International Journal of Learning, Teaching and Educational Research
Vol. 25, No. 8, pp. 289-312, August 2026
https://doi.org/10.26803/ijlter.25.8.11
Received May 31, 2026; Revised Jul 21, 2026; Accepted Jul 22, 2026

Authentic Project-Based Learning with Generative AI: Fostering Creative and Critical Thinking in Design Education

Lisa Indriati, Mai Neo* and Heidi Yeen-Ju Tan
Centre for Immersive and Innovative Technology,
COE for Immersive Experience, Faculty of Creative Multimedia,
Multimedia University, Cyberjaya, Malaysia

Abstract. The rapid integration of generative artificial intelligence (AI) into creative practices raises discussion about how design education can continue to develop creative and critical thinking skills. Project-based learning (PBL) grounded in Herrington's Authentic Learning framework helps design students' experience real-world contexts that mirror professional creative practice. However, effectiveness when supported by AI tools remains underexplored. This study investigates how AI-integrated authentic project-based learning, scaffolded through a structured AI dialogue approach based on the Devil's Advocate Reflective Thinking Framework (D.A.R.T. Framework). A convergent mixed-methods design was employed to 53 undergraduate students enrolled in a Design Management course in Indonesia. Drawing on a validated Likert-scale survey, individual reflection papers, and peer assessment scores, the 17-item AI-Integrated Authentic Learning Scale (AI-ALS) demonstrated excellent reliability (Cronbach's α = 0.949); students rated AI scaffolding and autonomy highest (D4, M = 4.29) and real-world AI application lowest (D1, M = 3.68). Qualitative analysis of reflection papers identified communication as the dominant coding category (366 occurrences), while conflict management was the least represented (32 occurrences), indicating a significant tendency toward conflict avoidance. Peer assessment scores were uniformly positive (all items M > 4.20), supporting the survey finding that AI functioned as a creative scaffold rather than a replacement for student agency, provided that lecturer scaffolding guided students' engagement with AI and prevented uncritical reliance on AI-generated outputs. These findings offer practical implications for design educators and curriculum developers seeking to integrate generative AI as a pedagogical scaffold that supports, rather than replaces, students' creative agency and higher-order thinking skills in preparation for professional practice.

Citation : Indriati, L., Neo, M., Tan, H, Y-J. (2026). Authentic Project-Based Learning with Generative AI: Fostering Creative and Critical Thinking in Design Education. *International Journal of Learning, Teaching and Educational Research, 25*(8), 289–312. https://doi.org/10.26803/ijlter.25.8.11

*Corresponding author: Mai Neo; *neo.mai@mmu.edu.my*

Keywords: authentic learning; project-based learning; AI-integrated learning; design education

1. Introduction

The emergence of artificial intelligence (AI) as a creative tool led to an urgent reconsideration of how design education develops and assesses higher-order thinking skills. Several studies report that students tend to adopt AI outputs with insufficient critical review, particularly when task complexity is low or when assessment criteria do not reward process-level reasoning (Nguyen & Nguyen, 2025), and that AI can function productively in formative assessment only when human reflection is deliberately maintained alongside AI feedback (Sperber et al., 2025). In design-specific contexts, unmediated reliance on generative technologies has been related to a homogenisation of creative outputs and a significant loss in original conceptualisation among student designers (García-Meza et al., 2026). These findings indicate that the risk of using AI in design education is not merely accelerated production, but a substitute for the research and ideation processes that design education is expressly built to promote.

The use of generative AI in education presents a duality, the same systems that help tailor learning and speed up production also risk damaging critical thinking, originality, and academic integrity (Kasneci et al., 2023; Dwivedi et al., 2023; Cotton et al., 2024). The fundamental principle is well recognized in cognitive psychology as cognitive offloading, wherein learners externalise mental effort to a tool and may compromise their competency (Risko & Gilbert, 2016; Sparrow et al., 2011). In design education, where the value of a solution lies in the reasoning behind it, preserving student agency while still benefiting from AI is a primary concern. One widely proposed response is to cultivate AI literacy such as the capacity to understand, critically evaluate, and collaborate with AI rather than passively consume its output (Long & Magerko, 2020; Ng et al., 2021).

However, how such literacy develops within authentic disciplinary practice, and how it varies across cultural and institutional contexts (Yusuf et al., 2024), remain underexplored. Prior interventions show that AI-supported design tools are most effective when students retain meaningful control over creative decisions (Ji et al., 2025), and pedagogical models have called for generative AI to be positioned as a scaffold rather than a replacement for human cognition (Zailuddin et al., 2024). However, these responses have largely been evaluated as isolated tool-adoption interventions or single-session studies; none situates AI scaffolding within a sustained, authentic, project-based learning environment in which students must apply AI-mediated creative decisions to a real-world brief over an extended timeframe.

This unresolved question corresponds to three distinct gaps in literature. Theoretically, Herrington's Authentic Learning framework (Herrington & Herrington, 1998; Herrington, 2014) was developed before generative AI existed, and it remains unclear how the framework's principles of articulation, reflection, and coaching and scaffolding should be adapted when the coaching agent is an AI system rather than a human mentor. Methodologically, prior studies have relied primarily on single-instrument, self-report measures of AI adoption (Nguyen & Nguyen, 2025; Sperber et al., 2025), leaving unresolved how AI-integrated authentic learning can be assessed reliably when self-reported experience is triangulated against observed collaborative behavior. Contextually, the majority of published studies on AI in design education originate from institutions in developed countries (Ji et al., 2025; Shyu & Tasir, 2026); how AI-integrated authentic PBL functions within Indonesian design education, where access to AI tools and institutional AI policy may differ substantially, remains largely undocumented. Despite growing interest in the educational applications of generative AI, research that examines how AI can be integrated within authentic project-based learning environments while preserving students' creative and critical thinking, across these theoretical, methodological, and contextual dimensions, remains limited.

This study responds to these three gaps through a convergent mixed-methods approach with 53 undergraduate design students enrolled in a Design Management course at a private university in Surabaya, East Java, Indonesia. The Visual Communication Design program structures the Design Management course around a brand identity brief tied to a genuine final-project exhibition, giving students an authentic client-facing task whose stakes, audience, and output format are identical to those encountered in professional creative practice; and Indonesian creative design education, as a developing-economy context with emerging institutional AI policy, provides a boundary condition under which the contextual gap identified above can be directly examined. Students completed this brief while integrating AI tools as learning assistants throughout the ideation, iteration, and refinement processes, guided by a structured AI dialogue approach grounded in Herrington's principles of articulation, reflection, and coaching and scaffolding through Devil's Advocate Reflective Thinking Framework (D.A.R.T. Framework).

The research problem addressed by this study is that although generative AI is increasingly adopted in design education, it remains unclear whether AI can be integrated within an authentic project-based learning environment as a scaffold for creative and critical thinking, rather than as a substitute for student agency, and how such integration can be reliably assessed. This study therefore pursues three research objectives: (1) to evaluate the reliability and dimensional structure of a purpose-built survey instrument (the AI-ALS) measuring students' experience of AI-integrated authentic learning; (2) to identify, through qualitative analysis of student reflection papers, the collaborative competencies most and least developed under AI-integrated authentic PBL; and (3) to examine, through convergent analysis of survey, reflection, and peer assessment data, the extent to which AI-supported authentic PBL preserves students' creative and critical thinking and collaborative agency.

The study makes two distinct contributions that are not, to the authors' knowledge, available in the existing design-education literature. Conceptually, it introduces the D.A.R.T. (Devil's Advocate Reflective Thinking) Framework as an operationalisation of Herrington's articulation, reflection, and coaching-and-scaffolding principles for AI-mediated authentic learning, repositioning generative AI from a content-producing assistant to a critical discussion partner. Methodologically, it develops and provides initial reliability evidence for the AI-Integrated Authentic Learning Scale (AI-ALS), a purpose-built instrument for assessing students' experience of AI-integrated authentic learning across four theoretically grounded dimensions—the first instrument explicitly aligned with an authentic-learning framework in this domain.

2. Literature Review

2.1 Authentic Learning in Design Education

Herrington and Herrington (1998) proposed a model of authentic learning with nine principles including real-world relevance, ill-defined tasks, multiple perspectives, collaborative learning, reflection, and authentic assessment. In design education, this framework has proven productive because design itself is an authentic discipline, professional designers work within client briefs, audience needs, and iterative feedback; the conditions that authentic PBL seeks to simulate (Hasan et al., 2023). The design management course represents one of the most authentic tasks available in the design curriculum of the Visual Communication Design study program used in this research: the stakes, audience, and output format are identical to those encountered in professional creative practice.

Herrington's framework, however, was formulated before generative AI existed, and several of its nine principles require theoretical extension in AI-mediated contexts. The principle of coaching and scaffolding at critical times assumes a human mentor capable of contingent, relational judgment; when an AI system participates in this scaffolding role, as in the present study, it remains unclear whether AI can perform the same function or introduces a qualitatively different form of support requiring its own theoretical treatment.

Studies applying authentic learning theory in technology-mediated environments have found that the reflection and collaborative learning principles are particularly vulnerable to erosion when digital tools reduce the visibility of student reasoning to instructors (Mercier et al., 2023; Hasan et al., 2023). The present study's central theoretical position is therefore that AI can function as a scaffolding agent within an authentic learning environment provided student agency is deliberately preserved, positioning this study as a deliberate extension of Herrington's framework to AI-mediated design education rather than a straightforward application of it.

2.2 Generative AI in Creative Education

Research on generative AI in higher education has expanded rapidly since the public release of large language models and image generation tools. AI can function as a productive partner in formative assessment when human reflection is maintained alongside AI feedback (Sperber et al., 2025). However, multiple studies caution that students tend to adopt AI outputs with insufficient critical evaluation, particularly when task complexity is low or when assessment criteria do not reward process-level reasoning (Nguyen & Nguyen, 2025).

In design education, the risks associated with AI adoption are critical because the study program values originality, stylistic voice, and creative problem-solving as core professional competencies. Technology adoption in creative education must be accompanied by metacognitive scaffolding that makes students aware of when and how AI is influencing their decisions (AlGerafi et al., 2023). The approach proposed as the D.A.R.T. (Devil's Advocate Reflective Thinking) Framework can be positioned against existing models of AI in pedagogy. Most reported approaches cast generative AI as a productivity or tutoring aid that accelerates ideation, supplies feedback, or scaffolds tasks (AlGerafi et al., 2023; Ji et al., 2025; Sperber et al., 2025).

The work that expressly frames AI as a scaffold rather than a replacement (Zailuddin et al., 2024), nonetheless views the system as a support for student. The present framework differs in the direction of the interaction. Instead of providing answers or feedback, the AI is prompted to adopt an adversarial, devil's-advocate stance that questions the student's reasoning and requires decisions to be defended and refined. It thereby reframes AI from a source of content or assistance to an instrument of structured critical dialogue that has not previously been operationalised within an authentic-learning framework in design education.

The authentic-learning perspective adopted here has deep theoretical roots. It draws on situated cognition, which holds that knowledge is inseparable from the activity, context, and culture in which it is developed and used (Brown, Collins, & Duguid, 1989), and on situated learning theory, which frames learning as participation in the practices of a community (Lave & Wenger, 1991). Herrington and colleagues translated these ideas into design principles for technology-supported environments, specifying characteristics such as real-world relevance, ill-defined tasks, sustained investigation, multiple perspectives, collaboration, reflection, articulation, coaching, and authentic assessment (Herrington & Oliver, 2000; Herrington & Herrington, 2006; Herrington, Reeves, & Oliver, 2010). These characteristics provide the pedagogical scaffolding within which the present study situates AI.

Project-based learning operationalises these principles through extended, product-oriented inquiry. Foundational reviews characterise it as student-centered instruction organised around a driving question and realistic artifacts, distinguished by autonomy, constructive investigation, collaboration, communication, and reflection (Thomas, 2000; Kokotsaki, Menzies, & Wiggins, 2016; Krajcik & Shin, 2014). Because these features make interpersonal processes visible and accessible, PBL is a natural setting in which to examine how AI reshapes not only individual creative and critical thinking but also the collaborative competencies through which design work is accomplished.

Two further bodies of theory inform this study. Collaborative-learning research, grounded in social interdependence theory, shows that interpersonal and metacognitive competencies develop through positive interdependence and visible, negotiated interaction, and that these gains depend on how group work is structured (Johnson & Johnson, 2009). AI literacy research, in turn, reframes the learner's relationship to AI as one of critical evaluation and collaboration rather than consumption (Long & Magerko, 2020; Ng et al., 2021). Positioning AI as a discussion partner within a collaborative, authentic project therefore sits at the intersection of these literatures: it treats critical engagement with AI as a competency developed through structured social interaction, not as an incidental by-product of tool use. Combining Herrington's Authentic Learning principles with the emerging literature on generative AI in education, this study conceptualises authentic task conditions as the antecedent, AI scaffolding as the mediating mechanism, and students' creative thinking, critical thinking, and collaborative competencies as the consequent outcomes as shown in Figure 1.

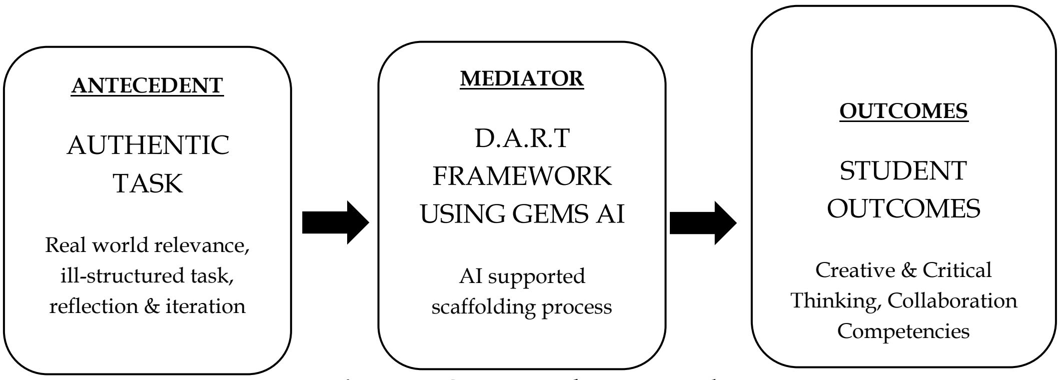

Figure 1: Conceptual Framework

Authentic, high-stakes task conditions are theorised to increase the motivational salience of AI-supported ideation, which in turn shapes how critically students evaluate AI-generated content before incorporating into their work; the D.A.R.T. framework's articulation, reflection, and coaching and scaffolding components is theorised to be the mechanism through which this critical evaluation is trained rather than assumed. This yields three theoretical propositions examined in this study:

P1: In an authentic PBL environment where AI is positioned as a scaffolding partner rather than a generative agent, students will demonstrate increasing critical evaluation of AI outputs as task authenticity and stakes increase.

P2: AI-integrated authentic PBL will support creative ideation (through broadened idea generation) while preserving students' evaluative agency, provided lecturer scaffolding guides the AI interaction.

P3: Collaborative competencies developed under AI-integrated authentic PBL will be uneven across dimensions, with competencies requiring direct interpersonal negotiation (such as conflict management) developing more slowly than competencies supported by task structure alone (such as role distribution).

3. Methodology

This study was conducted in a Design Management course at a private university in Surabaya, East Java, Indonesia. Within the Visual Communication Design curriculum, Design Management is the sub-discipline concerned with planning, coordinating, and delivering design projects against client requirements—operationalised here through a brand-identity brief for a genuine final-project exhibition—rather than with the production of individual design artifacts in isolation. The setting, a private university in Surabaya, East Java, Indonesia, is a Southeast Asian developing-economy context whose creative-design education and emerging institutional AI policy give the study its regional relevance. Participants were 53 undergraduate students in the third semester of a Visual Communication Design program, organised into 13 self-selected groups of four to five members. A sample size of 53 is considered adequate for reliability analysis, as estimates of internal consistency using Cronbach's alpha are generally stable with samples of this size (Nunnally & Bernstein, 1994).

Interpretation of the internal-consistency coefficients further followed the cautions of Taber (2018) and Tavakol and Dennick (2011), who note that a high alpha reflects item interrelatedness and scale length rather than unidimensionality. Therefore, the sample was deemed sufficient for assessing the internal consistency of the measurement instrument. Students were required to develop a comprehensive brand identity for the final project exhibition over a 16-week project. The course was designed around Herrington's Authentic Learning principles, incorporating complex ill-structured tasks, opportunities for interaction with practising designers through guest critique sessions, and embedded assessment requiring sustained reflection on the collaborative process.

All participants provided written informed consent prior to participation. The study received ethical approval from the Research Management Centre (RMC) Multimedia University; Reference No. EA0242023; approved on December 29, 2023. Participation was entirely voluntary, and students were informed that their decision to participate or withdraw at any time would not affect their course grades.

3.1 AI Supported Authentic Learning

This study proposes AI as a critical discussion partner rather than a generative assistant. The D.A.R.T. Framework (Devil's Advocate Reflective Thinking Framework), presented in Figure 2 is grounded in Herrington's Authentic Learning theory and utilises the articulation, reflection, and coaching and scaffolding principles. Students engage in a structured dialogue process that requires them to articulate their reasoning, receive critical questioning (a form of scaffolded coaching), and reflect on the adequacy of their justifications.

In operational terms, the roles within each dialogue cycle are distributed as follows: the student articulates a design decision (articulation); the GEMS-based AI assistant adopts a Devil's advocate perspective, posing critical counter-questions rather than supplying ready-made solutions (coaching and scaffolding); and the student then reflects on, and revises, that justification against the client brief (reflection). Figure 2 summarises these roles and their sequence. The class design incorporates an AI assistant using GEMS, a custom Google Gemini assistant that lecturers create to support students working on specific tasks. All materials, information sources, and data needed for the task were provided to the system allowing students to discuss with the AI assistant in a controlled environment.

The collaborative assessment model operationalises authentic assessment, collaboration, and reflection. Peer assessment and individual reflection together create an assessment that captures both the final outputs and the process of authentic collaborative design work. Constructivist, technology-mediated learning environments grounded in authentic principles enabled design students to develop not only technical competencies but also interpersonal and metacognitive skills such as communication, critical reflection, and collaborative negotiation (Neo & Neo, 2009).

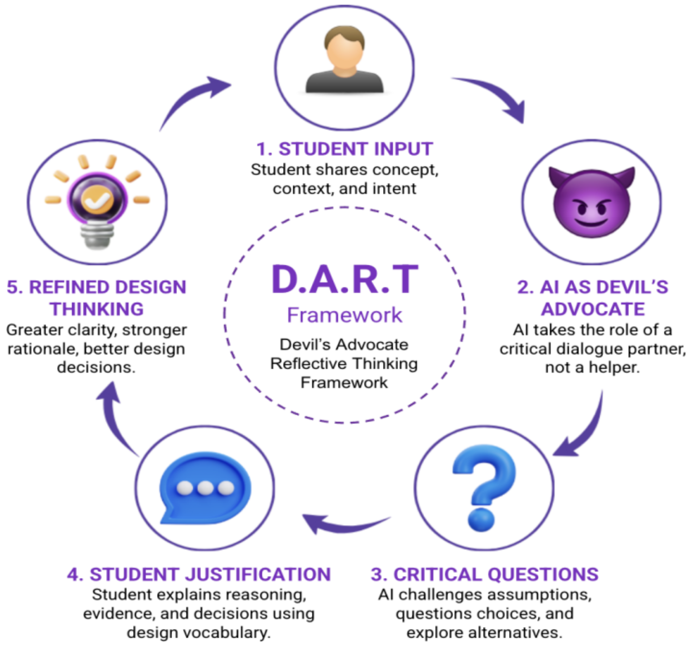

Figure 2: AI integration using D.A.R.T. Framework

3.2 Research Framework

This study adopts a convergent mixed-methods research design (Creswell & Creswell, 2018), in which quantitative and qualitative data are collected simultaneously, analysed independently, and merged to produce integrated findings. This design is particularly appropriate when research questions require both quantitative and qualitative depth within a single data collection window, within a project from the Design Management course. The theoretical framework integrating both data is Herrington's Authentic Learning model, which provides the rationale for (a) the design project as an authentic task context, (b) reflection and peer assessment as learning-embedded assessment instruments, and (c) AI tools as scaffolding that should not be used as student competence grows. Figure 3 is presented to provide a holistic understanding of the research design.

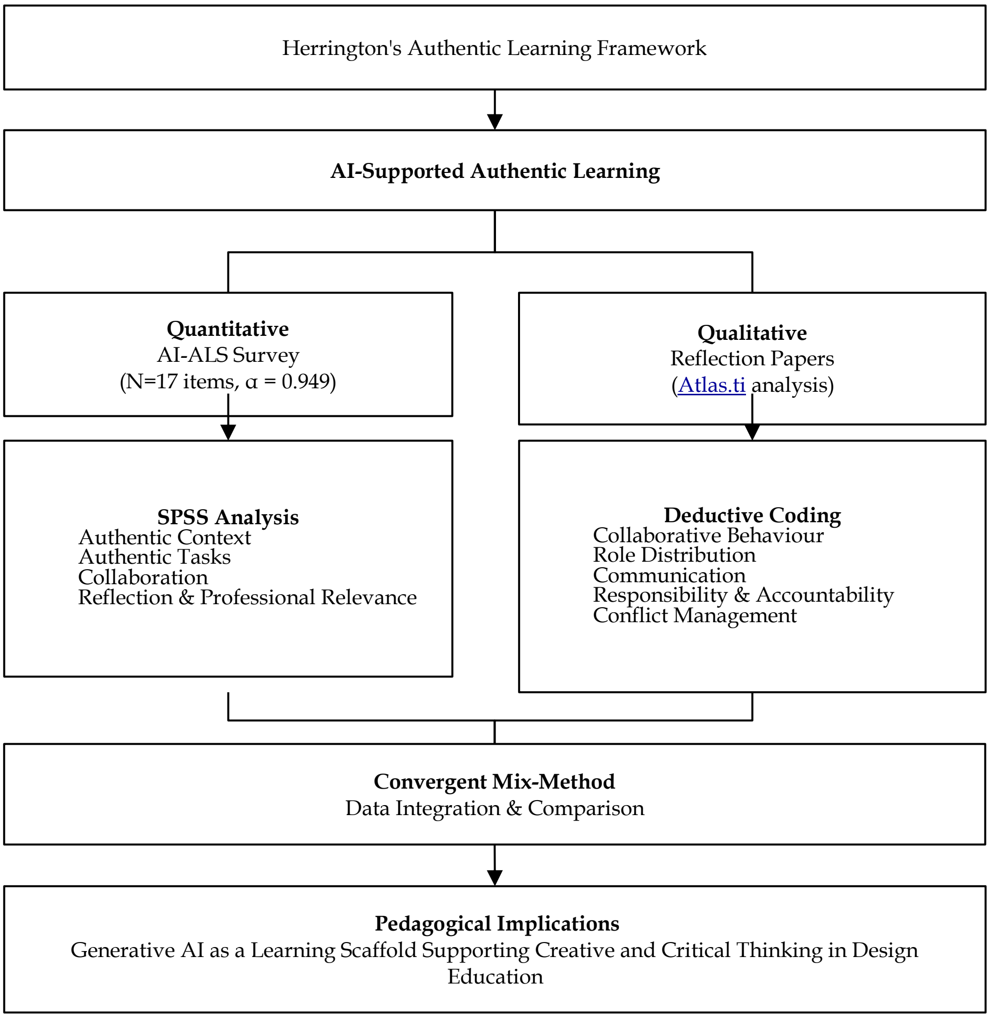

Figure 3: Research Methodology Design Framework

Data were collected through three instruments aligned with the convergent design. The AI-Integrated Authentic Learning Scale (AI-ALS) is a 17-item self-report questionnaire rated on a five-point Likert scale (1 = strongly disagree to 5 = strongly agree); its items were written to operationalise the four Herrington-derived dimensions (D1 real-world application, D2 creative thinking, D3 critical thinking, and D4 scaffolding and autonomy) and were reviewed for clarity prior to administration.

Individual reflection papers, written in Bahasa Indonesia, captured students' first-person accounts of their collaborative and AI-mediated design process in response to a standardised reflective prompt. A five-item peer-assessment rubric asked each student to rate every other group member on a five-point scale across five collaborative behaviors. The instruments were administered within a single data-collection window tied to the 16-week project, with the AI-ALS and peer assessment completed at the end of the project and the reflection papers submitted.

Quantitative data were analysed in IBM SPSS Statistics 30. Internal consistency was estimated with Cronbach's alpha, $\alpha = [k / (k - 1)] \times [1 - (\Sigma\sigma^2\_item / \sigma^2\_total)]$, where k is the number of items, $\Sigma\sigma^2\_item$ is the sum of the item variances, and σ^2_total is the variance of the total score. Item quality was examined through corrected item-total correlations—the correlation between each item and the sum of the remaining items—with values ≥ 0.30 treated as acceptable. For each item and dimension, means, standard deviations, and the percentage of respondents selecting 4 or 5 were computed. Qualitative reflection data were analysed through directed content analysis in Atlas.ti using the five-code, fifteen-sub-code framework, with category frequencies (occurrence counts) tabulated across the 53 papers. Finally, the quantitative and qualitative strands were merged in a convergence–divergence matrix (Table 6) to yield integrated findings.

4. Results and Findings

4.1 Survey Instrument: AI-Integrated Authentic Learning Scale (AI-ALS)

Addressing Research Objective 1 (evaluating the reliability and dimensional structure of the AI-ALS): A 17-item survey instrument was developed to measure students' experience of AI integration within the authentic PBL environment across four dimensions corresponding to Herrington's Authentic Learning elements. All items were rated on a five-point Likert scale (1 = strongly disagree, 2 = disagree, 3 = neutral, 4 = agree, 5 = strongly agree). Internal consistency was analysed using IBM SPSS Statistics 30. The overall AI-ALS demonstrated excellent reliability: Cronbach's α = 0.949, N items = 17, n = 53. Scale mean = 67.26 (SD = 10.73). Table 1 presents the complete instrument with item means, standard deviations, and corrected item-total correlations.

Table 1: AI-ALS Survey Results

Dim.	Items (N= 17) Cronbach Alpha .949 (n=53)	Mean (M)	Std. Dev. (SD)	Percent (%)
D1	AI assists in organizing our group's creative ideas more systematically.	4.02	0.75	73.6
	AI facilitates better critical thinking and discussion during collaboration.	3.81	0.92	69.8
	AI encourages more collaborative behavior among team members.	3.51	1.15	52.8
	Using AI during teamwork helps synthesize multiple perspectives effectively.	4.06	0.93	73.6
	I feel more motivated to engage in class activities when AI-supported tools are used.	3.75	0.81	64.2
D2	AI-enhanced instruction helps me generate a wider range of creative ideas.	4.00	0.73	77.4
	AI tools contribute to faster idea generation during group projects.	4.38	0.69	92.5
	AI applications support real-world problem-solving skills in design tasks.	3.91	0.84	71.7

	Integrating AI into teamwork activities improves overall project quality.	4.04	0.85	79.2
	AI tools help lecturers deliver clearer and more structured instructions.	3.96	0.71	77.4
D3	I feel that AI tools supplement my creativity rather than replace it.	4.15	0.74	79.2
	AI-enhanced collaboration leads to more cohesive final project outcomes.	3.79	0.86	62.3
	AI helps reduce misunderstandings among team members regarding project tasks.	3.53	1.05	47.2
	AI suggestions sometimes introduce new, unexpected design perspectives.	4.19	0.71	83.0
	AI-enhanced instruction improves my understanding of project requirements.	4.08	0.83	73.6
D4	- Using AI in class helps me feel more confident in contributing ideas to my team.	3.94	0.86	71.7
	AI suggestions stimulate more diverse brainstorming discussions in teams.	4.15	0.89	79.2

Note: Dim = Dimension; D1 = Real-world AI application, D2 = Creative thinking with AI, D3 = Critical thinking & AI evaluation, D4 = AI scaffolding. Percent (%) represents the proportion of respondents (n = 53) who selected 4 ("agree") or 5 ("strongly agree") on the 5-point Likert scale for each item. Corrected item-total correlations of 0.30 or above are conventionally interpreted as acceptable item discrimination within a dimension.

Table 2 presents descriptive statistics and the reliability of the dimensions. AI scaffolding and autonomy (D4: M = 4.29) recorded the highest dimension mean, followed by critical thinking and AI evaluation (D3: M = 4.10) and creative thinking with AI (D2: M = 3.97). The real-world AI application (D1: M = 3.68) recorded the lowest mean dimension. This lower rating was driven mainly by the item on AI encouraging collaborative behavior among team members (M = 3.51); the D1 items on AI's usefulness for organising creative ideas (M = 4.02) and synthesising multiple perspectives (M = 4.06) were rated close to the 'agree' anchor.

Table 2: AI-ALS Dimension Summary

Dimension (Herrington Principle)	Dim. Mean	α	Interpretation
D1 — Real-world AI application (Authentic context)	3.68	0.842	M = 3.68, between the 'neutral' (3) and 'agree' (4) anchors and the lowest of the four dimensions. This rating was driven mainly by the item on AI encouraging collaborative behaviour (M = 3.51); items on organising creative ideas (M = 4.02) and synthesising multiple perspectives (M = 4.06) were rated close to 'agree', indicating the reservation concerned AI's role in fostering collaboration rather than its real-world relevance.
D2 — Creative thinking with AI (Ill-defined tasks)	3.97	0.858	M = 3.97, approaching the 'agree' anchor (4). Students used AI as a creative scaffold, generating broader idea ranges while preserving personal design decisions.
D3 — Critical thinking & AI evaluation (Higher-order cognition)	4.10	0.861	M = 4.10, at the 'agree' anchor (4). Students critically evaluated AI outputs before adoption.
D4 — AI scaffolding & autonomy (Scaffolding)	4.29	0.771	M = 4.29, above the 'agree' anchor (4). Students reported reduced AI dependence on the project output.

4.2 Qualitative Content Analysis

Addressing Research Objective 2 (identifying the collaborative competencies most and least developed under AI-integrated authentic PBL): Individual reflection papers were analysed using Atlas.ti (free web version, as shown in Figure 4) through directed qualitative content analysis. A coding framework comprising five main codes and fifteen sub-codes as shown in Table 3 was developed from collaborative learning theory and prior research. Initial coding was conducted by the first author. Reflection papers written in Bahasa Indonesia were translated using ChatGPT Go with a standardised prompt specifying semantic accuracy and cross-document consistency; all translations were verified by a bilingual researcher.

To ensure translation accuracy and avoid interpretive bias, verification followed a structured process: (1) AI-generated English translation was compared to Bahasa Indonesia by a bilingual researcher fluent in both languages and familiar with the design domain; (2) inconsistencies were resolved and coding was conducted on the verified English texts; and (3) every passage quoted in the Results was individually re-verified against its source. A second independent coder was not employed, and inter-rater reliability was not established; this single-coder approach, together with the first author's dual role as course instructor and researcher, is acknowledged as a methodological limitation that may bias interpretation toward positive collaborative outcomes.

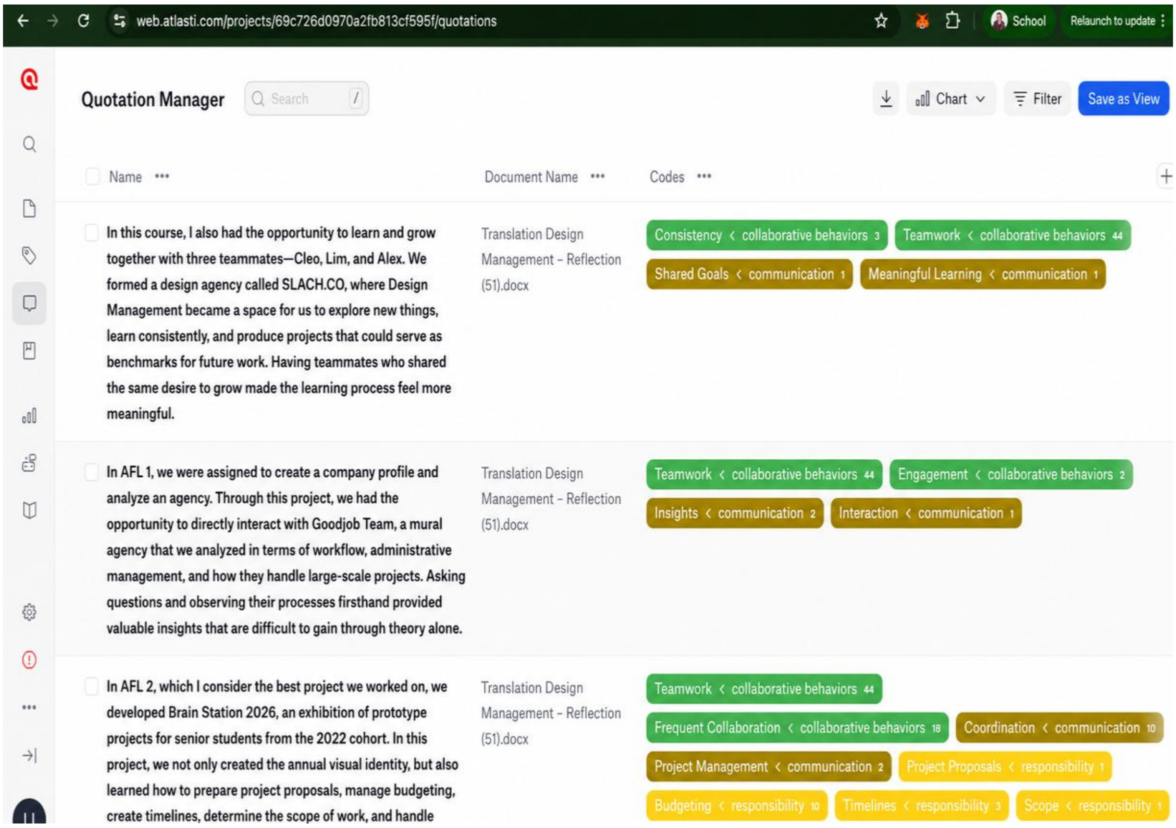

Figure 4: Atlas.ti web app preview

Table 3: Coding Framework with Sub-codes and Definitions

Main Code	Sub-code	Code Label	Definition
Collaborative Behaviour	CB1	Active Collaboration	Students actively engage in joint work, shared decision-making, or co-creation processes.
	CB2	Passive Participation	Participation characterised by following instructions without initiative or leadership.
	CB3	Unequal Participation	Perceived imbalance in contribution levels among group members.
Role Distribution	RD1	Clearly Defined Roles	Roles and responsibilities are explicitly assigned and understood by all members.
	RD2	Flexible Roles	Roles adjusted based on group needs or individual strengths.
	RD3	Overlapping Roles	Confusion regarding responsibilities among group members.
Communication	COM1	Effective Communicatio n	Clear, timely, and constructive exchange of ideas and feedback among group members.
	COM2	Ineffective Communicatio n	Communication breakdowns leading to misunderstandings, delays, or coordination failures.

	COM 3	Communication Strategies	Deliberate use of digital tools or structured systems to facilitate group communication.
Responsibility & Accountability	RA1	High Responsibility	Strong sense of individual ownership and accountability toward assigned tasks.
	RA2	Shared Accountability	Collective responsibility for the quality and outcomes of the group project.
	RA3	Lack of Responsibility	Failure to meet obligations or over-reliance on other members to complete tasks.
Conflict Management	CM1	Constructive Resolution	Conflicts addressed through structured discussion, compromise, or negotiation.
	CM2	Conflict Avoidance	Deliberate avoidance of disagreement to maintain group harmony, at cost of critical discussion.
	CM3	Unresolved Conflict	Ongoing tension that negatively affects group performance or individual motivation.

Table 4 presents the frequency of coded items and total occurrences across 53 reflection papers. Communication was the most frequently coded category (366 occurrences), followed by Responsibility and Accountability (318 occurrences). Collaborative Behavior and Role Distribution were the third and fourth most frequent dimensions. Conflict Management was underrepresented (32 occurrences), representing only 2.8% of total occurrences. Within this category, the coded segments clustered around conflict avoidance (sub-code CM2) rather than constructive resolution (CM1) or unresolved conflict (CM3), reinforcing the avoidance reading rather than indicating an absence of tension.

Table 4: Total Occurrences per Category from Atlas.ti Content Analysis (Ranked)

Category (n=53)	Occurrences	Examples of Related Terms
Communication	366	discussion, feedback, pitching, presenting ideas, explain reasoning
Responsibility & Accountability	318	responsible, ownership, accountability, commitment, deadline
Collaborative Behaviour	275	teamwork, collaboration, collective effort, co-creation, joint work
Role Distribution	154	role, task distribution, divide tasks, job description, responsible for
Conflict Management	32	conflict, miscommunication, disagreement, compromise, resolution

The occurrences do not capture the quality of that communication. A closer reading of the reflections surfaced a nuanced and partly disconfirming pattern such as in several groups the high volume of communication was AI-mediated rather than peer-to-peer, with members exchanging and commenting on AI-generated concepts in place of direct discussion with one another. A high communication count did not always correspond to substantive collaborative dialogue, as one student made explicit as Student 13, Group 1 said:

> *"We shared the AI concepts in our group chat and started commenting on those instead of having proper discussions. I realised later that we were reacting to the AI, not to each other."*

4.3 Peer Assessment

Contributing to the convergent analysis for Research Objective 3: A five-item peer assessment rubric measured individual contributions across five collaborative dimensions on a five-point scale. Each student independently rated all other group members at project completion. Individual peer assessment mean scores were computed by averaging all ratings received from group members as presented in Table 5.

Table 5: Peer Assessment Results (Ranked)

Items	Means (M)	% negative response	% neutral	% positive response
Show a supportive and cooperative attitude in the group	4.32	8.45	8.45	83.10
Contribute to the successful execution of group projects	4.31	10.33	5.16	84.51
Completing the distribution of tasks in groups well	4.27	10.80	5.63	83.57
Always attend group discussions on time and be proactive in those discussions	4.23	10.33	10.33	79.34
Contribute to group discussions, especially in the process of research and idea development	4.20	11.27	11.27	77.46

Note. n = 53. M = group mean of peer-assigned scores (scale 1–5).
Negative = rated ≤ 2; Neutral = rated 3; Positive = rated ≥ 4.

All items achieved mean scores above 4.20, with the most positive evaluations for showing a supportive attitude (M = 4.32) and contributing to project success (M = 4.31). The item related to research and idea development received the lowest mean score (M = 4.20), reflecting unequal engagement during the concept-development phases. Each group was provided with a summary of the peer-assessment results to facilitate reflection on their strengths and weaknesses, thereby supporting continuous improvement throughout the project, as shown in Figure 5.

Group: Group 4

Class	: Design Management
Topic	: Pitching Presentation for Brainstation 2025

Peer Scoring & Comments

Development of Design Project Proposal	4.28
Delivery & Presentation Skills	3.87
Visual Aids / Slides	4.33
AVERAGE KU3.2 (35% of AFL 2)	4.16

- Lengkap, bagus, relevan
- Tema dan konsep unik dan design bagus
- konsep unik dan lucu namun beberapa anggota masih membaca slide saat presentasi
- very cute
- Presentasinya kurang interaktif
- Great presentation.
- sudah bagus
- well presented
- menurut saya warna-warna yang gunakan menarik, simple, dan mudah di pahami
- sudah BAGUSSSS
- keren
- maskotnya lucu ngestack2 jadi satu
- Bagus
- slide cukup menarik, penyampaian saat presentasi juga cukup
- beberapa masih terlalu membaca hp
- Ide dari base puzzle sebagai satu kesatuan sangat menarik dan luas untuk di jelajah, konsep dasarnya bagus
- Cukup menarik
- tidak ada
- oke
- Element graphic kurang, riset dan design layout bagus
- penjelasan sudah baik
- suara krg keras dikit
- Cukup lengkap materinya namun cara presentasi masih monoton
- sudah baik
- Unik dan Menarik konsepnya
- nice concept and cute visuals
- Kurang banyak assets, mascot lucu
- Very nice, cool concept.
- okee
- Tidak ada
- kerennn all good guys konsepnya lucuuu, buat sarannya maybe next presentnya bisa lebih jelas di cara penyampaian informasinya yaahh

Notes from Lecturer

- konsepnya menarik dan visualnya playfull walaupun simple hanya menggunakan elemen geometri
- logomarknya tidak langsung terlihat B dan S, mungkin karena warna-warni, bisa dicoba monokrom (gelap - terang) atau warm/cool tone per hurufnya mgkn akan lebih terbaca

Group: Group 7

Class	: Design Management
Topic	: Pitching Presentation for Brainstation 2025

Peer Scoring & Comments

Development of Design Project Proposal	4.83
Delivery & Presentation Skills	4.71
Visual Aids / Slides	4.88
AVERAGE KU3.2 (35% of AFL 2)	4.81

- Bagus
- Keren ada animasi logonya
- warna yang digunakan cukup basic namun penjelasannya lengkap dan proposalnya lengkap dan menarik
- Visual menarik
- bagus
- Sangat kereeenn dan lengkap wowwyeah
- Please pick this group. Thank you 🙏
- bagus pol
- ini bagus banget bahkan lengkap semua dri mockup, terus designnya yang very simple idenya ini membuat sy terpukau selama melihat presentasi kalian keren buangettt
- Perfect 👌
- sudah oke keren bgt
- Menarik
- perfect
- Very goodd
- Mengesankan karena ada animasi untuk logo mereka
- Sangat bagus dan inovatif, paparan materi juga sangat lengkap dan immersive
- perfect dude
- Output2 designnya sangat menarik dan banyak
- penyampaian jelas dan bagus
- gila luengkap, tapi ya idenya sama kayak tahun lalu brainstation like kinda alike
- amazing, lots of detail baik konsep maupun output
- tidak ada
- presentasinya terlalu cepat tidak sempat melihat visual dari ppt
- mgkin lembih dikasi jeda buat pas tampilin slide mockupny
- lengkap. Bagus, relevna
- Kereeeeeeeenn
- emezing
- Penjelasan sangat panjang, tetapi untung nya saya mengerti sampe "WAAWWW"
- Speechless, everything is so good and thought very well. It's like a proper event already. If there were any suggestion it would be to actually use this for the next brainstation since everything has been very well prepared
- super good wow
- Tidak ada

Notes from Lecturer

- taglinenya nice dan straightforward, membawa pesan/misi posiitif
- simple buth cathcy, very nice
- ada animasi +point
- ada 4 mascot yang melambangkan key value brainstation
- collateral lengkap

Figure 5: Peer assessment feedback Summary for Each Group

4.4 Data Integration and Comparison

For the purposes of this integration, convergence is defined as agreement in direction and substantive interpretation between the quantitative AI-ALS dimension score and the qualitative reflection theme addressing the same construct, including cases where the qualitative evidence is drawn from a related rather than identically labeled code (e.g., collaborative behavior occurrences used as a proxy for creative-process engagement, not a direct measure of creative thinking). Divergence is defined as a case where the two data strands point to different substantive conclusions about the same construct.

Table 6 presents a matrix that integrates quantitative AI-ALS findings, qualitative reflection paper themes, and peer assessment data. Three of the five dimensions showed convergent findings across data strands, supporting the interpretation that AI-integrated authentic PBL supports creative ideation while preserving critical agency; where a qualitative code does not map directly onto a quantitative dimension, this imprecision is noted in the discussion of that row. The conflict management dimension has the lowest occurrences, indicating a conflict-avoidance dominant pattern.

Table 6: Comparison Matrix (Convergent Mixed-Methods Integration)

Dimension	AI-ALS Survey (Quantitative)	Reflection Themes (Qualitative)	Convergence / Divergence
Creative Thinking	D2 M = 3.97; students used AI as scaffold while preserving agency	275 occurrences; students describe AI as idea-starter, not decision-maker	Convergent. Both data confirm AI augmented rather than replaced creative thinking.
Critical Thinking	D3 M = 4.10 (highest dimension); strong corrected item-total r = 0.839	Reflection papers reveal students questioned AI accuracy before use; improved evaluation capacity over time	Convergent. Quantitative data confirmed by qualitative evidence of evaluative reasoning.
Real-world Application	D1 M = 3.68 (lowest); hesitancy in early-project AI adoption	366 communication occurrences; portfolio context flagged as motivating but unfamiliar territory	Partially divergent. Survey reveals lower authentic integration; reflections reveal motivational benefit of real context.
Scaffolding/ Autonomy	D4 M = 4.29; declining AI dependence confirmed quantitatively	Students reflect on growing confidence; note instructor guidance as critical enabler	Convergent. Both data confirm the fading-scaffolding principle.
Conflict Management	Not a direct AI-ALS dimension; peer scores for proactiveness item M = 4.23	Only 32 occurrences; conflict avoidance dominant pattern	Divergent. Peer scores suggest positive performance; reflections reveal avoidance rather than constructive resolution.

5. Discussion

The strongest finding across both datasets is that AI functioned as a creative scaffold rather than a creative replacement within the authentic PBL context. The AI-ALS D2 (Creative Thinking, M = 3.97) and D3 (Critical Thinking, M = 4.10) scores confirm that students maintained evaluative control over AI outputs, while the qualitative data from 275 coded collaborative behavior instances showed that students consistently described AI as an idea-starter rather than a decision-maker.

The high critical thinking dimension score (D3 M = 4.10) suggests that the authentic, high-stakes nature of the task motivated students to evaluate AI outputs carefully. This finding is broadly consistent with Shyu and Tasir's (2026) observation that generative AI in project-based 3D modeling accelerated ideation while introducing new demands for critical evaluation of AI-generated content; the present study extends this by showing that authentic, client-facing stakes appear to strengthen rather than dilute students' evaluative engagement with AI outputs, confirming and extending rather than merely replicating this prior finding. Student 14, Group 7 expressed a view that aligns with Herrington's principles of authentic tasks:

> "*Using AI made me think more carefully about the reasoning behind my design choices, because I knew the output had to answer the client's needs according to the brief, not the AI.*"

Communication, responsibility and accountability became evident as the dominant themes in student reflections. Students who wrote more extensively about communication strategies were more likely to receive positive peer ratings, consistent with research - that link reflective articulation of collaborative competencies and observable teamwork quality (Francis et al., 2024). Student 8, Group 10 expressed that they had not really communicated with their peers because AI assistants helped them generate ideas faster; this later caused communication issues to appear in the group project. The AI-mediated communication pattern justifies further exploration as a primary finding, because it helps us understand the results about how people interact to each other.

Communication was the most-coded category initially reads as a strength; however, some groups are spending a lot of time interacting with AI, rather than talking to each other directly. This means the artificial intelligence system is affecting the way people interact with each other. Two implications follow. First, frequency-based indicators of collaboration can overstate its quality when an AI intermediary is present, so assessment instruments should distinguish peer-to-peer dialogue from AI-mediated exchange. Second, the pattern indicates a new competency for AI-integrated teamwork, leaning on AI to inform discussion without allowing it to substitute the interpersonal negotiation through which common understanding and accountability are built.

> "*We shared the AI concepts in our group chat and started commenting on those instead of having proper discussions. I realised later that we were reacting to the AI, not to each other.*"

The conflict management findings are consistent across both the original study and the extended AI-integrated analysis. With only 32 reflection occurrences

(2.8% of total), conflict management remains the most underdeveloped reflective competency in this course. The qualitative evidence confirms widespread avoidance: students chose harmony over critical confrontation, suppressing disagreements that could have improved decision quality. The integration of AI introduces a new dimension to this challenge. When AI provides a design suggestion that a student believes is suboptimal, the student must now navigate not only interpersonal conflict with human peers but also a form of authority conflict with an AI system—and do so in a group context where deference to AI may be perceived as less threatening to group harmony than opposing a peer.

Future curricula must explicitly address constructive AI critique as a professional competency, alongside constructive peer critique. This pattern is consistent with the broader collaborative learning literature on groupthink and social loafing, which finds that when task visibility is high, but confrontation carries social cost, group members frequently substitute harmony-preservation for critical negotiation (Karau & Williams, 1993; Luo et al., 2021). Student 5, Group 7 expressed an avoidance pattern: unlike loafing driven by low task visibility, the avoidance observed here occurred despite high task visibility, suggesting that AI-mediated group communication may introduce a distinct pathway to conflict suppression that existing social loafing frameworks do not fully account for.

> *"I knew the AI's direction was wrong for our client, but it was easier to go along with it than to explain why we should change it. The group seemed relieved to have an answer, even if it wasn't the right one."*

The use of AI as an as an idea starter is consistent with a way of thinking that says people and things can work together to be creative (Risko & Gilbert, 2016; Sparrow et al., 2011), in which routine mental work is handed over to a help while the high-level judgement that authentic assessment aims to build remains with the student; the high critical thinking level (D3, M = 4.10) indicates that this offloading did not take away the ability to make judgments, specifically the ability of the evaluative agency. The collaborative findings align with collaborative learning theory (Johnson & Johnson, 2009; Neo & Neo, 2009; Mercier et al., 2023), which holds that interpersonal and metacognitive competencies develop through visible, negotiated interaction; the emergence of an AI-mediated communication pattern, in which students reacted to shared AI outputs rather than to one another, suggests that this visibility can be diminished when a generative intermediary absorbs part of the interactional load.

Finally, the conflict-avoidance pattern is best read through group-dynamics theory: consistent with social-interdependence and social-loafing accounts (Karau & Williams, 1993; Luo et al., 2021), students preserved group harmony at the expense of constructive dissent, and the introduction of an AI authority appears to lower the perceived cost of deferring rather than debating, adding an AI-specific pathway to conflict suppression that extends these classical accounts. This study cautions against students' uncritical adoption of AI outputs while itself drawing on AI tools in the research process: generative AI for translating reflection papers, Atlas.ti for supporting qualitative coding, and language-editing software for the manuscript. The authors distinguish two modes of AI use. In the research workflow, AI functioned as a methodological instrument applied under deliberate human oversight, with its outputs verified by bilingual and disciplinary expertise; the practice cautioned against for students is the use of AI as a cognitive substitute for independent reasoning. Making this distinction explicit is itself an enactment of the critical AI literacy the study advocates. Nonetheless, the reliance on AI-assisted translation for primary qualitative data is acknowledged as a limitation, and future studies should adopt human-led translation protocols with documented discrepancy rates. The conflict-avoidance pattern is plausibly shaped by culturally specific norms. Indonesian collaborative practice, situated within a broader Southeast Asian orientation toward social harmony and the preservation of group cohesion, tends to discourage open confrontation; in such a setting, deferring to an AI-proposed direction may be experienced as a face-preserving alternative to challenging a peer, which would help explain why students suppressed disagreement even when task visibility was high.

On this reading, introducing an AI authority does not merely add a technical actor but interacts with harmony-oriented values, potentially lowering the social cost of deference. This regional interpretation also frames the study comparatively: whereas much of the reported evidence on AI in design education originates in developed-economy, and often more individualist, settings (Ji et al., 2025; Shyu & Tasir, 2026; García-Meza et al., 2026), the present findings from Surabaya, East Java, Indonesia, illustrate how a developing-country, harmony-oriented context can reshape AI-mediated collaboration. Cross-cultural replication that contrasts harmony-oriented and confrontation-tolerant educational cultures is therefore a priority for establishing which patterns are universal and which are context-bound.

6. Limitations

This study is limited to a single cohort (n = 53) at a single private Indonesian university, using a custom GEMS AI assistant and a specific 16-week brand identity brief in Design Management class. The finding that AI functioned as a creative scaffold when student agency was preserved (D2, D3) is theoretically grounded in Herrington's framework and is likely to generalise across contexts where similar authentic task conditions and lecturer scaffolding are present.

The real-world application hesitancy identified in D1 and the pronounced conflict-avoidance pattern identified in the qualitative data are more plausibly context-specific: both may reflect the particular institutional setting, the custom GEMS configuration, the specific project brief, and cultural norms around harmony and confrontation in Indonesian higher education collaborative practice, rather than a universal feature of AI-integrated authentic PBL.

Recommendations for design educators and curriculum developers should therefore be read as strongly supported for the scaffold-preservation finding and as hypotheses requiring cross-institutional replication and conflict-avoidance findings. Future research should extend the AI-ALS across multiple institutions and cultural contexts to test the generalisability of the five-dimensional structure and examine whether the hesitancy pattern is specific to the Indonesian context or reflects a broader pattern in in design education in developing-countries.

7. Conclusion

The findings establish the internal consistency of the AI-ALS as a reliable and theoretically grounded instrument for assessing AI-integrated learning experiences in design education (overall $\alpha = 0.949$). This should not be read as full psychometric validation: construct validity evidence through factor analysis with a larger and more diverse sample, content validity evidence through expert panel review, and cross-institutional and cross-cultural replication are all required before the AI-ALS can be recommended for broad adoption. With these conditions made explicit, the instrument may be used on a provisional basis by educators and researchers as part of standards-based end-of-project evaluation procedures, alongside other evidence of AI-integrated learning experience.

Beyond what was previously known from single-instrument studies of AI adoption in design education, this study establishes that survey, reflection, and peer-assessment data can diverge from one another in theoretically informative ways: the conflict-management divergence identified here, where peer scores suggested functional performance while reflections revealed active avoidance, demonstrates that triangulated assessment can surface collaborative dynamics that single-instrument evaluation would miss entirely.

These contributions point to three specific directions for future research arising from this study's limitations: confirmatory factor analysis of the AI-ALS with a larger, multi-institutional sample to establish construct validity; independent inter-rater verification of qualitative coding in replications of this study; and longitudinal designs tracking AI-ALS scores and collaborative competency development across multiple semesters to test whether the D1 hesitancy pattern and the conflict-avoidance dynamic diminish as students gain sustained experience with AI-integrated authentic PBL.

The integration of generative AI to design education is not a temporary disruption but a durable transformation of the professional landscape for which students are being prepared. The findings of this study suggest that this transformation can be navigated productively within authentic PBL contexts, provided that curriculum designers address the real-world application gap, build explicit instruction in critical AI evaluation, and develop assessment instruments capable of capturing the new collaborative dynamics that AI introduces, and provided that the validation caveats above are respected before the AI-ALS is adopted at scale.

8. Acknowledgments

The authors wish to acknowledge the use of ChatGPT, Atlas.ti and Grammarly in the writing of this paper. The ChatGPT tool assisted in translating students' reflection papers from Bahasa Indonesia to English. Atlas.ti was used to support the deductive qualitative analysis and to identify the frequency of coded themes. Grammarly was used to improve the paper's language and grammar. The authors would like to extend their sincere thanks to the students and staff of the School of

Creative Industry at Universitas Ciputra, Indonesia, for their valuable participation in this research study.

9. References

AlGerafi, M. A. M., Zhou, Y., Oubibi, M., & Wijaya, T. T. (2023). Unlocking the Potential: A Comprehensive Evaluation of Augmented Reality and Virtual Reality in Education. Electronics, 12(18), 3953. https://doi.org/10.3390/electronics12183953

Brown, J. S., Collins, A., & Duguid, P. (1989). *Situated cognition and the culture of learning. Educational Researcher, 18*(1), 32–42.

Chan, C. K. Y., & Hu, W. (2023). Students' voices on generative AI: Perceptions, benefits, and challenges in higher education. *International Journal of Educational Technology in Higher Education, 20*(1), 43. https://doi.org/10.1186/s41239-023-00411-8

Cotton, D. R. E., Cotton, P. A., & Shipway, J. R. (2024). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International, 61*(2), 228–239. https://doi.org/10.1080/14703297.2023.2190148

Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approach* (5th ed.). SAGE Publications.

Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., Baabdullah, A. M., Koohang, A., Raghavan, V., Ahuja, M., Albanna, H., Albashrawi, M. A., Al-Busaidi, A. S., Balakrishnan, J., Barlette, Y., Basu, S., Bose, I., Brooks, L., Buhalis, D., … Wright, R. (2023). *Opinion paper: "So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy*. *International Journal of Information Management, 71*, 102642. https://doi.org/10.1016/j.ijinfomgt.2023.102642

Fleischmann, K. R. (2026). The mutual alignment of AI and society: Cultivating complementary literacies in ethics and AI. In J. M. Lavista Ferres (Ed.), *Degrees of change: What AI means for education and the next generation* (pp. 69–78). John Wiley & Sons. https://doi.org/10.1002/9781394413096.ch6

García-Meza, I. M., González Bello, E. O., & Morales-Holguín, A. (2026). Inteligencia artificial y pensamiento creativo en la formación del diseño: Un análisis crítico. *Cuadernos del Centro de Estudios en Diseño y Comunicación*, (293), 105–118. https://doi.org/10.18682/cdc.vi293.13381

Hasan, R., Hussain, S., Malik, M. H., & Akhtar, N. (2023). *Group work assessment: Assessing peer assessment at Global College of Engineering and Technology.* In *International Conference on Teaching and Learning – Digital Transformation of Education and Employability (ICTL 2022). SHS Web of Conferences, 156*, 02001. https://doi.org/10.1051/shsconf/202315602001

Herrington, A., & Herrington, J. (Eds.). (2006). *Authentic learning environments in higher education*. Information Science Publishing.

Herrington, J. (2014). Introduction to authentic learning. In A. Hegarty, J. McDonald, & S. K. Lim (Eds.), *Activity theory, authentic learning and emerging technologies* (pp. 61–67). Routledge.

Herrington, J., & Herrington, A. (1998). Authentic assessment and multimedia: How university students respond to a model of authentic assessment. *Higher Education Research & Development, 17*(3), 305–322. https://doi.org/10.1080/0729436980170304

Herrington, J., & Oliver, R. (2000). An instructional design framework for authentic learning environments. *Educational Technology Research and Development, 48*(3), 23–48. https://doi.org/10.1007/bf02319856

Herrington, J., Reeves, T. C., & Oliver, R. (2010). *A guide to authentic e-learning*. Routledge.

Ji, S., Mokmin, N. A. M., & Wang, J. (2025). Evaluating the impact of augmented reality on visual communication design education: Enhancing student motivation, achievement, interest, and engagement. *Education and Information Technologies, 30*(5), 6617–6639. https://doi.org/10.1007/s10639-024-13050-x

Johnson, D. W., & Johnson, R. T. (2009). An educational psychology success story: Social interdependence theory and cooperative learning. *Educational Researcher, 38*(5), 365–379. https://doi.org/10.3102/0013189X09339057

Karau, S. J., & Williams, K. D. (1993). Social loafing: A meta-analytic review and theoretical integration. *Journal of Personality and Social Psychology, 65*(4), 681–706. https://doi.org/10.1037/0022-3514.65.4.681

Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., … Kasneci, G. (2023). *ChatGPT for good? On opportunities and challenges of large language models for education*. *Learning and Individual Differences, 103*, 102274. https://doi.org/10.1016/j.lindif.2023.102274

Kokotsaki, D., Menzies, V., & Wiggins, A. (2016). Project-based learning: A review of the literature. *Improving Schools, 19*(3), 267–277. https://doi.org/10.1177/1365480216659733

Krajcik, J. S., & Shin, N. (2014). Project-based learning. In R. K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (2nd ed., pp. 275–297). Cambridge University Press.

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge University Press.

Long, D., & Magerko, B. (2020). What is AI literacy? Competencies and design considerations. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1–16). Association for Computing Machinery. https://doi.org/10.1145/3313831.3376727

Luo, Z., Marnburg, E., Øgaard, T., & Okumus, F. (2021). Exploring antecedents of social loafing in students' group work: A mixed-methods approach. *Journal of Hospitality, Leisure, Sport & Tourism Education, 28*, 100314. https://doi.org/10.1016/j.jhlste.2021.100314

Mercier, E., Goldstein, M. H., Baligar, P., & Rajarathinam, R. J. (2023). Collaborative learning in engineering education. In A. Johri (Ed.), *International handbook of engineering education research* (pp. 402–432). Routledge. https://doi.org/10.4324/9781003287483-23

Neo, M., & Neo, T. K. (2009). Engaging students in multimedia-mediated constructivist learning: Students' perceptions. *Journal of Educational Technology & Society, 12*(2), 254–266.

Ng, D. T. K., Leung, J. K. L., Chu, S. K. W., & Qiao, M. S. (2021). Conceptualizing AI literacy: An exploratory review. *Computers and Education: Artificial Intelligence, 2*, 100041. https://doi.org/10.1016/j.caeai.2021.100041

Nguyen, P. B. T., & Nguyen, T. N. P. (2025). *Ethical and responsible AI in educational use of ChatGPT. TPM – Testing, Psychometrics, Methodology in Applied Psychology, 32*(S7), 1322–1335. https://tpmap.org/submission/index.php/tpm/article/view/2351

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). McGraw-Hill.

Risko, E. F., & Gilbert, S. J. (2016). Cognitive offloading. *Trends in Cognitive Sciences, 20*(9), 676–688. https://doi.org/10.1016/j.tics.2016.07.002

Shyu, W. E., & Tasir, Z. (2026). Generative AI in project-based 3D modeling: Effects on creativity, critical thinking, and workflow. *Online Learning, 30*(1), 531–555. https://doi.org/10.24059/olj.v30i1.4769

Sparrow, B., Liu, J., & Wegner, D. M. (2011). Google effects on memory: Cognitive consequences of having information at our fingertips. *Science, 333*(6043), 776–778. https://doi.org/10.1126/science.1207745

Sperber, L., MacArthur, M., Minnillo, S., Stillman, N., & Whithaus, C. (2025). Peer and AI Review + Reflection (PAIRR): A human-centered approach to formative assessment. *Computers & Composition, 76*, 102921. https://doi.org/10.1016/j.compcom.2025.102921

Taber, K. S. (2018). The use of Cronbach's alpha when developing and reporting research instruments in science education. *Research in Science Education, 48*(6), 1273–1296. https://doi.org/10.1007/s11165-016-9602-2

Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education, 2*, 53–55. https://doi.org/10.5116/ijme.4dfb.8dfd

Thomas, J. W. (2000). *A review of research on project-based learning*. The Autodesk Foundation.

Yusuf, A., Pervin, N., & Román-González, M. (2024). Generative AI and the future of higher education: A threat to academic integrity or reformation? Evidence from multicultural perspectives. *International Journal of Educational Technology in Higher Education, 21*(1), 21. https://doi.org/10.1186/s41239-024-00453-6

Zailuddin, M. F. N. O., Nik Harun, N. A., Abdul Rahim, H. A., Kamaruzaman, A. F., Berahim, M. H., Harun, M. H., & Ibrahim, Y. (2024). Redefining creative education: A case study analysis of AI in design courses. *Journal of Research in Innovative Teaching & Learning, 17*(2), 282–296. https://doi.org/10.1108/jrit-01-2024-0019