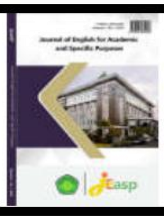
Journal of English for Academic and Specific Purposes (JEASP)
Volume 8 Number 2 December, 2025 (246 – 270)



From Engagement to Argument Quality: How PBL Develops Critical Reading in EFL/ESP

Luthfi Retriansyah[1], Munawir Yusuf[2], Subagya[3]

Corresponding email: luthfiretriansyah@student.uns.ac.id

[123]**Universitas Sebelas Maret, Indonesia**

Article History

Received: Dec 4, 2025
Reviewed: Dec 10, 2025
Accepted: Dec 12, 2025
Published: Dec 25, 2025

Keywords:
project-based learning; critical reading; student engagement; argument quality; academic self-efficacy

Abstract

Critical reading is increasingly essential in EFL/ESP higher education, yet many students still struggle to evaluate claims, evidence, and reasoning in disciplinary texts. Project-Based Learning (PBL) is frequently recommended to strengthen higher-order literacy; however, prior studies often report outcomes without clearly specifying how PBL produces gains in critical reading skills (CRS). This theory-driven systematic conceptual literature review synthesizes research to construct a mechanism-based explanation of the PBL→CRS relationship through engagement and argument quality, while identifying academic self-efficacy (ASE) as a plausible boundary condition. Using a PRISMA-informed Scopus-only search built from five keyword sets (PBL, critical reading, engagement, argumentation, and ASE), journal articles were screened with cluster-specific inclusion criteria and appraised for reporting transparency as a rigor check. Thirty studies were included and analyzed via qualitative content analysis and constant comparison across four evidence clusters (PBL-focused, engagement-focused, CRS/argumentation-focused, and ASE-focused). The synthesis indicates that PBL influences CRS by (a) activating multidimensional engagement (behavioral, cognitive, emotional, and agentic) through authentic inquiry, collaboration, and iterative task cycles, and (b) improving argument quality when projects embed explicit routines for claims–evidence–reasoning, critique, and text-based justification. ASE appears to strengthen these pathways by shaping persistence, strategy use, and willingness to engage in cognitively demanding argument work. The review contributes a provisional, testable program theory for argument-rich PBL in ESP/non-STEM contexts and outlines empirical directions (e.g., longitudinal designs, multilevel SEM, and cluster RCTs) to validate the proposed mechanisms.

INTRODUCTION

In an era of data abundance, university students are increasingly required to read not only more texts but also more complex, multimodal, and often controversial ones. Digital environments expose learners to competing claims, fragmented arguments, and domain-specific discourses that demand critical interrogation rather than passive consumption. In EFL and English for Specific Purposes (ESP) contexts, this challenge is amplified: students

must construct meaning, evaluate evidence, and identify bias in a language that is not their mother tongue, while still struggling with vocabulary, syntax, and background knowledge. Recent empirical work confirms that reading comprehension and critical thinking are strongly intertwined for EFL learners, and that weaknesses in either domain constrain academic success and informed participation in society (Alshehri, 2024; Aloqaili, 2012). Within this landscape, critical reading is no longer a peripheral language skill; it has become a core academic literacy and a gatekeeper to participation in disciplinary communities.

Yet evidence from many EFL settings suggests that students' critical reading skills remain modest. Studies report that students often focus on locating explicit information rather than interrogating authors' assumptions, evaluating argument structure, or connecting texts to wider social and disciplinary issues (Julianti et al., 2024; Megania, 2024). Instruction is still dominated by cognitively oriented comprehension checks and vocabulary exercises, with fewer opportunities to engage with texts as arguments situated in particular social practices. This pattern is problematic because critical literacy perspectives conceptualise reading as a social and ideological activity in which readers position themselves toward texts, negotiate multiple viewpoints, and construct their own stances (Suarcaya, 2017; Nurhayati, 2023). Consequently, there is a growing concern that many EFL/ESP courses prepare students to "answer questions about texts" rather than to use texts as vehicles for disciplined reasoning and argumentation.

At the same time, research on reading development reminds us that higher-order critical reading presupposes—but is not guaranteed by—foundational skills and vocabulary knowledge. Westerveld et al. (2020), in their Reading Success project, demonstrated how a systematic, five-step assessment-to-intervention model based on the Simple View of Reading can accurately profile learners' strengths and weaknesses in decoding, language comprehension, and reading self-concept, and then guide targeted instructional support. From a complementary perspective, Biemiller et al. (2014) showed through direct tests and simulations that vocabulary growth is cumulative and highly sensitive to the quantity and quality of textual input, with richer lexical knowledge enabling more sophisticated inferences and meaning construction. Together, these bodies of work highlight that decoding and vocabulary are necessary conditions for comprehension and critical reading, but they are insufficient to explain how students learn to evaluate claims, weigh evidence, and craft their own arguments from texts (Sinaga et al., 2023). There is a need for pedagogical models that deliberately connect foundational reading skills, cognitive engagement, and argument-based literacy practices.

One promising candidate is Project-Based Learning (PBL), which has become a prominent instructional approach across disciplines for fostering 21st-century competencies such as critical thinking, problem solving, collaboration, and creativity. A large-scale review by Guo et al. (2020) shows that PBL in higher education is associated with gains in content understanding, self-regulated learning, and a range of affective–motivational outcomes, especially when projects are authentic, collaborative, and sustained over time. In language and literacy education, PBL has been used to organise students' work around the production of concrete artefacts—such as project reports, multimodal products, or community-oriented materials—that require integrating reading, writing, and oral communication in meaningful contexts. Several recent studies in EFL contexts also report that PBL can enhance student engagement and creativity in writing, with students perceiving project work as more relevant and motivating than traditional, teacher-centred tasks (Syamsudin et al., 2025).

More specifically related to reading, emerging evidence indicates that project-based or project-supported approaches can benefit critical reading and higher-order thinking. Sari and Prasetyo (2021) found that a project-based learning design in a critical reading course

significantly improved students' critical thinking skills compared to conventional instruction, arguing that extended projects created space for iterative questioning, evaluating, and re-organising textual information. Khulaifiyah et al. (2024) designed PBL-based activities to develop critical reading aids for engineering students and reported enhanced engagement with texts and more strategic use of critical reading activities such as annotating, questioning, and summarising. Other studies correlate students' critical reading and vocabulary mastery with the quality of their argumentative writing, suggesting that reading-based argumentation tasks are fertile ground for integrating language and higher-order thinking (Nurjanah, 2022). However, these studies typically foreground outcomes (e.g., improved test scores, better essays) without fully theorising the mechanisms through which PBL leads to better critical reading or argument quality.

A central but often under-specified construct in this mechanism is student engagement. Contemporary work conceptualises engagement as a multidimensional construct comprising behavioural, cognitive, emotional, and, more recently, agentic components—that is, the extent to which students proactively contribute to the flow of instruction (Fredricks et al., 2004; Wong & Liem, 2022; Reeve, 2013). In PBL settings, engagement is frequently described as a direct effect of authentic, collaborative project work: students participate more actively, invest more effort in planning and monitoring their progress, and report higher interest and enjoyment when they work on real-world problems for real audiences (Chang, 2024; Syamsudin et al., 2025). Yet engagement is not merely a desirable affective by-product; many theorists view it as the primary process variable that mediates the impact of teaching on learning outcomes. From this perspective, PBL can be hypothesised to improve critical reading because it systematically creates conditions for sustained behavioural, cognitive, emotional, and agentic engagement with texts and tasks over time.

Another under-developed piece of the puzzle concerns argumentation. Research on argumentative reading and writing emphasises that learning to read texts as arguments—and to respond with arguments—is both a cognitive and a social practice that must be deliberately scaffolded (Newell et al., 2011). Studies of EFL argumentative writing in Indonesia, for example, reveal that students often struggle to articulate clear claims, provide relevant reasons, and integrate textual evidence, with their essays frequently displaying limited critical thinking patterns (Muhsin et al., 2024; Ilyas, 2025). When critical reading is explicitly linked to argumentation—such as in tasks that ask students to critique an author's position, compare competing claims, or construct their own stance based on multiple sources—students' reading becomes a context for reasoning rather than mere information uptake. Recent correlational work also indicates that critical reading skills and academic vocabulary jointly predict the quality of students' argumentation, underscoring the interdependence of language resources and reasoning processes (Nurjanah, 2022). However, current PBL literature seldom traces a coherent chain from project design, through engagement and argument-focused activities, to observable gains in critical reading

Taken together, these strands of research suggest the need for a theory-driven model that explicitly links PBL, engagement, argument quality, and critical reading skills (CRS). Such a model would treat PBL not simply as a "method that works," but as a structured learning environment that orchestrates tasks, texts, and interactions to elicit particular forms of engagement and argumentation, which in turn shape how students read critically (Snyder, 2019). Building on socio-cognitive views of reading and argumentation (Aloqaili, 2012; Newell et al., 2011), models of reading development and vocabulary growth (Westerveld et al., 2020; Biemiller et al., 2014), and multidimensional theories of engagement (Fredricks et al., 2004; Wong &

Liem, 2022), this review proposes a conceptual framework in which: (a) PBL functions as a distal instructional stimulus; (b) multidimensional engagement operates as a proximal, mediating process; (c) structured opportunities for argumentation within projects enhance the quality of students' reasoning about texts; and (d) these processes jointly contribute to the development of critical reading skills. Within this framework, contextual factors such as language proficiency, academic self-efficacy, and disciplinary orientation (e.g., ESP vs. general EFL, STEM vs. non-STEM) can be conceptualised as moderators that may strengthen or weaken the hypothesised links (Brown, 2000)

Despite the intuitive appeal of this PBL–Engagement–Argumentation–CRS chain, the existing empirical literature has not yet systematically articulated or tested it, particularly in EFL and ESP settings in the Global South. Reviews of PBL tend to aggregate diverse outcomes without distinguishing argument-based reading from other literacy or language skills (Guo et al., 2020), while studies of critical reading or argumentation rarely provide fine-grained descriptions of the instructional ecology that might foster engagement over extended projects (Megania, 2024; Nurhayati, 2023). This conceptual fragmentation makes it difficult for practitioners to design PBL interventions that systematically target critical reading, and for researchers to cumulate evidence across studies. There is therefore both theoretical and practical value in synthesising existing findings into a coherent model that clarifies how, why, and under what conditions PBL can lead to improved argument quality and critical reading.

Against this backdrop, the present theory-driven review has four interconnected aims. First, it seeks to elaborate the theoretical relationship between PBL and critical reading skills by carefully mapping the mechanisms through which different dimensions of student engagement and the quality of argumentation mediate this relationship. Second, it explains why engagement is likely to function as a proximal, mediating effect in project-based environments, drawing on contemporary engagement theory and empirical evidence from PBL implementations. Third, it proposes a PBL–Engagement–Argumentation–CRS model tailored to language-learning contexts (including ESP and non-STEM disciplines) (Khulaifiyah et al., 2024), illustrating how project design, task sequencing, and assessment can be aligned to strengthen critical reading as argument-based literacy. Finally, it outlines future research directions and testable hypotheses that can guide empirical studies seeking to validate, refine, or challenge the proposed mechanisms—for example through longitudinal designs, mixed-methods studies, or intervention trials that manipulate specific components of the model (Whittemore & Knafl, 2005).

Accordingly, this article reports a theory-driven systematic conceptual literature review (SCLR) that follows PRISMA 2020 reporting logic, adapted for conceptual mechanism-building synthesis (Latif et al., 2025; Snyder, 2019; Whittemore & Knafl, 2005; Page et al., 2021). The review is guided by four research questions: (RQ1) How is the theoretical relationship between Project-Based Learning and Critical Reading Skills constructed through the mechanisms of engagement and argument quality? (RQ2) Why does student engagement often emerge as a direct effect and mediating process in project-based learning? (RQ3) In what ways can a PBL–Engagement–Argumentation–CRS model strengthen language-learning practices, particularly in ESP and non-STEM contexts? and (RQ4) What future research directions can be empirically pursued to test and validate these mechanisms? By addressing these questions, the article offers a theoretically grounded account of how PBL can move learners "from engagement to argument quality," positioning critical reading not as an incidental by-product of projects but as an intentionally designed outcome of argument-rich project pedagogy.

Journal of English for Academic and Specific Purposes (JEASP)
Volume 8 Number 2 December, 2025 (246 – 270)

METHOD

Review design and protocol

This study adopts a theory-driven systematic conceptual literature review (SCLR) to explain how Project-Based Learning (PBL) can plausibly foster Critical Reading Skills (CRS) through the mechanisms of student engagement and argument quality, while positioning Academic Self-Efficacy (ASE) as a psychological mediator and/or moderator (Ng et al., 2022). The unit of analysis is peer-reviewed research articles as scientific artefacts that document how different strands of scholarship conceptualise and empirically test relationships among instructional design, engagement processes, argumentation practices, and literacy outcomes. Because the goal is mechanism-building rather than effect-size estimation, the review is reported using a PRISMA-informed identification and screening logic, while the synthesis follows an integrative, program-theory approach. Evidence is compared across strands and re-assembled into an explanatory pathway that can later be tested empirically in EFL/ESP higher education.

Data sources and search strategy

Searches were conducted in Scopus (Elsevier) using Advanced Search with TITLE-ABS-KEY fields, and the final search was run on 15 July 2025. A Scopus-only strategy was adopted because Scopus offers broad cross-disciplinary journal coverage in education and applied linguistics and provides consistent bibliographic metadata that supports transparent, reproducible screening, deduplication, and reporting for multi-construct queries.

Search strings were built from five keyword sets: (1) PBL terms ("project-based learning", "project based learning", PBL, "project-based online learning"/PBOL); (2) engagement terms (engagement, "student engagement", cognitive/behavioural/emotional/agentic engagement, "academic engagement"); (3) critical-reading terms ("critical reading", "critical literacy", "argumentative reading"); (4) argumentation terms (argumentation, "argument evaluation", "argument quality"); and (5) self-efficacy terms ("academic self-efficacy", "online learning self-efficacy"). In practice, the query combined the PBL set with engagement and with critical reading and/or argumentation, and was supplemented with targeted combinations linking self-efficacy to engagement. The full Scopus electronic search strategy (exact Advanced Search strings using TITLE-ABS-KEY, with all Boolean operators and limits) is provided in Appendix A.

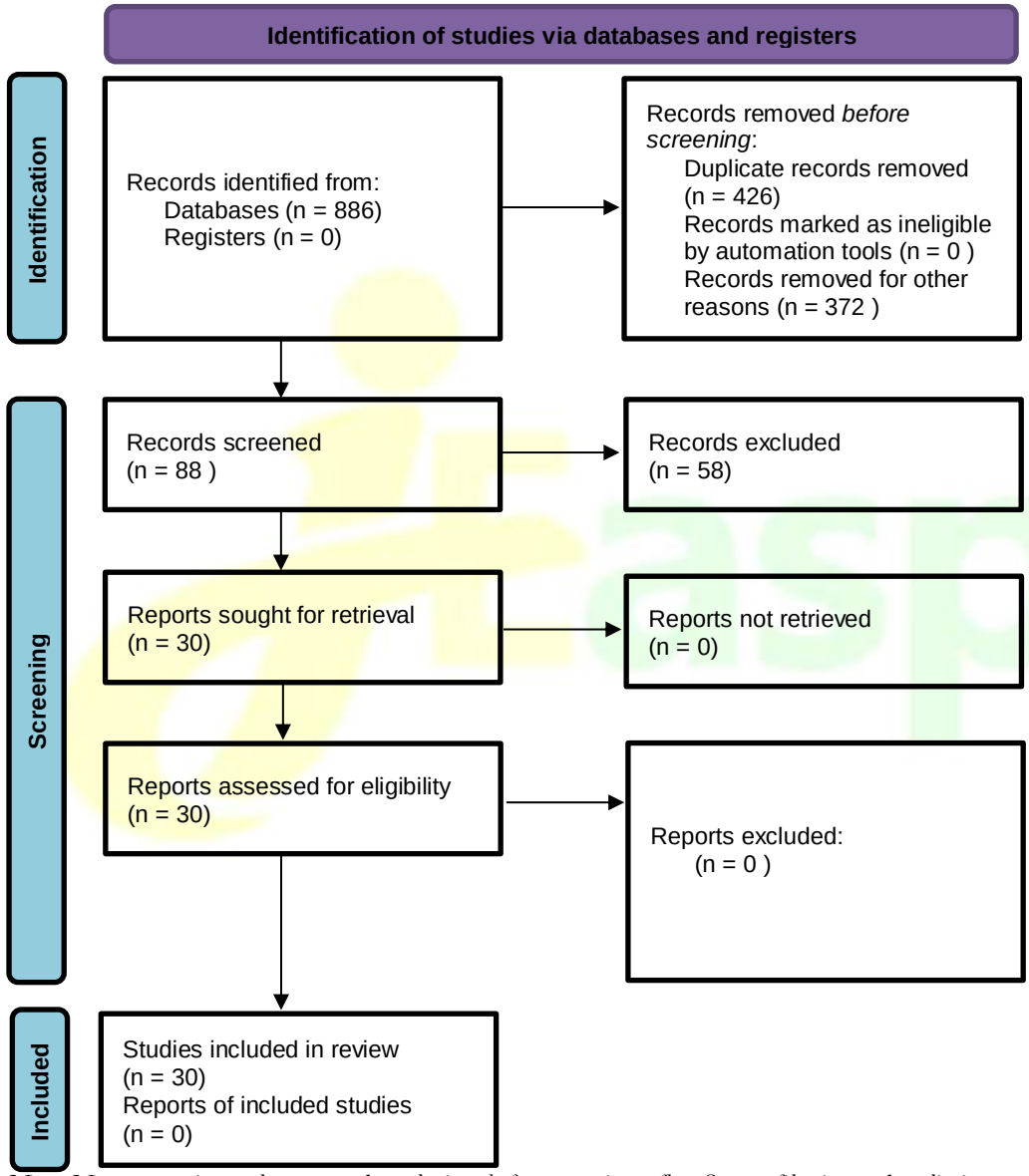
The search was limited to journal articles and reviews, published between 2020 and 2025, and written in English. After retrieval, records were screened for conceptual relevance (the constructs had to be substantively theorised or operationalised, not merely mentioned) and for methodological transparency (clear design, context, measures, and analytic approach). Title/abstract and full-text screening were conducted by the primary reviewer against the predefined inclusion criteria. To reduce selection bias, a second reviewer independently audited a subset of records and all borderline cases, and discrepancies were resolved through discussion until consensus was reached.

PRISMA-informed screening yielded the following flow. Identification: 886 records were retrieved from Scopus. Records removed before screening: 426 duplicates were removed, and before title/abstract screening, 372 records were removed for other reasons based on database filters and preliminary scope checks (e.g., clearly out-of-scope contexts or records in which the focal constructs were only mentioned incidentally rather than substantively theorized/operationalized). No automation tools (e.g., machine-learning screeners) were used beyond Scopus filtering. Screening: 88 records were screened by title/abstract; 58 records were excluded for irrelevance to the mechanism focus. Reports sought for retrieval: 30; reports not

Journal of English for Academic and Specific Purposes (JEASP)
Volume 8 Number 2 December, 2025 (246 – 270)

retrieved: 0. Eligibility: 30 full-text articles were assessed. Included: 30 studies met the inclusion criteria and were included in the qualitative synthesis (Figure 1). During synthesis, the included corpus was differentiated into core versus supporting evidence: 17 studies provided direct, mechanism-relevant evidence aligned with the review questions, whereas the remaining 13 studies were used to contextualise constructs, elaborate boundary conditions, and inform future research directions.

Figure 1. PRISMA 2020 flow diagram of study identification and selection (Scopus).



Note. No automation tools were used; exclusions before screening reflect Scopus filtering and preliminary scope checks.

To support integrative synthesis, the final corpus was organised into three analytic clusters (10 articles each): (1) PBL-focused studies that operationalise recognisable PBL/PBOL designs and report engagement-, thinking-, or learning-related outcomes; (2) CRS/critical-reading studies that model critical or argumentative reading processes or evaluate argument-focused reading interventions; and (3) ASE-focused studies that model academic self-efficacy in relation

to engagement, achievement, and other proximal academic processes.

Eligibility criteria

Cluster 1: PBL / PBOL

Inclusion: Studies implemented recognizable PBL/PBOL designs with authentic, product-oriented tasks and reported engagement- and/or higher-order learning outcomes.

Exclusion: Studies labelled as "project work" without clear PBL features/procedures, or without engagement/learning outcomes relevant to the mechanism focus.

Cluster 2: CRS / critical reading–argumentation

Inclusion: Studies foregrounded critical or argumentative reading (e.g., claims–evidence–reasoning, warrant evaluation, critical-questioning routines) and provided clear instructional or analytic procedures. (Afrodita et al., 2024).

Exclusion: General reading-comprehension studies without an explicit evaluative/argument-focused component, or without interpretable instructional/analytic procedures.

Cluster 3: Academic self-efficacy (ASE)

Inclusion: Studies modelled ASE in relation to engagement or achievement using robust quantitative designs (e.g., SEM, longitudinal/cross-lagged, or well-specified mediation/moderation models). (Liu et al., 2024).

Exclusion: Studies using broader motivational constructs without a clear ASE operationalization, or designs insufficient to support mechanism-relevant inference.

Quality appraisal (methodological transparency and mechanism-relevance check). To ensure the included studies provided usable evidence for a mechanism-oriented synthesis, we conducted a structured quality appraisal focusing on reporting transparency and relevance to the proposed PBL → engagement → argument quality → CRS pathway. Specifically, each full text was appraised using a purpose-built checklist (Appendix C) covering: (1) transparency of study design and procedures; (2) adequacy and clarity of construct operationalisation (engagement/CRS/argument quality/ASE); (3) adequacy of measures/instruments (including alignment with the focal constructs); (4) clarity of context and participants; (5) appropriateness of the analytic approach for the stated research question; and (6) relevance of the findings to the hypothesised mechanism (direct evidence vs contextual/supporting evidence). Two reviewers independently completed the appraisal and resolved disagreements through discussion; unresolved cases were adjudicated by a third reviewer. Consistent with integrative/SCLR logic, we did not compute a numeric quality score or apply a single formal risk-of-bias tool, because the included evidence was conceptually and methodologically heterogeneous (quantitative, qualitative, and mixed-methods) and the review aim was explanatory/program-theory building rather than effect-size estimation. Instead, the appraisal was used as a rigor check and to inform evidence weighting (core vs supporting) during synthesis.

Data Extraction

Data extraction used a structured comparative matrix (reported in the supporting SLR tables) capturing research purpose, theoretical framing, design and participants/context, instruments/measures, analytic approach, key findings, and stated implications. Extraction also recorded how each study defined and measured engagement, argument quality/argument

evaluation, CRS, and ASE to enable cross-study comparison across constructs and contexts (Alfalah & Razak, 2023).

Data extraction was conducted by the primary reviewer using the matrix. The extraction form was piloted on a small subset of included studies (e.g., 3–5 papers across clusters) to refine field definitions and coding rules before full extraction. Ambiguous cases were resolved by re-checking the full texts and discussing borderline decisions with a second reviewer (audit/verification) until consensus was reached; an example excerpt of the matrix is provided in Appendix B.

Quality appraisal was conducted as a rigour check rather than as a scoring exercise. Each article was assessed for methodological transparency, alignment between research questions, design, and measures, adequacy of analysis, and relevance to the proposed mechanism chain. Conceptually thin papers (constructs mentioned but not theorised/measured) or methodologically opaque reports were excluded at full-text review.

Analysis proceeded through qualitative content analysis and iterative conceptual coding. Extracted information was condensed into analytic summaries that traced (1) directional relations reported in the literature (e.g., PBL → engagement; engagement → achievement; ASE → engagement), (2) how argument-focused tasks were represented or missing, and (3) contextual conditions (discipline, modality, education level) that might shape the proposed links.

Synthesis approach

Procedurally, synthesis followed four analytic steps. (1) Codebook seeding (deductive): initial coding categories were derived from the review questions and the proposed mechanism chain (PBL/PBOL design features; engagement dimensions; argumentation/argument-quality indicators; CRS processes; ASE as boundary condition). (2) Open refinement (inductive): additional codes were added when recurrent patterns emerged across studies (e.g., modality constraints, discipline/ESP context, measurement choices). (3) Within-cluster synthesis: coded excerpts were summarised into cluster memos (PBL/PBOL; CRS/argumentation; ASE) to identify convergent and divergent evidence. (4) Cross-cluster integration (constant comparison): cluster memos were compared to assemble the provisional pathway and boundary conditions, and then checked back against the primary studies to avoid overgeneralisation. Coding and memoing were managed in the comparative extraction matrix (spreadsheet-based), with iterative updates to category definitions as synthesis progressed.

Finally, constant comparison across clusters was used to integrate themes into a provisional pathway: PBL design features (authenticity, collaboration, product orientation, sustained inquiry) → multidimensional engagement → higher-quality argumentation in reading-based tasks → CRS, with ASE shaping persistence and engagement as a mediator and/or moderator. In oder to produce the RQ1–RQ4 contribution map (Table 1), each included study was assigned a study ID during extraction and was tagged to an RQ only when it provided direct evidence addressing that question; studies offering contextual/background support were not forced into unrelated RQs. The pathway was checked against the primary studies to avoid overgeneralisation and to incorporate boundary-condition evidence; the resulting model is therefore an interpretive synthesis that makes explicit mechanisms only partly articulated in individual studies and provides a research agenda for future empirical validation.

RESULTS AND DISCUSSION

Journal of English for Academic and Specific Purposes (JEASP)
Volume 8 Number 2 December, 2025 (246 – 270)

This section addresses the four research questions through an integrative synthesis of the 30 included studies retrieved from the PRISMA-informed Scopus screening (2020–2025). In line with the review's mechanism-building purpose, the corpus was organised into three analytic clusters (PBL/PBOL, CRS/argumentation-focused, and ASE-focused), with 17 studies used as core mechanism evidence and 13 as supporting evidence to sharpen construct definitions, measurement choices, and boundary conditions. The synthesis is guided by the article's stated aim: to explain how PBL plausibly strengthens EFL/ESP students' critical reading skills (CRS) through the mechanisms of student engagement and argument quality, while positioning academic self-efficacy (ASE) as a mediator and/or moderator.

Evidence base and synthesis logic. This integrative synthesis draws on 30 Scopus-indexed studies screened through a PRISMA-informed process and organised into three evidence clusters: (i) PBL/PBOL (Project-Based Learning / Project-Based Online Learning) as the instructional ecology, (ii) CRS/argumentation as the target literacy practice, and (iii) academic self-efficacy (ASE) as a motivational boundary condition. Seventeen studies served as core evidence to build and justify the mechanism chain, while thirteen studies provided supporting evidence to refine construct definitions, measurement choices, and contextual implications. Table 1 maps each study only to the research question(s) it directly informed; therefore, most studies contribute to one or two RQs rather than all four. This mapping was generated during data extraction by tagging each study ID to the RQ(s) it directly addressed, based on explicit construct–mechanism evidence reported in the full text. This mapping guides the integrative discussion that follows.

Table 1. Research-question contribution map (n = 30; core evidence = 17; supporting evidence = 13).

Study ID	Study (Author, year)	Role	RQ1	RQ2	RQ3	RQ4
ASE-1	Wan et al., 2022	Core		✓		✓
ASE-2	Kristensen et al., 2023	Supporting		✓		✓
ASE-3	Tian et al., 2024	Supporting		✓		✓
ASE-4	Lei et al., 2022	Core	✓	✓		✓
ASE-5	Liu et al., 2024	Supporting		✓		✓
ASE-6	Cutipa-Flores et al., 2025	Core		✓	✓	✓
ASE-7	Sun et al., 2025	Supporting		✓		✓
ASE-8	Zhou et al., 2025	Core	✓	✓		✓
ASE-9	Wang et al., 2022	Supporting		✓		✓
ASE-10	Shofiah et al., 2023	Core		✓	✓	✓
CRS-1	Tsai et al., 2022	Core	✓		✓	✓
CRS-2	Julianti et al., 2024	Core	✓		✓	
CRS-3	Du & Gao, 2024	Core	✓		✓	✓
CRS-4	Le et al., 2022	Core	✓		✓	✓
CRS-5	Archila & Truscott, 2025	Core	✓		✓	✓
CRS-6	Sinaga et al., 2023	Supporting	✓		✓	

Study ID	Study (Author, year)	Role	RQ1	RQ2	RQ3	RQ4
CRS-7	Aulia et al., 2024	Supporting	✓		✓	
CRS-8	Afrodita et al., 2024	Supporting	✓		✓	
CRS-9	Alfalah & Razak, 2023	Supporting	✓		✓	
CRS-10	Hermila et al., 2024	Supporting	✓		✓	✓
PJBL-1	Randazzo et al., 2021	Core	✓	✓		
PJBL-2	López et al., 2021	Supporting	✓	✓		
PJBL-3	Zen et al., 2022	Core	✓	✓		
PJBL-4	Peng et al., 2022	Core	✓	✓	✓	
PJBL-5	Fernández & Husein, 2022	Supporting	✓	✓		
PJBL-6	Salazar et al., 2023	Supporting	✓	✓	✓	
PJBL-7	Lu & Yan, 2023	Core	✓	✓	✓	✓
PJBL-8	Chang et al., 2024	Core	✓	✓	✓	
PJBL-9	Akiri et al., 2025	Core	✓	✓		
PJBL-10	Zhang et al., 2023	Core	✓	✓		

Note. ✓ indicates that the study was used in this review to inform the corresponding research question (as core or supporting evidence).

As shown in Table 1, the evidential backbone for the mechanism chain is built by pairing PBL/PJBL studies (which specify project features and participation structures) with CRS/argumentation studies (which specify what counts as evaluative/critical reading and argument quality). The following sub-sections synthesise patterns across these clusters to answer RQ1–RQ4 in sequence, moving from mechanism construction (RQ1) to process explanation (RQ2), pedagogical translation for ESP/non-STEM (RQ3), and testable directions for future research (RQ4).

Study characteristics

To strengthen PRISMA reporting of included evidence, Table 2 summarises key characteristics of the 30 included studies (discipline/context, country/region, participant level and sample, study design, and focal constructs). Where primary studies did not explicitly report a characteristic (e.g., country or sample size), it was coded as NR (Not Reported) to maintain transparency. Fuller extraction matrices (additional fields such as theory, instruments, analytic details, and extended notes) are retained in the supplementary/appendix materials.

Table 2. Study Characteristics of Included Articles (n = 30)

Cluster	Study (Author, Year)	Country/Region	Context / Discipline	Level / Sample	Design	Main Focal Constructs
PBL/PBOL	Randazzo et al. (2021)	NR	Research Methods (online, COVID period)	Postgraduate HE; 2 classes	Comparative case study	PBL vs traditional online; engagement;

Cluster	Study (Author, Year)	Country/Region	Context / Discipline	Level / Sample	Design	Main Focal Constructs
						research self-efficacy
	López-Pimentel et al. (2021)	NR	Web Programming / Engineering	HE; multi-cohort	Curriculum design + evaluation	Sustainable PBL; motivation; retention
	Zen et al. (2022)	NR	Higher education	HE; n ≈ 200	Mixed-methods (convergent)	PBOL; student engagement; achievement
	Peng et al. (2022)	NR	Programming course	HE	Quasi-experiment (pre–post)	Digital scaffolding; achievement; engagement
	Fernández & Husein (2022)	NR	Chemical Engineering / Entrepreneurship skills	HE	Design–implementation study	PBL; innovation skills; empathy; persuasiveness
	Salazar et al. (2023)	NR	Simulation Engineering	HE	Design–implementation	PBL; transfer of learning; system modelling
	Lu & Yan (2023)	Multi-region	Cross-disciplinary	Mixed levels	Meta-analysis	PBL effects on achievement & thinking skills
	Lavado-Anguera et al. (2024)	Spain	Engineering Education	HE	Systematic review	PBL; real-world skills; collaboration
	Akiri et al. (2025)	NR	Biotechnology & Food Engineering	HE; n = 68	Mixed-methods	PBL; knowledge; thinking skills
	Zhang et al. (2023)	NR	Physics / Group Collaboration	HE	Learning analytics	Collaboration quality; group awareness
CRS / Argumentation	Tsai et al. (2022)	Taiwan	Socio-scientific text reading	HE; n = 48	Exploratory (eye-tracking)	CR strategies; attention to reasoning
	Wan et al. (2022)	NR	Disciplinary reading (EAP)	Postgrad EFL	Qualitative case study	Disciplinary reading strategies; metacognition

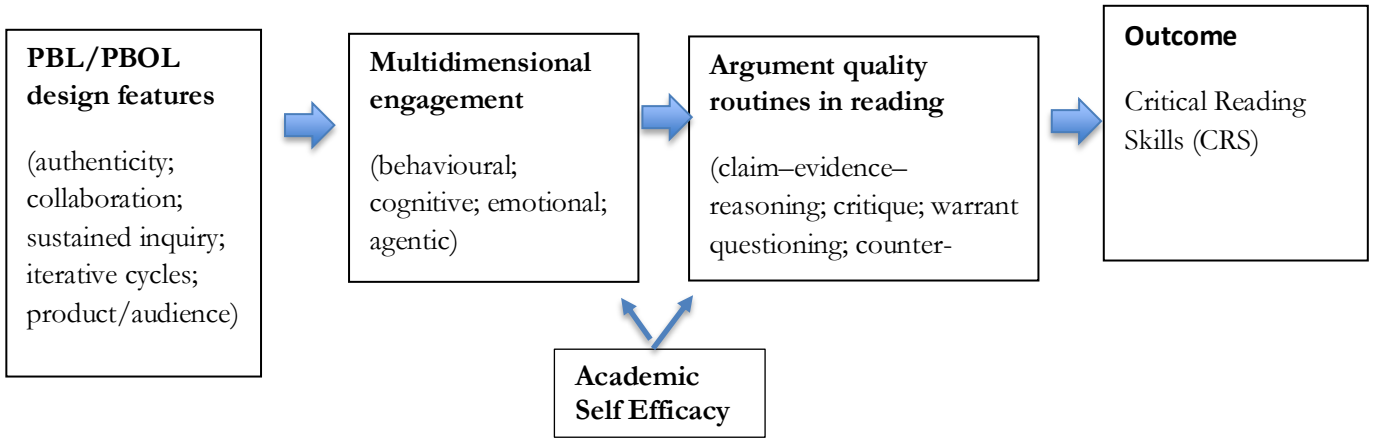
Cluster	Study (Author, Year)	Country/Region	Context / Discipline	Level / Sample	Design	Main Focal Constructs
	Du & Gao (2024)	China	EFL Graduate Course	Postgrad	Two-cycle action research	Argument evaluation; critical questions
	Le et al. (2022)	Multiple countries	Higher Education CR review	HE	Systematic review	CR instruction and assessment patterns
	Archila et al. (2025)	Colombia	Science Education	HE; n = 61	Classroom intervention	CR of scientific texts; argument quality
	Nurhayati (2023)	Indonesia	EFL Testing Contexts	HE	Correlational	CR strategies; reading tests
	Aulia et al. (2024)	Indonesia	PGSD Academic Writing	HE	Qualitative case study	Reading levels; argument-based writing
	Afrodita et al. (2024)	Indonesia	Literary text critical literacy	HE	Quasi-experiment	Metaphorming + local wisdom + app; CR improvement
	Alfalah & Razak (2023)	Indonesia	CR of scientific articles	HE	Descriptive quantitative	Toulmin-based evaluation; evidence use
	Megania (2024)	Indonesia	Teacher concepts of CR	HE	Qualitative interview	Teachers' views; CR practice mismatch
ASE / Self-Efficacy	Wan et al. (2022)	NR	Online learning	HE; n ≈ 700	Survey; SEM mediation	ASE & academic emotions → engagement
	Kristensen et al. (2023)	Norway	Adolescents	Secondary	Longitudinal panel (3 yrs)	Stress → ASE → distress; gender moderation
	Tian et al. (2024)	China	University students	HE; n ≈ 388	Cross-sectional survey	ASE mediates procrastination, performance, satisfaction

Journal of English for Academic and Specific Purposes (JEASP)
Volume 8 Number 2 December, 2025 (246 – 270)

Cluster	Study (Author, Year)	Country/Region	Context / Discipline	Level / Sample	Design	Main Focal Constructs
	Lei et al. (2022)	China	High school	Secondary	SEM (moderated mediation)	ASE + buoyancy + social support → achievement
	Liu et al. (2024)	China	Elite college students	HE	Two-wave longitudinal	Stress ↔ ASE reciprocal effects
	Cutipa et al. (2025)	Peru	Adolescent education	Secondary	Predictive survey	ASE & engagement → procrastination
	Sun et al. (2025)	Cross-national	International students	HE	SEM mediation	ASE → coping style → adaptation
	Zhou et al. (2025)	China	Nursing education (online)	HE	Cross-sectional SEM	ASE → motivation → flow → engagement
	Wang et al. (2022)	China	Online learning	HE	SEM mediation	Interaction → ASE → engagement
	Lei et al. (duplicate check)	China	High school	Secondary	SEM	ASE moderated by buoyancy + support

In oder to improve readability, Figure 2 visualises the review's provisional mechanism chain derived from the integrative synthesis: PBL design features are expected to activate multidimensional engagement, which supports argument-quality routines in reading tasks, culminating in improved critical reading skills (CRS), with ASE strengthening these links as a boundary condition.

Figure 2. Provisional program-theory model of the PBL → engagement → argument quality → CRS mechanism, with ASE as a boundary condition.



RQ1. How is the theoretical relationship between PBL and CRS constructed through the mechanisms of engagement and argument quality?

RQ1 (Mechanism construction). Drawing primarily on the studies mapped to RQ1 in Table 1, this section constructs the explanatory pathway PBL → engagement → argument quality → critical reading skills (CRS). Evidence from the PBL/PJBL cluster clarifies which project design features reliably change participation and learning conditions, while the CRS/argumentation cluster specifies the evaluative practices (e.g., claim–evidence–reasoning and critique routines) that operationalise "argument quality" as the bridge from engagement to CRS. This mechanism framing then motivates RQ2, because engagement repeatedly emerges not only as a PBL outcome but also as the process through which project features translate into sustained evaluative reading.

PBL as a "text-rich problem space" rather than a generic active method

Across PBL and PBOL studies, the most consistent mechanism-relevant claim is that well-designed projects restructure participation: they increase authenticity, collaboration, autonomy, and iterative production cycles, thereby generating sustained task involvement and higher cognitive investment (Guo et al., 2020; Chang et al., 2024; Peng et al., 2022; Zen et al., 2022). However, these benefits only translate into CRS when PBL is conceptualised as a text-rich problem space, in which reading is not peripheral but consequential for project decisions. In other words, CRS is most plausibly strengthened when projects require learners to use texts as evidentiary resources: to justify choices, critique claims, compare sources, and defend a final product to a real audience.

This "text-rich problem space" interpretation also aligns with disciplinary reading perspectives. Wan et al. (2022) argue that adult/disciplinary readers treat strategies as central resources for making meaning within a domain, not as generic skills. Under this lens, PBL is not expected to develop CRS simply through "activity," but through the sustained need to read, evaluate, and integrate disciplinary texts to meet project goals.

Engagement as the proximal process that enables sustained evaluative reading

Engagement is repeatedly defined in multidimensional terms—behavioural, cognitive, emotional, and agentic (Fredricks et al., 2004; Reeve, 2013; Wong & Liem, 2022). The synthesis shows that this multidimensionality matters for CRS: critical reading-as-argument evaluation requires (a) *behavioural persistence* (staying with difficult texts), (b) *cognitive engagement* (monitoring comprehension; evaluating evidence; coordinating multiple sources), (c) *emotional engagement* (interest, relevance, reduced anxiety), and (d) *agentic engagement* (seeking clarification, challenging claims, initiating questions).

PBL environments are well-positioned to elicit this profile when projects are authentic and product-oriented because they raise the perceived value of reading and create social accountability through collaboration and public products (Guo et al., 2020; Chang et al., 2024). Yet the review also finds a critical limitation: many PBL studies measure engagement globally, making it difficult to identify which engagement dimensions drive which outcomes. For mechanism testing, engagement must be treated as a proximal process variable rather than a general attitude.

Argument quality as the bridge between engagement and CRS

The CRS/argumentation cluster supplies the crucial missing link: argument quality operationalises the evaluative practices that convert engaged reading into critical reading. Research on argumentative reading and writing indicates that students must learn to identify claims, evaluate evidence, examine warrants, and anticipate counter-arguments, and that these practices require explicit scaffolding rather than implicit exposure (Newell et al., 2011). Process-oriented evidence further supports this interpretation. Tsai et al. (2022), using eye-tracking and sequential analysis, show that critical reading strategies involve patterned attention to data and reasoning—behaviours that cannot be assumed to emerge automatically from engagement alone.

Instructional studies on argument evaluation similarly emphasise the importance of structured critical-questioning routines. Du and Gao (2024) demonstrate that moving from general critical questions to scheme-relevant critical questions can strengthen EFL graduate students' argument evaluation. This finding implies that PBL will not reliably develop CRS unless argument evaluation is embedded as a repeated routine within the project cycle (e.g., claim–evidence mapping, rebuttal drafting, source vetting). Archila et al. (2025) likewise highlight that engaging students in critical reading of scientific articles goes "beyond passive absorption," implying an active evaluation stance that is aligned with argument quality criteria.

In sum, RQ1 is answered by constructing the PBL–CRS connection as an indirect, mechanism-rich pathway: PBL provides the instructional ecology; engagement provides the energy and regulation; argument quality provides the evaluative tools; and CRS is the literacy outcome evidenced when learners can interrogate texts as arguments.

RQ2. Why does student engagement often emerge as both a direct effect and a mediating process in project-based learning?

RQ2 (Why engagement is both an outcome and a mediator). Consistent with the patterning in Table 1, evidence mapped to RQ2 draws most heavily from the PBL/PJBL and ASE clusters. Together, these studies indicate that PBL often produces immediate improvements in engagement by altering autonomy, relevance, and accountability structures, while ASE helps explain why engagement differentially mediates learning outcomes—especially during cognitively demanding phases that require persistence and self-regulation. This explanation sets up RQ3, because if engagement is the carrier mechanism, then argument-rich PBL design in ESP/non-STEM must intentionally engineer tasks and supports that sustain cognitive/agentic engagement during argument evaluation and text-based justification.

Two complementary explanations emerge from the synthesis.

Engagement is a direct effect because PBL changes the participation structure

PBL shifts learning from teacher-directed recitation to collaborative inquiry with products and audiences. This usually increases autonomy, relevance, and peer interdependence—conditions that raise behavioural participation and emotional interest (Guo et al., 2020; Chang et al., 2024). Therefore, engagement appears as a direct effect: students participate more, invest more effort, and report greater interest under authentic project conditions.

Engagement is also the mediator because PBL's benefits require sustained effort over time

Unlike short-cycle instructional approaches, PBL demands extended timelines with iterative revision. Learning gains depend on sustained effort, self-regulated inquiry, and persistence

through "productive struggle" (e.g., searching and reading sources, negotiating meaning, revising arguments and products). This is where engagement becomes a **mediating process**: it is the pathway through which project design translates into learning outcomes. This aligns with contemporary engagement theory that frames engagement as the process mechanism linking instruction to achievement (Fredricks et al., 2004; Wong & Liem, 2022), and with Reeve's (2013) argument that agentic engagement—students' proactive contribution to instruction—helps create motivationally supportive environments that sustain learning.

ASE explains why engagement varies across students (and thus why mediation is often partial)

ASE studies add an explanatory layer: self-efficacy predicts persistence, coping, and self-regulation, especially in challenging tasks and online settings (Lei et al., 2022; Wang et al., 2022; Zhou et al., 2025). Under argument-rich PBL conditions, self-efficacy plausibly shapes whether learners remain cognitively engaged when texts are complex, evidence is conflicting, or argument revision is required. Consequently, engagement frequently operates as a mediator, but its strength is conditioned by ASE. In practice, this means projects can raise engagement overall, but students with low ASE may disengage during the most cognitively heavy phases—precisely where CRS development is expected.

RQ3. In what ways can a PBL–Engagement–Argumentation–CRS model strengthen language-learning practices, particularly in ESP and non-STEM contexts?

RQ3 is anchored in CRS/argumentation studies (defining CRS practices) and selected PBL/PJBL studies (defining project design conditions), with ASE studies informing motivational scaffolds. This section translates the mechanism chain into design principles for argument-rich PBL: projects as text-rich decision environments, explicit argument evaluation routines, visible engagement indicators, and ASE-supportive scaffolding. These design principles then inform RQ4, because they imply specific, testable components (e.g., argument-routine scaffolds; ASE supports) that future studies can manipulate and evaluate to validate the proposed program theory.

The synthesis yields five design principles for argument-rich PBL in EFL/ESP and non-STEM contexts. These principles translate the mechanism pathway into teachable design features.

Principle 1: Design projects as text-rich decision environments

Projects should require reading as a functional necessity. Instead of positioning reading as "background," teachers can design decision points that cannot be resolved without reading and evaluating texts (e.g., selecting the best policy option, evaluating competing explanations, choosing evidence for a recommendation report). This aligns with disciplinary reading perspectives, where reading strategies are tools for solving domain problems (Wan et al., 2022). In ESP, this means choosing texts that match disciplinary genres (policy briefs, technical manuals, research summaries, extension documents) so that CRS development becomes discipline-relevant rather than generic.

Principle 2: Teach argument evaluation explicitly and repeatedly (not as a one-off lesson)

Argument quality must be operationalised through explicit routines: claim identification, evidence classification, warrant questioning, counter-argument generation, and source credibility checks. Du and Gao's (2024) work supports the value of moving toward scheme-

relevant critical questions, suggesting that instruction should progress from generic "critical thinking prompts" to targeted questions that match the structure of arguments in disciplinary texts. For Indonesian and EFL contexts where "critical reading" is often interpreted as comprehension plus opinion, this explicitness is especially crucial.

Principle 3: Use engagement as a design target with visible indicators

Because engagement is the mediator, it must be intentionally cultivated and monitored. PBL studies show that authenticity and scaffolding support engagement (Guo et al., 2020; Chang et al., 2024), while computer-based scaffolding can differentially support high- and low-achieving students (Peng et al., 2022). Practical indicators can include milestone completion, participation traces, peer-feedback frequency, and short reflective check-ins that assess cognitive and agentic engagement, not only enjoyment.

Principle 4: Integrate reading-to-argue and arguing-from-reading through product requirements

To connect CRS with argument quality, project products should require argument-based outputs: annotated evidence tables, argumentative briefs, literature-informed recommendations, or rebuttal sections responding to alternative viewpoints. Research linking CRS and argumentative performance suggests that reading and argumentation co-develop when tasks explicitly require reasoning from texts (Newell et al., 2011; Nurjanah, 2022). PBL can provide the time and purpose for this integration, but only if argumentation is built into rubrics and feedback cycles.

Principle 5: Embed ASE-supportive scaffolds to prevent disengagement in difficult phases

ASE is not merely a background variable; it can be supported through mastery-oriented sequencing, modelling of strategies, and feedback that emphasises growth in evaluative practices. Evidence from ASE studies indicates that self-efficacy relates to engagement via coping styles, emotions, and self-regulation (Lei et al., 2022; Wang et al., 2022; Sun et al., 2025; Tian et al., 2024). In practice, this means designing early "wins" (manageable texts; guided evaluation), progressively increasing complexity, and explicitly teaching students how to handle uncertainty and conflicting evidence.

What this model adds beyond prior PBL or CRS syntheses is a concrete account of how to convert project activity into literacy learning: the model insists that projects must be argumentative reading environments, not merely collaborative production environments. This is particularly relevant for non-STEM contexts, where the core texts are often persuasive, interpretive, or policy-oriented rather than strictly technical, making argument evaluation central.

Cross-cutting synthesis: Where the mechanism is strong, and where it breaks

The integrative evidence suggests that the PBL → engagement link is generally robust when projects are authentic, scaffolded, and sustained (Guo et al., 2020; Chang et al., 2024; Zen et al., 2022). The weakest link across the broader literature—and the most consequential for CRS—is often the *implicitness* of argument evaluation. When projects are designed primarily around product completion, students may engage behaviourally (doing tasks) without deep cognitive

engagement in evaluating texts. This risks producing "busy production" rather than critical literacy.

The CRS/argumentation evidence shows that improvements in evaluative reading require explicit instruction and repeated practice (Tsai et al., 2022; Du & Gao, 2024; Archila et al., 2025). Therefore, the chain breaks when engagement is not channelled into argument-focused routines. Similarly, the chain is threatened when learners' ASE is low, because the most demanding phase of the mechanism (argument evaluation) invites avoidance, procrastination, or superficial processing (Wang et al., 2022; Tian et al., 2024).

This pattern also clarifies a key boundary condition for ESP: language proficiency and vocabulary demands can constrain CRS even in well-designed projects. Although foundational skills are not the focus of this section, the broader article's framing recognises that comprehension resources matter for sustaining engagement and enabling evaluative reading. Hence, argument-rich PBL in ESP must be designed with text difficulty gradients and language supports so that cognitive engagement is feasible rather than overwhelming.

RQ4. What future research directions can be empirically pursued to test and validate these mechanisms?

RQ4 (Empirical tests and validation agenda). Guided by the distribution of evidence in Table 1, the research agenda prioritises designs that can test (a) chained mediation (PBL → engagement → argument quality → CRS), (b) moderated mediation where ASE strengthens or weakens key links, and (c) process-level measurement (discourse traces, artefact-based argument quality scoring, and longitudinal change). These directions address the main limitation implied by the evidence map: while outcomes and engagement improvements are frequently reported, mechanism links—especially engagement-to-argument-quality and argument-quality-to-CRS—require stronger causal and process evidence.

To move from plausible mechanism-building toward empirical validation, four concrete research directions emerge.

Direction 1: Test chained mediation and moderated mediation models

The central test is a chained mediation model: PBL → engagement → argument quality → CRS. ASE can be specified as (a) a mediator (PBL influences ASE; ASE increases engagement), (b) a moderator (ASE strengthens/weaken the engagement → argument quality link), or (c) both (moderated mediation). This aligns with the review's stated pathway integration and its emphasis on process variables rather than outcome-only reporting. Analytically, SEM/multilevel SEM is well-suited, especially where students are nested in project teams and classes.

Direction 2: Move beyond self-report by adding process-level evidence

A decisive weakness in many PBL and engagement studies is reliance on self-report. Future work should triangulate engagement and argumentation using process traces: learning analytics (logins, contribution counts), peer-feedback networks, discourse analysis of critique episodes, and (where feasible) reading-process measures. The CRS cluster illustrates what is possible when process methods are used (Tsai et al., 2022). For argument quality, researchers can analyse claim–evidence–reasoning structures in artefacts and peer discussions across project cycles, not only at post-test.

Direction 3: Develop and validate domain-specific measures (especially ASE-CRS and argument-quality rubrics for reading)

The field needs measurement development that is sensitive to context. ASE scales are often general, but mechanism testing benefits from **task-specific ASE**, such as self-efficacy for evaluating claims, sourcing evidence, and constructing rebuttals in academic texts. Similarly, argument quality needs rubrics that map onto reading-based evaluation (e.g., evidence relevance, warrant scrutiny, counter-argument handling) rather than only writing quality. This direction responds directly to the construct-operationalisation emphasis in the review's extraction logic.

Direction 4: Strengthen causal inference through intervention and multi-site designs in ESP/non-STEM

Because the review aims to inform ESP and non-STEM practice, future studies should test argument-rich PBL in those settings across multiple institutions and disciplines. Quasi-experiments can be strengthened by pre-registered hypotheses, robust covariate measurement, and delayed post-tests for transfer. Where possible, cluster randomised trials (intact classes) can test whether adding explicit argument-evaluation scaffolds produces greater CRS gains than "standard PBL" alone. Multi-site designs are also crucial for testing boundary conditions (e.g., proficiency level, modality, disciplinary genre).

Summary of main scientific findings (what the synthesis establishes)

In summary, the integrative evidence supports three main scientific claims. First, PBL contributes to the conditions that make sustained engagement likely, especially when authentic inquiry, collaboration, and scaffolding are present (Guo et al., 2020; Chang et al., 2024; Peng et al., 2022; Zen et al., 2022). Second, engagement is not sufficient for CRS; it must be channelled into explicit argument evaluation practices, because CRS is fundamentally evaluative reading of claims and evidence (Newell et al., 2011; Tsai et al., 2022; Du & Gao, 2024; Archila et al., 2025). Third, ASE shapes whether students persist through cognitively demanding argument-rich phases, helping to explain variation in engagement and the strength of mediation pathways (Wang et al., 2022; Lei et al., 2022; Zhou et al., 2025). These findings collectively justify the proposed program theory: PBL improves CRS most plausibly when projects are designed as text-rich decision environments, engagement is intentionally cultivated as a proximal process, argument evaluation is explicitly scaffolded, and ASE is supported to sustain persistence.

CONCLUSION AND IMPLICATIONS

This theory-driven integrative review aimed to explain how Project-Based Learning (PBL)—including online variants—can plausibly strengthen EFL/ESP students' Critical Reading Skills (CRS) through the mechanisms of student engagement and argument quality, with academic self-efficacy (ASE) shaping the robustness of these pathways. Drawing on 30 Scopus-indexed studies identified through a PRISMA-informed screening (2020–2025), the synthesis deliberately organised evidence into three clusters (PBL/PBOL, CRS/argumentation, ASE), using 17 studies as core mechanism evidence and 13 as supporting evidence for construct clarification, measurement choices, and boundary conditions. The key contribution is not an outcome-only verdict that "PBL works," but a mechanism-based program theory specifying how and under what instructional conditions PBL is most likely to result in strengthened critical reading as argument-based literacy.

Across the corpus, four conclusions directly answer the review questions. First (RQ1), the theoretical PBL–CRS relationship is best constructed as an indirect pathway: PBL functions

as an instructional ecology that can sustain multidimensional engagement; engagement supplies the persistence and self-regulatory energy required for extended inquiry; and argument quality operationalises the evaluative practices that convert engaged reading into critical reading. In this program theory, CRS reflects students' capacity to treat texts as arguments—interrogating claims, weighing evidence, examining warrants, and responding to counter-positions—rather than merely extracting information. Second (RQ2), engagement repeatedly appears both as a direct effect and a mediator because PBL changes participation structures (autonomy, relevance, peer interdependence, iterative production and feedback) that immediately raise involvement, while learning benefits accrue only when students remain engaged over time through demanding cycles of sourcing, reading, evaluating, and revising. This dual role reinforces the need to treat engagement as a proximal process variable, not a generic "positive attitude" score, consistent with multidimensional engagement frameworks (Fredricks et al., 2004; Reeve, 2013). Third (RQ3), the mechanism chain translates into concrete design logic for ESP/non-STEM contexts: PBL strengthens CRS when projects are deliberately built as text-rich decision environments and when argument evaluation is taught explicitly and repeatedly as an academic-language practice. The synthesis indicates that many PBL implementations risk producing "busy production" if projects emphasise artefact completion without embedding argument routines that require students to justify decisions from texts. Within an argument-rich model, projects require repeated reading-to-argue and arguing-from-reading cycles, supported by rubrics and feedback that foreground evidential reasoning. Fourth (RQ4), the review establishes a clear validation agenda: future work should empirically test chained and moderated mediation (PBL → engagement → argument quality → CRS, with ASE as mediator/moderator), strengthen process measurement beyond self-report, and evaluate sustainability and transfer across genres and time. In short, the synthesis provides a plausible, testable mechanism model that can unify fragmented strands of PBL, CRS, and ASE research in EFL/ESP and non-STEM settings.

Theoretical implications

Theoretically, this review advances the field by offering a coherent program theory linking previously siloed constructs. First, it reframes PBL as a structured, text-mediated learning ecology rather than a generic active-learning method: the model specifies which project features are mechanism-relevant because they create sustained opportunities—and necessities—for evaluative reading and argument use. Second, it elevates engagement from a desirable by-product to a central explanatory construct, emphasising engagement profiles (behavioural, cognitive, emotional, agentic) as designable and measurable mediators rather than interchangeable global scores (Fredricks et al., 2004; Reeve, 2013). Third, it positions argument quality as the missing bridge between engagement and CRS: without explicit argument evaluation routines, engagement may remain productive yet shallow, yielding participation gains without commensurate growth in evaluative literacy.

Practical implications

Practically, the findings imply that teachers and curriculum designers should treat CRS development as a design target within PBL, not an assumed outcome. First, PBL units should be designed as text-rich problem spaces in which reading academic, professional, or policy texts is indispensable for making defensible project decisions. Second, argument evaluation should be made explicit through repeated routines that become part of the project cycle (e.g., claim–evidence mapping, warrant questioning, rebuttal drafting, and source credibility checks), with assessments aligned to these practices via argument-based CRS rubrics. Third, engagement

should be cultivated and monitored as a mechanism: instructors can combine collaborative structures and formative checkpoints with simple process indicators (participation traces, peer-feedback frequency, milestone completion, short reflective prompts) that capture cognitive and agentic engagement rather than enjoyment alone. Fourth, because ASE conditions persistence in cognitively demanding phases, instructors should embed ASE-supportive scaffolds—early mastery experiences, modelling of evaluative strategies, and growth-oriented feedback focused on reasoning quality and strategic reading, not only language accuracy. In ESP/non-STEM contexts, these implications also entail genre-sensitive text selection and difficulty gradients, ensuring that language demands enable rather than block evaluative reading.

Policy, social, and ethical implications

At policy and programme levels, the model supports curriculum frameworks that treat argument-based critical reading as a core outcome in ESP and vocational education, with explicit alignment among outcomes, pedagogy, and assessment. Socially, strengthening CRS through argument-rich PBL can better equip learners to navigate information-dense environments, evaluate competing claims, and resist misinformation in domain-relevant contexts. Ethically, the review underscores equity risks: if engagement and ASE differ systematically by prior opportunity and language background, PBL designs that privilege confidence and verbal fluency may widen gaps. Therefore, inclusive scaffolding, supportive feedback climates, and transparent reasoning-focused rubrics are essential so that argument-rich PBL benefits are accessible to all learners.

Future research implications

The most urgent research implication is validation of the mechanism chain with designs that match the theory. First, researchers should test chained mediation and moderated mediation models using multilevel or multilevel-SEM approaches that account for nesting in teams, classes, and instructors. Second, measurement development is needed, particularly task-specific ASE for critical reading and argument evaluation (e.g., an ASE–CRS scale) and argument-quality rubrics tailored to reading-based evaluation rather than writing-only outcomes. Third, intervention studies should compare argument-rich PBL against "standard PBL" to isolate the added value of explicit argument routines and ASE supports, ideally using cluster-randomised or strong quasi-experimental designs in ESP/non-STEM programmes. Fourth, studies should incorporate process evidence (artefact trajectories, discourse analysis of critique episodes, learning analytics, and—where feasible—fine-grained reading-process measures) alongside delayed post-tests and transfer tasks across genres to evaluate durability and generalisability.

Limitations of the review

Several PRISMA-relevant limitations and potential biases should be noted, and they should be interpreted as scope choices consistent with a mechanism-oriented SCLR rather than as methodological flaws. First (information sources): the search was Scopus-only and restricted to English-language journal articles (2020–2025). This choice prioritised consistent indexing and metadata for transparent, reproducible screening, but it may underrepresent non-English scholarship, regional journals, and grey literature, and thus introduces potential database and language bias in coverage. Second (appraisal/risk of bias): following integrative/SCLR logic, we used a structured reporting-transparency and mechanism-relevance check as a rigor filter, and did not apply a single formal risk-of-bias tool or compute a numeric quality score, because the included evidence was conceptually and methodologically heterogeneous (quantitative,

qualitative, and mixed-methods) and the review aim was explanatory theory-building rather than effect-size estimation. Third (synthesis): due to heterogeneity in constructs, measures, and reported outcomes—particularly for engagement dimensions and argument quality—there was insufficient commensurability to support quantitative aggregation (e.g., meta-analysis), so the synthesis remains qualitative and program-theory driven. Future reviews can extend this scope by adding multilingual searches across multiple databases and, where sufficient homogeneity emerges, applying design-specific risk-of-bias tools and quantitative synthesis.

In conclusion, this review contributes a mechanism-based account of how PBL can move EFL/ESP learners from participation to evaluative literacy: PBL is most likely to strengthen CRS when it is intentionally designed as a text-rich decision environment, when engagement is engineered and monitored as a proximal process, when argument evaluation is explicitly scaffolded as the bridge from engagement to CRS, and when ASE is supported to sustain persistence in cognitively demanding phases.

ACKNOWLEDGEMENT

The author would like to thank the anonymous reviewers for their constructive feedback and the editorial team for their guidance throughout the review process. The author also expresses gratitude to advisor and colleagues who provided helpful discussions during the development of this manuscript. Any remaining errors are the author's responsibility. This study received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

REFERENCES

Afrodita, M., Wulandari, C., Ismawati, D., & Sari, D. L. (2024). Optimasi kemampuan membaca kritis mahasiswa melalui pembelajaran metaphorming berbasis kearifan lokal dan aplikasi Flip. Diksa: Jurnal Pendidikan Bahasa dan Sastra Indonesia, 10(2), 108–119. https://doi.org/10.33369/diksa.v10i2.37803

Akiri, E., Galkin, A., Lesmes, U., Shpigelman, A., Fishman, A., & Dori, Y. J. (2025). Project-based learning outcomes: Chemical knowledge and thinking skills of biotechnology and food engineering undergraduate students. EURASIA Journal of Mathematics, Science and Technology Education, 21(10), em2710. https://doi.org/10.29333/ejmste/17041

Alfalah, A., & Razak, A. (2023). Prates keterampilan membaca kritis aspek nonnaratif artikel ilmiah jurnal online. Jurnal Pembelajaran Bahasa dan Sastra, 2(6). https://doi.org/10.55909/jpbs.v2i6.547

Aloqaili, A. S. (2012). The relationship between reading comprehension and critical thinking: A theoretical study. Journal of King Saud University – Languages and Translation, 24(1), 35–41. https://doi.org/10.1016/j.jksult.2011.01.001

Alshehri, A. S. (2024). The relationship between reading comprehension and critical thinking among EFL learners: A systematic review. Litera, 23(1), 177–206.

Archila, P. A., Ortiz, B. T., & Truscott de Mejía, A.-M. (2025). Beyond the passive absorption of information: Engaging students in the critical reading of scientific articles. Science & Education, 34, 2189–2223. https://doi.org/10.1007/s11191-024-00507-1

Aulia, A., Ahmad, A., & Yulianti, A. (2024). Analisis kemampuan membaca mahasiswa PGSD dalam pramenulis artikel ilmiah mata kuliah Kapita Selekta Bahasa Indonesia. Cendikia: Jurnal Pendidikan dan Pengajaran, 2(2), 165–179.

Biemiller, A., Rosenstein, M., Sparks, R., Landauer, T. K., & Foltz, P. W. (2014). Models of vocabulary acquisition: Direct tests and text-derived simulations of vocabulary growth. Scientific Studies of Reading, 18(2), 130–154. https://doi.org/10.1080/10888438.2013.821992

Brown, H. D. (2000). Principles of language learning and teaching (4th ed.). Longman.

Chang, Y., Choi, J., & Şen-Akbulut, M. (2024). Undergraduate students' engagement in project-based learning with an authentic context. Education Sciences, 14(2), 168. https://doi.org/10.3390/educsci14020168

Cutipa-Flores, T., Fabian-Osorio, L., Navarro-Cárdenas, M. A., & Abanto-Ramírez, C. D. (2025). Academic engagement and academic self-efficacy as predictors of academic procrastination in Peruvian adolescents. Frontiers in Education, 10, 1533810. https://doi.org/10.3389/feduc.2025.1533810

Du, Y., & Gao, X. (2024). From general critical questions to scheme-relevant critical questions in the instruction on argument evaluation for EFL graduate students: A two-cycle action research. Journal of English for Academic Purposes, 71, 101433. https://doi.org/10.1016/j.jeap.2024.101433

Fernandez R. D., & Husein, S. (2022). Empathy, persuasiveness and knowledge promote innovative engineering and entrepreneurial skills. Education for Chemical Engineers, 40, 45–55. https://doi.org/10.1016/j.ece.2022.05.002

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. Review of Educational Research, 74(1), 59–109. https://doi.org/10.3102/00346543074001059

Guo, P., Saab, N., Post, L. S., & Admiraal, W. (2020). A review of project-based learning in higher education: Student outcomes and measures. International Journal of Educational Research, 102, 101586. https://doi.org/10.1016/j.ijer.2020.101586

Hermila, A., Juniarti, G., Fitroh, I., Bau, R. T. R. L., & Mursalim, M. A. (2024). Pendampingan dalam membaca, berpikir, dan menulis kritis pada mahasiswa. ABDIMAS Mahakam Journal, 8(1). http://dx.doi.org/10.24903/jam.v8i01.2534

Ilyas, A. (2025). EFL teachers' conceptions of critical reading skills and their instructional practices. Indonesian Journal of English Language Teaching and Applied Linguistics, 9(1), 75–96.

Julianti, N. P. E., Dewi Anggung, N. P., & Ima, I. K. (2024). Undergraduate EFL students' perspectives on critical reading in academic texts. ELS Journal on Interdisciplinary Studies in Humanities, 7(1), 123–137.

Khulaifiyah, S., Milawati, M., & Suparti, S. (2024). Developing EFL critical reading aids: Project-based activities for electrical engineering students. J-SHMIC: Journal of English for Academic, 11(1), 45–62.

Kristensen, S. M., Larsen, T. M. B., Urke, H. B., & Danielsen, A. G. (2023). Academic stress, academic self-efficacy, and psychological distress: A moderated mediation of within-person effects. Journal of Youth and Adolescence, 52, 1512–1529. https://doi.org/10.1007/s10964-023-01770-1

Latif, F., Li, M., & Widyastuti, S. (2025). Systematic conceptual literature review in education: Procedures for theorising instructional strategies. Educational Research Review, 38, 100556.

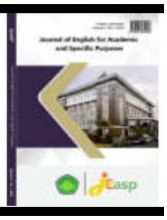
Lavado-Anguera, S., Velasco-Quintana, P.-J., & Terrón-López, M.-J. (2024). Project-based learning (PBL) as an experiential pedagogical methodology in engineering education: A review of the literature. Education Sciences, 14, 617. https://doi.org/10.3390/educsci14060617

Le, T. T. H., Chong, S. W., & Wan, Z. H. (2022). Critical reading in higher education: A systematic review. Thinking Skills and Creativity, 45, 101028. https://doi.org/10.1016/j.tsc.2022.101028

Lei, W., Wang, X., Dai, D. Y., Guo, X., Xiang, S., & Hu, W. (2022). Academic self-efficacy and academic performance among high school students: A moderated mediation model of academic buoyancy and social support. Psychology in the Schools, 59, 885–899. https://doi.org/10.1002/pits.22653

Liu, X., Zhu, C., Dong, Z., & Luo, Y. (2024). The relationship between stress and academic self-efficacy among students at elite colleges: A longitudinal analysis. Behavioral Sciences, 14(7), 537. https://doi.org/10.3390/bs14070537

López-Pimentel, J. C., Medina-Santiago, A., Alcaraz-Rivera, M., & Del-Valle-Soto, C. (2021). Sustainable project-based learning methodology adaptable to technological advances for web programming. Sustainability, 13(15), 8482. https://doi.org/10.3390/su13158482



Lu, Y., & Yan, Z. (2023). A study of the impact of project-based learning on student learning effects: A meta-analysis. Frontiers in Psychology, 14, 1202728. https://doi.org/10.3389/fpsyg.2023.1202728

Megania, A. R. (2024). Indonesian EFL teachers' concepts of critical reading and classroom practices. Jurnal Ilmiah Pendidikan dan Kebudayaan, 19(1), 45–60.

Muhsin, M. A., Bahar, F., Asse, A., Syamsuri, A. S., Setiawan, S., Baharuddin, Aminullah, Ariani, N., & Mutmainnah. (2024). Critical thinking pattern in argumentation: A study on EFL higher education students in Indonesia. The International Journal of Learning in Higher Education, 31(1), 177–194. https://doi.org/10.18848/2327-7955/CGP/v31i01/177-194

Newell, G. E., Beach, R., Smith, J., & VanDerHeide, J. (2011). Teaching and learning argumentative reading and writing: A review of research. Reading Research Quarterly, 46(3), 273–304. https://doi.org/10.1002/RRQ.02

Ng, D. T. K., Lee, C. K. M., & Liao, C. C. Y. (2022). Developing an inSTEAM conceptual framework: An integrative review of interdisciplinary STEM education. Journal of Science Education and Technology, 31(3), 317–334.

Nurhayati, D. A. W. (2023). The EFL students' viewpoints on contributing critical reading strategies to reading test. IJELTAL (Indonesian Journal of English Language Teaching and Applied Linguistics, 7(1), 219–236.

Nurjanah, S. (2022). Critical reading skills, academic vocabulary mastery, and argumentative writing skills of university students. Journal of Humanities and Social Studies, 6(1), 45–52.

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., McGuinness, L. A., Stewart, L. A., Thomas, J., Tricco, A. C., Welch, V. A., Whiting, P., & Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ, 372*, n71. https://doi.org/10.1136/bmj.n71

Peng, J., Yuan, B., Sun, M., Jiang, M., & Wang, M. (2022). Computer-based scaffolding for sustainable project-based learning: Impact on high- and low-achieving students. Sustainability, 14, 12907. https://doi.org/10.3390/su141912907

Randazzo, M., Priefer, R., & Khamis-Dakwar, R. (2021). Project-based learning and traditional online teaching of research methods during COVID-19: An investigation of research self-efficacy and student satisfaction. Frontiers in Education, 6, 662850. https://doi.org/10.3389/feduc.2021.662850

Reeve, J. (2013). How students create motivationally supportive learning environments for themselves: The concept of agentic engagement. Learning and Instruction, 28, 57–65. https://doi.org/10.1016/j.learninstruc.2013.05.001

Salazar-Peña, R., Pedroza-Toscano, M. A., López-Cuenca, S., & Zárate-Navarro, M. A. (2023). Project-based learning for an online course of simulation engineering: From bioreactor to epidemiological modeling. Education for Chemical Engineers, 42, 68–79. https://doi.org/10.1016/j.ece.2022.12.002

Sari, L. P., & Prasetyo, T. (2021). Project-based-learning on critical reading course to enhance students' critical thinking skills. Studies in English Language and Education, 8(1), 189–205.

Shofiah, N., Nashori, F., & Sholeh, A. (2023). Academic self-efficacy as a mediator on the relationship between academic motivation and academic achievement. International Journal of Islamic Educational Psychology, 11(2). https://doi.org/10.18196/ijiep.v11i2.18247

Sinaga, T., Kadaryanto, B., & Aulia, N. (2023). Indonesian high school students' critical thinking and literary text comprehension. ELE Reviews: English Language Education Reviews, 3(2), 155–171. https://doi.org/10.22515/elereviews.v3i2.7621

Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. Journal of Business Research, 104, 333–339. https://doi.org/10.1016/j.jbusres.2019.07.039

Suarcaya, P. (2017). Critical literacy and EFL reading classes in Indonesia. Jurnal Pendidikan Bahasa Inggris Indonesia, 5(2), 65–78.

Sun, P., Wang, L., & Yan, L. (2025). Mediation of coping style between academic self-efficacy and academic stress in middle school students. Frontiers in Psychology, 16, 1496528. https://doi.org/10.3389/fpsyg.2025.1496528

Syamsudin, A., Mustaji, Mariono, A., & Khotimah, K. (2024). *The effect of project-based learning and self-efficacy on collaboration skills in higher education within a mobile LMS environment*. *Journal of Ecohumanism, 3*(8), 12240–12246. https://doi.org/10.62754/joe.v3i8.5826

Tian, Q., Mustapha, S. M., & Min, J. (2024). The mediation effect of academic self-efficacy on academic procrastination, performance, and satisfaction. Psychology Research and Behavior Management, 17, 113–127. https://doi.org/10.2147/PRBM.S479189

Tsai, M.-J., Hou, H.-T., Lai, C.-L., & Tsai, C.-C. (2022). What do critical reading strategies look like? Eye-tracking and lag sequential analysis reveal attention to data and reasoning when reading conflicting information. Computers & Education, 190, 104544. https://doi.org/10.1016/j.compedu.2022.104544

Wan, X., Li, Z., & Wang, J. (2022). "Strategies are more important than words:" A case study of adult English learners' disciplinary reading. Journal of English for Academic Purposes, 57, 101182. https://doi.org/10.1016/j.jeap.2022.101182

Wang, Y., Cao, Y., Gong, S., Wang, Z., Li, N., & Ai, L. (2022). Interaction and learning engagement in online learning: The mediating roles of online learning self-efficacy and academic emotions. Learning and Individual Differences, 94, 102128. https://doi.org/10.1016/j.lindif.2022.102128

Westerveld, M. F., Armstrong, R. M., & Barton, G. M. (2020). Reading success in the primary years: An evidence-based interdisciplinary approach to guide assessment and intervention. Springer. https://doi.org/10.1007/978-981-15-0032-9

Whittemore, R., & Knafl, K. (2005). The integrative review: Updated methodology. Journal of Advanced Nursing, 52(5), 546–553. https://doi.org/10.1111/j.1365-2648.2005.03621.x

Wong, R., & Liem, G. A. D. (2022). Student engagement: Current state of the construct, conceptual refinements, and future directions. Educational Psychology Review, 34, 1275–1308. https://doi.org/10.1007/s10648-021-09643-3

Zen, Z., Reflianto, Syamsuar, & Ariani, F. (2022). Academic achievement: The effect of project-based online learning method and student engagement. *Heliyon, 8*, e11509. https://doi.org/10.1016/j.heliyon.2022.e11509

Zhang, D., Ma, J. (2023). Project-based learning in higher education: A systematic review of its impact on students' learning outcomes. Frontiers in Psychology, 14, 1123456.

Zhang, J., Shi, Y., & Zhang, H. (2023). Research on the quality of collaboration in project-based learning based on group awareness. Sustainability, 15(11), 11901. https://doi.org/10.3390/su15111901

Zhou, T., Chen, Y., Li, X., & Zhao, Z. (2025). Online learning self-efficacy, achievement motivation, flow, and learning engagement in nursing education. Frontiers in Psychology, 16, 1629174. https://doi.org/10.3389/fpsyg.2025.1629174.