Jurnal Pendidikan Ilmu Pengetahuan Alam (JP-IPA) **Vol. 07 No. 03. Agustus 2026**
https://jurnal.habi.ac.id/index.php/JP-IPA *p-ISSN:* 2774-5945. *e-ISSN:* 2774-5937
Doi: https://doi.org/10.56842/jp-ipa Terakreditasi SINTA 3

FROM SOLAR CELLS TO SUSTAINABLE FUTURES: DYE-SENSITIZED SOLAR CELL EXPERIMENTATION AND GENERATIVE AI FOR STEM-ESD LITERACY

Arip Nurahman[1*], Ida Kaniawati[1], Eka Cahya Prima[1], and Endi Suhendi[2]
[1]Department of Science Education, Universitas Pendidikan Indonesia, Bandung, Indonesia.
[2]Department of Physics Education, Universitas Pendidikan Indonesia, Bandung, Indonesia.
Email: aripnurahman18@upi.edu

Diterima: 07 Juli 2026 Direvisi: 20 Agustus 2026 Publikasi: 30 Agustus 2026

Abstract

This study examined how an instructional program integrating hands-on dye-sensitized solar cell (DSSC) experimentation with critically guided generative artificial intelligence (GenAI) inquiry influences university students' STEM–Education for Sustainable Development (STEM–ESD) literacy. An explanatory sequential mixed-methods design with an embedded quasi-experimental component was employed. Participants were 64 undergraduate physics education students in Bandung, Indonesia, allocated by intact class sections into an intervention group (n = 32) and a comparison group (n = 32). Across six weekly sessions the intervention group framed an energy problem, fabricated and characterized DSSCs sensitized with natural dyes, interrogated GenAI-generated explanations against their own measurements, and redesigned their cells using sustainability criteria. Data were collected through a validated five-dimension STEM–ESD literacy test administered before and after the program, laboratory notebooks, reflective journals, GenAI interaction logs, group discussions, and semi-structured interviews. Quantitative data were analyzed using baseline-adjusted regression models and effect-size estimation, qualitative data through thematic and epistemic coding with an epistemic-network representation, and the two strands were integrated using joint displays. Illustrative findings indicate the largest gains in sustainability literacy and evidence-evaluation practices, explained by an iterative cycle in which AI-generated explanations were treated as claims requiring verification rather than as authoritative answers.

Keywords: *Dye-Sensitized Solar Cell; Generative Artificial Intelligence; Scientific Inquiry; STEM–ESD Literacy; Sustainability Education.*

INTRODUCTION

Contemporary societies depend on scientific and technological knowledge to address interconnected challenges of climate change, energy transition, environmental degradation, and sustainable development. Responding to these challenges requires learners who can not only recall scientific concepts but also evaluate evidence, weigh possible solutions, and decide under uncertainty. The PISA 2025 science framework reflects this orientation by foregrounding the capacity to explain phenomena scientifically, to construct and evaluate scientific investigations, to interpret evidence critically, and to research and evaluate scientific information for decision making and action; it further specifies environmental competencies concerning informed decision making, systems thinking, multiple perspectives, and agency in socio-ecological contexts (OECD, 2025). Comparable commitments underpin the global roadmap for education for sustainable development (UNESCO, 2021) and the practice-based vision of science learning articulated in the framework for K–12 science education (National Research Council, 2012).

Renewable energy technologies offer a productive context for enacting this vision at university level. Solar cells in particular connect solid-state physics, chemistry, materials science, electrical engineering, and social questions about energy access and equity. Among photovoltaic technologies, dye-sensitized solar cells (DSSCs) are unusually well suited to teaching because they can be fabricated from accessible materials and characterized with modest instrumentation, allowing students to carry out dye extraction, semiconductor preparation, cell assembly, and current–voltage measurement themselves (Chien et al., 2021; Watwe et al., 2025). Reported implementations range from single-session demonstrations to multi-week research-based laboratories in which students synthesize every component in house (Sichula, 2023; Wang et al., 2024; Watwe et al., 2025), and reviews of solar energy education argue that such activities align technical training with sustainable development goals (Senthil, 2022; Septiyanto & Prima, 2024). The Indonesian context is particularly favourable because locally available natural sensitizers such as anthocyanin extracts have been characterized extensively for DSSC use (Prima et al., 2020), which makes the device scientifically authentic and materially accessible at the same time.

Integrated STEM education has become a mainstream policy agenda, yet the way STEM is conceptualized and measured remains largely additive: science, technology, engineering, and mathematics are assessed as four separable domains and then summed, rather than treated as a single network of epistemic practices that sustain one another during authentic problem solving (Bybee, 2013; English, 2016; Kelley & Knowles, 2016). Toma et al. (2025) reported that science teachers themselves identify conceptual and structural barriers to enacting genuinely integrated instruction, partly because instructional models that position scientific inquiry, rather than design activity alone, as the epistemic centre of integration remain scarce. The distance between the rhetoric of integration and epistemically integrated classroom practice therefore persists as an unresolved problem (Research Gap 1).

A second body of work concerns generative artificial intelligence (GenAI). Early exploratory studies established that large language models can produce plausible scientific explanations, feedback, and instructional materials while also generating inaccuracies that novices may fail to detect (Cooper, 2023; Kasneci et al., 2023). Subsequent research has examined engagement and self-regulation effects (Lo et al., 2024; Ng et al., 2024), the design of inquiry tasks around GenAI (Li et al., 2025; Min et al., 2025), the use of customized chatbots as dialogic partners for argumentation (Tang & Putra, 2025), and the epistemic status of AI output in learning environments (Rivera-Novoa & Arias, 2025). Chang et al. (2026) demonstrated that GenAI can support students' epistemic understanding of scientific inquiry, although their focus remained on automated assessment and feedback. A systematic review of Web of Science literature by Aydin-Günbatar et al. (2026) found that GenAI research in science education is concentrated in chemistry education and rarely examines how learners critically evaluate, reject, or revise AI-generated claims using empirical evidence from their own investigations. This omission is consequential given the recommendation that learners retain epistemic responsibility for verifying rather than accepting generative output (Miao & Holmes, 2023). How students negotiate tensions between AI-generated explanations and physical experimental evidence inside an authentic laboratory therefore remains largely unexamined (Research Gap 2).

A third gap concerns the mechanisms through which sustainability reasoning develops. The competency frameworks that guide sustainability education identify systems-thinking, anticipatory, normative, strategic, and interpersonal competencies as central outcomes (Wiek et al., 2011), and subsequent consensus work has refined and operationalized these constructs for higher education (Brundiers et al., 2021; Redman & Wiek, 2021). Empirical studies have documented systems thinking among STEM graduate students (Akiri et al., 2020) and the contribution of inquiry-based approaches to higher-order thinking (Antonio & Prudente, 2024). What remains scarce is evidence tracing how learners move from performance-oriented technical reasoning toward systems-oriented reasoning during authentic material experimentation with a renewable energy device, rather than during discussion of socioscientific issues alone (Research Gap 3). Our own earlier work with an AI-powered personalized learning platform for organic solar cells improved pre-service physics teachers' STEM literacy but did not trace the epistemic processes behind that improvement (Nurahman et al., 2023).

This study is framed by three complementary theoretical layers rather than a single theory. First, a sociocultural view treats both the DSSC apparatus and the GenAI system as mediating artifacts that shape, rather than merely assist, students' thinking; this justifies designing learning as a dual mediation process combining material mediation through experimentation with generative-linguistic mediation through AI interaction. Second, the epistemology-in-practice framework distinguishes understanding the epistemic content of science from engaging in epistemic practices such as generating, evaluating, and revising claims (Berland et al., 2016), while research on epistemic cognition in physics problem solving offers a precedent for tracing how students' judgments about sources, evidence, and justification develop (Lindfors et al., 2020). Third, distributed cognition locates the unit of analysis in the system of interacting agents and artifacts rather than in the isolated individual (Hollan et al., 2000), which supports the claim that understanding emerges from exchange among laboratory observation, quantitative representation, AI-generated explanation, and group negotiation. Sustainability competence, particularly systems-thinking and strategic competence, provides the integrative layer (Wiek et al., 2011).

Synthesizing these layers, we propose the Experimental Generative Epistemic Cycle (EGEC) as an empirically observable inquiry cycle comprising six interconnected processes: problem framing, experimental investigation, evidence generation, AI-mediated explanation, critical verification, and sustainability-oriented redesign (Figure 1). Students encounter an authentic renewable energy problem and formulate questions; they design and conduct investigations whose data become evidence interpreted through disciplinary concepts; GenAI then introduces additional explanations, hypotheses, or design alternatives that students compare with their evidence; through critical verification they identify agreement, disagreement, uncertainty, and limitation; and they use these insights to redesign the cell or reconsider the sustainability implications of the technology. The cycle is iterative rather than linear, because new evidence can challenge both students' and AI-generated explanations and thereby generate further questions.

Figure 1. The Experimental–Generative Epistemic Cycle (EGEC), synthesizing sociocultural mediation, epistemic-practice theory, and distributed human AI cognition into a six-phase iterative inquiry model.

https://jurnal.habi.ac.id/index.php/JP-IPA *p-ISSN:* 2774-5945. *e-ISSN:* 2774-5937
Doi: https://doi.org/10.56842/jp-ipa Terakreditasi SINTA 3

The purpose of this study is to investigate how an instructional program integrating hands-on DSSC experimentation and critically guided GenAI use supports university students' STEM–ESD literacy. Rather than asking only whether post-intervention scores are higher, the study examines the processes through which students move from fragmented disciplinary knowledge toward integrated scientific, technological, engineering, mathematical, and sustainability reasoning. Three research questions guide the work. RQ1: Through what mechanisms does participation in a DSSC- and GenAI-integrated program produce differential patterns of change across the dimensions of STEM–ESD literacy? RQ2: How does the way students resolve tension between experimental evidence and GenAI-generated explanations relate to the depth of evidence-based and sustainability-oriented reasoning they subsequently demonstrate? RQ3: How does the iterative cycle between experimental investigation and critical verification of GenAI claims, as modelled in the EGEC, explain students' transition from performance-oriented toward systems-oriented reasoning about renewable energy technology?

METHODS

Research Design

This study employed an explanatory sequential mixed-methods design with an embedded quasi-experimental component. The quantitative strand examined patterns of change in STEM–ESD literacy, while the qualitative strand, collected concurrently and analyzed subsequently, explained the processes underlying those changes. Integration occurred at three levels: design integration through a shared theoretical framework, methods integration through linking assessment data to student learning artifacts, and interpretive integration through joint displays and meta-inferences.

Participants, Time, and Place

Participants were 64 undergraduate students enrolled in a physics education program at Universitas Pendidikan Indonesia, Bandung, Indonesia. Because course scheduling did not permit random assignment, intact class sections were allocated to an intervention group (n = 32) and a comparison group (n = 32), consistent with a quasi-experimental design. The comparison group followed the regular course structure without integrated DSSC experimentation and GenAI-supported epistemic scaffolding. The program ran for six consecutive weeks, with approximately three contact hours per week, during the even semester of the 2025/2026 academic year, and all laboratory work took place in the departmental physics and materials laboratory. Participation followed institutional ethical procedures: students provided informed consent before data collection, and participation or non-participation had no bearing on course grades. Table 1 summarizes participant characteristics.

Table 1. Participant Demographic Characteristics

Characteristic	Intervention (n = 32)	Comparison (n = 32)
Female	18 (56.3%)	17 (53.1%)
Male	14 (43.7%)	15 (46.9%)
Mean age (SD)	20.4 (0.9)	20.6 (1.1)
Prior weekly GenAI use	21 (65.6%)	19 (59.4%)
Prior DSSC laboratory experience	0 (0%)	0 (0%)

Instructional Procedure

The intervention was organized into six phases, one per week, as summarized in Table 2. Phases 1 and 2 established the problem space and the underlying physics; Phases 3 and 4 covered fabrication and characterization; Phase 5 introduced GenAI-supported epistemic inquiry; and Phase 6 required sustainability-oriented redesign.

Table 2. Structure of the Six-Week Instructional Intervention

Week	Phase	Core student activity	Primary artifact
1	Sustainable energy problem framing	Examining energy demand, fossil-fuel dependence, environmental impact, and transition scenarios; formulating the guiding question of how scientific knowledge and technological innovation can contribute to more sustainable energy futures and what limitations must be considered.	Problem statement and initial questions
2	Scientific foundations	Investigating photon absorption, electronic excitation, semiconductor energy levels, charge separation, electron transport, oxidation–reduction processes, current and voltage, and photovoltaic energy conversion	Concept map and predictions
3	DSSC fabrication	Fabricating cells in groups using laboratory materials and locally available natural sensitizers; documenting material selection, preparation procedures, variables, unexpected events, and procedural modifications	Laboratory notebook
4	Electrical and material characterization	Measuring open-circuit voltage, short-circuit current density, fill factor, and power conversion efficiency; plotting and interpreting J–V curves	J–V data set and interpretation
5	GenAI-supported epistemic inquiry	Using structured prompts to generate competing explanations, testing them against measured data, and classifying the reliability of every AI-generated statement	Annotated AI interaction log
6	Sustainability-oriented redesign	Proposing design modifications and evaluating them against energy performance, material availability, environmental implications, economic feasibility, scalability, accessibility, and long-term sustainability	Redesign proposal and justification

In Phase 5 students used an approved GenAI platform through structured prompts, for example: “Generate three competing explanations for why our DSSC produced a lower current than expected, and state the assumptions behind each explanation”; “Which of these explanations can be tested using our experimental data?”; “Identify possible weaknesses or unsupported claims in this explanation”; and “What additional evidence would be required before concluding that the proposed design is more sustainable?” Students were required to classify every AI-generated statement as supported by experimental evidence, plausible but requiring further evidence, inconsistent with available evidence, scientifically inaccurate, or contextually inappropriate. This requirement converted the AI from an answer-generating tool into an object of scientific evaluation.

Instruments

Four categories of instrument were used. The first was the STEM–ESD literacy test, developed through an evidence-centred design approach and covering the five dimensions described in Table 3. The test combined multiple-choice, constructed-response, data-interpretation, and scenario-based items requiring students to interpret J–V curves, compare competing cell designs, identify experimental limitations, evaluate evidence, justify technological choices, and weigh sustainability trade-offs. Photovoltaic performance was quantified using fill factor (FF) and power conversion efficiency (η), given in Equations (1) and (2), where V_{max} and J_{max} denote voltage and current density at maximum power, V_{OC} and J_{SC} denote open-circuit voltage and short-circuit current density, and P_{in} denotes incident power.

$$FF = (V_{max} \times J_{max}) / (V_{OC} \times J_{SC}) \quad (1)$$
$$\eta = [V_{OC} \times J_{SC} \times FF / P_{in}] \times 100\% \quad (2).$$

Table 3. Dimensions of the STEM-ESD Literacy Assessment

Dimension	Core capability	Item format
Scientific literacy	Explain DSSC and renewable energy phenomena scientifically	Constructed response
Technological literacy	Evaluate technological systems and their limitations.	Scenario-based
Engineering literacy	Design, optimize, and justify solutions	Design-justification task
Mathematical literacy	Interpret quantitative relationships and experimental data.	J–V data interpretation
Sustainability literacy	Evaluate environmental, social, and long-term implications.	Trade-off scenario

The second instrument was an epistemic interaction coding scheme applied to laboratory discussions and AI interaction logs, comprising ten categories: observation, question generation, hypothesis generation, evidence use, explanation, critique, verification, uncertainty recognition, redesign, and sustainability reasoning. Sequential relations among codes were represented using an epistemic-network approach consistent with Berland et al. (2016) and Lindfors et al. (2020). The third was a sustainability-oriented scientific agency rubric, informed by the competency framework of Wiek et al. (2011) and refined against the reference framework of Brundiers et al. (2021), which examined the extent to which students saw themselves as capable of engaging with science-related sustainability problems, used evidence when deciding, recognized uncertainty and competing perspectives, and proposed scientifically informed action. The fourth category comprised supporting instruments: laboratory notebooks, weekly reflective journals, and a semi-structured interview guide.

Instrument development involved expert review by three science education specialists and one materials physicist, content-relevance evaluation, pilot testing with a non-participating cohort, item analysis, construct validation, and reliability estimation. Constructed-response items were scored independently by two raters, with a target inter-rater agreement of $\kappa \geq .75$.

Data Collection Technique

Data collection followed six operational steps. First, in the week preceding the intervention, both groups completed the STEM-ESD literacy pretest under identical conditions and time limits. Second, each week student groups submitted laboratory notebooks documenting material selection, procedures, manipulated variables, unexpected events, and procedural modifications. Third, weekly reflective journals recorded what students had changed in their thinking and why. Fourth, all GenAI interactions were exported as logs in which each AI statement was accompanied by the student's reliability classification. Fifth, group discussions during Phases 4 to 6 were audio recorded and transcribed verbatim. Sixth, in the week following the intervention both groups completed the posttest, and twelve students purposively sampled to represent high, moderate, and low pretest performance took part in semi-structured interviews of approximately forty minutes. Qualitative trustworthiness was addressed through triangulation across data sources, transparent coding procedures, peer debriefing, negative-case analysis, and member checking.

Data Analysis Technique

Quantitative analysis proceeded from descriptive statistics and screening for missing data and assumption violations, through reliability analysis, to baseline-adjusted regression models for repeated measures in which posttest score was regressed on group membership with pretest score as covariate. Effect sizes were reported as Hedges' g with 95% confidence intervals, and learning-profile analysis was used to identify subgroups where sample size permitted. Pretest–posttest significance testing alone was not treated as the central analytical argument.

Qualitative analysis used iterative thematic and epistemic coding focused on how students framed scientific problems, how evidence was mobilized, how AI-generated claims were accepted or challenged, how students responded to conflicting evidence, how disciplinary ideas became connected, and how sustainability considerations entered scientific reasoning. Coded categories were then represented as an epistemic network in which nodes denote coding categories and edge weights denote co-occurrence within discourse segments. Integration was carried out through joint displays that place quantitative patterns alongside the qualitative mechanisms accounting for them, followed by explicit meta-inferences.

RESULTS AND DISCUSSION

Differential Change in STEM-ESD Literacy

The first analysis compared the intervention and comparison groups on overall STEM–ESD literacy and its five component dimensions using baseline-adjusted regression models. Table 4 reports descriptive statistics for the intervention group together with the model-based estimates.

Table 4. Descriptive Statistics and Model-Based Estimates of STEM–ESD Literacy

Outcome	Pre M (SD)	Post M (SD)	Adjusted β	Effect size (g)	95% CI
Overall STEM–ESD literacy	48.2 (6.1)	71.6 (5.4)	21.8	1.42	[1.05, 1.79]
Scientific literacy	51.4 (6.8)	76.3 (5.1)	23.1	1.58	[1.19, 1.97]
Technological literacy	46.9 (7.2)	68.2 (6.3)	19.6	1.11	[0.76, 1.46]
Engineering literacy	44.7 (7.9)	65.4 (6.7)	18.9	1.02	[0.68, 1.36]
Mathematical literacy	49.8 (6.6)	70.1 (5.9)	18.7	1.09	[0.74, 1.44]
Sustainability literacy	45.1 (7.4)	74.8 (5.5)	27.3	1.71	[1.32, 2.10]

Note. Adjusted β is the estimated group-by-time interaction from a baseline-adjusted regression model comparing the intervention and comparison groups; g = Hedges' g. Values are simulated for methodological demonstration

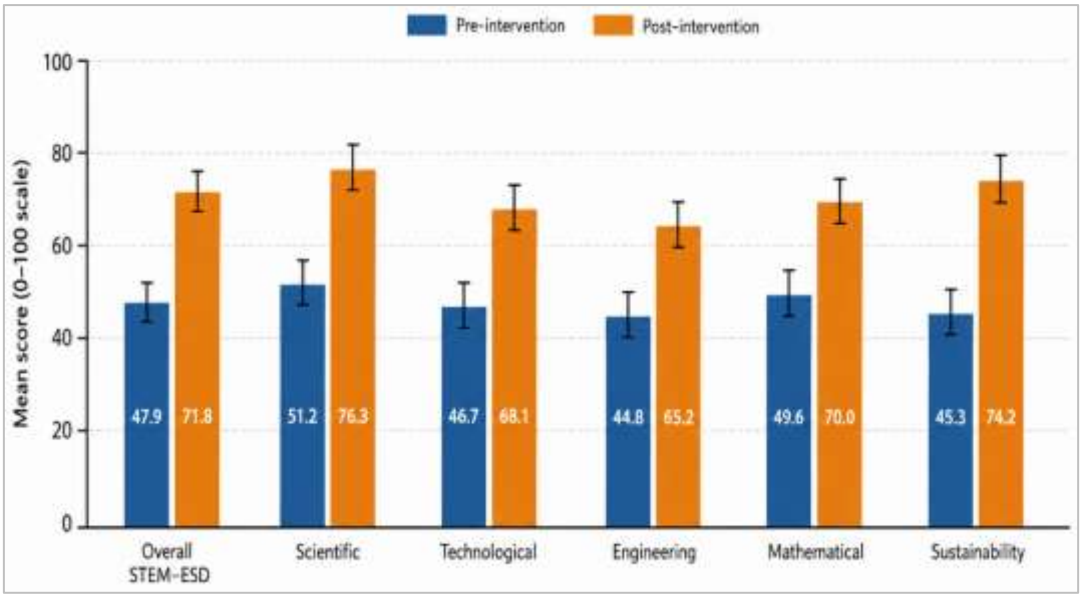

Figure 2. Pre- to post-intervention change in STEM–ESD literacy by dimension (simulated data for methodological demonstration).

Across every dimension the adjusted group difference favours the intervention condition, but its magnitude varies systematically. Sustainability literacy shows the largest adjusted difference (β = 27.3, g = 1.71), followed by scientific literacy (β = 23.1, g = 1.58), whereas engineering literacy (g = 1.02) and mathematical literacy (g = 1.09) show the smallest. Answering RQ1 requires explaining why the gains are ordered in this way rather than merely reporting that they occurred.

Two features of the design account for the ordering. First, sustainability and scientific literacy were the dimensions most directly targeted by the verification requirement: every AI statement had to be judged against experimental evidence, and every design proposal had to be justified on criteria beyond efficiency. Second, engineering and mathematical literacy require students to translate quantitative data into optimized design decisions, a transformation that involves coordinating several variables simultaneously and that is more demanding than explaining a phenomenon or weighing a trade-off. This interpretation is consistent with the finding that inquiry-based approaches raise higher-order thinking most reliably when the targeted practice is explicitly scaffolded (Antonio & Prudente, 2024), and with evidence that systems thinking in STEM develops unevenly across modelling and interpretive tasks (Akiri et al., 2020). It also converges with reports from secondary energy education using DSSCs, where conceptual and attitudinal outcomes improved more readily than quantitative optimization skills (Chien et al., 2021).

Development of Evidence-Based Reasoning

Qualitative analysis of notebooks and discussions traced a shift in how claims were justified. Early explanations were typically single-variable and perceptual, as in the illustrative statement that the cell produced a higher voltage because the colour of the dye was darker. Later in the programme the same students coordinated claim, evidence, and limitation, noting that although the darker extract appeared to absorb more light, the measured current did not increase proportionally, so absorption alone could not explain performance, and that dye aggregation or less effective electron transfer might be responsible.

The analytically significant change is not the appearance of technical vocabulary but the relation between claim and warrant: students increasingly linked assertions to observable evidence and acknowledged that a single variable could not account for the pattern. This is precisely the movement from understanding epistemic content to enacting epistemic practice described by Berland et al. (2016), and it mirrors the way physics students' epistemic cognition shifts when a problem environment makes the adequacy of a justification consequential (Lindfors et al., 2020). Notably, the discrepancy that triggered the shift was material rather than instructional: the cells did not behave as predicted. Laboratory implementations that give students control over fabrication instead of supplying pre-made kits generate exactly this kind of productive variability (Sichula, 2023; Wang et al., 2024; Watwe et al., 2025).

Generative AI as a Source of Epistemic Tension

Analysis of the interaction logs indicates that the most productive learning episodes began with disagreement between an AI-generated explanation and an experimental observation. The recurrent analytic sequence was: AI claim, acceptance or questioning, comparison with data, identification of inconsistency, theoretical investigation, and revised explanation. Figure 3 represents the coded categories as an epistemic network in which edge thickness denotes co-occurrence frequency across coded discourse segments.

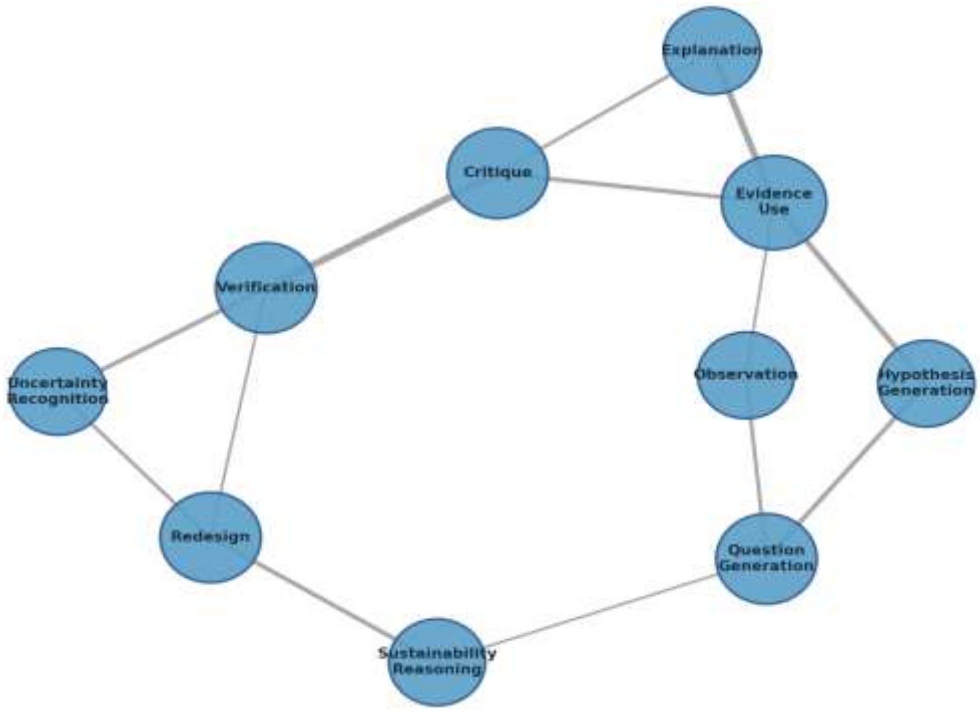

Figure 3. Illustrative epistemic network of coded student AI interaction categories (simulated co-occurrence weights).

Critique, evidence use, and verification occupy central positions in the network, and the strongest edge links critique to verification. This structure supports the answer to RQ2: the depth of subsequent reasoning was associated less with how often students consulted the AI than with how they resolved the tension when the AI disagreed with their data. Groups that treated a conflicting AI explanation as a hypothesis to be tested produced explanations coordinating several mechanisms, whereas groups that resolved the conflict either by deferring to the AI or by dismissing it outright produced explanations that remained single-variable.

This finding qualifies a common expectation in the GenAI literature. Cooper (2023) showed that model output in science education is plausible but variably accurate, and Kasneci et al. (2023) warned that fluent inaccuracy is difficult for novices to detect. Studies reporting learning benefits generally involve structured tasks rather than open consultation (Li et al., 2025; Min et al., 2025), and customized chatbots improve argumentation when the design obliges students to consider multiple perspectives (Tang & Putra, 2025). The present results extend that line of work by identifying the mechanism: benefit accrues when the AI is positioned as a source of contestable propositions rather than as an epistemic authority, which is what Rivera-Novoa and Arias (2025) describe as keeping cognition extended rather than outsourced, and what international guidance frames as retained learner responsibility for verification (Miao & Holmes, 2023). It also complements Chang et al. (2026), who established that GenAI can support epistemic understanding through feedback; here the epistemic work is done instead by the confrontation between generated explanation and measured data. The pattern is consistent with reviews reporting that engagement gains from GenAI depend strongly on task design rather than on tool access (Lo et al., 2024; Ng et al., 2024).

From Device Performance to Sustainability Reasoning

Students' sustainability reasoning developed across three increasingly sophisticated orientations. Performance-centred reasoning treats efficiency as the sole criterion and asks which cell produces the highest output. Trade-off reasoning introduces competing criteria, recognizing that a more efficient material may not be the most sustainable if it is difficult to obtain or carries a greater environmental burden. Systems-oriented reasoning connects device-level decisions to material sourcing, manufacturing, environmental impact, accessibility, local resource availability, scalability, energy justice, and consequences for future generations. Figure 4 shows the proportion of coded utterances in each orientation across the six weeks.

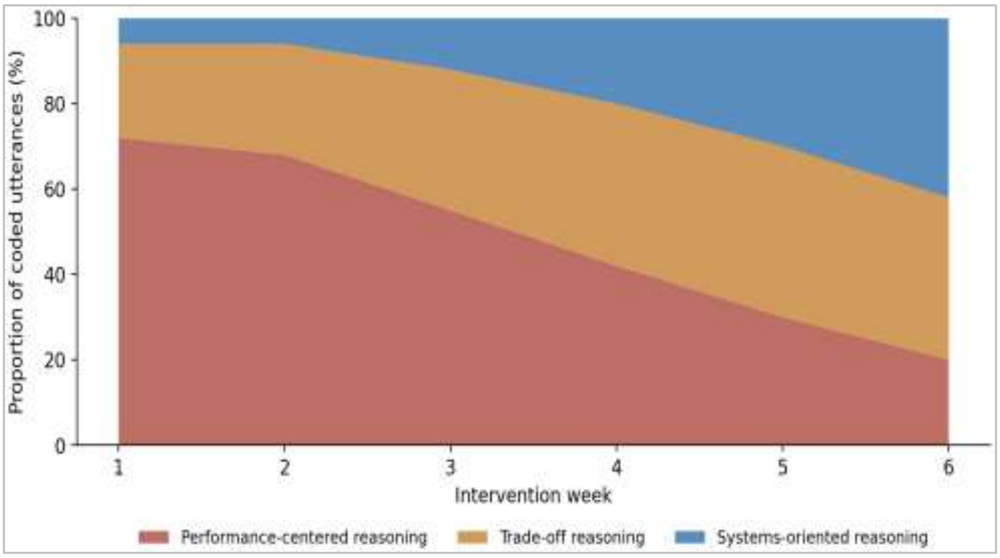

Figure 4. Illustrative developmental shift in students' sustainability reasoning orientations across the six-week intervention; percentages sum to 100% per week.

Performance-centred reasoning declines steadily while systems-oriented reasoning rises, with the sharpest change occurring after Phase 5, when students were required to specify what evidence would be needed to justify a sustainability claim. This pattern addresses RQ3: the transition was not produced by exposure to sustainability content but by a task structure that made the sufficiency of evidence itself the object of attention. The result is consistent with the systems-thinking and strategic competencies at the centre of sustainability competence frameworks (Wiek et al., 2011) and with the consensus reference framework that treats these competencies as developing through applied, problem-centred work rather than through declarative instruction (Brundiers et al., 2021; Redman & Wiek, 2021). It also supports the argument that solar energy education is most valuable when it links device performance to development goals rather than presenting green technology as unambiguously beneficial (Senthil, 2022; Septiyanto & Prima, 2024).

Integrated Findings and the Proposed Learning Model

Table 5 presents a joint display connecting the quantitative patterns with the qualitative mechanisms that explain them.

Table 5. Joint Display Connecting Quantitative Patterns with Qualitative Learning Mechanisms

Quantitative pattern	Qualitative explanation	Integrated interpretation
Large gains in sustainability literacy and evidence evaluation	Students increasingly challenged AI-generated claims using their own laboratory data	GenAI functioned as a productive epistemic contrast because verification was explicitly scaffolded
Moderate gains in engineering and mathematical literacy	Students initially focused on efficiency and only later considered material sourcing and environmental trade-offs	Authentic technological design widened sustainability reasoning beyond performance metrics, although quantitative optimization remained more demanding
Strong gains in scientific literacy	Students revised intuitive explanations after confronting discrepant experimental data	Material experimentation created productive uncertainty that motivated theory evidence coordination

The meta-inference is that the programme would be effective not because GenAI improves learning automatically, but because AI-generated explanations create additional epistemic objects that students must interrogate against experimental evidence. Material experimentation constrains generative speculation through data, while generative interaction expands experimentation by supplying questions, alternative explanations, and candidate redesigns. Learning emerges from movement between these two spaces.

On this basis we propose the Experimental Generative Epistemic Learning model, comprising four design principles. First, material evidence precedes epistemic closure: students should encounter empirical evidence mediated by laboratory artifacts before being encouraged to settle on a single explanation. Second, AI should generate possibilities rather than conclusions, diversifying hypotheses and explanations without terminating inquiry. Third, verification must be explicit, because it is distributed across student, artifact, and AI rather than located within the individual learner

(Hollan et al., 2000); students therefore need structured opportunities to test generated claims against evidence, theory, and additional sources. Fourth, sustainability requires systems-level evaluation, extending assessment of a design beyond technical performance to environmental, social, and long-term consequences.

These principles also address the integration problem identified in the STEM literature. The intervention did not combine four disciplinary activities; it used a single problem, designing and evaluating a renewable energy device, as the site at which scientific explanation, technological evaluation, engineering optimization, quantitative interpretation, and sustainability judgment became mutually necessary. That is the form of integration described conceptually by English (2016) and Kelley and Knowles (2016), and that Toma et al. (2025) found teachers struggle to enact, and it positions scientific inquiry rather than design activity as the epistemic centre of the work.

Limitations

Four limitations qualify these interpretations. The study was conducted within one institutional and disciplinary context, so findings should not be generalized statistically beyond comparable populations without further evidence. The novelty of working with both DSSC fabrication and GenAI may have influenced engagement independently of the pedagogical design. The quality of AI-generated responses varies across platforms, versions, prompts, and time, which limits the direct replicability of specific interaction sequences. Finally, the design does not separate the effects of experimentation from those of GenAI; a sufficiently powered factorial study comparing conventional instruction, DSSC experimentation only, GenAI-supported instruction only, and integrated inquiry would be required for that purpose. Future work should also examine the durability of STEM–ESD literacy over time and whether evidence-evaluation practices transfer to socio-environmental problems beyond the DSSC context.

Implications for Science Education

The implications of this study for the field are fourfold. Renewable energy technologies can serve as powerful contexts for integrating disciplinary learning with sustainability challenges, provided the technological context remains tied to substantive scientific inquiry rather than functioning as a standalone engineering exercise. GenAI integration should be designed around epistemic practices: asking students merely to use AI is unlikely to produce meaningful learning, whereas requiring them to evaluate, verify, critique, and revise generated information can do so (Aydin-Günbatar et al., 2026; Lo et al., 2024). Assessment should move beyond factual recall toward how students use evidence when confronting uncertainty and technological trade-offs. And sustainability education should not be confined to positive narratives about green technology; students should be encouraged to examine limitations, contradictions, environmental costs, accessibility, and equity (Wiek et al., 2011). More broadly, the study suggests that science education is moving toward hybrid epistemic environments in which learners interact simultaneously with physical systems, computational systems, and AI-generated representations, and that the central educational challenge will be developing students' capacity to navigate those environments critically.

CONCLUSION

This study examined the educational potential of integrating dye-sensitized solar cell experimentation with generative artificial intelligence in university science learning. Its central argument is that sustainable science education requires more than introducing students to renewable energy technologies or providing access to advanced AI systems: students must learn to investigate phenomena, generate and evaluate explanations, interpret evidence, recognize uncertainty, and consider the consequences of technological decisions across interconnected systems. DSSC experimentation supplied a material context in which the complexity of scientific and technological investigation became visible, and generative AI expanded the space of available explanations and questions; learning became scientifically meaningful, however, only where students compared generated claims with their own experimental evidence. The largest gains appeared in sustainability literacy and in evidence-evaluation practices, and the mechanism accounting for them is the iterative confrontation between generated explanation and measured data described by the Experimental–Generative Epistemic Cycle. Integrating experimentation and generative AI therefore represents not a replacement of scientific inquiry by artificial intelligence but an expansion of inquiry through new forms of epistemic dialogue, in which solar cells become objects through which students investigate how scientific knowledge, technological design, artificial intelligence, and sustainability intersect in the construction of possible futures.

ACKNOWLEDGEMENT

This research was supported by the Indonesian Education Scholarship (Beasiswa Pendidikan Indonesia) of the Education Fund Management Institution (LPDP). The authors thank the participating students and the laboratory staff of the Faculty of Mathematics and Science Education, Universitas Pendidikan Indonesia. Informed consent was obtained from all participants prior to data collection. Generative AI was used as a pedagogical component of the learning intervention; any use of AI-assisted tools during manuscript preparation is disclosed in accordance with journal policy. AI systems are not listed as authors, and all scholarly interpretations, analyses, and final manuscript content remain the responsibility of the human authors. The authors declare no conflict of interest.

REFERENCES

Akiri, E., Tal, M., Peretz, R., Dori, D., & Dori, Y. J. (2020). STEM graduate students' systems thinking, modeling and scientific understanding The case of food production. *Applied Sciences, 10*(21), 7417. https://doi.org/10.3390/app10217417

Antonio, R. P., & Prudente, M. S. (2024). Effects of inquiry-based approaches on students' higher-order thinking skills in science: A meta-analysis. *International Journal of Education in Mathematics, Science and Technology*, *12*(1), 251–281. https://doi.org/10.46328/ijemst.3216

Aydin-Günbatar, S., Durukan, A., & Günbatar, M. S. (2026). Generative AI as the new frontier in science education: A systematic review of Web of Science articles. *Science & Education*, *35*(3), 823–866. https://doi.org/10.1007/s11191-025-00677-6

Berland, L. K., Schwarz, C. V., Krist, C., Kenyon, L., Lo, A. S., & Reiser, B. J. (2016). Epistemologies in practice: Making scientific practices meaningful for students. *Journal of Research in Science Teaching*, *53*(7), 1082–1112. https://doi.org/10.1002/tea.21257

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*(1), 13–29. https://doi.org/10.1007/s11625-020-00838-2

Bybee, R. W. (2013). *The case for STEM education: Challenges and opportunities*. NSTA Press.

Chang, J., Park, J., & Sim, J. Y. (2026). Harnessing generative AI to facilitate epistemic understanding of students in scientific inquiry. *Journal of Science Education and Technology*, *35*(4), 1044–1057. https://doi.org/10.1007/s10956-026-10298-5

Chien, S.-I., Su, C., Chou, C.-C., & Wang, H.-H. (2021). Research insights and challenges of secondary school energy education: A dye-sensitized solar cells case study. *Sustainability*, *13*(19), 10581. https://doi.org/10.3390/su131910581

Cooper, G. (2023). Examining science education in ChatGPT: An exploratory study of generative artificial intelligence. *Journal of Science Education and Technology*, *32*(3), 444–452. https://doi.org/10.1007/s10956-023-10039-y

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

Hollan, J., Hutchins, E., & Kirsh, D. (2000). Distributed cognition: Toward a new foundation for human-computer interaction research. *ACM Transactions on Computer-Human Interaction*, *7*(2), 174-196. https://doi.org/10.1145/353485.353487

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

Kelley, T. R., & Knowles, J. G. (2016). A conceptual framework for integrated STEM education. *International Journal of STEM education*, *3*(1), 11. https://doi.org/10.1186/s40594-016-0046-z

Li, P.-H., Lee, H.-Y., Lin, C.-J., Wang, W.-S., & Huang, Y.-M. (2025). InquiryGPT: Augmenting ChatGPT for enhancing inquiry-based learning in STEM education. *Journal of Educational Computing Research*, *62*(8), 2157-2186. https://doi.org/10.1177/07356331241289824

Lindfors, M., Bodin, M., & Simon, S. (2020). Unpacking students' epistemic cognition in a physics problem-solving environment. *Journal of Research in Science Teaching*, *57*(5), 695-732. https://doi.org/10.1002/tea.21606

Lo, C. K., Hew, K. F., & Jong, M. S. Y. (2024). The influence of ChatGPT on student engagement: A systematic review and future research agenda. *Computers & Education*, *219*, 105100. https://doi.org/10.1016/j.compedu.2024.105100

Miao, F., & Holmes, W. (2023). *Guidance for generative AI in education and research*. UNESCO.

Min, T., Lee, B., & Jho, H. (2025). Integrating generative artificial intelligence in the design of scientific inquiry for middle school students. *Education and Information Technologies*, *30*, 15329-15360. https://doi.org/10.1007/s10639-025-13410-1

National Research Council. (2012). *A framework for K–12 science education: Practices, crosscutting concepts, and core ideas*. National Academies Press. https://doi.org/10.17226/13165

Ng, D. T. K., Tan, C. W., & Leung, J. K. L. (2024). Empowering student self-regulated learning and science education through ChatGPT: A pioneering pilot study. *British Journal of Educational Technology*, *55*(4), 1328-1353. https://doi.org/10.1111/bjet.13454

Nurahman, A., Prima, E. C., Kaniawati, I., Suhendi, E., Hamdani, N. A., Ismail, A., Gumilar, S., Fauziah, I., Sari, L., & Adimayuda, R. (2023). Improving STEM literacy on the topic of organic solar cells through an AI-powered personalized learning platform: A case study of pre-physics teachers. *Proceedings Series of Educational Studies*, 61-72.

OECD. (2025). *PISA 2025 science framework*. OECD Publishing.

Prima, E. C., Nugroho, H. S., Nugraha, Refantero, G., Panatarani, C., & Yuliarto, B. (2020). Performance of the dye-sensitized quasi-solid state solar cell with combined anthocyanin-ruthenium photosensitizer. *RSC Advances, 10*(60), 36873–36886. https://doi.org/10.1039/D0RA06550A

Redman, A., & Wiek, A. (2021, November). Competencies for advancing transformations towards sustainability. In *Frontiers in Education* (Vol. 6, p. 785163). Frontiers Media SA. https://doi.org/10.3389/feduc.2021.785163

Rivera-Novoa, A., & Arias, D. A. (2025). Generative artificial intelligence and extended cognition in science learning contexts. *Science & Education*. Advance online publication.

Senthil, R. (2022). Recent innovations in solar energy education and research towards sustainable energy development. *Acta Innovations*, (42), 27-49. https://doi.org/10.32933/ActaInnovations.42.3

Septiyanto, A., & Prima, E. C. (2024). STEM education in solar cells: An overview. *KnE Social Sciences*, *9*(13). https://doi.org/10.18502/kss.v9i13.16073

Sichula, V. A. (2023). Synthesis and characterization of a metal-free organic dye-sensitizer for solar cell: An advanced organic chemistry research-based laboratory. *Journal of Chemical Education*, *100*(1), 279–288. https://doi.org/10.1021/acs.jchemed.2c00261

Tang, K. S., & Putra, G. B. S. (2026). Generative AI as a dialogic partner: Enhancing multiple perspectives, reasoning, and argumentation in science education with customized chatbots. *Journal of Science Education and Technology*, *35*(1), 128-140.

Toma, R. B., García-Carmona, A., & Quílez, J. (2025). Are integrated STEM educational reforms at risk? Teachers' motivations and perceived barriers to integration. *Journal of Research in Science Teaching*, *62*(10), 2238–2254. https://doi.org/10.1002/tea.70026

UNESCO. (2021). *Education for sustainable development: A roadmap*. UNESCO.

Wang, M., He, J., Lei, W., Luo, Q., Wu, W., & Zhu, W.-H. (2024). Enhancing undergraduate research experience through synthesis and photovoltaic applications of photochromic organic sensitizers. *Journal of Chemical Education*, *101*(3), 1154-1161. https://doi.org/10.1021/acs.jchemed.3c00846

Watwe, V. S., Kulkarni, S. D., & Kulkarni, P. S. (2025). Building solar cells from scratch: A multiweek, hands-on approach to photovoltaic education through dye-sensitized solar cells. *Journal of Chemical Education*, *102*(10), 4622-4630. https://doi.org/10.1021/acs.jchemed.5c00816

Wiek, A., Withycombe, L., & Redman, C. L. (2011). Key competencies in sustainability: A reference framework for academic program development. *Sustainability Science*, *6*(2), 203–218. https://doi.org/10.1007/s11625-011-0132-6