behind the fragile pathways, and we intend to pursue this as a distinct, appropriately resourced study once fieldwork is complete. Third, future instrument development in this space would benefit from explicitly separating SRL's sub-processes rather than treating SRL as a unidimensional composite, and from distinguishing frequency, breadth, and manner of AI use as conceptually and empirically separate constructs. Fourth, an integrated model that tests SRL alongside AI self-efficacy and AI-related anxiety within a single structural framework would help clarify which psychological mechanism, or combination of mechanisms, is most consequential for creative outcomes in design education.

■ CONCLUSION

This study examined whether self-regulated learning helps explain the relationship between generative AI use and creativity among Indonesian design students, a pathway proposed on theoretical grounds but never previously tested in this population or field of study. Among 120 visual communication design students, the capacity to plan, monitor, and evaluate one's own creative process proved to be the single strongest correlate of creative performance identified in this study, a relationship robust enough to survive every alternative specification tested, even as the role of generative AI use in shaping that capacity

remained, on the present evidence, an open question rather than a settled one: the AI-usage pathways were adequately powered to detect the effect sizes observed, yet unstable under sensitivity analysis, so they are best read as preliminary estimates awaiting confirmation in a larger, independent sample rather than as settled findings. Theoretically, this extends self-regulated learning's established relevance to creative outcomes into AI-assisted visual design and Indonesian design students, a population where the relationship had not previously been documented. Methodologically, the value of this study lies less in any individual coefficient than in the practice of reporting them alongside deliberate stress tests: by showing which findings held and which did not across independent perturbations of the sample, this study aims to model a standard of transparency in which "significant" and "robust" are treated as distinct claims. For design educators, the practical implication is to invest confidently in cultivating students' self-regulatory skills, since their link to creative outcomes does not depend on resolving the more uncertain questions surrounding AI's specific contribution. As generative AI becomes further embedded in creative education, the more durable pedagogical lever may be the human capacity for self-directed reflection within which AI is used, rather than the tool itself, a claim offered as a starting hypothesis for the larger, multi-institutional, and qualitatively enriched research this topic now warrants.

■ DECLARATION OF GENERATIVE AI USAGE IN THE WRITING PROCESS

During the preparation of this manuscript, the author(s) used Claude (Anthropic) for data cleaning and analysis code development, including implementing the custom PLS-SEM procedure described in the Analytic Strategy and the second-order robustness check reported above, as well as drafting and revising manuscript text

and formatting. After using this tool, the author(s) reviewed, verified, and revised all content, including the statistical results and their interpretation, and accept full responsibility for the final content of the article.

■ REFERENCES

- Aiken, L. R. (1980). Content validity and reliability of single items or questionnaires. *Educational and Psychological Measurement*, 40(4), 955–959. <https://doi.org/10.1177/001316448004000419>
- Alghamdi, A. A. (2025). University students' perceptions of generative AI for critical thinking and creativity: The influence of self-efficacy and disciplinary differences. *Innovations in Education and Teaching International*, 1–15.
- Amusa, L. B., & Hossana, T. (2024). An empirical comparison of some missing data treatments in PLS-SEM. *Plos one*, 19(1), e0297037.
- Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51(6), 1173–1182. <https://doi.org/10.1037/0022-3514.51.6.1173>
- Chauliac, M., Willems, J., Gijbels, D., & Donche, V. (2023). The prevalence of careless response behaviour and its consequences on data quality in self-report questionnaires on student learning. *Frontiers in Education*, 8. <https://doi.org/10.3389/feduc.2023.1197324>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences (2nd ed.)*. Lawrence Erlbaum Associates.
- Dijkstra, T. K., & Henseler, J. (2015). Consistent partial least squares path modeling. *MIS Quarterly*, 39(2), 297–316. <https://doi.org/10.25300/MISQ/2015/39.2.02>
- Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2), 175–191. <https://doi.org/10.3758/BF03193146>
- Gu, W. (2026). Does generative AI dependence foster creativity? Evidence from sports undergraduates. *PLoS One*, 21(2), Article e0341277.
- Guilford, J. P. (1967). *The nature of human intelligence*. McGraw-Hill.
- Hair, J. F., Jr., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2022). *A primer on partial least squares structural equation modeling (PLS-SEM) (3rd ed.)*. Sage Publications.
- Helmiatin, Hidayat, A., & Kahar, M. R. (2024). Investigating the adoption of AI in higher education: A study of public universities in Indonesia. *Cogent Education*, 11(1), Article 2380175. <https://doi.org/10.1080/2331186X.2024.2380175>
- Henseler, J., Ringle, C. M., & Sarstedt, M. (2015). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Science*, 43(1), 115–135. <https://doi.org/10.1007/s11747-014-0403-8>
- Higher Education Policy Institute. (2026). *Student generative AI survey 2026 (HEPI Report 199)*. <https://www.hepi.ac.uk/reports/student-generative-ai-survey-2026/>
- Hwang, Y., & Wu, Y. (2025). The influence of generative artificial intelligence on creative cognition of design students: A chain mediation model of self-efficacy and anxiety. *Frontiers in Psychology*, 15, Article 1455015. <https://doi.org/10.3389/fpsyg.2024.1455015>

- Jabeen, F., Awan, R. N., & Anwar, M. N. (2025). Mediating role of self-directed learning in the relationship between digital skills and creative mindset among university students. *Pakistan Journal of Law, Analysis and Wisdom*, 4(8), 131–144.
- Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika*, 39(1), 31–36. <https://doi.org/10.1007/BF02291575>
- Kock, N., & Hadaya, P. (2018). Minimum sample size estimation in PLS-SEM: The inverse square root and gamma-exponential methods. *Information Systems Journal*, 28(1), 227–261. <https://doi.org/10.1111/isj.12131>
- Lin, H., & Xie, L. (2026). Stimulating or constraining creativity? Traditional vs. generative AI on divergent thinking in product design. *Frontiers in Psychology*, 17, Article 1839565. <https://doi.org/10.3389/fpsyg.2026.1839565>
- Liu, Y., Chin, W. W., Cheah, J.-H., Hair, J. F., & Lyu, C. (2025). Tackling missing data in PLS-SEM: Strategies and insights for business research. *Journal of Business Research*, 201, Article 115739.
- Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, 88(5), 879–903. <https://doi.org/10.1037/0021-9010.88.5.879>
- Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods*, 40(3), 879–891. <https://doi.org/10.3758/BRM.40.3.879>
- Rahiem, M. D. H. (2026). Generative AI in higher education in Indonesia: Patterns of use and learning impact. *Social Sciences & Humanities Open*, 13, Article 102672. <https://doi.org/10.1016/j.ssaho.2026.102672>
- Sarstedt, M., Hair, J. F., Cheah, J.-H., Becker, J.-M., & Ringle, C. M. (2019). How to specify, estimate, and validate higher-order constructs in PLS-SEM. *Australasian Marketing Journal*, 27(3), 197–211. <https://doi.org/10.1016/j.ausmj.2019.05.003>
- Torrance, E. P. (1974). *Torrance tests of creative thinking: Norms-technical manual*. Personnel Press.
- Ward, M. K., & Meade, A. W. (2023). Dealing with careless responding in survey data: Prevention, identification, and recommended best practices. *Annual Review of Psychology*, 74, 577–596.
- Wiradarmo, A. A. (2026). Designing with Generative AI: Nurturing or Numbing Student Creativity?. *Meraki: Journal of Creative Industries*, 3(02), 57–68.
- Xia, Q., Liu, Q., Tlili, A., & Chiu, T. K. F. (2026). A systematic literature review on designing self-regulated learning using generative artificial intelligence and its future research directions. *Computers & Education*, 240, Article 105465. <https://doi.org/10.1016/j.compedu.2025.105465>
- 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, 1808183.
- Xu, J., Luo, Y., & Wu, Y. (2026). AI support in self-regulated learning: A decade of technological evolution and meta-analysis. *British Journal of Educational Technology*. <https://doi.org/10.1111/bjet.70058>
- Xu, X., Qiao, L., Cheng, N., Liu, H., & Zhao, W. (2025). Enhancing self-regulated learning and learning experience in

generative AI environments: The critical role of metacognitive support. *British Journal of Educational Technology*, 56(5), 1842-1863. <https://doi.org/10.1111/bjet.13599>

Zhou, Z., Guo, H., Ma, F., Yang, C., & Gao, Y. (2025). The chain mediating role of critical thinking and AI self-efficacy in GenAI usage competence and engineering students' creativity. *Scientific Reports*, 15, Article 35945. <https://doi.org/10.1038/s41598-025-21132-0>

Zhuang, Z., & Li, X. (2025). The influence of collaborative music creation supported by generative artificial intelligence on students' creativity. *Frontiers in Psychology*, 16, 1709513.

Zimmerman, B. J. (2020). Attaining self-regulation: *A social cognitive perspective*. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 13–39). Academic Press.