Pendas : Jurnal Ilmiah Pendidikan Dasar,
ISSN Cetak : 2477-2143 ISSN Online : 2548-6950
Volume 11 Nomor 03, September 2026

DEEP LEARNING ACROSS DISCIPLINES: EMPIRICAL EVIDENCE AND A CROSS-DISCIPLINARY FRAMEWORK FROM INDONESIAN SECONDARY SCHOOLS

Andi Muhammad Irfan Taufan Asfar[1], Ahmad Yani[2], Andi Aviva[3], Astiani Safitri[4], Indri Dwi Meilan[5], Andi Zulqadri[6], Jumardi[7], Asridal[8], Gita Rahman[9], Andi Muhammad Iqbal Akbar Asfar[10]

[1,2]Pendidikan Matematika, Universitas Muhammadiyah Bone
[3]Pendidikan Ekonomi, Universitas Muhammadiyah Bone
[4]Pendidikan Biologi, Universitas Muhammadiyah Bone
[5,6,8,9]Pendidikan Keolahragaan, Universitas Muhammadiyah Bone
[7]Pendidikan Bahasa Inggris, Universitas Muhammadiyah Bone
[10]Teknik Kimia, Politeknik Negeri Ujung Pandang
tauvanlewis00@gmail.com, andiifalasfar@gmail.com

ABSTRACT

Deep learning represents a transformative agenda in twenty-first-century education, emphasizing cognitive engagement, conceptual connections, metacognitive reflection, knowledge transfer, and collaboration. However, evidence from secondary-school classroom practices, particularly across subjects, remains limited. This study examined the implementation of deep learning in Grades 10 and 11 at SMAN 19 Bone and developed a model grounded in findings. An embedded mixed-method design was employed, with a dominant qualitative approach supported by observations. Participants included eight teachers of Mathematics, Indonesian Language, Science, and Social Studies and 64 students. Data were collected through observations, semi-structured interviews, document analysis, and a five-indicator deep learning rubric. Quantitative analyses included descriptive statistics, one-way ANOVA, post hoc tests, and an independent-samples t-test. Results indicated high to very high levels of deep learning. Indonesian Language achieved the highest score, while Social Studies recorded the lowest. Cognitive engagement was the strongest indicator, whereas metacognitive reflection was the weakest. Significant differences were found across subjects and between Grades 10 and 11. Qualitative findings showed that problem-based, inquiry-based, and project-based learning enhanced cognitive engagement, although reflection was not systematically integrated. The study proposed the Deep Learning Cross-Disciplinary Framework, positioning reflection as a moderating variable with implications for curriculum, teacher development, assessment, and instruction.

Keywords: *deep learning, cross-disciplinary learning, metacognitive reflection, cognitive engagement.*

ABSTRAK

Abstrak Pembelajaran mendalam merupakan agenda transformatif pendidikan abad ke-21 yang menekankan keterlibatan kognitif, koneksi konseptual, refleksi metakognitif, transfer pengetahuan, dan kolaborasi. Namun, bukti empiris berbasis praktik kelas nyata di sekolah menengah khususnya lintas mata pelajaran masih terbatas. Penelitian ini menganalisis implementasi pembelajaran mendalam pada kelas X dan XI di SMAN 19 Bone serta merumuskan model konseptual berbasis temuan empiris. Penelitian menggunakan desain embedded mixed-method dengan pendekatan kualitatif dominan yang diperkuat kuantifikasi observasi terstruktur. Partisipan terdiri atas delapan guru Matematika, Bahasa Indonesia, IPA, dan IPS serta 64 siswa. Data dikumpulkan melalui observasi, wawancara semi-terstruktur, analisis dokumen, dan rubrik lima indikator pembelajaran mendalam. Analisis kuantitatif mencakup statistik deskriptif, ANOVA satu arah, uji post hoc, dan independent sample t-test. Hasil menunjukkan skor pembelajaran mendalam berada pada kategori tinggi hingga sangat tinggi. Bahasa Indonesia memperoleh capaian tertinggi, sedangkan IPS terendah. Keterlibatan kognitif menjadi indikator terkuat, sementara refleksi metakognitif merupakan aspek terlemah. Terdapat perbedaan signifikan antar mata pelajaran dan antara kelas X dan XI. Temuan kualitatif menunjukkan bahwa problem-based learning, inquiry learning, dan project-based learning meningkatkan keterlibatan kognitif, tetapi refleksi belum terintegrasi secara sistematis. Penelitian merumuskan Deep Learning Cross Disciplinary Framework yang menempatkan refleksi sebagai variabel moderator yang berimplikasi pada kurikulum, guru, asesmen, dan pembelajaran.

Kata Kunci: pembelajaran mendalam, lintas mata Pelajaran, refleksi metakognitif, keterlibatan kognitif.

A. Introduction

The transformation of twenty-first-century education demands a fundamental paradigm shift from knowledge transmission toward meaning construction, higher-order thinking, knowledge transfer, collaboration, and reflective practice. Educational systems worldwide are increasingly challenged to move beyond surface-level outcomes and cultivate learners capable of understanding, interpreting, evaluating, and applying knowledge across diverse and novel contexts (Zaman & Akhter, 2023). This imperative has intensified in recent years as converging global evidence points to a persistent gap between access to schooling and the actual quality of learning achieved.

The UNESCO Global Education Monitoring Report (2024) documents that the global secondary education completion rate rose from 53% in 2015 to 59% in 2023, yet approximately 650 million children and adolescents still leave school without attaining secondary credentials. This statistic underscores that the central challenge of contemporary education lies not merely in expanding access, but in ensuring that learning is substantive, transferable, and genuinely transformative (UNESCO, 2024). The OECD's PISA 2022 results reinforce this concern: Indonesian students performed significantly below the OECD average in mathematics, reading, and science with only 18%, 25%, and 34% of students reaching Level 2 or above in these three domains, respectively, compared to OECD averages of 69%, 74%, and 76% (OECD, 2023). These figures suggest a systemic orientation toward surface learning, characterized by rote memorization, procedural drill, and recall-based assessment, rather than the deep comprehension and cognitive flexibility demanded by contemporary life and work.

Within this context, *deep learning* has emerged as a pedagogically and theoretically compelling response. Rooted in the seminal distinction between deep and surface approaches to learning proposed by Baihaqi et al. (2025), deep learning refers to learners' intentional engagement with content to construct meaning, establish conceptual connections, reflect on thinking processes, and transfer understanding to new situations. Contemporary frameworks have elaborated this conception significantly. Adeoye, Baharun & Munawwaroh (2025) articulate deep learning as the development of six global competencies character, citizenship, collaboration, communication, creativity, and critical thinking positioned as essential outcomes of transformative education. Nafi'ah & Faruq (2025) emphasizes that deep approaches to learning are characterized by an intention to understand, active integration of new and prior knowledge, and a reflective orientation toward the learning process.

Recent empirical work has further refined understanding of the conditions that enable or constrain deep learning. Fawzia and Karim (2024) demonstrate that deep learning is meaningfully shaped by learning environment design, instructional content, and, critically, the nature of assessment practices both formative and summative. Their cross-disciplinary engineering education analysis reveals that assessment alignment with deep learning goals constitutes one of the most powerful levers available to educators. Paleenud (2024) extends this analysis by distinguishing the differential contributions of lecture-based and project-based instructional environments to deep and surface learning approaches, emphasizing that the *quality* of pedagogical design not merely the type of activity determines learning depth. Complementarily, Nieminen and Pesonen (2022) demonstrate that assessment design exerts significant influence over student learning approaches, with authentic assessment tasks more reliably eliciting deep engagement than conventional testing formats.

The challenge of achieving deep learning is further complicated by the disciplinary fragmentation characteristic of secondary schooling. In most secondary school systems, mathematics, language arts, natural sciences, and social studies are taught as discrete, bounded subjects with limited intentional connections drawn between them. This organizational structure risks producing compartmentalized knowledge that students cannot integrate or apply holistically (Neville et al., 2023). Yet real world problems the very contexts in which twenty-first-century competencies must be exercised are inherently cross-disciplinary, demanding the synthesis of quantitative reasoning, communicative skill, scientific understanding, and social analysis. Cross-disciplinary learning, understood as intentional integration of knowledge, methods, and perspectives from multiple disciplines, has been proposed as a structural response to this fragmentation (Crichton, Crichton & Colville, 2022).

Research on cross-disciplinary learning at the secondary level remains, however, comparatively underdeveloped. Ellingsen et al. (2021) examined problem-based projects in cross-disciplinary student teams and found that cross field experiences can promote students' reflection on their own learning engagement and deepen conceptual understanding. Soini et al. (2022) similarly demonstrated that cross disciplinary project-based approaches foster knowledge integration and collaborative competencies, though they also identify significant implementation challenges related to teacher coordination and curriculum structure. These findings point to the potential of cross-disciplinary deep learning while simultaneously revealing its practical complexity.

Project-based learning (PjBL) has received particular attention as an instructional strategy capable of supporting deep learning across disciplines. Systematic reviews have consistently identified PjBL as an effective model for enhancing learning quality, twenty-first-century skills, and content integration (Fitriyani et al., 2026). However, as Al-Shawabkeh et al. (2026) caution, PjBL does not automatically produce deep learning outcomes; it requires deliberate scaffolding, iterative reflection, formative feedback, and explicit attention to knowledge transfer to achieve its potential. This nuance that active learning methods are necessary but not sufficient conditions for deep learning represents a critical theoretical gap that the present study addresses empirically.

Despite the growth of deep learning scholarship, several significant research gaps persist. First, the majority of deep learning research has been conducted in higher education or professional training contexts, with comparatively little rigorous empirical investigation at the secondary school level (Prasittisopin, 2025). Second, most investigations focus on a single discipline, leaving underexplored the question of how deep learning manifests and varies across subjects within the same institutional context. Third, while reflection has been theorized as a constitutive element of deep learning (Samtani et al., 2023), its functional role as a potential moderating variable one that amplifies or attenuates the relationship between active instructional strategies and deep learning outcomes has not been empirically examined. Fourth, in the Indonesian context specifically, research on deep learning has been predominantly conceptual and normative, with limited classroom-level empirical evidence from authentic teaching practice.

The present study addresses these gaps by making four distinct contributions. First, it investigates deep learning empirically across multiple subjects within a single secondary school institution, enabling direct cross-disciplinary comparison.

Second, it employs an embedded mixed-method design that integrates qualitative richness of classroom observation with quantitative precision of indicator scoring, producing a methodologically robust evidence base. Third, it advances the theoretical position that metacognitive reflection functions not merely as a learning outcome but as a *moderating variable* that determines whether active instructional engagement translates into genuine deep learning. Fourth, it proposes the Deep Learning Cross Disciplinary Framework (DL-CDF), an empirically derived conceptual model that explicates the structural relationships among pedagogical input, mediating processes, reflective moderation, and deep learning output in cross-disciplinary secondary education.

The following research questions guide this investigation:

1. How is deep learning implemented in cross-disciplinary classroom practices among Grade X and XI students at SMAN 19 Bone?
2. What is the level of achievement across the five deep learning indicators cognitive engagement, conceptual connection, reflection, knowledge transfer, and collaboration?
3. Are there significant differences in deep learning scores across subjects and between grade levels?
4. What conceptual model can adequately explain the implementation of cross-disciplinary deep learning at the secondary school level?

B. Method

Research Design

This study employed an embedded mixed-method design (Creswell, 2024), in which a dominant qualitative strand was enhanced by quantitative observational scoring. This design was selected because understanding the implementation of deep learning requires both interpretive richness to capture instructional dynamics, teacher reasoning, and student experience and numerical precision to measure and compare indicator achievement across subjects and grade levels. The qualitative strand provided depth of explanation; the quantitative strand provided comparative evidence and effect size estimation. Methodological integration occurred at the interpretation stage, where qualitative themes were used to explain patterns in quantitative findings.

Research Setting and Participants

The study was conducted at SMAN 19 Bone, South Sulawesi, Indonesia, a state secondary school with heterogeneous student demographics. The school was selected purposively on the basis of its recent engagement with active learning reforms aligned with the *Merdeka Belajar* curriculum framework, making it an appropriate site for examining the naturalistic implementation of deep learning practices.

Participants consisted of eight teachers across four subject areas and 64 students from Grade X and Grade XI. Teacher participants were selected purposively to represent the four core disciplinary areas: Mathematics (n = 2), Bahasa Indonesia (n = 2), Natural Sciences (n = 2), and Social Studies (n = 2), with each pair teaching both Grade X and Grade XI classes. Student participants (32 from Grade X; 32 from Grade XI) were selected based on their enrollment in the observed classes.

Participant Group	n	Description
Mathematics teachers	2	Teaching Grade X and XI
Bahasa Indonesia teachers	2	Teaching Grade X and XI
Natural Sciences teachers	2	Teaching Grade X and XI
Social Studies teachers	2	Teaching Grade X and XI
Grade X students	32	Enrolled in observed classes
Grade XI students	32	Enrolled in observed classes

Data Collection

Four complementary data collection techniques were employed.

Classroom observation was the primary data source. Observers used a structured deep learning rubric to score each of the five indicators while simultaneously recording field notes documenting instructional strategies, student behavior, teacher-student interaction patterns, collaborative activities, and the presence or absence of reflective practice. Each class was observed across multiple sessions to ensure representativeness.

Semi-structured interviews were conducted with all eight teachers and a purposively selected subsample of 16 students (four per subject, two per grade level). Teacher interviews explored lesson planning rationale, strategy selection, assessment practices, interdisciplinary integration intentions, and perceived barriers to deep learning. Student interviews probed learning experience, conceptual understanding, collaborative engagement, and reflective awareness.

Learning document analysis examined lesson plans (*modul ajar*), student worksheets (LKPD), assessment instruments, project rubrics, and student work samples. Documents were analyzed for evidence of cognitive demand, conceptual integration, and reflective task design.

Structured observational scoring operationalized deep learning achievement through a validated five-indicator rubric, enabling quantitative comparison across subjects and grade levels.

Instrumentation: Deep Learning Observation Rubric

The observation rubric assessed five deep learning indicators using a four-point Likert-type scale, yielding a total score range of 5–20. Indicator descriptors were developed from established deep learning frameworks (Arya et al., 2026; Carosia, Silva & Coelho, 2025) and adapted to the secondary classroom context.

Indicator	Score 1	Score 2	Score 3	Score 4
Cognitive engagement	Passive, following instructions	Answering lower-order questions	Analyzing problems	Evaluating and synthesizing ideas
Conceptual connection	No connections made	Limited cross-concept links	Connecting several concepts	Integrating concepts across disciplines
Reflection	No reflective activity	Descriptive reflection	Analytical reflection	Critical and metacognitive reflection
Knowledge transfer	No concept application	Applying to similar examples	Applying to new contexts	Applying flexibly and creatively
Collaboration	Passive in group	Limited participation	Active contribution	Leading discussion and co-constructing ideas

Score categories were established as follows:

Score Range	Percentage	Category
17–20	85%–100%	Very High
13–16	65%–84%	High
9–12	45%–64%	Moderate
5–8	25%–44%	Low

Validity and Reliability

Content validity was established through expert review by two educational psychology scholars and two senior teachers, who assessed indicator alignment, descriptor clarity, score measurability, and theoretical grounding in deep learning literature. Revisions were made iteratively based on reviewer feedback.

Internal consistency reliability was calculated using Cronbach's alpha ($\alpha = 0.86$), indicating good internal consistency. Inter-rater reliability was assessed through independent parallel scoring sessions, yielding an inter-rater agreement coefficient of 0.81, which exceeds the 0.75 threshold recommended for observational research (Creswell, 2024). These psychometric properties indicate that the rubric produced stable and internally consistent measurements.

Data Analysis

Qualitative analysis followed the interactive model of Miles, Huberman, and Saldaña (2014), comprising three iterative stages: data condensation (coding and thematic categorization), data display (thematic matrices and narrative accounts), and conclusion drawing and verification. Codes were organized into eight primary thematic categories: instructional strategy, cognitive engagement, conceptual integration, reflection, knowledge transfer, collaboration, assessment practice, and implementation constraints.

Quantitative analysis employed descriptive statistics (mean, standard deviation, percentage achievement), one-way ANOVA to test for significant differences across subjects, Tukey's HSD post hoc test to identify pairwise differences, and independent sample t-test to compare Grade X and Grade XI. The percentage achievement formula was:

$$P = \frac{\text{Score obtained}}{\text{Maximum score}} \times 100\%$$

Effect sizes were reported as η^2 (ANOVA) and Cohen's *d* (t-test). Statistical significance was set at $\alpha = 0.05$. All quantitative analyses were conducted using SPSS.

Research Ethics

This study adhered to established ethical principles for educational research. Institutional approval was obtained from SMAN 19 Bone. All participants provided informed consent, and participation was voluntary with guaranteed right of withdrawal. Participant identities are anonymized in all reporting. Data were used exclusively for academic purposes and stored securely.

C. Results and Discussion

Overview of Deep Learning Implementation

Classroom observations revealed that deep learning practices were emerging across all four subject areas, primarily through three instructional strategies: problem-based learning (PBL), inquiry learning, and project-based learning (PjBL). In Mathematics, teachers presented contextual problems requiring data analysis and multi-step reasoning, engaging students in argument construction and solution justification. In Bahasa Indonesia, structured discussion, text interpretation, argumentative writing, and collaborative report production provided sustained opportunities for discursive engagement. In Natural Sciences, students engaged in structured observation, hypothesis formulation, simple experimentation, and data interpretation. In Social Studies, students analyzed social phenomena and drew connections to contemporary community contexts.

These patterns indicate a meaningful shift away from transmission-oriented instruction toward exploratory, inquiry driven pedagogy. Nonetheless, learning depth was not uniform across subjects, and the integration of explicit metacognitive reflection at the close of learning activities was inconsistent a finding that proved central to interpreting indicator-level and cross subject differences.

Deep Learning Scores by Subject

Table 1 presents mean deep learning scores, standard deviations, percentage achievement, and achievement categories by subject.

Table 1 *Descriptive Statistics of Deep Learning Scores by Subject*

Subject	*n*	*M*	*SD*	Achievement (%)	Category
Mathematics	16	15.80	2.10	79.00	High
Bahasa Indonesia	16	16.50	1.95	82.50	High
Natural Sciences	16	14.70	2.30	73.50	High
Social Studies	16	13.90	2.55	69.50	High

All four subjects achieved scores in the High category. Bahasa Indonesia obtained the highest mean score, reflecting its structural openness to interpretive, argumentative, and reflective activities forms of engagement that closely align with deep learning's conceptual requirements. Social Studies achieved the lowest score, not due to inherent content limitations, but because its instructional design relied more heavily on explanation-confirmation patterns, with fewer opportunities for data-based inquiry, critical reflection, and authentic project integration.

These findings are consistent with the discipline-sensitivity of deep learning observed in broader literature. Adipat et al. (2023) demonstrate that disciplinary affordances the inherent opportunities within subject-matter structures for particular forms of engagement significantly influence the depth of learning achievable within that subject. Subjects that structurally invite dialogue, interpretation, and reflective articulation tend to yield higher deep learning scores than those oriented primarily toward content coverage and procedural consolidation (Rapanta & Felton, 2022).

Achievement by Deep Learning Indicator

Table 2 *Mean Scores and Percentage Achievement per Deep Learning Indicator*

Indicator	*M*	Achievement (%)	Category
Cognitive engagement	3.40	85.00	Very High
Collaboration	3.30	82.50	High
Conceptual connection	3.20	80.00	High
Knowledge transfer	3.00	75.00	High
Reflection	2.90	72.50	High

Cognitive engagement achieved the highest score, indicating that students were regularly engaged in analytical tasks comparing information, constructing arguments, evaluating solutions, and making judgments. This finding aligns with growing evidence that active instructional strategies reliably promote higher-order cognitive engagement when implemented with adequate task complexity (Gradini, Noviani & Ulya, 2025).

Collaboration also scored strongly, reflecting the prevalence of group discussion, cooperative projects, and peer-mediated learning activities across all four subjects. This pattern is consistent with findings by yadav & Singh (2026), who demonstrate that collaborative structures, when organized around substantive learning tasks, produce both social and cognitive benefits aligned with deep learning goals.

Reflection constituted the lowest-scoring indicator (72.50%), revealing a critical structural gap in the learning process. Despite high cognitive activity and social engagement, students were not systematically guided toward examining their own thinking, identifying conceptual difficulties, evaluating their reasoning strategies, or articulating how their understanding had changed. Ambaryani & Putranta (2022) emphasize, metacognitive reflection the capacity to monitor and regulate one's own cognitive processes is not a spontaneous by-product of activity-based learning; it requires deliberate pedagogical scaffolding, including explicitly designed reflective prompts, structured self-assessment tasks, and reflective journaling. The absence of such scaffolding explains why active learning activity did not fully translate into metacognitive depth.

Knowledge transfer scored above average (75.00%) but remained below the levels achieved for cognitive engagement and collaboration. This suggests that while students could analyze problems and work collaboratively, the explicit connection of disciplinary knowledge to novel, cross-contextual situations required stronger pedagogical scaffolding than was consistently provided.

One-Way ANOVA: Differences Across Subjects

Table 3 *One-Way ANOVA Results for Deep Learning Scores Across Subjects*

Source of Variation	*SS*	*df*	*MS*	*F*	*p*	η²
Between subjects	63.82	3	21.27	4.27	0.018	0.176
Within groups	298.80	60	4.98			
Total	362.62	63				

One-way ANOVA revealed a statistically significant difference in deep learning scores across subjects, $F(3, 60) = 4.27$, $p = 0.018$, $\eta^2 = 0.176$. The effect size of $\eta^2 = 0.176$ indicates a medium effect (Cohen, 1988), signifying that disciplinary context accounts for approximately 17.6% of the variance in deep learning scores a practically meaningful proportion.

Tukey's HSD post hoc comparisons identified the most significant pairwise differences:

- Bahasa Indonesia vs. Social Studies: $MD = 2.60$, $p = 0.014$
- Mathematics vs. Social Studies: $MD = 1.90$, $p = 0.048$
- Bahasa Indonesia vs. Natural Sciences: $MD = 1.80$, $p = 0.067$ (approaching significance)

These results demonstrate that the gap between subjects is not trivial. The underperformance of Social Studies relative to Bahasa Indonesia and Mathematics underscores the need for targeted instructional redesign, particularly emphasizing problem-based inquiry using social data, reflective assessment instruments, and cross-disciplinary project integration. This finding extends Paleenud, Tanprasert & Waleeittipat (2024) work by showing that instructional approach differences across disciplines not merely between project-based and lecture-based formats produce measurable variation in deep learning achievement.

Independent Sample t-Test: Grade X versus Grade XI

Table 4 *Independent Sample t-Test: Deep Learning Scores by Grade Level*

Grade	*n*	*M*	*SD*	*t*	*p*	Cohen's *d*
Grade X	32	14.80	2.40	2.42	0.021	0.61
Grade XI	32	16.20	2.20			

The independent sample t-test indicated a statistically significant difference between Grade X and Grade XI, $t\,(62) = 2.42$, $p = 0.021$, with a medium effect size of Cohen's $d = 0.61$. Grade XI students achieved substantially higher deep learning scores, indicating that deep learning capacity develops progressively with accumulated learning experience.

This finding provides important longitudinal evidence: Grade XI students demonstrated more sophisticated argumentative skills, more mature conceptual connections, and greater reflective orientation than their Grade X counterparts. This progression aligns with Budhiarti, Mytra & Slow (2025) developmental account of deep learning, in which prior knowledge structures, metacognitive awareness, and dispositional orientation toward understanding accumulate over time through sustained exposure to appropriately designed learning environments. It also resonates with recent evidence from Schneider and Preckel (2023), who demonstrate that student approaches to learning deep versus surface are partly shaped by cumulative educational experience and the sustained quality of instructional environments.

The significant grade level difference reinforces the argument that deep learning cannot be treated as an outcome achievable within isolated instructional episodes; it requires longitudinal curriculum design, consistent pedagogical commitment across subject areas, and developmental scaffolding that progressively deepens students' metacognitive capacity.

Qualitative Findings

Instructional Strategies

Three dominant instructional strategies were identified across the four subject areas.

Problem-based learning was observed most prominently in Mathematics and Social Studies. In Mathematics, teachers presented contextual problems requiring multi-step analysis and evidence-based justification. In Social Studies, teachers framed discussions around social phenomena with community relevance. However, PBL in Social Studies was seldom accompanied by structured reflection rubrics or authentic assessment, limiting its contribution to deep learning beyond the cognitive engagement phase.

Inquiry learning was most prevalent in Natural Sciences, where students formulated hypotheses, conducted structured observations, and interpreted data. This approach effectively stimulated cognitive engagement and conceptual connection but rarely extended to explicit cross-disciplinary transfer for instance, connecting scientific findings to social or environmental implications which would have enhanced the depth of learning further.

Project-based learning was most developed in Bahasa Indonesia and, partially, in Social Studies. In Bahasa Indonesia, students produced reports, conducted presentations, and connected their outputs to contextual issues. PjBL provided robust structures for collaboration and communication but, as observed, the quality of reflection generated was highly dependent on the nature and specificity of the guiding questions provided by the teacher. This mirrors the caution raised that PjBL without deliberate reflective scaffolding risks reducing to product-oriented activity rather than sustained conceptual inquiry.

Cross-Disciplinary Integration

Cross-disciplinary connections emerged in three observable patterns. Mathematics supported data analysis practices within Natural Sciences and Social Studies. Bahasa Indonesia supported scientific writing, argumentative structuring, and communicative presentation across subjects. Natural Sciences and Social Studies provided factual and contextual grounding for understanding environmental and social phenomena encountered in other disciplines.

Despite these emergent connections, cross disciplinary integration remained largely spontaneous and incidental rather than systematically planned. Teachers had not yet developed coordinated cross-subject project designs at the curriculum planning stage. This represents a significant implementation gap that echoes the findings of Christiansen et al., (2023), who identify teacher coordination, shared planning time, and institutional policy support as critical preconditions for effective cross-disciplinary pedagogical implementation. Without structural mechanisms such as interdisciplinary project weeks, co-planning frameworks, or thematic curriculum integration cross-disciplinary learning remains dependent on individual teacher initiative rather than becoming an institutionalized feature of the learning experience.

Reflection as the Central Weakness

Reflection emerged as the most critical and consistently underdeveloped dimension of deep learning practice. While several teachers invited students to state conclusions at the end of lessons, systematic metacognitive reflection structured around questions such as "What thinking strategies did you use?", "What was most challenging and why?", "How could this concept apply in a different context?", or "How has your understanding changed through this activity?" was rarely observed.

This finding confirms that cognitive activity and social engagement, while necessary, are insufficient to produce deep learning without deliberate reflective mediation. Zarepour et al. (2024) demonstrate, metacognitive reflection must be explicitly scaffolded through structured prompts, regular reflective journaling, and formative feedback cycles not assumed to emerge spontaneously from active learning participation. The low reflection scores despite high cognitive engagement and collaboration scores constitute the central empirical paradox of this study and the primary theoretical motivation for the DL-CDF model.

Discussion

The composite findings of this study reveal a clear and theoretically significant pattern: deep learning implementation at SMAN 19 Bone has achieved a meaningfully high level, particularly in cognitive engagement and collaboration, but has not reached its optimal potential because metacognitive reflection remains structurally underdeveloped. This pattern is not incidental it reflects a systemic gap in pedagogical design that active learning methods alone cannot bridge.

The finding that cognitive engagement achieved the highest indicator score while reflection achieved the lowest provides strong empirical support for a theoretically important distinction: *being active* in learning is not equivalent to *learning deeply*. Deep learning, as theorized by Marton and Säljö (1976) and elaborated by Entwistle (2018), is constituted not merely by engagement with tasks but by an intentional orientation toward constructing meaning, which requires the metacognitive dimension of reflecting on, evaluating, and reorganizing one's understanding. Without systematic reflective practice, high cognitive engagement risks remaining at the level of elaborate surface processing sophisticated in form but limited in conceptual depth.

The significant between-subject differences revealed by ANOVA demonstrate that deep learning is *discipline-sensitive*: the affordances and conventions of different subject areas create differential conditions for deep learning to emerge and be expressed. Bahasa Indonesia's structural openness to interpretive dialogue, argumentative writing, and reflective articulation of meaning created conditions closely aligned with deep learning's requirements. Social Studies' relative reliance on explanation and confirmation patterns produced lower deep learning achievement, not because the subject's content is less conducive to depth indeed, social analysis is inherently interpretive but because its pedagogical operationalization did not yet fully exploit this potential.

These results extend the findings of Fawzia and Karim (2024) from an engineering education context to a secondary school, multi-subject Indonesian context, confirming that assessment design and learning environment quality are cross-disciplinarily significant determinants of deep learning. They also resonate with Soini et al. (2022), whose investigation of cross-disciplinary project-based learning identifies curriculum structure and teacher collaboration as key enabling conditions conditions that this study reveals to be currently underdeveloped in the research setting.

The significant grade-level difference (d = 0.61) carries important longitudinal implications. Grade XI students' consistently higher deep learning scores suggest that the capacity for deep learning is developmental, accumulating with experience, prior knowledge consolidation, and repeated exposure to appropriately demanding instructional environments. This finding provides empirical grounding for the recommendation that deep learning be designed as a longitudinal curricular trajectory rather than a collection of isolated instructional events.

The qualitative findings regarding cross-disciplinary integration reinforce this developmental argument: where cross-disciplinary connections emerged, they appeared to deepen conceptual understanding and enrich contextual relevance. However, their incidental character limited their pedagogical impact. Systematic cross-disciplinary curriculum design with shared planning, co-constructed projects, and intentional thematic integration remains an institutional imperative identified but not yet achieved in the current setting.

The Deep Learning Cross-Disciplinary Framework (DL-CDF)

Based on the convergent evidence of qualitative and quantitative findings, this study proposes the Deep Learning Cross-Disciplinary Framework (DL-CDF) as a conceptual model explicating the structural conditions and processes through which deep learning is achieved in cross-disciplinary secondary education contexts.

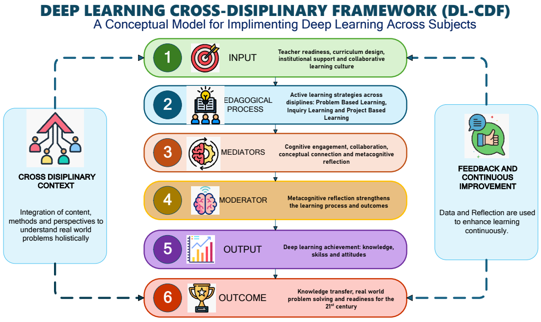

Figure 1. The Deep Learning Cross-Disciplinary Framework (DL-CDF).

The DL-CDF comprises five structural layers:

<u>Input Layer</u> identifies the preconditions that shape pedagogical possibility: teacher professional readiness, curriculum design quality, institutional support structures, and the school's prevailing learning culture. These preconditions determine the range of pedagogical strategies available and the sustainability of deep learning implementation.

<u>Pedagogical Process Layer</u> captures the three dominant instructional approaches observed: PBL, inquiry learning, and PjBL. These constitute the active

instructional mechanisms through which deep learning is initiated but are insufficient on their own to guarantee depth.

Mediator Layer identifies the three process variables cognitive engagement, collaboration, and conceptual connection that mediate the relationship between instructional strategy and deep learning output. High-quality instructional strategies are effective to the extent that they successfully activate these mediating processes.

Moderator Layer positions metacognitive reflection as the critical moderating variable. This is the most theoretically novel contribution of the model. Reflection does not merely add to deep learning output; it determines the *efficiency* with which active engagement translates into depth of understanding. High reflection amplifies the effect of strong mediation; low reflection attenuates it, explaining why high cognitive engagement and collaboration do not automatically produce high overall deep learning scores when reflection is absent.

Output and Outcome Layers distinguish between the observable learning achievement score (output) and the broader educational goals it serves (outcome): critical thinking, knowledge transfer, problem-solving capacity, and twenty-first-century competency development, all aligned with the targets of SDG 4 on quality education.

The DL-CDF advances the theoretical understanding of deep learning in two principal ways. First, it provides a structural account of deep learning as a *layered process* rather than a binary state, identifying the specific mechanisms through which instructional quality is translated or fails to be translated into learning depth. Second, by positioning reflection as a *moderator* rather than merely an outcome or add-on, it explains empirical patterns that simpler models cannot account for: why high engagement without reflection fails to produce commensurate learning depth.

Policy Implications

The findings and the DL-CDF model carry substantive implications for educational policy and practice at multiple levels.

At the instructional level, metacognitive reflection must be deliberately and systematically embedded in every learning session not treated as optional or appended to the end of lessons. Structured reflective tools such as learning journals, metacognitive prompt cards, exit tickets anchored to conceptual understanding, and structured self-assessment rubrics should be institutionalized as standard instructional components. As Jiang et al. (2022) and Nieminen and Pesonen (2022) demonstrate, reflection-oriented pedagogical design reliably enhances deep learning quality.

At the curriculum level, cross-disciplinary integration must transition from incidental to intentional. Schools should develop thematic cross-disciplinary project frameworks that formalize collaboration between Mathematics, Bahasa Indonesia, Natural Sciences, and Social Studies teachers. Shared planning time, co-designed learning modules, and interdisciplinary assessment rubrics are structural requirements for this transition (Neville et al., 2023; Soini et al., 2022).

At the assessment level, evaluation systems must be reformed to move beyond answer-correctness toward process-oriented assessment. Portfolio assessment, reflective journals, project-based rubrics, peer assessment, and self-evaluation instruments must be formally incorporated into grading and reporting frameworks (Nieminen & Pesonen, 2022; Fawzia & Karim, 2024).

At the professional development level, teacher capacity building must be reoriented toward *deep pedagogical competency* specifically, teachers' ability to design learning experiences that scaffold metacognitive reflection, facilitate cross-disciplinary integration, and use formative assessment data to deepen instructional responsiveness. Superficial training on activity methods is insufficient; sustained professional learning communities focused on instructional quality are required.

These implications align directly with SDG 4 Target 4.1 (ensuring that all youth achieve relevant and effective learning outcomes), Target 4.4 (substantially increasing youth with relevant skills), and the broader global agenda for quality education that is genuinely transformative rather than merely credential-conferring.

D. Conclusion

This study provides empirical evidence that deep learning implementation in cross-disciplinary classroom practice at SMAN 19 Bone has achieved a high level, with cognitive engagement and collaboration as the strongest indicators and metacognitive reflection as the most critical structural gap. Statistically significant differences across subjects (F (3, 60) = 4.27, p = 0.018, η^2 = 0.176) and between grade levels (t (62) = 2.42, p = 0.021, d = 0.61) demonstrate that deep learning is both discipline-sensitive and developmentally progressive.

The principal theoretical contribution of this study is the Deep Learning Cross-Disciplinary Framework (DL-CDF), which positions metacognitive reflection as a *moderating variable* rather than merely a learning outcome. This repositioning resolves a theoretically important empirical paradox high engagement without commensurate depth and offers a structurally explicit account of how pedagogical strategies, mediating processes, and reflective moderation jointly determine deep learning achievement.

Practically, the study recommends the systematic integration of reflective scaffolding into instructional design, the institutionalization of cross-disciplinary curriculum planning, the reform of assessment systems toward process-oriented evaluation, and the reorientation of teacher professional development toward deep pedagogical competency.

REFERENCES

Adeoye, M. A., Baharun, H., & Munawwaroh, I. (2025). Transformational leadership in education: harmonising accountability, innovation and global citizenship. Kharisma: Jurnal Administrasi dan Manajemen Pendidikan, 4(1), 14-30. https://doi.org/10.59373/kharisma.v4i1.68

Adipat, S., Chotikapanich, R., Laksana, K., Busayanon, K., Piatanom, P., Ausawasowan, A., & Elbasouni, I. (2023). Technological pedagogical content knowledge for professional teacher development. *Academic Journal of Interdisciplinary Studies*, *12*(1), 173-182. https://doi.org/10.36941/ajis-2023-0015

Al-Shawabkeh, H., Al-Khouri, B., Hamdan, S., & Mawaddah, I. A. (2026). Project-Based Learning

And 21st-Century Skills: Developing Critical Thinking And Collaboration In Primary Classrooms. *International Journal of Educatio Elementaria and Psychologia*, *3*(1), 1-13. https://doi.org/10.70177/ijeep.v3i1.3383

Ambaryani, S. E., & Putranta, H. (2022). Improving learners' metacognitive skills with self-regulated learning based problem-solving. *International Journal of Instruction*, *15*(2), 139-154. https://doi.org/10.29333/iji.2022.1528a

Arya, V., Saraf, A., Chichkanov, N., Papa, A., & Romano, M. (2026). AI-enhanced competency transfer hubs: a conceptual framework for university-industry engagement and knowledge sharing. *The Journal of Technology Transfer*, *51*(2), 682-712. https://doi.org/10.1007/s10961-025-10233-7

Baihaqi, M. A., Prayitno, H. J., Ishartono, N., Hamranani, S. S. T., Putra, C. A., & Nizaar, M. (2025). Reimagining 21st-Century Higher Education: A Literature-Based Exploration of Deep Learning for Cognitive, Social, and Emotional Skill Development. *Journal of Deep Learning*, 139-168. https://doi.org/10.23917/jdl.v1i2.11600

Budhiarti, Y., Mytra, P., & Slow, L. (2025). The role of deep learning in elementary education: Pedagogical insights from a literature study. *Jurnal Pedagogi Dan Inovasi Pendidikan*, *1*(2).

Carosia, A. E. D. O., da Silva, A. E. A., & Coelho, G. P. (2025). Predicting the Brazilian stock market with sentiment analysis, technical indicators and stock prices: A deep learning approach. *Computational Economics*, *65*(4), 2351-2378. https://doi.org/10.1007/s10614-024-10636-y

Christiansen, S. H., Juebei, C., & Xiangyun, D. (2023). Cross-institutional collaboration in engineering education–a systematic review study. *European Journal of Engineering Education*, *48*(6), 1102-1129. https://doi.org/10.1080/03043797.2023.2228727

Creswell, J. W. (2024). My 35 years in mixed methods research. *Journal of Mixed Methods Research*, *18*(3), 203-215. https://doi.org/10.1177/15586898241253892

Crichton, M., Crichton, H., & Colville, G. (2022). Students' Perceptions of Problem-Based Learning in Multidisciplinary Groups When Seeking to Solve an Engineering Grand Challenge. *Journal of problem based learning in higher education*, *10*(1), 20-35.

Ellingsen, P., Tonholm, T., Johansen, F. R., & Andersson, G. (2021). Learning from problem-based projects in cross-disciplinary student teams. *Education Sciences*, *11*(6), 259.

Entwistle, N. (2018). *Student learning and academic understanding: A research perspective with implications for teaching.* Academic Press.

Fawzia, S., & Karim, A. (2024). Exploring the connection between deep learning and learning assessments: a cross-disciplinary engineering education perspective. *Humanities and Social Sciences Communications*, *11*(1), 1-10.

Fitriyani, R., Nurlela, N., Aqil, N., Marba'i, M. I., Amelliyah, R., Khairani, D. R., & Rifai, D. S. (2026). Integration of Innovative Learning Methods in 21st Century Skills Development: A Systematic Review. *Paedagogie*, *21*(1), 545-562.

Gradini, E., Noviani, J., & Ulya, K. (2025). Fostering higher-order thinking skills in mathematics education: strategies, challenges, and classroom practices. *Prisma Sains: Jurnal Pengkajian Ilmu Dan Pembelajaran Matematika Dan IPA IKIP Mataram*, *13*(2), 135-163.

Jiang, Y., Ma, L., & Gao, L. (2022). Assessing teachers' metacognition in teaching: The Teacher Metacognitive Knowledge and Awareness Inventory. *Teaching and Teacher Education*, *116*, 103736. https://doi.org/10.1016/j.tate.2022.103736

Marton, F., & Säljö, R. (1976). on qualitative differences in learning: I. Outcome and process. *British Journal of Educational Psychology*, *46*(1), 4–11. https://doi.org/10.1111/j.2044-8279.1976.tb02980.x

Miles, M. B., Huberman, A. M., & Saldaña, J. (2014). Qualitative data analysis: A methods sourcebook (3rd ed.). SAGE Publications.

Nafi'ah, J., & Faruq, D. J. (2025). Conceptualizing deep learning approach in primary education: Integrating mindful, meaningful, and joyful. *Journal of Educational Research and Practice*, *3*(2), 225-237.

Neville, I. A., Petrass, L. A., & Ben, F. (2023). Cross disciplinary teaching: A pedagogical model to support teachers in the development and implementation of outdoor learning opportunities. *Journal of Outdoor and Environmental Education*, *26*(1), 1-21.

Nieminen, J. H., & Pesonen, H. V. (2022). Politicising inclusive assessment: How do alternative assessment formats redefine disability in higher education? *Teaching in Higher Education*, *27*(6), 776–791

OECD. (2023). *PISA 2022 results (Volume I): The state of learning and equity in education.* OECD Publishing. https://doi.org/10.1787/53f23881-en

Paleenud, I., Tanprasert, K., & Waleeittipat, S. (2024). Lecture-based and project-based approaches to instruction, classroom learning environment, and deep learning. *European Journal of Educational Research*, *13*(2), 531-539. https://doi.org/10.12973/eu-jer.13.2.531

Prasittisopin, L. (2025). Machine learning (ML) and deep learning

(DL) in sustainable concrete construction: review, trend and gap analyses. *Journal of Asian Architecture and Building Engineering*, 1-29. https://doi.org/10.1080/13467581.2025.2574571

Rapanta, C., & Felton, M. K. (2022). Learning to argue through dialogue: A review of instructional approaches. *Educational Psychology Review*, *34*(2), 477-509. https://doi.org/10.1007/s10648-021-09637-2

Samtani, S., Zhu, H., Padmanabhan, B., Chai, Y., Chen, H., & Nunamaker Jr, J. F. (2023). Deep learning for information systems research. *Journal of Management Information Systems*, *40*(1), 271-301. https://doi.org/10.1080/07421222.2023.2172772

Schneider, M., & Preckel, F. (2023). Variables associated with achievement in higher education: A systematic review of meta-analyses. *Psychological Bulletin*, *143*(6), 565–600. https://doi.org/10.1037/bul0000207

Soini, T., Pietarinen, J., Toom, A., & Pyhältö, K. (2022). Student teachers' sense of professional agency in the classroom during teacher education. *Teaching and Teacher Education*, *115*, 103729.

UNESCO. (2024). *Global education monitoring report 2024/5: Leadership in education Lead for learning.* UNESCO Publishing.

Yadav, A. R., & Singh, C. K. (2026). An Exploration of Teaching Strategies That Promote Problem-Solving and Higher-Order Thinking Skills. *International Journal of Engineering Science & Humanities*, *16*(1), 100-109.

Zaman, M. A. U., & Akhter, E. (2023). Adaptive learning systems for English literature classrooms: A review of AI-integrated education platforms. *International Journal of Scientific Interdisciplinary Research*, *4*(3), 56-86. https://doi.org/10.63125/a30ehr12

Zarepour, B., Soleimani, H., & Sanjari, M. (2024). Exploring assessment strategies used to promote metacognitive development in science education. *Assessment and Practice in Educational Sciences*, *2*(2), 1-9. https://doi.org/10.61838/japes.2.2.3