Journal of Curriculum and Teaching; Vol. 15, No. 4; 2026
ISSN 1927-2677 E-ISSN 1927-2685
Published by Sciedu Press

Advancing Sustainable STEM Education through Artificial Intelligence–Enhanced Project-Based Learning in Middle School Physics

Shiping Wang[1,*] & Hashimah Mohd Yunus[1]

[1]Taylor's University, 47500 Subang Jaya, Selangor, Malaysia

*Correspondence: Shiping Wang, Taylor's University, 47500 Subang Jaya, Selangor, Malaysia, Email Address: 0363501@sd.taylors.edu.my

Received: September 17, 2026 Accepted: October 22, 2026 Online Published: November 12, 2026

doi:10.5430/jct.v15n4p283 URL: https://doi.org/10.5430/jct.v15n4p283

Abstract

We examine how the perceived quality of artificial intelligence-enhanced project-based learning (AIPJBL) relates to sustainable STEM learning outcomes in middle-school physics, with physics learning engagement as an indirect pathway and student-perceived teacher AI pedagogical competence as a moderator. The experimental design combines a 12-week unit covering energy efficiency, heat transfer, renewable energy conversion, and smart lighting with a post-unit survey of 426 students and 24 subsequent interviews. The survey contains 22 five-point items and is analysed using partial least squares structural equation modelling with 5,000 bootstrap resamples. The retained results show positive associations between AIPJBL and outcomes ($\beta = .368$) and between AIPJBL and engagement ($\beta = .527$). The indirect association through engagement is .217, and the interaction with teacher competence is .080. Reported explained variance is 27.8% for engagement and 58.3% for outcomes. The thematic summary concerns inquiry, verification, agency, teacher guidance, and responsible-AI tensions. We contribute a theory-guided framework distinguishing pedagogical quality, engagement, and perceived capability.

Keywords: artificial intelligence in education, project-based learning, sustainable STEM education, middle school physics, student engagement, teacher AI pedagogical competence, PLS-SEM

1. Introduction

Education for sustainable development connects scientific understanding with systems thinking, collaboration, and reasoned action on environmental and community problems (UNESCO, 2017, 2020, 2021). Therefore, school STEM education needs to connect disciplinary knowledge with opportunities to examine how energy, materials, and technologies are used. For middle-school physics, this connection can be made through measurable questions about heat transfer, electricity, and energy conversion. Research on secondary STEM engagement and integrated instruction provides a rationale for this focus, rather than establishing that middle school is uniquely consequential for every learner (Struyf et al., 2019; Thibaut et al., 2018; Wang et al., 2016).

Project-based learning provides an organisational structure for sustained inquiry, collaboration, testing, and revision. Reviews associate well-designed projects with favourable learning outcomes while identifying substantial variation in implementation and assessment (C.-H. Chen & Yang, 2019; Condliffe et al., 2017; Kokotsaki et al., 2016). In integrated STEM, the challenge is to connect the product that students construct to the disciplinary ideas that explain its performance (English, 2016; Holmlund et al., 2018; Kelley & Knowles, 2016). Thus, a functioning prototype is not sufficient evidence of conceptual understanding.

Artificial intelligence (AI) offers additional ways to scaffold project-based physics inquiry through questions, hints, explanations, and formative feedback (Hwang et al., 2020; Zhang & Aslan, 2021). Generative AI is one category of AI that produces content in response to prompts; it is not a synonym for all AI-enabled tools. Its educational use also raises concerns about inaccurate explanations, fabricated sources, privacy, and the outsourcing of reasoning (Akgun & Greenhow, 2022; Kasneci et al., 2023; UNESCO, 2023). Consequently, the relevant issue is how AI support is integrated into investigation and verification, not simply whether a tool is available.

Against this background, we define AIPJBL as the perceived quality of an authentic project in which AI-assisted

inquiry and feedback remain subject to student investigation and teacher guidance. The instructional context described in this manuscript is a 12-week Grade 7–9 physics or integrated-science unit in public middle schools, with four pathways: an energy-efficient classroom, a solar cooker, a micro-wind turbine, and smart lighting. These pathways connect physics explanations to sustainability decisions.

Two relationships guide the inquiry into AIPJBL and sustainable STEM learning outcomes. Physics learning engagement provides a proposed indirect pathway, whereas teacher AI pedagogical competence provides a proposed condition under which the AIPJBL–outcome association may vary. We distinguish student involvement in physics reasoning from demonstrated learning, and student perceptions of teacher practice from independently assessed teacher competence.

Specifically, we examine associations among AIPJBL, physics learning engagement, student-perceived teacher AI pedagogical competence, and self-reported sustainable STEM learning outcomes. The outcomes concern perceived physics application, evidence-based design, trade-off reasoning, communication, and readiness for sustainability-oriented action, not measured achievement gains or long-term environmental behaviour (Kioupi & Voulvoulis, 2019; Redman et al., 2021; Sass et al., 2020). This focus connects pedagogical design, learner participation, and teacher guidance without attributing educational effectiveness to AI alone.

2. Literature Review and Hypothesis Development

2.1 AI-Enhanced Project-Based Learning and Constructivist Theory

We define AI-enhanced project-based learning (AIPJBL) as the perceived quality of a sustained, authentic project in which AI tools scaffold inquiry, modelling, feedback, comparison, and revision while students retain responsibility for evidence, explanation, and design decisions. This operational definition synthesises project-based learning principles with critical AI literacy (Condliffe et al., 2017; Kokotsaki et al., 2016; Ng et al., 2021). The six item domains are AI-supported testable questions, an authentic sustainability problem, comparison of alternatives, evidence-based revision, integrated STEM practices, and verification of AI output. Student ownership informs the pedagogy but is not claimed to be an additional measured dimension.

The constructivist logic of AIPJBL holds that learners build and reorganise knowledge through purposeful interaction with problems, representations, materials, peers, and feedback. Integrated STEM frameworks similarly emphasise situated learning, scientific inquiry, engineering design, technological literacy, and mathematical thinking as mutually reinforcing practices (Kelley & Knowles, 2016; Martín-Páez et al., 2019; Ortiz-Revilla et al., 2020; Perignat & Katz-Buonincontro, 2019; Quigley et al., 2017; Thibaut et al., 2018). In a physics project, for example, students designing an energy-efficient classroom must connect heat transfer, electrical power, measurement uncertainty, cost, user behaviour, and environmental impact. The learning task is not simply to recall a formula; it is to decide which model is appropriate, collect defensible data, interpret discrepancies, and justify a design under constraints.

Within this constructivist account, AI support is valuable only insofar as learners examine and use it in their own reasoning. Reviews identify personalisation, adaptive feedback, assessment support, and learning analytics as possible affordances, but they do not establish that presenting an explanation ensures its comprehension (Chassignol et al., 2018; L. Chen et al., 2020; Zawacki-Richter et al., 2019). Human-centred accounts likewise distinguish the provision of assistance from learners' control over inquiry (Luckin et al., 2016; Ouyang & Jiao, 2021; Roll & Wylie, 2016).

Project-based pedagogy also makes the social construction of knowledge visible. Teams negotiate roles, explain competing models, document design decisions, and expose assumptions to critique. STEM integration research identifies the need to connect collaboration and engineering design explicitly to disciplinary understanding rather than assume those connections emerge automatically (Dare et al., 2018; Estapa & Tank, 2017; Margot & Kettler, 2019). In addition, interviews with teachers and administrators identify professional-development needs for integrated STEM implementation (Shernoff et al., 2017). In AIPJBL, we therefore treat AI-generated proposals as resources for comparison with peer reasoning and physical evidence.

AIPJBL also offers a route to sustainable STEM education because project constraints can require students to consider environmental, technical, social, and economic consequences together. A solar cooker, for instance, can be evaluated for thermal performance, material availability, safety, cost, usability, and waste. Sustainability education frameworks caution against reducing complex issues to one technically optimised variable (Leicht et al., 2018; Lozano et al., 2017; Wals & Benavot, 2017). Accordingly, the outcome of an AIPJBL unit should be more than a functional prototype. It should include the capacity to explain mechanisms, interpret evidence, compare trade-offs, anticipate consequences, and communicate an ethically defensible recommendation.

2.2 Physics Learning Engagement and Self-Determination Theory

Physics learning engagement refers to students' active, affective, cognitive, and agentic involvement in physics learning. Behavioural engagement is visible in attention, participation, persistence, and completion; emotional engagement involves interest, enjoyment, curiosity, or frustration; cognitive engagement involves strategy use, conceptual elaboration, monitoring, and willingness to invest effort; agentic engagement occurs when students ask questions, express preferences, suggest procedures, and help shape instruction. Contemporary engagement research treats these dimensions as related but non-identical and emphasises that engagement is responsive to the design of the learning environment (Bond et al., 2020; Fredricks et al., 2016; Wong & Liem, 2022).

Self-determination theory explains why support for autonomy, competence, and relatedness can encourage autonomous motivation and engagement (Ryan & Deci, 2020). Autonomy concerns ownership of action, competence concerns a sense of effectiveness, and relatedness concerns connection with others. In a project, choices about tests, feedback on a prototype, and shared responsibility may support these needs, but the presence of such activities does not establish that students experience them as supportive. Consistent with this account, Jang et al. (2016) examine need-supportive and need-thwarting processes in engagement. Furthermore, students can contribute to their learning conditions by requesting clarification or proposing project directions (Reeve & Shin, 2020). We use self-determination theory as a rationale for H2 and H4.

AI can influence all three needs, but its effect is not automatically positive. Immediate hints and differentiated explanations can strengthen competence, especially when students are stuck at a productive but solvable impasse. Conversational interaction may make it easier for quieter students to rehearse explanations before speaking to peers. Translation and multimodal support can widen access. At the same time, opaque recommendations, excessive correction, or answer completion can reduce autonomy and ownership. If some teams have better devices, subscriptions, or prompt knowledge, AI may also weaken relatedness and perceived fairness. Research on generative AI therefore calls for careful task design, transparency, and critical literacy rather than unrestricted adoption (Farrokhnia et al., 2024; Grassini, 2023; Lo, 2023; Rahman & Watanobe, 2023).

In the proposed model, engagement is a possible process linking learning opportunities to perceived outcomes, not a substitute for learning. Concentration, persistence, interest, explanation to peers, and initiative can make disciplinary work more likely, yet a student may remain busy or enthusiastic while retaining a misconception. Therefore, the engagement construct concerns involvement in physics inquiry, whereas SSLO concerns self-reported capability. Neither measure independently establishes conceptual change. This distinction follows the separation between engagement and educational outcomes in engagement research (Fredricks et al., 2016; Wong & Liem, 2022).

2.3 Teacher AI Pedagogical Competence and Sustainable STEM Learning Outcomes

Teacher AI pedagogical competence (TAIPC) is the capacity to select, explain, orchestrate, monitor, and evaluate AI-supported learning in ways that are pedagogically purposeful, technically informed, ethically responsible, and appropriate to specific physics content. It builds on digital competence frameworks and technological pedagogical content knowledge, but it adds knowledge of model limitations, data practices, prompt design, explainability, output verification, academic integrity, and human oversight. DigCompEdu positions professional engagement, digital resources, teaching, assessment, learner empowerment, and learners' digital competence as interconnected areas of teacher practice (Redecker, 2017). Research on TPACK likewise shows that effective technology integration requires teachers to connect tool affordances with content representations and pedagogical decisions rather than treat technology as a separate skill (Schmid et al., 2020; Valtonen et al., 2017; Willermark, 2018).

The AI context intensifies this requirement. Teachers must anticipate that generative systems can produce plausible but incorrect physics explanations, inappropriate levels of complexity, hidden bias, and unverifiable references. They must decide when AI feedback is useful, when direct instruction or peer dialogue is preferable, and how to preserve students' responsibility for reasoning. Intelligent-TPACK extends conventional TPACK by incorporating ethical and intelligent-system knowledge into teachers' professional judgement (Celik, 2023). Systematic reviews similarly conclude that AI can assist planning, implementation, and assessment, but teacher agency and professional learning remain critical to its classroom value (Celik et al., 2022; UNESCO, 2024b).

TAIPC is expected to strengthen AIPJBL in at least five ways. First, competent teachers align AI use with a precise physics learning objective. Second, they design verification routines that require students to compare AI output with measurements, textbooks, simulations, and causal models. Third, they regulate cognitive load by deciding when to provide prompts, worked examples, or direct explanation. Fourth, they protect privacy and equity by minimising personal data, using approved accounts, and organising shared access. Fifth, they use formative evidence from both AI

interactions and physical project work to adapt instruction. Falloon (2020) argues that teacher digital competence must be understood as situated professional action, not merely operational proficiency; this principle is even more important when the technology can generate content and recommendations.

Sustainable STEM learning outcomes (SSLO) are defined as students' perceived ability to apply physics and integrated STEM practices to sustainability-oriented problems. The construct includes conceptual application, evidence-based design, systems thinking, trade-off evaluation, solution communication, and action orientation. It therefore extends conventional achievement without replacing it. Physics concepts such as energy conservation, efficiency, thermal transfer, forces, and electrical power remain central, but students must use them to interpret real systems and make justified decisions. Education for sustainable development emphasises competencies that enable learners to understand interdependence, anticipate consequences, collaborate across perspectives, and act strategically (Brundiers et al., 2021; UNESCO, 2020). Studies of sustainability consciousness and competence assessment further suggest that knowledge, attitudes, and action-related capacities should be considered together while remaining empirically distinguishable (Olsson et al., 2016; Redman et al., 2021).

The SSLO measure does not capture objective physics achievement, verified transfer, or long-term environmental behaviour. Rather, it records students' perceived capability to explain sustainable technologies, analyse energy and resource data, compare alternatives, design evidence-based responses, and use STEM knowledge in future sustainability problems. This bounded definition connects disciplinary application with sustainability-oriented decision making without equating confidence or willingness with demonstrated competence (Brundiers et al., 2021; UNESCO, 2024c).

2.4 AI-Enhanced Project-Based Learning and Sustainable STEM Outcomes

Project-based learning research provides a rationale for connecting authentic inquiry to disciplinary application, but not direct evidence for the added contribution of AI. Reviews emphasise teamwork, scaffolding, assessment, and sustained inquiry as conditions for successful implementation (Condliffe et al., 2017; Kokotsaki et al., 2016), while meta-analytic evidence identifies variation across contexts and designs (C.-H. Chen & Yang, 2019). For AI-enhanced projects, the unresolved issue is how feedback and generated alternatives fit into these same inquiry and revision processes. The present hypothesis therefore concerns perceived quality within AIPJBL, not a comparison demonstrating that AI outperforms otherwise equivalent non-AI projects.

Similarly, integrated STEM frameworks connect disciplinary knowledge to situated design problems, rather than treating technology as an independent learning objective (English, 2016; Kelley & Knowles, 2016). The K-12 STEM literacy framework of Falloon et al. (2020) maps curriculum and pedagogical characteristics that can guide this integration. For example, the energy-efficient classroom pathway connects an AI-generated suggestion about insulation to a testable heat-transfer claim; the relevant learning demand is explaining and evaluating that claim. Studies of STEM implementation reinforce the need to make the connection between an artefact and its scientific mechanism explicit (Dare et al., 2018; Estapa & Tank, 2017). The example illustrates the proposed design logic and is not an additional classroom observation.

Accordingly, we use AI consistently as the umbrella term for AI-enabled instructional functions, including adaptive support and content generation; "intelligent systems" is not a separate construct. Generative AI refers specifically to systems producing new responses, such as explanations or design suggestions. Within AIPJBL, these functions are embedded in a project decision: a suggestion is compared with evidence, feedback is considered during revision, and an explanation is checked against physics principles (L. Chen et al., 2020; Hwang et al., 2020). This human-guided arrangement is consistent with the AI-supported and AI-empowered paradigms discussed by Ouyang and Jiao (2021). Its intended advantage is additional support for inquiry, not automatic learning from exposure to AI output.

Nevertheless, broad reviews of AI and STEM education do not, on their own, establish how an integrated AI-enhanced physics project relates to sustainability-oriented capability (Hwang et al., 2020; Li et al., 2020; Zawacki-Richter et al., 2019). This is the gap addressed by H1: quality-focused construct links AI assistance to evidence use, physics application, and sustainability trade-offs in one learning environment. We therefore propose a positive association with SSLO while recognising that a self-report association cannot establish achievement gains or an AI-specific effect.

H1. AI-enhanced project-based learning is significantly and positively associated with sustainable STEM learning outcomes.

2.5 AI-Enhanced Project-Based Learning and Physics Learning Engagement

Authentic projects can give students a reason to persist with difficult physics questions because their decisions affect a shared product. Project-based learning research links collaborative, iterative, and authentic activity to engagement,

while also identifying uneven participation and workload (Almulla, 2020; Guo et al., 2020). However, the higher-education focus of Guo et al. limits direct generalisation to middle-school learners. For AIPJBL, the pertinent question is whether AI-supported questioning and revision are associated with involvement in physics inquiry, rather than whether learners enjoy using a new tool.

Engagement scholarship explains why AIPJBL may be influential. Environmental complexity, defined as the simultaneous presence of challenge and support, is associated with stronger momentary engagement in secondary classrooms (Shernoff et al., 2016). AIPJBL can create this balance by presenting a consequential problem while providing responsive hints, visualisations, and feedback. It can also support autonomy when teams choose design pathways, competence when progress is visible, and relatedness when students collaborate around a shared sustainability goal. These conditions are consistent with the self-determination process described by Jang et al. (2016).

AI literacy introduces a further condition: learners need to recognise limitations and evaluate output quality rather than accept an answer because it is fluent (Ng et al., 2021; UNESCO, 2024a). AI prompts can provide an opportunity to rehearse questions or compare explanations, but these are affordances rather than observed motivational effects. Reviews of generative AI identify both potential support and risks to independent reasoning (Farrokhnia et al., 2024; Grassini, 2023; Lo, 2023; Rahman & Watanobe, 2023). Thus, the hypothesised association concerns engagement within evidence-centred project routines, not prompt-and-copy activity.

Furthermore, evidence from higher-education AI reviews cannot simply be transferred to Grade 7–9 physics (Bond et al., 2024; Lo, 2023), and broader AI-in-education research calls for closer attention to pedagogical context (Chiu et al., 2023). This contextual limitation motivates H2. We expect perceived AIPJBL quality to be positively associated with concentration, persistence, interest, explanation, and initiative; these indicators do not constitute separate tests of four engagement dimensions.

H2. AI-enhanced project-based learning is significantly and positively associated with physics learning engagement.

2.6 Physics Learning Engagement and Sustainable STEM Outcomes

Engagement describes how students relate to a learning activity; learning concerns changes in what they understand or can do. These concepts are related but not interchangeable. Mathematics and science engagement research identifies behavioural, cognitive, emotional, and social indicators (Fredricks et al., 2016; Wang et al., 2016), while conceptual work recommends specifying the object of engagement, such as physics inquiry (Wong & Liem, 2022). In the present model, active participation is therefore a possible condition for learning, not sufficient evidence of it.

For example, asking why two temperature measurements differ may open an opportunity to reason about heat loss, but the question itself does not show that the student understands the mechanism. Similarly, copying a correct AI-generated calculation can produce a successful product without revealing independent understanding. These examples illustrate why AIPJBL engagement must be distinguished from SSLO and why conceptual assessments are needed beyond self-report. The educational-interface account of engagement offers a related interpretive perspective, although it was developed in higher education rather than middle-school physics (Kahu & Nelson, 2018).

Agentic engagement can create opportunities for explanation and feedback, but those opportunities should not be confused with demonstrated transfer (Reeve & Shin, 2020). In a design project, a student might propose a different sensor location or ask why two efficiency calculations disagree. Whether that contribution reflects or produces understanding requires evidence from the explanation, measurement, and subsequent performance. We therefore treat transfer to a new sustainability problem as a question for assessment rather than an outcome established by engagement alone.

The remaining issue is whether students who report stronger physics engagement also report greater sustainability-oriented application and design capability. Research on agentic engagement suggests a rationale for this association, but does not establish transfer in this particular unit (Reeve & Shin, 2020). H3 therefore links two separately defined self-report constructs. Prior knowledge and teaching quality remain plausible alternative influences, and observed performance would be needed to distinguish perceived capability from demonstrated learning.

H3. Physics learning engagement is significantly and positively associated with sustainable STEM learning outcomes.

2.7 The Mediating Role of Physics Learning Engagement

The rationale for an indirect pathway combines the preceding arguments: project conditions may be associated with engagement, and engagement may in turn be associated with perceived sustainability-oriented capability. Research on

classroom challenge and support and on need-supportive teaching makes this sequence theoretically plausible (Jang et al., 2016; Shernoff et al., 2016). However, the proposed sequence is not established merely because two associations are positive. We use H4 to examine the indirect association while retaining the direct AIPJBL–SSLO path.

This distinction is particularly important when AI is embedded in project work. A hint may be ignored, misunderstood, or used productively; a complete answer may be copied or critically evaluated. Tool output alone cannot reveal the learner's response, and participation alone cannot demonstrate learning. Human-centred AI accounts support attention to learners' agency and educational context rather than treating automation as an educational outcome (Holmes et al., 2019; Williamson & Eynon, 2020; Zhang & Aslan, 2021). Consequently, the gap motivating H4 concerns the relationship among perceived project quality, engagement, and perceived capability, not a proven psychological transmission mechanism.

Because feedback, representations, and sustainability framing could also relate to outcomes independently of engagement, we retain a direct path alongside the indirect path. This permits a pattern consistent with partial statistical mediation without presupposing that it will occur. A non-zero indirect association would be compatible with the proposed explanation, but a single post-unit survey cannot establish temporal ordering, rule out reverse associations, or show that engagement caused learning. Longitudinal and performance evidence would be needed for those stronger interpretations.

H4. The indirect association between AI-enhanced project-based learning and sustainable STEM learning outcomes through physics learning engagement is significant and positive.

2.8 The Moderating Role of Teacher AI Pedagogical Competence

Teacher AI pedagogical competence concerns the decisions through which a tool is connected to a physics task: selecting a suitable response, checking a generated explanation, adjusting support, and deciding when a non-AI approach is preferable. Digital-competence and TPACK research provides a rationale for expecting these pedagogical decisions to matter beyond operational proficiency (Falloon, 2020; Redecker, 2017; Schmid et al., 2020). We therefore examine whether the association between AIPJBL quality and perceived outcomes differs with student-perceived teacher guidance; we do not claim that a student rating directly measures a teacher's professional knowledge.

In the AI context, reviews identify both instructional opportunities and challenges involving trust, workload, competence, and ethics (Celik et al., 2022). Intelligent-TPACK subsequently specifies professional knowledge for ethically integrating AI-based tools (Celik, 2023). Applied to project-based physics, this reasoning concerns guidance that connects generated feedback to a concept, asks for verification, preserves students' responsibility for explanations, and provides equitable access. These are the enacted practices represented by TAIPC, rather than an inventory of platform skills.

Because teacher decisions help enact the perceived quality of AIPJBL, we expect its association with SSLO to vary with TAIPC. Guidance that connects an AI response to a physics principle and a project milestone may make the support more usable, whereas poorly aligned assistance may leave misconceptions unresolved (Celik, 2023; Celik et al., 2022). This argument motivates an interaction between perceived project quality and perceived teacher practice. It does not establish that the proposed guidance occurred or that identical technologies necessarily produced different learning outcomes.

The gap addressed by H5 is therefore the conditional association between pedagogical quality and perceived outcomes, alongside teacher competence's separate main association. We retain TAIPC → SSLO so that the interaction is not confused with the main contribution of teacher guidance. We make no a priori claim about the interaction's size or practical importance.

H5. Teacher AI pedagogical competence significantly and positively moderates the association between AI-enhanced project-based learning and sustainable STEM learning outcomes.

2.9 Research Questions and Theoretical Model

The five hypotheses now follow the specific issues developed: integrated pedagogical quality and sustainable STEM capability (H1), middle-school physics engagement in AI-supported projects (H2), the distinction between engagement and perceived capability (H3), their indirect association (H4), and variation associated with perceived teacher practice (H5).

RQ1. What is the direction and magnitude of the association between perceived AIPJBL quality and self-reported SSLO?

RQ2. How does reported physics learning engagement vary with perceived AIPJBL quality?

RQ3. What is the association between physics learning engagement and self-reported SSLO after accounting for AIPJBL and TAIPC?

RQ4. What indirect association between AIPJBL and SSLO is present through physics learning engagement?

RQ5. How does the AIPJBL–SSLO association vary across levels of student-perceived TAIPC?

Figure 1 shows the complete hypothesised structural specification. H1 concerns AIPJBL → SSLO; H2 and H3 form the AIPJBL → PLE → SSLO indirect association in H4. H5 concerns the product term AIPJBL × TAIPC → SSLO. Dashed arrows into the product term denote its construction, not additional estimated causal paths. TAIPC → SSLO is a main-effect adjustment, not a sixth hypothesis.

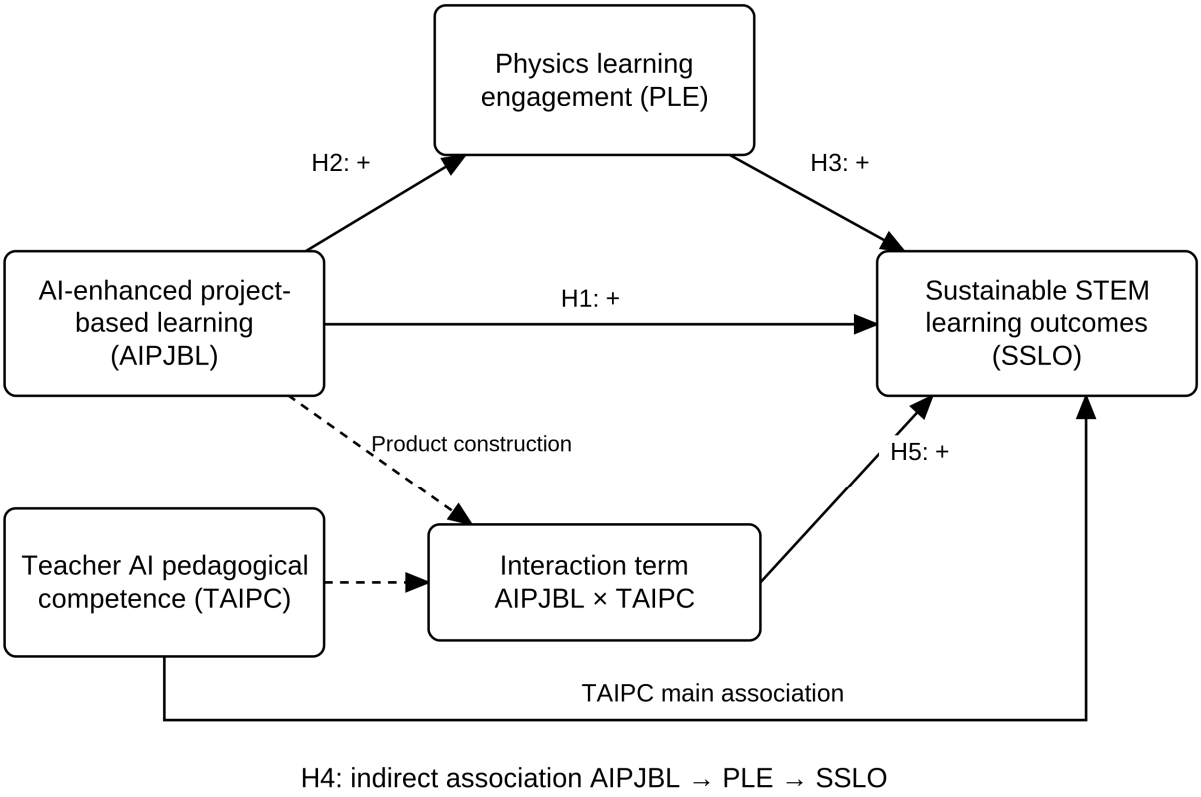


Figure 1. Theoretical Model

3. Methodology

3.1 Participants and Research Setting

The manuscript reports 426 complete student responses from Grade 7–9 public middle-school classes and describes school-level stratification with intact-class sampling. Table 1 retains the supplied profile: 212 girls, 210 boys, and four students who preferred not to disclose gender, with urban, county-town, and rural school-location categories. The reported qualitative component involves 18 students and six teachers. Sample size alone does not establish adequate power for a small interaction or stable bootstrap inference, particularly when class-level dependence is unmodelled (Hair et al., 2022; Kock & Hadaya, 2018).

Table 1. Participant Profile (N = 426)

Characteristic	Category	n	%
Gender	Girl	212	49.8
	Boy	210	49.3
	Prefer not to say	4	0.9
Grade	Grade 7	132	31.0
	Grade 8	146	34.3
	Grade 9	148	34.7
School location	Urban	198	46.5
	County-town	127	29.8
	Rural	101	23.7
Prior educational AI use	None	104	24.4
	Occasional	163	38.3
	Monthly	108	25.4
	Weekly or more	51	12.0
Project pathway	Energy-efficient classroom	123	28.9
	Solar cooker and heat transfer	116	27.2
	Micro-wind turbine	108	25.4
	Smart lighting and motion	79	18.5

3.2 Research Design

The reported design is explanatory sequential: a 12-week AIPJBL unit is followed by one student survey, then interviews intended to contextualise the quantitative pattern. The quantitative component addresses RQ1–RQ5, while the qualitative component concerns how participants describe inquiry, engagement, teacher guidance, and responsible-AI difficulties. Interpretation links the two components rather than treating interviews as another statistical test of the structural paths.

3.3 Instructional Context and AIPJBL Unit

The instructional account describes a Grade 7–9 physics or integrated-science unit lasting 12 weeks, with two 45-minute lessons per week and additional team work where school policy permitted. Four project pathways are specified: an energy-efficient classroom plan, a solar cooker, a micro-wind turbine, and a smart lighting system. Each pathway addresses core physics concepts while requiring consideration of sustainability constraints. The energy-efficient classroom pathway connects heat transfer, thermal insulation, electrical power, and energy audits; the solar cooker connects radiation, conduction, convection, temperature measurement, and material properties; the wind turbine connects forces, rotational motion, energy conversion, circuits, and efficiency; and smart lighting connects electricity, sensing, power, user behaviour, and cost. Figure 2 displays these pathways and their common instructional sequence; it is a visualisation of the existing account, not new implementation evidence.

All pathways share six specified phases. During problem framing, students analyse a local sustainability need and define success criteria. During concept preparation, the teacher provides direct instruction, demonstrations, and worked examples. During AI-assisted inquiry, teams generate questions, compare explanations, and plan tests. During prototyping, they collect physical measurements and document design choices. During critique and revision, they compare AI feedback with experimental evidence and peer review. During public communication, teams present the prototype, a physics explanation, a sustainability trade-off analysis, and an AI-use log.

The protocol specifies school-approved accounts, avoidance of personal data in prompts, disclosure of AI assistance, and logs of prompts and revisions. Its four verification questions are: What claim is being made? Which physics principle supports it? What evidence can test it? What uncertainty or alternative explanation remains? These routines are consistent with critical AI literacy and human oversight (Akgun & Greenhow, 2022; UNESCO, 2023, 2024a).

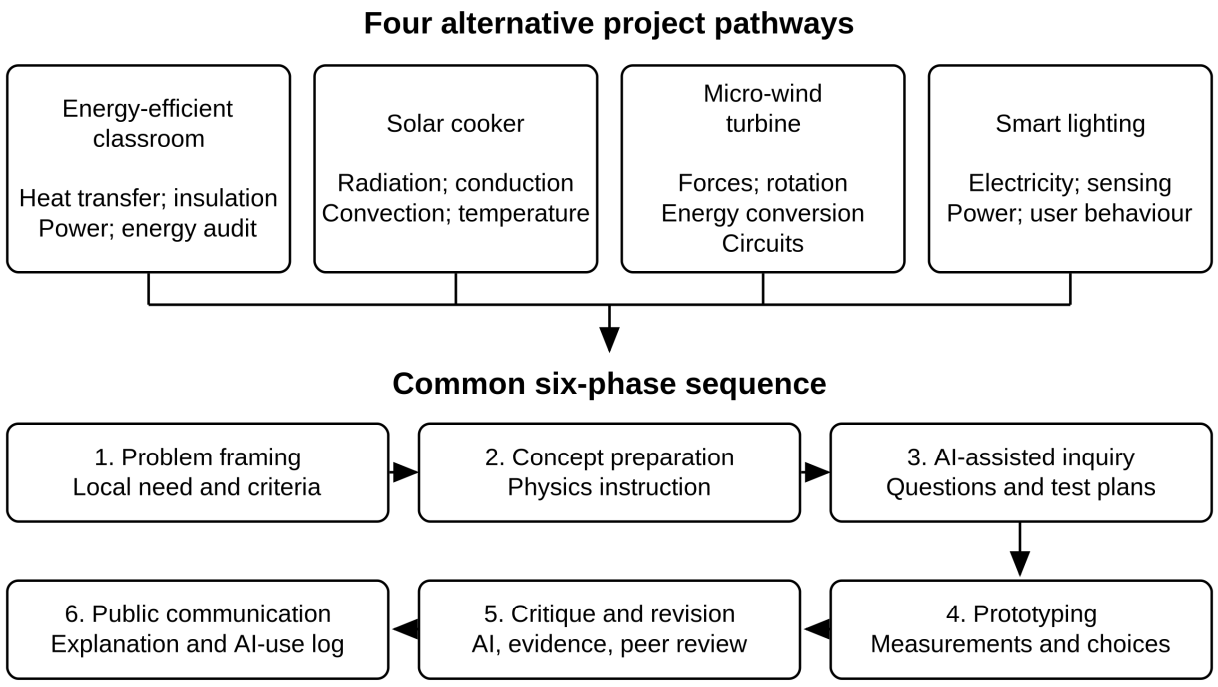


Figure 2. Four Project Pathways and the Shared Six-Phase Instructional Sequence

3.4 Quantitative Strand: Measurement Development and Questionnaire Profile

All constructs are modelled reflectively and measured on a five-point scale ranging from 1 (strongly disagree) to 5 (strongly agree). AIPJBL contains six items representing authentic sustainability problems, AI-supported question generation, feedback and alternative comparison, evidence-based revision, STEM integration, and verification. The item pool synthesises established project-based learning principles and AI literacy requirements (Condliffe et al., 2017; Kokotsaki et al., 2016; Ng et al., 2021). Physics learning engagement contains five items adapted conceptually from mathematics and science engagement research and covers concentration, persistence, enjoyment, explanation to peers, and initiative (Fredricks et al., 2016; Struyf et al., 2019; Wang et al., 2016).

TAIPC is measured from the student perspective because the moderator represents competence as experienced in the enacted classroom. Five items assess whether the teacher explains appropriate AI use, helps students evaluate output, links AI feedback to physics, protects privacy and access, and uses evidence to adjust instruction. The dimensions are informed by DigCompEdu, TPACK measurement, and Intelligent-TPACK (Celik, 2023; Redecker, 2017; Schmid et al., 2020). SSLO contains six items assessing conceptual application, energy and systems analysis, evidence-based design, trade-off evaluation, communication, and future action orientation, drawing on ESD competency frameworks (Brundiers et al., 2021; Redman et al., 2021; UNESCO, 2017).

The instrument is described as a contextual adaptation administered once after the AIPJBL unit. Its 22 indicators are specified reflectively; this is an analytical assumption, not evidence that the different item domains are interchangeable. Reliability and validity diagnostics are reported before structural paths. No subgroup comparisons are interpreted, and the results do not establish differences by grade, gender, location, or prior AI experience.

Table 2 summarises the four constructs, item counts, operational domains, and conceptual sources after their measurement rationale. Appendix A shows the complete 22-item questionnaire.

Table 2. Questionnaire Profile

Variable	Items	Operational focus	Conceptual sources
AI-enhanced project-based learning (AIPJBL)	6	Authenticity, AI-supported inquiry, iterative design, evidence use, STEM integration, verification	Condliffe et al. (2017); Kokotsaki et al. (2016); Ng et al. (2021)
Physics learning engagement (PLE)	5	Concentration, persistence, enjoyment, peer explanation, initiative	Fredricks et al. (2016); Wang et al. (2016)
Teacher AI pedagogical competence (TAIPC)	5	Purposeful use, output evaluation, physics alignment, privacy/equity, adaptive teaching	Celik (2023); Redecker (2017); Schmid et al. (2020)
Sustainable STEM learning outcomes (SSLO)	6	Concept application, energy analysis, design, trade-offs, communication, action orientation	Brundiers et al. (2021); Redman et al. (2021); UNESCO (2017)

3.5 Quantitative Strand: Data Analysis

The analysis specifies AIPJBL and TAIPC as exogenous constructs, physics learning engagement (PLE) as the mediator, and SSLO as the outcome. An AIPJBL × TAIPC product term represents moderation. The 22 questionnaire indicators use five-point response scales.

At the construct-score level, the specified relationships can be written as Equations (1)–(2). X denotes AIPJBL, M denotes PLE, W denotes TAIPC, Y denotes SSLO, and Z denotes the constructed interaction term; ε denotes residual variation.

$$M = \alpha_1 + aX + \varepsilon_1 (1)$$

$$Y = \alpha_2 + c'X + bM + dW + gZ + \varepsilon_2 (2)$$

The reflective measurement model was evaluated using Cronbach's alpha, composite reliability (ρC), ρA, average variance extracted (AVE), indicator loadings, and indicator variance inflation factors. Discriminant validity was assessed using the Fornell–Larcker criterion and the heterotrait–monotrait ratio (HTMT). Reliability values above .700, AVE values above .500, and HTMT values below .850–.900 were interpreted in conjunction with substantive judgement (Hair et al., 2019, 2022; Henseler et al., 2016).

The analytical account describes 5,000 nonparametric bootstrap resamples, standard errors, t statistics, p values, and percentile confidence intervals. It evaluates H4 through the specific indirect association and H5 through the product term, with simple slopes intended at the TAIPC mean and one standard deviation below and above it. Explanatory summaries are R^2 and adjusted R^2; f^2 describes incremental explanatory contribution. The reported predictive procedure is ten-fold cross-validation with Q^2predict, RMSE, and MAE (Shmueli et al., 2019). Neither R^2 nor Q^2predict is a global model-fit test.

3.6 Bias Assessment, Unit of Analysis, and Implementation Fidelity

All four constructs are represented by individual students' questionnaire responses. AIPJBL and TAIPC refer to experienced classroom practices, not independently measured classroom-level exposures or teacher-level attributes. Students are nevertheless nested in classes and schools, and the reported model does not account for that nesting. Accordingly, the analysis concerns student-level associations; it is not a cross-level model.

The specified fidelity criteria are an authentic driving problem, explicit physics objectives, iterative testing, evidence verification, equitable access, and formative teacher intervention across the six instructional phases. We distinguish perceived AIPJBL quality from AI-use frequency: frequent use could reflect productive inquiry, confusion, or answer outsourcing.

The study reports construct-level full-collinearity VIFs below its stated warning value of 3.300. These values are retained as collinearity diagnostics; they neither establish the absence of common-method bias nor validate the data's provenance (Fuller et al., 2016; Jordan & Troth, 2020). Same-source ratings and the single measurement occasion remain limitations. Likewise, the reported predictive summaries cannot establish transportability to different schools, and should be checked for separation of model fitting from held-out prediction (Hair et al., 2022; Shmueli et al., 2019).

3.7 Qualitative Strand: Sampling and Interview Procedure

The submitted account reports 18 students selected for variation in engagement and project pathway and six participating teachers. It describes 24 interviews lasting 25–40 minutes, audio-recording with consent, and the use of

prompt logs, project notebooks, data tables, and prototypes to support recall. The semi-structured interview guide is found in Appendix B. It addresses perceived assistance, difficulties, verification, engagement, teacher guidance, inclusion, and sustainability trade-offs.

3.8 Qualitative Strand: Thematic Analysis

The account describes hybrid deductive–inductive thematic analysis: the four constructs and responsible-AI issues provide initial organising categories, while inductive coding allows accounts outside that framework to inform interpretation. This approach is consistent with Xu and Zammit (2020), who explain how theory-led and data-led coding can be combined in educational research. The two researchers coded an initial subset independently, discussed differences, refined a codebook, and linked themes to interview records and artefacts.

Here, negative or discrepant cases mean accounts that contradict a proposed explanation—for example, evidence that ready access to AI coincided with reduced physics reasoning, or that an ostensibly supportive routine excluded a learner. Their purpose is to challenge theme boundaries and prevent selecting only accounts that fit the quantitative pattern.

4. Results

4.1 Preliminary Analysis and Reporting Conventions

The results describe 426 complete cases and no missing values. Indicator standard deviations range from 1.168 to 1.271 in Table 4. Reported full-collinearity VIFs are 1.810 for AIPJBL, 1.790 for PLE, 1.200 for TAIPC, and 2.360 for SSLO. These diagnostics do not eliminate same-source bias. The complete-case count is not a response rate and does not show whether cases were excluded before analysis.

4.2 Construct Reliability and Convergent Validity

Table 3 reports internal consistency and convergent validity. All Cronbach's alpha values ranged from .896 to .903, composite reliability values ranged from .920 to .928, and the ρA values ranged from .909 to .917. AVE values ranged from .658 to .721, exceeding the .500 criterion. These results indicate that the indicators cohered strongly around their intended constructs and that each construct explained more than half of its indicators' variance.

Table 3. Construct Reliability and Convergent Validity

Construct	ρA	ρC	Cronbach's alpha	AVE
AI-enhanced project-based learning	.914	.924	.901	.670
Physics learning engagement	.917	.928	.903	.721
Teacher AI pedagogical competence	.914	.926	.900	.715
Sustainable STEM outcomes	.909	.920	.896	.658

Note. ρA = construct reliability; ρC = composite reliability; AVE = average variance extracted.

4.3 Estimated PLS-SEM Model and Indicator Assessment

Figure 3 displays the same structural specification as Figure 1, together with the reported loading ranges and explained variance. The six AIPJBL loadings range from .750 to .874, the five PLE loadings from .832 to .875, the five TAIPC loadings from .800 to .870, and the six SSLO loadings from .758 to .837. Table 4 preserves all item-specific loadings. The product term is shown as constructed from both AIPJBL and TAIPC, not as an additional independent questionnaire construct.

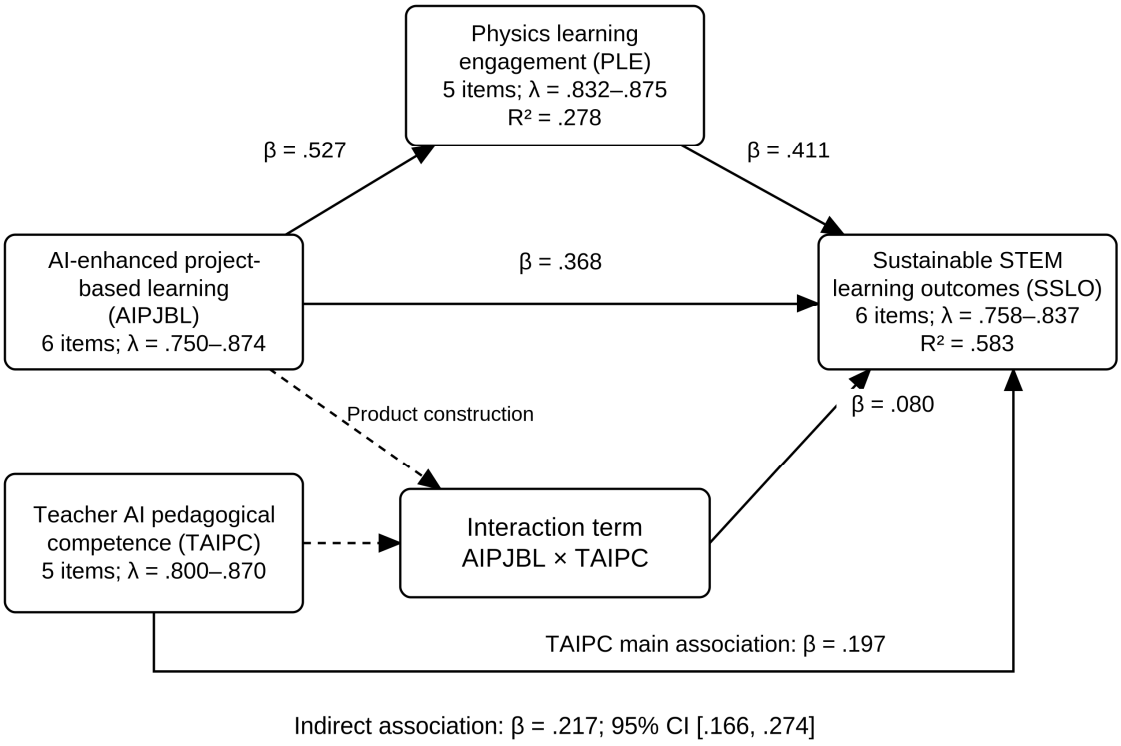


Figure 3. Estimated PLS-SEM Model with Reported Loading Ranges

4.4 Indicator Reliability and Collinearity

Table 4 provides the retained item-level results. Every reported loading exceeds .700, and indicator VIFs range from 1.750 to 2.854. AIPJBL4, representing evidence-based revision, has the strongest AIPJBL loading (.874), while AIPJBL6, representing AI-output verification, has the lowest (.750). The strongest PLE loading concerns peer explanation (PLE4, .875); PLE2 represents persistence after difficulty. These summaries concern indicator association and collinearity, not global model fit or the demonstration of learning.

Table 4. Indicator Loadings, Collinearity, and Descriptive Statistics

Indicator	Item focus	Loading	VIF	Mean	SD
AIPJBL1	AI-supported testable questions	0.827	2.240	3.169	1.192
AIPJBL2	Authentic sustainability problem	0.826	2.299	3.033	1.192
AIPJBL3	Comparison of design alternatives	0.814	2.121	3.070	1.192
AIPJBL4	Evidence-based revision	0.874	2.854	3.169	1.248
AIPJBL5	Integrated STEM practices	0.816	2.150	3.056	1.241
AIPJBL6	Verification of AI output	0.750	1.750	3.183	1.196
PLE1	Concentration	0.863	2.546	3.094	1.231
PLE2	Persistence after difficulty	0.840	2.250	3.082	1.223
PLE3	Enjoyment/interest	0.836	2.224	3.115	1.261
PLE4	Explanation to peers	0.875	2.719	3.106	1.271
PLE5	Initiative in planning	0.832	2.182	3.148	1.192
TAIPC1	Appropriate-use guidance	0.861	2.505	3.092	1.213
TAIPC2	Evaluation of AI output	0.870	2.628	3.122	1.182
TAIPC3	Physics-content alignment	0.800	1.945	3.063	1.209

Table 4. Indicator Loadings, Collinearity, and Descriptive Statistics (continued)

Indicator	Item focus	Loading	VIF	Mean	SD
TAIPC4	Privacy and equitable access	0.868	2.607	3.211	1.201
TAIPC5	Adaptive use of evidence	0.828	2.185	3.120	1.216
SSLO1	Physics concept application	0.826	2.259	3.122	1.244
SSLO2	Energy/system analysis	0.830	2.238	3.141	1.229
SSLO3	Evidence-based design	0.837	2.331	3.096	1.213
SSLO4	Sustainability trade-offs	0.758	1.777	3.143	1.214
SSLO5	Solution communication	0.837	2.358	3.056	1.205
SSLO6	Future STEM action orientation	0.777	1.917	3.148	1.168

4.5 Discriminant Validity and Construct Correlations

Table 5 separates the construct correlations from the square roots of AVE. Each square root of AVE exceeds the relevant pairwise correlations; for SSLO, .811 exceeds its correlations of .647 with AIPJBL and .644 with PLE. Table 6 then presents a full symmetric correlation matrix, with the diagonal set to 1.000. It is a clearer display of existing values, not a new correlation analysis. The reported HTMT values in Table 7 range from .207 to .720. These values are compatible with discriminant validity under the specified measurement model but do not establish its overall fit or rule out common-method influence.

Table 5. Fornell–Larcker Matrix: Square Roots of AVE and Construct Correlations

Construct	1	2	3	4
AIPJBL	.819			
PLE	.527	.849		
TAIPC	.319	.186	.846	
SSLO	.647	.644	.382	.811

Table 6. Construct Correlation Matrix

Construct	AIPJBL	PLE	TAIPC	SSLO
AIPJBL	1.000	.527	.319	.647
PLE	.527	1.000	.186	.644
TAIPC	.319	.186	1.000	.382
SSLO	.647	.644	.382	1.000

Table 7. Discriminant Validity: Heterotrait–Monotrait Ratio (HTMT)

Construct	1	2	3	4
AIPJBL				
PLE	.584			
TAIPC	.354	.207		
SSLO	.720	.714	.425	

Note. Construct order: 1 = AIPJBL; 2 = PLE; 3 = TAIPC; 4 = SSLO.

4.6 Structural Restrictions and Model Saturation

The outcome equation includes the three other constructs and their focal interaction; consequently, a substantial R^2 cannot by itself establish a well-fitting or distinctive theory. At the level of a conventional four-variable linear path model without the product term, allowing X and W to covary but fixing W → M to zero imposes one covariance restriction. There are 10 unique variance/covariance moments and nine free parameters: three exogenous variances/covariance, four paths, and two residual variances. This one-degree-of-freedom count assumes uncorrelated disturbances and is not the degrees of freedom of the 22-indicator PLS model or its moderation extension.

4.7 Path Coefficients and Effect Sizes

Table 8 summarises the retained direct, indirect, and total coefficients and f^2 values. AIPJBL has a reported f^2 of .385 for PLE. For SSLO, the reported incremental contributions are .292 for PLE, .218 for AIPJBL, .083 for TAIPC, and .015 for the product term. The reported AIPJBL–SSLO total coefficient is .585, with direct and indirect coefficients of .368 and .217. Because the direct association is moderated, this total is interpreted at the moderator reference value, not as a constant total association across all TAIPC values.

Table 8. Path Coefficients and Effect Sizes

Effect	Direct β	Indirect β	Total β	Cohen's f^2
AIPJBL → PLE	.527	-	.527	.385
AIPJBL → SSLO	.368	.217	.585	.218
PLE → SSLO	.411	-	.411	.292
TAIPC → SSLO	.197	-	.197	.083
AIPJBL × TAIPC → SSLO	.080	-	.080	.015

4.8 Explanatory and Predictive Assessment

Table 9 reports R^2 = .278 for PLE (adjusted R^2 = .277) and R^2 = .583 for SSLO (adjusted R^2 = .579). The reported Q^2predict values are .268 and .572, respectively. PLE prediction errors are RMSE = .856 and MAE = .721; SSLO errors are RMSE = .654 and MAE = .534.

Table 9. Explanatory and Predictive Assessment

Construct	R^2	Adjusted R^2	Q^2predict	RMSE	MAE
Physics learning engagement	.278	.277	.268	.856	.721
Sustainable STEM outcomes	.583	.579	.572	.654	.534

4.9 Structural Model and Conditional Associations

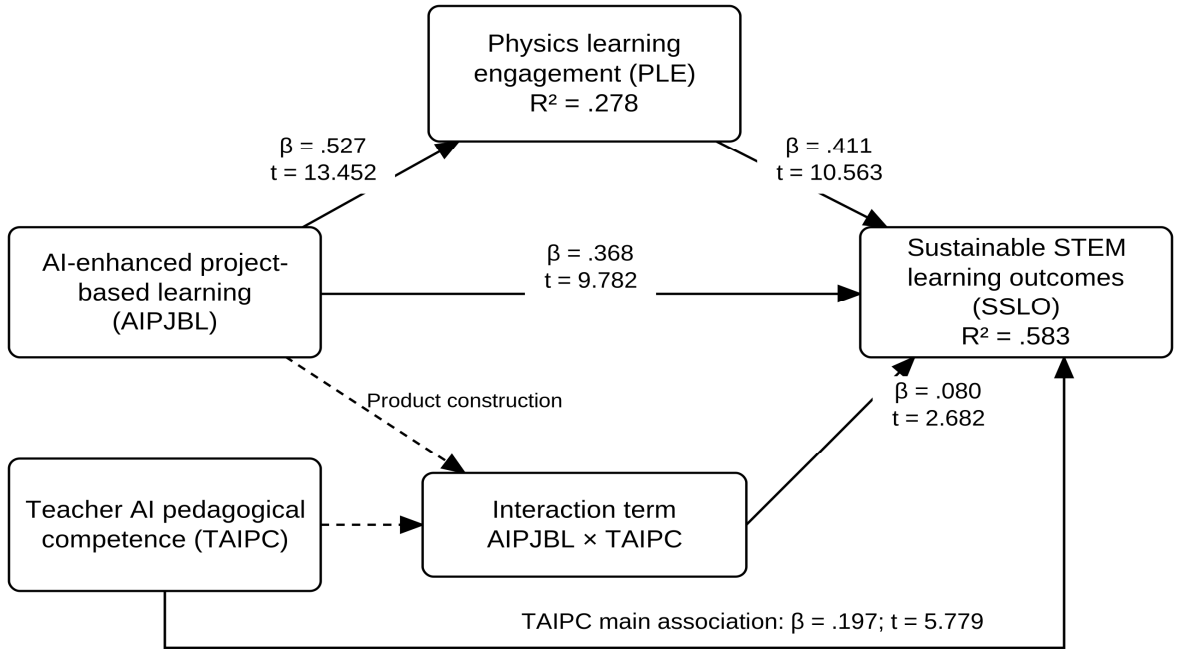


Figure 4. Structural Model with Reported Coefficients and Bootstrap T Statistics

Figure 4 displays the standardised coefficients, t statistics, and R^2 values with the same paths as Figures 1 and 3. The simple slopes are .288, .368, and .448 at the stated low, mean, and high TAIPC values. These are consistent with adding and subtracting the interaction coefficient from the AIPJBL coefficient.

4.10 Hypothesis Testing

Table 10 preserves the reported path estimates, bootstrap summaries, and hypothesis decisions. H1–H3 are positive, with reported coefficients of .368, .527, and .411. H4 has a reported indirect coefficient of .217 (95% CI [.166, .274]), and H5 has a reported interaction coefficient of .080 (p = .007; 95% CI [.022, .139]). TAIPC's main association is .197. "Supported" records the decision in the supplied analysis, not independently verified empirical support. The point estimates and bootstrap statistics have not been re-estimated.

Table 10. Structural Paths and Hypothesis Decisions

Path	β	SE	t	p	95% CI	Reported decision
H1: AIPJBL → SSLO	.368	.038	9.782	< .001	[.291, .440]	Supported
H2: AIPJBL → PLE	.527	.039	13.452	< .001	[.450, .605]	Supported
H3: PLE → SSLO	.411	.039	10.563	< .001	[.335, .488]	Supported
H4: AIPJBL → PLE → SSLO	.217	.027	7.950	< .001	[.166, .274]	Supported
H5: AIPJBL × TAIPC → SSLO	.080	.030	2.682	.007	[.022, .139]	Supported
Control/main effect: TAIPC → SSLO	.197	.034	5.779	< .001	[.129, .262]	-

Note. SE = reported bootstrap standard error; CI = confidence interval.

4.11 Qualitative Themes and Evidential Limits

The thematic summary contains five themes: just-in-time conceptual scaffolding; iteration that makes physics relationships visible; agency through local sustainability relevance; teacher orchestration; and responsible-AI tensions. Their content concerns restating concepts, comparing alternatives, interpreting measurements, participating in project decisions, evaluating generated claims, and managing copying, access, privacy, and prompt dependence.

Table 11 shows the thematic counts for traceability. The first three themes provide possible interpretations of the AIPJBL–PLE and AIPJBL–SSLO associations; the teacher-guidance theme concerns the proposed TAIPC condition, while responsible-AI tensions describe possible limitations.

Table 11. Retained Qualitative Themes and Reported Interview Counts

Theme	Thematic content	Reported records	Model connection
Just-in-time conceptual scaffolding	Restating concepts, question generation, comparing alternatives	21/24	AIPJBL → engagement/outcomes
Iteration made physics visible	Measurement, error recognition, redesign, causal explanation	20/24	Engagement → outcomes
Agency through local sustainability relevance	Choice, ownership, public purpose, team responsibility	19/24	AIPJBL → engagement
Teacher orchestration as a boundary condition	Verification, timing of support, inclusion, conceptual correction	18/24	TAIPC moderation
Responsible-AI tensions	Copying, device inequality, privacy, source uncertainty, prompt dependence	16/24	Negative cases/conditions

5. Discussion

5.1 AIPJBL and Perceived Sustainable STEM Capability

For RQ1, the reported pattern is compatible with students perceiving greater sustainability-oriented capability when they also perceive their projects as authentic, integrated, iterative, and evidence-centred. Project-based learning reviews and integrated STEM frameworks offer a rationale for connecting these features to disciplinary application (C.-H. Chen & Yang, 2019; Condliffe et al., 2017; English, 2016; Kelley & Knowles, 2016; Kokotsaki et al., 2016).

The sustainability dimension concerns the reasoning demanded by the task: explaining a mechanism, comparing alternatives, and considering environmental, technical, social, and cost consequences. These demands align with competency-oriented sustainability education (Brundiers et al., 2021; Kioupi & Voulvoulis, 2019). Yet the SSLO items record perceived capability, not independently assessed success at those tasks. A practical implication is to pair self-report with explanation-based assessment and evidence-based project rubrics; it is not a finding that students already demonstrated transfer or sustainability competence.

Nor should the reported pattern be read as evidence that more AI use produces better learning. This distinction matters because fluent generated text can conceal unresolved misconceptions or uncertainty about sources (Farrokhnia et al., 2024; Kasneci et al., 2023; Tlili et al., 2023). The relevant object of evaluation is students' reasoning within the designed learning process.

5.2 AIPJBL and Physics Engagement

For RQ2, the association is consistent with students being more involved when a physics project has a meaningful purpose and provides opportunities to make choices, receive feedback, and contribute to a team. Self-determination theory offers a rationale for considering these opportunities as potentially supportive of motivation; it does not make their psychological effects automatic (Jang et al., 2016; Ryan & Deci, 2020).

The themes of question development and iterative measurement provide possible contextual interpretations. They are compatible with research linking classroom engagement to a combination of challenge and support (Shernoff et al., 2016). For practice, the implication is to check whether AI assistance returns students to investigating a physics question or simply helps them finish an output.

Equity also requires attention to whose participation is supported. The same class may contain learners with different access to devices, language support, or confidence in questioning generated responses. The present model does not test these differences, so it cannot establish that AI reduced or increased inequality.

5.3 Engagement is Not Equivalent to Learning

For RQ3, students described by the reported pattern as more concentrated, persistent, and active in explaining physics also report greater sustainability-oriented capability. This interpretation is compatible with engagement as the learner's involvement in educational activity (Kahu & Nelson, 2018; Wong & Liem, 2022). It does not show that involvement caused conceptual understanding: knowledgeable students may find it easier to engage, and shared response tendencies may influence both ratings. The distinction between engagement and learning therefore remains central.

An instructional response would be to examine the quality of the explanation behind an activity. For example, a learner who proposes another sensor location creates an opportunity to reason about measurement, but understanding is shown through a justified explanation and defensible interpretation of the evidence, not the proposal alone. Research on agentic engagement motivates attention to such student contributions (Reeve & Shin, 2020).

Accordingly, assessment should distinguish task completion from disciplinary understanding. A student can contribute to a product or experiment with prompts without explaining why a physical system behaves as it does. Rubrics should therefore attend to causal explanation, measurement quality, uncertainty, and revision as well as the final artefact.

5.4 Interpreting the Indirect Association

For RQ4, the reported pattern is compatible with engagement linking perceived project quality and perceived capability. This interpretation accords with accounts of engagement within classroom support and motivation (Jang et al., 2016; Shernoff et al., 2016). However, all survey measures concern one post-unit occasion, so the pattern cannot establish which experience came first. It could also reflect reciprocal relationships or omitted influences. We therefore describe an indirect association, not proof that AI-enhanced projects improved learning by increasing engagement.

For future evaluation, tool-use records should be connected to what students do with the information. Two learners can spend comparable time with AI yet differ in whether they compare models, check a calculation, revise a test, or copy a conclusion. Prompt sequences, revision histories, dialogue, and observation could help distinguish these processes. Explainable learning analytics offers a related research direction, provided its interpretation and use of student data remain transparent (Khosravi et al., 2022).

The direct association also leaves room for explanations beyond engagement, including prior knowledge, feedback quality, representation quality, and the framing of sustainability problems. Those explanations were not measured as mediators here. Self-efficacy, epistemic vigilance, collaborative regulation, and AI literacy remain possible constructs for future study rather than mechanisms established by the current model.

5.5 The Role of Perceived Teacher Guidance

For RQ5, the reported pattern suggests that the relationship between project quality and perceived capability may be stronger when students also perceive more purposeful teacher guidance. The reported interaction is small; it does not establish a substantial educational benefit, a teacher-level effect, or the effectiveness of professional development. This bounded interpretation is consistent with Intelligent-TPACK and reviews that locate AI use within teacher planning, judgement, assessment, and ethical practice (Celik, 2023; Celik et al., 2022).

The pedagogical meaning concerns guidance that connects a generated suggestion to a physics concept and to evidence. Teachers may ask students to explain why a response is plausible, decide when direct teaching is preferable, and require physical evidence to be considered when it conflicts with generated text. Such practices are reasonable design implications of teacher-AI frameworks, not actions independently verified by the supplied records (Celik, 2023; UNESCO, 2024b). Student ratings of these practices must not be equated with direct assessments of teachers' expertise.

The separate association with perceived teacher guidance also cautions against treating AI as autonomous instruction. Professional learning could combine physics-task design, analysis of AI errors, verification dialogue, privacy scenarios, and assessment of student reasoning. UNESCO's teacher AI competency framework similarly connects human-centred practice, ethics, foundations, pedagogy, and professional learning (UNESCO, 2024b). Whether such development improves student outcomes requires evidence beyond the present self-report model.

5.6 Practical Meaning of the Framework

Taken together, the reported associations organise three questions for educational evaluation: what the project makes possible, how students engage with those opportunities, and how teacher guidance supports their reasoning. They do not rank independent interventions or demonstrate that increasing one questionnaire score will produce a corresponding learning gain.

For school improvement, these questions support examining the project, learner participation, and teacher guidance together. A review might consider whether the task is conceptually demanding, whether students justify revisions with evidence, and whether teachers have resources for appropriate support. This is a proposed diagnostic approach, not an intervention tested here. It is consistent with research identifying teacher and institutional conditions for integrated STEM implementation (Margot & Kettler, 2019; Shernoff et al., 2017).

5.7 Theoretical Contribution and Level of Measurement

We make a conceptual contribution by specifying AIPJBL as pedagogical quality rather than AI access, distinguishing engagement from perceived capability, and representing teacher guidance through both a main association and an interaction. The framework connects constructivist project design, engagement, Intelligent-TPACK, and sustainability education.

The framework also distinguishes the referent of a construct from its level of measurement. AIPJBL and TAIPC concern experiences of classroom practice, but both are measured through student reports, as are PLE and SSLO. Consequently, the model cannot establish classroom-level implementation effects or independently measured teacher competence. Recognising these boundaries clarifies which additional observations, assessments, and multilevel analyses would be needed.

5.8 Mixed-Methods Interpretation

The qualitative themes offer possible meanings for the reported associations: questioning and iteration concern the use of project opportunities, local relevance concerns reasons to participate, and teacher guidance concerns the evaluation of generated information.

The responsible-AI theme also preserves issues that could challenge an exclusively positive interpretation, including copying, unequal access, privacy, source uncertainty, and prompt dependence. A negative or discrepant case is relevant when it contradicts a proposed explanation, not merely because a participant mentions a difficulty.

6. Conclusion

We specify a model connecting AIPJBL quality, physics engagement, student-perceived teacher AI pedagogical competence, and self-reported sustainable STEM capability. The numerical pattern is consistent with direct, indirect, and conditional associations, but it does not establish objective learning gains, causal mediation, or AI-specific effectiveness. The educational proposition is nevertheless clear: AI assistance should be integrated into physics inquiry

in ways that leave students responsible for questions, evidence, explanations, and design decisions. Sustainability-oriented projects provide contexts in which these responsibilities can be made explicit, while teacher guidance can support verification and inclusion.

7. Implications of the Study

Theoretical implications. The framework combines constructivist project design, engagement theory, Intelligent-TPACK, and sustainability competence to specify distinct pedagogical, learner, and teacher-related relationships. It does not establish why identical tools produce different outcomes. Future theories and empirical designs should test these proposed indirect and conditional relationships while distinguishing perceived capability from demonstrated learning.

Pedagogical implications. Middle-school physics teachers should begin with a sustainability problem and a physics objective, then decide whether AI adds a useful scaffold. Productive uses include generating alternative hypotheses, comparing explanations, organising data, and critiquing drafts. Students should maintain prompt and verification logs, identify AI contributions, and justify every claim with physical evidence. Assessment should combine conceptual questions, project performance, notebooks, design rationales, and reflection so that polished language does not conceal weak understanding.

Professional-development implications. Teacher learning should include subject-specific AI error analysis, prompt design, formative assessment, privacy, bias, accessibility, and strategies for maintaining productive struggle. Professional learning communities can co-design project pathways and build shared libraries of verified prompts, common misconceptions, and non-AI alternatives. Competence should be evaluated through classroom planning and student reasoning, not through confidence with a particular platform.

Leadership and policy implications. Schools need approved tools, clear data-governance rules, device-sharing plans, technical support, and protected collaboration time. Equity audits should examine who has access, who participates in group decisions, and whose language or cultural knowledge is represented in project criteria. Procurement should consider transparency, age appropriateness, data retention, accessibility, and the ability to disable unnecessary data collection. Policies should preserve teacher and student agency rather than mandate usage quotas.

Sustainability implications. Project criteria should include environmental and social consequences alongside technical performance. Students can be asked to compare embodied resources, energy savings, affordability, maintenance, and user behaviour. Such tasks support the transition from isolated environmental facts to systems-based decision making and align physics learning with the broader goals of education for sustainable development (UNESCO, 2020, 2024c; Vare et al., 2019).

8. Limitations and Future Research Directions

Several limitations constrain interpretation. First, the cross-sectional post-unit specification does not establish temporal ordering or causality and cannot rule out prior achievement, selection, school resources, or pre-existing engagement. Second, all quantitative constructs use student self-report. Third, the analysis does not model students' nesting in classes and schools. Finally, neither a global fit export nor the information needed to verify interaction scaling and held-out prediction was supplied. These limitations are not resolved by significant coefficients or favourable reliability statistics.

Stronger causal inference would require longitudinal, quasi-experimental, or cluster-randomised evidence with pre-test and post-test measures. Repeated measurement could also examine how engagement changes across project phases and whether novelty effects persist. These extensions are recommended as future research rather than treated as procedures completed in the present study.

The reliance on student self-report also limits conclusions about objective physics achievement and observed teacher competence. Future research should combine surveys with physics concept assessments, performance tasks, observations, project-artefact rubrics, prompt logs, and teacher measures. TAIPC could be assessed through multiple perspectives, including teacher self-assessment, lesson-plan analysis, and observation, while longer-term follow-up could examine transfer to new problems.

The structural specification is simple but only weakly restrictive in its four-construct linear component. Future analysis should compare theoretically justified alternatives rather than equate parsimony with confirmed fit. AI literacy, self-efficacy, epistemic vigilance, collaborative regulation, prior achievement, socioeconomic resources, and school climate are possible additional influences. Qualitative work should also include accounts that challenge expected

benefits, including refusal, restricted access, and experiences of controlling assistance.

Finally, tools and policies change rapidly. Findings tied to one platform may become obsolete, so studies should describe functional affordances, data practices, and pedagogical routines rather than only brand names. Comparative research could test different forms of AI support, including conversational systems, simulations, learning analytics, and teacher-facing tools. The most important future question is not whether AI is present, but which human-AI arrangements strengthen disciplinary reasoning, sustainability competence, and educational justice.

References

Akgun, S., & Greenhow, C. (2022). Artificial intelligence in education: Addressing ethical challenges in K–12 settings. *AI and Ethics, 2*, 431–440. https://doi.org/10.1007/s43681-021-00096-7

Almulla, M. A. (2020). The effectiveness of the project-based learning (PJBL) approach as a way to engage students in learning. *SAGE Open, 10*(3), Article 2158244020938702. https://doi.org/10.1177/2158244020938702

Bond, M., Buntins, K., Bedenlier, S., Zawacki-Richter, O., & Kerres, M. (2020). Mapping research in student engagement and educational technology in higher education: A systematic evidence map. *International Journal of Educational Technology in Higher Education, 17*, Article 2. https://doi.org/10.1186/s41239-019-0176-8

Bond, M., Khosravi, H., De Laat, M., Bergdahl, N., Negrea, V., Oxley, E., Pham, P., Chong, S. W., & Siemens, G. (2024). A meta systematic review of artificial intelligence in higher education: A call for increased ethics, collaboration, and rigour. *International Journal of Educational Technology in Higher Education, 21*, Article 4. https://doi.org/10.1186/s41239-023-00436-z

Brundiers, K., Barth, M., Cebrián, G., Cohen, M., Diaz, L., Doucette-Remington, S., Dripps, W., Habron, G., Harré, N., Jarchow, M., Losch, K., Michel, J., Mochizuki, Y., Rieckmann, M., Parnell, R., Walker, P., & Zint, M. (2021). Key competencies in sustainability in higher education: Toward an agreed-upon reference framework. *Sustainability Science, 16*, 13–29. https://doi.org/10.1007/s11625-020-00838-2

Celik, I. (2023). Towards Intelligent-TPACK: An empirical study on teachers' professional knowledge to ethically integrate artificial intelligence-based tools into education. *Computers in Human Behavior, 138*, Article 107468. https://doi.org/10.1016/j.chb.2022.107468

Celik, I., Dindar, M., Muukkonen, H., & Järvelä, S. (2022). The promises and challenges of artificial intelligence for teachers: A systematic review of research. *TechTrends, 66*, 616–630. https://doi.org/10.1007/s11528-022-00715-y

Chassignol, M., Khoroshavin, A., Klimova, A., & Bilyatdinova, A. (2018). Artificial intelligence trends in education: A narrative overview. *Procedia Computer Science, 136*, 16–24. https://doi.org/10.1016/j.procs.2018.08.233

Chen, C.-H., & Yang, Y.-C. (2019). Revisiting the effects of project-based learning on students' academic achievement: A meta-analysis investigating moderators. *Educational Research Review, 26*, 71–81. https://doi.org/10.1016/j.edurev.2018.11.001

Chen, L., Chen, P., & Lin, Z. (2020). Artificial intelligence in education: A review. *IEEE Access, 8*, 75264–75278. https://doi.org/10.1109/ACCESS.2020.2988510

Chiu, T. K. F., Xia, Q., Zhou, X., Chai, C. S., & Cheng, M. (2023). Systematic literature review on opportunities, challenges, and future research recommendations of artificial intelligence in education. *Computers and Education: Artificial Intelligence, 4*, Article 100118. https://doi.org/10.1016/j.caeai.2022.100118

Condliffe, B., Quint, J., Visher, M. G., Bangser, M. R., Drohojowska, S., Saco, L., & Nelson, E. (2017). *Project-based learning: A literature review*. MDRC.

Dare, E. A., Ellis, J. A., & Roehrig, G. H. (2018). Understanding science teachers' implementations of integrated STEM curricular units through a phenomenological multiple case study. *International Journal of STEM Education, 5*, Article 4. https://doi.org/10.1186/s40594-018-0101-z

English, L. D. (2016). STEM education K–12: Perspectives on integration. *International Journal of STEM Education, 3*, Article 3. https://doi.org/10.1186/s40594-016-0036-1

Estapa, A. T., & Tank, K. M. (2017). Supporting integrated STEM in the elementary classroom: A professional development approach centered on an engineering design challenge. *International Journal of STEM Education, 4*, Article 6. https://doi.org/10.1186/s40594-017-0058-3

Falloon, G. (2020). From digital literacy to digital competence: The teacher digital competency framework. *Educational Technology Research and Development, 68*, 2449–2472. https://doi.org/10.1007/s11423-020-09767-4

Falloon, G., Hatzigianni, M., Bower, M., Forbes, A., & Stevenson, M. (2020). Understanding K–12 STEM education: A framework for developing STEM literacy. *Journal of Science Education and Technology, 29*, 369–385. https://doi.org/10.1007/s10956-020-09823-x

Farrokhnia, M., Banihashem, S. K., Noroozi, O., & Wals, A. (2024). A SWOT analysis of ChatGPT: Implications for educational practice and research. *Innovations in Education and Teaching International, 61*(3), 460–474. https://doi.org/10.1080/14703297.2023.2195846

Fredricks, J. A., Wang, M.-T., Schall Linn, J., Hofkens, T. L., Sung, H., Parr, A., & Allerton, J. (2016). Using qualitative methods to develop a survey measure of math and science engagement. *Learning and Instruction, 43*, 5–15. https://doi.org/10.1016/j.learninstruc.2016.01.009

Fuller, C. M., Simmering, M. J., Atinc, G., Atinc, Y., & Babin, B. J. (2016). Common methods variance detection in business research. *Journal of Business Research, 69*(8), 3192–3198. https://doi.org/10.1016/j.jbusres.2015.12.008

Grassini, S. (2023). Shaping the future of education: Exploring the potential and consequences of AI and ChatGPT in educational settings. *Education Sciences, 13*(7), Article 692. https://doi.org/10.3390/educsci13070692

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*, Article 101586. https://doi.org/10.1016/j.ijer.2020.101586

Hair, J. F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2022). *A primer on partial least squares structural equation modeling (PLS-SEM)* (3rd ed.). SAGE.

Hair, J. F., Risher, J. J., Sarstedt, M., & Ringle, C. M. (2019). When to use and how to report the results of PLS-SEM. *European Business Review, 31*(1), 2–24. https://doi.org/10.1108/EBR-11-2018-0203

Henseler, J., Hubona, G., & Ray, P. A. (2016). Using PLS path modeling in new technology research: Updated guidelines. *Industrial Management & Data Systems, 116*(1), 2–20. https://doi.org/10.1108/IMDS-09-2015-0382

Holmes, W., Bialik, M., & Fadel, C. (2019). *Artificial intelligence in education: Promises and implications for teaching and learning*. Center for Curriculum Redesign.

Holmlund, T. D., Lesseig, K., & Slavit, D. (2018). Making sense of "STEM education" in K–12 contexts. *International Journal of STEM Education, 5*, Article 32. https://doi.org/10.1186/s40594-018-0127-2

Hwang, G.-J., Xie, H., Wah, B. W., & Gašević, D. (2020). Vision, challenges, roles and research issues of artificial intelligence in education. *Computers and Education: Artificial Intelligence, 1*, Article 100001. https://doi.org/10.1016/j.caeai.2020.100001

Jang, H., Kim, E.-J., & Reeve, J. (2016). Why students become more engaged or more disengaged during the semester: A self-determination theory dual-process model. *Learning and Instruction, 43*, 27–38. https://doi.org/10.1016/j.learninstruc.2016.01.002

Jordan, P. J., & Troth, A. C. (2020). Common method bias in applied settings: The dilemma of researching in organizations. *Australian Journal of Management, 45*(1), 3–14. https://doi.org/10.1177/0312896219871976

Kahu, E. R., & Nelson, K. (2018). Student engagement in the educational interface: Understanding the mechanisms of student success. *Higher Education Research & Development, 37*(1), 58–71. https://doi.org/10.1080/07294360.2017.1344197

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

Kelley, T. R., & Knowles, J. G. (2016). A conceptual framework for integrated STEM education. *International Journal of STEM Education, 3*, Article 11. https://doi.org/10.1186/s40594-016-0046-z

Khosravi, H., Shum, S. B., Chen, G., Conati, C., Tsai, Y.-S., Kay, J., Knight, S., Martinez-Maldonado, R., Sadiq, S., & Gašević, D. (2022). Explainable artificial intelligence in education. *Computers and Education: Artificial Intelligence, 3*, Article 100074. https://doi.org/10.1016/j.caeai.2022.100074

Kioupi, V., & Voulvoulis, N. (2019). Education for sustainable development: A systemic framework for connecting the SDGs to educational outcomes. *Sustainability, 11*(21), Article 6104. https://doi.org/10.3390/su11216104

Kock, N., & Hadaya, P. (2018). Minimum sample size estimation in PLS-SEM: The inverse square root and gamma-exponential methods. *Information Systems Journal, 28*(1), 227–261. https://doi.org/10.1111/isj.12131

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

Leicht, A., Heiss, J., & Byun, W. J. (Eds.). (2018). *Issues and trends in education for sustainable development*. UNESCO.

Li, Y., Wang, K., Xiao, Y., & Froyd, J. E. (2020). Research and trends in STEM education: A systematic review of journal publications. *International Journal of STEM Education, 7*, Article 11. https://doi.org/10.1186/s40594-020-00207-6

Lo, C. K. (2023). What is the impact of ChatGPT on education? A rapid review of the literature. *Education Sciences, 13*(4), Article 410. https://doi.org/10.3390/educsci13040410

Lozano, R., Merrill, M. Y., Sammalisto, K., Ceulemans, K., & Lozano, F. J. (2017). Connecting competences and pedagogical approaches for sustainable development in higher education: A literature review and framework proposal. *Sustainability, 9*(10), Article 1889. https://doi.org/10.3390/su9101889

Luckin, R., Holmes, W., Griffiths, M., & Forcier, L. B. (2016). *Intelligence unleashed: An argument for AI in education*. Pearson.

Margot, K. C., & Kettler, T. (2019). Teachers' perception of STEM integration and education: A systematic literature review. *International Journal of STEM Education, 6*, Article 2. https://doi.org/10.1186/s40594-018-0151-2

Martín-Páez, T., Aguilera, D., Perales-Palacios, F. J., & Vílchez-González, J. M. (2019). What are we talking about when we talk about STEM education? A review of literature. *Science Education, 103*(4), 799–822. https://doi.org/10.1002/sce.21522

Miao, F., Holmes, W., Huang, R., & Zhang, H. (2021). *AI and education: Guidance for policy-makers*. UNESCO.

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

Olsson, D., Gericke, N., & Chang Rundgren, S.-N. (2016). The effect of implementation of education for sustainable development in Swedish compulsory schools: Assessing pupils' sustainability consciousness. *Environmental Education Research, 22*(2), 176–202. https://doi.org/10.1080/13504622.2015.1005057

Ortiz-Revilla, J., Adúriz-Bravo, A., & Greca, I. M. (2020). A framework for epistemological discussion on integrated STEM education. *Science & Education, 29*, 857–880. https://doi.org/10.1007/s11191-020-00131-9

Ouyang, F., & Jiao, P. (2021). Artificial intelligence in education: The three paradigms. *Computers and Education: Artificial Intelligence, 2*, Article 100020. https://doi.org/10.1016/j.caeai.2021.100020

Perignat, E., & Katz-Buonincontro, J. (2019). STEAM in practice and research: An integrative literature review. *Thinking Skills and Creativity, 31*, 31–43. https://doi.org/10.1016/j.tsc.2018.10.002

Quigley, C. F., Herro, D., & Jamil, F. M. (2017). Developing a conceptual model of STEAM teaching practices. *School Science and Mathematics, 117*(1–2), 1–12. https://doi.org/10.1111/ssm.12201

Rahman, M. M., & Watanobe, Y. (2023). ChatGPT for education and research: Opportunities, threats, and strategies. *Applied Sciences, 13*(9), Article 5783. https://doi.org/10.3390/app13095783

Redecker, C. (2017). *European framework for the digital competence of educators: DigCompEdu*. Publications Office of the European Union. https://doi.org/10.2760/159770

Redman, A., Wiek, A., & Barth, M. (2021). Current practice of assessing students' sustainability competencies: A review of tools. *Sustainability Science, 16*, 117–135. https://doi.org/10.1007/s11625-020-00855-1

Reeve, J., & Shin, S. H. (2020). How teachers can support students' agentic engagement. *Theory Into Practice, 59*(2), 150–161. https://doi.org/10.1080/00405841.2019.1702451

Roll, I., & Wylie, R. (2016). Evolution and revolution in artificial intelligence in education. *International Journal of Artificial Intelligence in Education, 26*, 582–599. https://doi.org/10.1007/s40593-016-0110-3

Ryan, R. M., & Deci, E. L. (2020). Intrinsic and extrinsic motivation from a self-determination theory perspective: Definitions, theory, practices, and future directions. *Contemporary Educational Psychology, 61*, Article 101860. https://doi.org/10.1016/j.cedpsych.2020.101860

Sass, W., Boeve-de Pauw, J., De Maeyer, S., & Van Petegem, P. (2020). Redefining action competence: The case of sustainable development. *The Journal of Environmental Education, 51*(4), 292–305. https://doi.org/10.1080/00958964.2020.1765132

Schmid, M., Brianza, E., & Petko, D. (2020). Developing a short assessment instrument for technological pedagogical content knowledge (TPACK.xs) and comparing the factor structure of an integrative and a transformative model. *Computers & Education, 157*, Article 103967. https://doi.org/10.1016/j.compedu.2020.103967

Shernoff, D. J., Kelly, S., Tonks, S. M., Anderson, B., Cavanagh, R. F., Sinha, S., & Abdi, B. (2016). Student engagement as a function of environmental complexity in high school classrooms. *Learning and Instruction, 43*, 52–60. https://doi.org/10.1016/j.learninstruc.2015.12.003

Shernoff, D. J., Sinha, S., Bressler, D. M., & Ginsburg, L. (2017). Assessing teacher education and professional development needs for the implementation of integrated approaches to STEM education. *International Journal of STEM Education, 4*, Article 13. https://doi.org/10.1186/s40594-017-0068-1

Shmueli, G., Sarstedt, M., Hair, J. F., Cheah, J.-H., Ting, H., Vaithilingam, S., & Ringle, C. M. (2019). Predictive model assessment in PLS-SEM: Guidelines for using PLSpredict. *European Journal of Marketing, 53*(11), 2322–2347. https://doi.org/10.1108/EJM-02-2019-0189

Struyf, A., De Loof, H., Boeve-de Pauw, J., & Van Petegem, P. (2019). Students' engagement in different STEM learning environments: Integrated STEM education as promising practice? *International Journal of Science Education, 41*(10), 1387–1407. https://doi.org/10.1080/09500693.2019.1607983

Thibaut, L., Ceuppens, S., De Loof, H., De Meester, J., Goovaerts, L., Struyf, A., Boeve-de Pauw, J., Dehaene, W., Deprez, J., De Cock, M., Hellinckx, L., Knipprath, H., Langie, G., Struyven, K., Van de Velde, D., Van Petegem, P., & Depaepe, F. (2018). Integrated STEM education: A systematic review of instructional practices in secondary education. *European Journal of STEM Education, 3*(1), Article 2. https://doi.org/10.20897/ejsteme/85525

Tlili, A., Shehata, B., Adarkwah, M. A., Bozkurt, A., Hickey, D. T., Huang, R., & Agyemang, B. (2023). What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. *Smart Learning Environments, 10*, Article 15. https://doi.org/10.1186/s40561-023-00237-x

UNESCO. (2017). *Education for sustainable development goals: Learning objectives*.

UNESCO. (2020). *Education for sustainable development: A roadmap*.

UNESCO. (2021). *Reimagining our futures together: A new social contract for education*.

UNESCO. (2023). *Guidance for generative AI in education and research*.

UNESCO. (2024a). *AI competency framework for students*.

UNESCO. (2024b). *AI competency framework for teachers*.

UNESCO. (2024c). *Greening curriculum guidance: Teaching and learning for climate action*.

Valtonen, T., Sointu, E., Kukkonen, J., Kontkanen, S., Lambert, M. C., & Mäkitalo-Siegl, K. (2017). TPACK updated to measure pre-service teachers' twenty-first century skills. *Australasian Journal of Educational Technology, 33*(3), 15–31. https://doi.org/10.14742/ajet.3518

Vare, P., Arro, G., de Hamer, A., Del Gobbo, G., de Vries, G., Farioli, F., Kadji-Beltran, C., Kangur, M., Mayer, M., Millican, R., Nijdam, C., Réti, M., & Zachariou, A. (2019). Devising a competence-based training program for educators of sustainable development: Lessons learned. *Sustainability, 11*(7), Article 1890. https://doi.org/10.3390/su11071890

Wals, A. E. J., & Benavot, A. (2017). Can we meet the sustainability challenges? The role of education and lifelong learning. *European Journal of Education, 52*(4), 404–413. https://doi.org/10.1111/ejed.12250

Wang, M.-T., Fredricks, J. A., Ye, F., Hofkens, T. L., & Schall Linn, J. (2016). The math and science engagement scales: Scale development, validation, and psychometric properties. *Learning and Instruction, 43*, 16–26. https://doi.org/10.1016/j.learninstruc.2016.01.008

Willermark, S. (2018). Technological pedagogical and content knowledge: A review of empirical studies published from 2011 to 2016. *Journal of Educational Computing Research, 56*(3), 315–343. https://doi.org/10.1177/0735633117713114

Williamson, B., & Eynon, R. (2020). Historical threads, missing links, and future directions in AI in education. *Learning, Media and Technology, 45*(3), 223–235. https://doi.org/10.1080/17439884.2020.1798995

Wong, Z. Y., & Liem, G. A. D. (2022). Student engagement: Current state of the construct, conceptual refinement, and future research directions. *Educational Psychology Review, 34*, 107–138. https://doi.org/10.1007/s10648-021-09628-3

Xu, W., & Zammit, K. (2020). Applying thematic analysis to education: A hybrid approach to interpreting data in practitioner research. *International Journal of Qualitative Methods, 19*, 1–9. https://doi.org/10.1177/1609406920918810

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education: Where are the educators? *International Journal of Educational Technology in Higher Education, 16*, Article 39. https://doi.org/10.1186/s41239-019-0171-0

Zhang, K., & Aslan, A. B. (2021). AI technologies for education: Recent research and future directions. *Computers and Education: Artificial Intelligence, 2*, Article 100025. https://doi.org/10.1016/j.caeai.2021.100025

Appendix A: Questionnaire Items

Table A1. Questionnaire Items

Code	Questionnaire item
AIPJBL1	AI tools helped my group turn the project problem into testable physics questions.
AIPJBL2	The project was connected to a real sustainability problem in our school or community.
AIPJBL3	AI feedback helped us compare alternative explanations or design ideas.
AIPJBL4	Our group used measurements and evidence to revise the project.
AIPJBL5	The project integrated physics, engineering design, mathematics, and digital tools.
AIPJBL6	We verified AI-generated information using physics principles, measurements, or reliable sources.
PLE1	I concentrated carefully during the physics project.
PLE2	I kept working when our design or explanation did not work at first.
PLE3	I enjoyed investigating the physics problem.
PLE4	I explained physics ideas and evidence to other members of my group.
PLE5	I took initiative by suggesting questions, tests, or improvements.
TAIPC1	My teacher clearly explained when and how AI should be used in the project.
TAIPC2	My teacher helped us judge whether AI output was accurate and trustworthy.
TAIPC3	My teacher connected AI feedback to the physics concepts we were learning.
TAIPC4	My teacher protected privacy and ensured fair access to AI tools.
TAIPC5	My teacher used our project evidence and questions to adjust instruction.
SSLO1	I can use physics concepts to explain how a sustainability technology works.
SSLO2	I can analyse energy use, efficiency, or resource flows in a real system.
SSLO3	I can design and improve a solution using measurements and scientific evidence.
SSLO4	I can compare technical, environmental, social, and cost trade-offs.
SSLO5	I can communicate and defend a sustainability solution using STEM evidence.
SSLO6	I am prepared to use STEM knowledge to address future sustainability problems.

Response scale: 1 = strongly disagree, 2 = disagree, 3 = neither agree nor disagree, 4 = agree, 5 = strongly agree.

Appendix B: Semi-Structured Interview

The ten questions below are student-facing prompts.

Table B1. Semi-Structured Interview Guide

No.	Interview question
1	Describe the project problem and why it mattered to your school or community.
2	At which points did AI help your group move forward, and what exactly did it help you do?
3	Describe a time when an AI response was unclear, inaccurate, or not useful. How did you respond?
4	How did your group use measurements, physics principles, or other sources to verify AI-generated information?
5	Which project activities made you most and least engaged in learning physics? Why?
6	How did working with AI affect your confidence, persistence, or willingness to ask questions?
7	How did your teacher guide AI use, physics reasoning, teamwork, and revision?
8	Did all students have a fair opportunity to use the tools and contribute to decisions? Explain.
9	What did the project teach you about sustainability trade-offs, such as cost, energy, materials, or user needs?
10	What should be changed in a future AI-enhanced physics project, and what should remain the same?

Author Statements

Acknowledgments

The authors sincerely thank the participating schools, teachers, and students for their cooperation and contributions to this study. The authors also appreciate the support provided during the implementation of the AI-enhanced project-based learning activities and data collection.

Author Contributions

Conceptualization: Shiping Wang (https://orcid.org/0009-0005-2859-2997) and Hashimah Mohd Yunus (https://orcid.org/0009-0009-8200-2053)

Methodology: Shiping Wang (https://orcid.org/0009-0005-2859-2997) and Hashimah Mohd Yunus (https://orcid.org/0009-0009-8200-2053)

Formal analysis: Shiping Wang (https://orcid.org/0009-0005-2859-2997)

Investigation: Shiping Wang (https://orcid.org/0009-0005-2859-2997)

Data curation: Shiping Wang (https://orcid.org/0009-0005-2859-2997)

Visualization: Shiping Wang (https://orcid.org/0009-0005-2859-2997)

Writing – original draft: Shiping Wang (https://orcid.org/0009-0005-2859-2997)

Writing – review & editing: Shiping Wang (https://orcid.org/0009-0005-2859-2997) and Hashimah Mohd Yunus (https://orcid.org/0009-0009-8200-2053)

Supervision: Hashimah Mohd Yunus (https://orcid.org/0009-0009-8200-2053)

Project administration: Shiping Wang (https://orcid.org/0009-0005-2859-2997)

All authors have read and approved the final version of the manuscript.

Consent for Publication

All authors have read and approved the published version of the manuscript.

Funding

This research received no external funding.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Informed Consent

Informed consent procedures were implemented for all participants involved in the study. For student participants who were minors, parental or legal guardian consent and student assent were obtained prior to participation. Participation was voluntary, and participants were informed that they could withdraw from the study without penalty.

Ethics Approval

Ethical approval for this study was granted under Reference No. HEC 2026/585 on 19 October 2026, with approval valid until April 2028. The study was conducted in accordance with the approved ethical procedures and applicable principles for research involving human participants.

Data Availability Statement

The datasets analyzed during the current study are not publicly available because they contain research data involving school-age participants and are subject to ethical, confidentiality, and privacy restrictions. The data may be made available by the corresponding author upon reasonable request, subject to applicable ethical, institutional, and data-protection requirements.

Use of Artificial Intelligence

The authors used ChatGPT 5.6 solely for language editing and grammar improvement during manuscript preparation. All AI-assisted revisions were reviewed and verified by the authors, who take full responsibility for the accuracy, integrity, and final content of the manuscript.

Supplementary Materials

No supplementary materials are available.

Journal Information

Ethics Statement

The journal's policies adhere to the Core Practices established by the Committee on Publication Ethics (COPE).

Preservation Policy

This journal is archived by Library and Archives Canada (LAC).

Open Access

This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC BY 4.0). Please find details at https://creativecommons.org/licenses/by/4.0.

Copyright © The Author(s)

Copyright remains with the author(s). The author(s) grant the journal the right of first publication, and the article is published under the terms of the Creative Commons Attribution (CC BY 4.0) License.