

Future Faculty Development for AI-Mediated Higher-Order Learning: The PEDAL Model for Higher Education

Tendai Charles, Marwa Elkhoully

Faculty of Education, The British University in Dubai, Dubai 345015, United Arab Emirates

© Higher Education Press 2026

Abstract Generative artificial intelligence (GenAI) has intensified the debate in higher education about academic integrity, student learning, and the future of academic work. Faculty development, however, often remains concentrated on tool demonstrations, compliance advice, and risk management rather than on the redesign of intellectually demanding learning. This conceptual review develops the PEDAL model—purpose, epistemic challenge, dialogue, accountability, and learning evidence—as a practical synthesis for faculty development in AI-rich higher education. The model was derived through an integrative, abductive synthesis of scholarship on higher-order thinking, GenAI, constructive alignment, cognitive apprenticeship, co-regulation, AI literacy, assessment, and professional learning, supplemented by recent empirical work from resource-diverse settings. PEDAL positions AI as a bounded pedagogical resource that can generate alternatives, simulations, misconceptions, and prompts for critique without replacing disciplinary judgment. The paper explains the derivation and relationships of the five principles, provides worked redesigns in the social sciences and engineering, and proposes evidence strategies intended to be feasible for large and diverse classes. It also distinguishes minimum program-level conditions from more mature institutional support. The contribution is not a new general theory of learning but a meso-level design heuristic that translates established insights into a coherent sequence of decisions for faculty development, assessment redesign, and empirical testing.

Keywords artificial intelligence, assessment design, faculty development, higher education, higher-order thinking, PEDAL model

1 Introduction

Generative artificial intelligence (GenAI) has moved from novelty to infrastructural presence in higher education. Systems capable of producing essays, summaries, explanations, code, images, feedback, and multimodal artifacts now sit alongside search engines, learning management systems, and digital libraries as part of the everyday study environment. Public debate frequently frames this development as a problem of academic integrity, risk, or regulation. Pedagogically, however, the more consequential question is whether GenAI will amplify student thinking or displace it. Miao and Holmes (2023) argued for a human-centered approach to GenAI in education, while scholarship on large language models has similarly suggested that their educational significance depends on how they are integrated into teaching, learning, and assessment rather than on the tools alone (Kasneci et al., 2023; Sharples, 2023).

The pressure on faculty is immediate. Students can obtain plausible responses to many conventional academic tasks within seconds, including tasks that previously served as proxies for comprehension, synthesis, argumentation, calculation, or design. This does not make higher-order thinking obsolete. However, it does expose how often higher education has assessed polished products without adequately capturing the reasoning that produced them. In this sense, GenAI is less a wholly new disruption of learning than a stress test of pedagogy. It reveals weaknesses in constructive alignment, task authenticity, assessment validity, and the professional support available to faculty (Charles & Hill, 2026).

Evidence on educational use is deliberately mixed. Students have reported using GenAI for explanation, brainstorming, drafting, and personalization, while also expressing concerns about accuracy, ethics, dependency, and future learning (Chan & Hu, 2023).

Received May 26, 2026; revised July 29, 2026; accepted August 31, 2026

Tendai Charles (✉)

E-mail: tendai.charles@buid.ac.ae

GenAI can support complex sense-making and metacognition when it is used as a simulation or dialogic partner (Borge et al., 2024), and collaborative activities can prompt comparison and critique (Ruiz-Rojas et al., 2024). However, work with knowledge workers has suggested that confident reliance on GenAI may reduce self-reported cognitive effort in some tasks (Lee et al., 2025). Reviews of AI and critical thinking have reached a similar conclusion: AI can scaffold reasoning or foster premature closure, depending on task design, learner regulation, and the evidentiary demands of assessment (Salido et al., 2025). The instructional question is therefore not simply whether students use AI but whether faculty can design conditions in which its use may support analysis, evaluation, creation, metacognitive regulation, and transfer rather than short-circuit those processes.

The challenge is especially consequential in resource-diverse and multilingual higher education systems. Institutions differ sharply in access to approved platforms, academic-development capacity, data-governance expertise, assessment flexibility, and student connectivity. Recent work in Palestinian and wider Middle Eastern higher education has illustrated this unevenness. Studies have documented educators' interest in using GenAI for assessment and content creation alongside concerns about academic integrity, policy ambiguity, professional readiness, dependency, and technostress (Alhur et al., 2025; Hamamra et al., 2024; Hamamra et al., 2026; Khlaif et al., 2024; Khlaif et al., 2026). These studies are important because they move the discussion beyond generic claims about global applicability. They show that pedagogical innovation is mediated by institutional precarity, workload, language, infrastructure, and the availability of locally credible guidance.

This problem is fundamentally a faculty-development problem. A systematic review found that the AI-in-education literature has concentrated more heavily on teaching applications than on teacher professional development (Tan et al., 2025). An integrative review similarly identified AI literacy and hands-on design activity as core ingredients of meaningful professional learning for a future with GenAI (Brandão et al., 2024). Competency frameworks add important baselines. UNESCO's *AI competency framework for teachers* links human-centered values, ethics, foundational knowledge, pedagogy, and professional learning (Miao & Cukurova, 2024), while its student framework foregrounds responsible engagement and co-creation (Miao et al., 2024). *The AI fluency framework* likewise emphasizes delegation, description, discernment, and diligence (Dakan & Feller, 2025). These resources clarify what responsible capability may involve, but faculty still

need support in translating competency statements into the design of particular courses, assignments, interactions, and evidence.

Existing frameworks illuminate substantial parts of this design problem. Constructive alignment links outcomes, activities, and assessment (Biggs & Tang, 2011); teaching as a design science foregrounds structured learning interactions and feedback (Laurillard, 2012); cognitive apprenticeship makes expert reasoning visible through modeling, coaching, articulation, reflection, and exploration (Collins et al., 1989); co-regulation explains how learners, teachers, peers, and technologies can jointly regulate learning (Lodge et al., 2023a); AI literacy frameworks specify critical, ethical, and functional competencies (Long & Magerko, 2020; Ng et al., 2021; Zhou & Schofield, 2024); and assessment guidance addresses acceptable AI use and trustworthy judgment (Lodge et al., 2023b; Perkins et al., 2024). The gap is not the absence of relevant theory. It is the absence of a concise, meso-level design architecture that organizes these insights around the decisions faculty must make when redesigning AI-mediated learning for higher-order outcomes.

This paper responds by proposing PEDAL, comprising purpose, epistemic challenge, dialogue, accountability, and learning evidence, as a practical synthesis and design heuristic for future faculty development. Its originality lies in the integration and sequencing of five decisions that are often separated across adjacent literatures: specifying the exact intellectual outcome; constructing a challenge in which human judgment remains indispensable; organizing interaction as inquiry rather than answer retrieval; maintaining ethical and epistemic responsibility; and collecting valid, feasible evidence of learning. The model is deliberately positioned with restraint. It is not presented as a wholly new theory of learning or as an empirically validated instrument. It is a theory-informed design proposition intended to support course redesign, faculty-development studios, assessment review, and subsequent empirical testing.

This paper makes four contributions. First, it provides a transparent account of the conceptual-review logic and the derivation of the five principles. Second, it maintains distinctions among analysis, evaluation, creation, metacognition, and transfer rather than treating higher-order thinking as a single undifferentiated outcome. Third, it operationalizes PEDAL through two worked disciplinary redesigns and a feasibility framework for workload, privacy, accessibility, and fairness. Fourth, it connects task design to differentiated institutional conditions, distinguishing what program teams might pilot with minimal infrastructure from supports requiring central governance, procurement, and sustained academic development.

2 Conceptual Approach and Review Logic

2.1 | Review Purpose and Search Routes

This article is a conceptual review and framework-development paper rather than a systematic review or meta-analysis. It follows the logic of integrative review work, in which the purpose is to synthesize relevant bodies of scholarship, identify tensions and omissions, and generate an interpretive framework that can guide research and practice (Torraco, 2005). The aim is disciplined synthesis rather than exhaustive coverage. Accordingly, the paper does not claim a complete census of publications, does not present a PRISMA flow diagram, and does not use study counts as evidence of effect. Instead, its methodological obligation is to make explicit the selection logic, the analytic questions, and the pathway from literature to model so that readers can inspect and challenge them.

The conceptual corpus was assembled and updated through purposive search routes rather than a single reproducible database protocol. Searches used ERIC, Google Scholar, Crossref-linked publisher platforms, and official repositories maintained by UNESCO, the OECD, national quality agencies, and professional or educational organizations. Backward and forward citation chaining was used to connect recent GenAI studies to established work on constructive alignment, higher-order thinking, cognitive apprenticeship, co-regulation, assessment validity, and professional learning. Search strings combined terms such as “generative AI,” “large language model,” “higher education,” “faculty development,” “teacher professional development,” “AI literacy,” “higher-order thinking,” “critical thinking,” “assessment redesign,” “metacognition,” “transfer,” “trust,” “technostress,” and “resource-constrained.” The contemporary search emphasis was 2020–July 2026, while earlier sources were retained when they provided a theoretical construct that remained necessary for the synthesis.

Sources were included when they met at least one of four functions. They defined a central construct; supplied evidence about the affordances, risks, or conditions of AI-mediated learning; specified faculty or student capabilities; or illuminated professional development and institutional conditions. Peer-reviewed empirical studies, reviews, conceptual papers, and authoritative policy or competency documents were prioritized. Seminal pre-GenAI work was included, where it anchored a construct that could not be understood adequately through recent literature alone. Tool tutorials, promotional claims, and generic productivity

advice were excluded because they did not address educational reasoning or design. School-sector and domain-specific sources were used selectively when they offered transferable insight into teacher capability, assessment, or ethical implementation; they were not treated as direct evidence for all higher education contexts. The Anthropic AI fluency framework, for example, is included as an explicitly identified industry competency framework for comparison and not as independent evidence of educational effectiveness. Selected comparative sources also included an OECD review of higher-order thinking in online higher education (Hague, 2024), an AI-across-the-curriculum model (Southworth et al., 2023), and a BEME guide to AI in medical education (Gordon et al., 2024).

Recent empirical studies from Palestine and the wider Middle East were included because these studies contribute evidence about faculty adoption, assessment redesign, policy ambiguity, dependency, and technostress under conditions that differ from highly resourced Anglophone institutions (Alhur et al., 2025; Hamamra et al., 2024; Hamamra et al., 2026; Khlaif et al., 2024; Khlaif et al., 2026). They do not establish that all resource-diverse systems are alike. Their role is to challenge universalizing assumptions and to identify contextual variables that a portable faculty-development model must leave open to local interpretation.

2.2 | Analytic Strands and Model Derivation

The corpus was organized into four analytic strands: higher-order thinking and durable learning; GenAI in higher education, including assessment implications; AI literacy and professional knowledge; and faculty professional development and institutional support. Across these strands, the synthesis asked four recurrent questions. What form of learning is intended? What does GenAI change about the visibility and ownership of that learning? What capabilities do faculty require to design an appropriate response? What institutional conditions support or constrain enactment? Table 1 makes this analytic structure explicit.

The synthesis proceeded iteratively and abductively. Concepts from one strand were used to interrogate claims in another rather than being summarized in isolation. Constructive alignment and taxonomies of higher-order thinking established the need to begin with outcome specificity, which generated the principle of “purpose.” Cognitive apprenticeship, assessment-validity concerns, and evidence about fluent but unreliable AI output converged on the need for tasks that preserve uncertainty, disciplinary standards, and visible judgment, generating an “epistemic challenge.” Co-regulation, cognitive presence, and dialogic learning

Table 1 Analytic strands, selection emphases, and their contribution to PEDAL

Analytic strand	Selection emphasis and representative sources	Recurrent analytic question	Contribution to PEDAL
Higher-order thinking and durable learning	Analysis, evaluation, creation, metacognition, co-regulation, and transfer (Anderson & Krathwohl, 2001; Brookhart, 2010; Facione, 1990; Garrison et al., 2001; Hague, 2024)	What form of learning is intended, and how should it be distinguished from adjacent outcomes	Establishes outcome specificity for purpose and outcome-aligned learning evidence
Generative AI, dialogue, and assessment	Affordances, cognitive offloading, simulation, collaborative critique, assessment reform, validity, and fairness (Borge et al., 2024; Kasneci et al., 2023; Lee et al., 2025; Lodge et al., 2023b; Perkins et al., 2024; Ruiz-Rojas et al., 2024; Zacharis & Papadakis, 2025)	What does AI change about the visibility, ownership, and validity of learning	Generates epistemic challenge, dialogue, and the need for richer evidence ecologies
AI literacy and professional knowledge	AI literacy, TPACK/Intelligent-TPACK, competency frameworks, fluency, and domain-specific guidance (Celik, 2023; Dakan & Feller, 2025; Gordon et al., 2024; Long & Magerko, 2020; Miao & Cukurova, 2024; Miao et al., 2024; Ng et al., 2021; Southworth et al., 2023)	What must faculty and students understand, evaluate, and take responsibility for	Defines bounded AI roles and the ethical, epistemic, and privacy dimensions of accountability
Faculty development and institutional context	Professional-learning design, trust, networks, adoption, dependency, technostress, and resource diversity (Alhur et al., 2025; Brandão et al., 2024; Garet et al., 2001; Hamamra et al., 2024; Hamamra et al., 2026; Khlaif et al., 2024; Khlaif et al., 2026; Lelescu et al., 2025; Nazaretsky et al., 2022; Tan et al., 2025; Trust et al., 2016)	What capabilities and organizational conditions enable sustainable enactment	Translates the five principles into faculty-development activities and differentiated institutional support

Notes. TPACK: technological pedagogical content knowledge; PEDAL: purpose, epistemic challenge, dialogue, accountability, and learning evidence.

indicated that interaction must be structured as iterative inquiry rather than one-shot answer retrieval, generating “dialogue.” AI literacy, ethics guidance, disclosure practices, and studies of dependency and trust generated accountability. Finally, assessment scholarship and the problem of product fluency established the need for multiple proportionate traces of reasoning, revision, metacognition, and transfer, generating “learning evidence.”

The order of the principles is functional rather than merely mnemonic. Purpose precedes tool selection because a learning outcome cannot be inferred from an AI affordance. Epistemic challenge follows because a nominally higher-order outcome will not be enacted if the task can be completed through generic prompting. Dialogue then specifies how learners engage with the challenge over time. Accountability governs the boundaries, verification, authorship, privacy, and ethical responsibilities of that engagement. Learning evidence tests whether the intended outcome was actually demonstrated and feeds information back into the redesign. The cycle is therefore iterative: evidence can reveal that the stated purpose was too broad, that the challenge was answerable without judgment, that dialogue encouraged acceptance rather than interrogation, or that accountability requirements were burdensome without being informative.

PEDAL is presented as a pragmatic minimum set for this particular design problem, not as a claim that five dimensions are universally necessary and sufficient for all learning with AI. Other constructs— affect, identity, motivation, accessibility, or disciplinary culture—may require explicit extension in particular contexts. Within the delimited problem of faculty development for AI-mediated higher-order learning, however, the five principles capture the recurring design requirements found across the selected literature while remaining concise enough for use in design studios and assessment review.

2.3 | Scope and Boundaries

The paper focuses on higher education faculty and academic-development practice. Selected school-sector and professional-education studies were used only when they illuminated transferable features of teacher learning, competency formation, or applied AI integration. GenAI refers principally to large language models and multimodal systems that can generate text, code, images, explanations, feedback, or simulations in response to prompts. The concern was not to catalog tools, which change rapidly, but to theorize durable design decisions that remain relevant across platforms.

Two boundaries are central. First, the model does not assume that AI should always be used. Initial recall, routine procedural fluency, time-pressured independent performance, and safety-critical competence may require restricted or absent AI use. Purposeful non-use is therefore compatible with PEDAL. Second, the paper does not offer empirical validation of the model. The worked examples are design propositions, not reports of implemented interventions or student outcomes. PEDAL should be treated as a sensitizing and testable architecture whose value depends on future enactment, comparison, and refinement across disciplines and institutional contexts.

3 Higher-Order Thinking as a Design and Evidentiary Problem

3.1 | Distinguishing Intellectual Operations, Regulation, and Transfer

Higher-order thinking is frequently invoked as an unqualified educational good, but the construct becomes analytically weak when every nonroutine

activity is placed under the same label. For the purposes of this paper, it is treated as a layered construct. The first layer comprises core intellectual operations: analysis, evaluation, and creation, as foregrounded in Anderson and Krathwohl's (2001) revision of Bloom's taxonomy. Analysis involves differentiating and relating components of a claim, problem, or system. Evaluation involves applying disciplinary criteria to evidence and alternatives. Creation involves producing a defensible synthesis, design, explanation, or artifact rather than merely recombining fluent language.

The second layer comprises regulatory processes. Facione (1990) defined critical thinking as purposeful, self-regulatory judgment, thereby locating monitoring and revision inside rather than outside intellectual performance. Metacognition concerns awareness of how one's understanding and strategy are changing, whereas co-regulation concerns how that monitoring is distributed across peers, teachers, tasks, and technological agents. Garrison et al. (2001) described cognitive presence as movement through triggering, exploration, integration, and resolution; Lodge et al. (2023a) extended the regulatory perspective to networks that include GenAI. These accounts imply that a learner's capacity to interrogate an AI response, recognize uncertainty, revise a position, and explain why a suggestion was rejected can be as educationally significant as the final product.

The third layer concerns durable learning indicators, particularly transfer. Brookhart (2010) treated transfer as the ability to use knowledge in a new situation rather than reproduce it in a familiar format. This distinction matters because an activity can support evaluation without demonstrating transfer. Students may accurately critique a flawed AI-generated explanation after extensive prompting yet fail to recognize the same misconception in a novel case. Similarly, a prompt log may show activity without showing disciplinary judgment. A credible design must therefore match evidence to the intended outcome: Comparative annotations may evidence analysis; a criteria-based recommendation may evidence evaluation; a new design under constraint may evidence creation; a reflective contrast between initial and revised reasoning may evidence metacognition; and a short, unfamiliar application may evidence transfer.

This specificity constrains inflated claims about AI-enhanced higher-order thinking. The presence of dialogue, creativity language, or critical-evaluation instructions does not guarantee the corresponding learning outcome. Faculty need to identify the intellectual verb, the disciplinary object on which it operates, the criteria by which performance will be judged, and the evidence through which ownership of the judgment becomes visible. PEDAL begins with this requirement.

3.2 | GenAI and the Visibility of Learning

GenAI complicates assessment because it can perform many actions that resemble academic work at the surface product level. It can summarize readings, draft arguments, generate code, propose designs, translate text, and imitate disciplinary discourse with considerable fluency. However, the production of an answer is not identical to the production of understanding. If higher-order learning includes weighing evidence, resolving ambiguity, justifying criteria, monitoring uncertainty, and transferring knowledge, then polished output may conceal rather than demonstrate learning (Azzam & Charles, 2025).

The resulting problem is epistemic as well as technical. Traditional take-home tasks often infer competence from a finished artifact. GenAI weakens that inference because the relationship between output quality and learner reasoning is now less direct. This does not mean that every assessment must be invigilated or conducted orally. This means that educators require a richer evidence ecology in which the final product is interpreted alongside proportionate traces of judgment, revision, verification, and application. Assessment reform discussions therefore emphasize trustworthy judgments about learning and students' capacity to work responsibly in an AI-rich world (Lodge et al., 2023b). Perkins et al. (2024) introduce a language for communicating permitted levels of AI use, but permission alone does not establish whether a task remains intellectually demanding.

GenAI is pedagogically indeterminate. Under well-designed conditions, it can externalize alternatives, simulate a stakeholder, generate a plausible misconception, provide a contrasting explanation, or create an imperfect artifact for diagnosis. Borge et al. (2024), for example, demonstrated the potential of a generative simulation to support meta- and socio-metacognitive sense-making. Under weak conditions, the same system can encourage cognitive offloading, automation bias, generic synthesis, and premature closure. PEDAL proposes that educational effects depend on the relationships among learner purpose, task challenge, interaction protocol, accountability, and evidence. These relationships, rather than the mere presence of the tool, are the object of faculty development.

3.3 | Faculty Capability and Contextual Variation

Competency and professional-knowledge frameworks establish that responsible AI integration is multidimensional. Technological pedagogical content knowledge (TPACK) positions technology integration as an interaction among technological, pedagogical, and content

knowledge (Mishra & Koehler, 2006), while Intelligent-TPACK makes ethical knowledge more explicit in AI-supported teaching (Celik, 2023). AI literacy frameworks add understanding, use, evaluation, communication, creation, and ethical awareness (Long & Magerko, 2020; Ng et al., 2021). UNESCO's teacher framework broadens the professional horizon to include human-centered values, AI foundations, pedagogy, and professional learning (Miao & Cukurova, 2024). The AI fluency framework's emphasis on delegation, discernment, and diligence is likewise relevant to deciding which intellectual work should remain human-led (Dakan & Feller, 2025).

These frameworks are necessary but insufficient for course redesign. A faculty member may understand hallucination, privacy, and prompt construction, yet still design an assessment that rewards generic output. Conversely, an expert assessor may recognize disciplinary judgment but lack protocols for transparent AI use. Faculty capability therefore needs an enactment layer: the capacity to convert professional knowledge into bounded AI roles, cognitively demanding tasks, dialogic sequences, fair accountability requirements, and feasible evidence plans.

Context determines how that capability can be enacted. Khlaif et al. (2024) showed that educators' adoption of GenAI for assessment is shaped by expectations of usefulness, social influence, institutional policy, and professional-development needs. Palestinian case studies indicate that AI can disrupt rote-learning conventions and open new forms of engagement but also intensify uncertainty about authorship, academic integrity, and educational purpose (Hamamra et al., 2024). Educators described both promise and precarity when innovation was pursued under constrained institutional conditions (Hamamra et al., 2026). Work on AI

dependency and technostress added a further warning: Efficiency gains can coexist with reduced professional autonomy, cognitive overreliance, workload pressure, and a need for clearer policy and training (Alhur et al., 2025; Khlaif et al., 2026).

Global sensitivity, therefore, cannot be achieved by declaring a model universally applicable. It requires explicit design options for contexts in which students do not share equal access to paid systems, institutions cannot mandate external accounts, bandwidth is unreliable, or privacy law limits uploading course material. It also requires multilingual and disciplinary adaptation. PEDAL is intended to travel through its questions rather than through standardized technological answers. A program can use lecturer-prepared AI outputs, locally approved tools, shared demonstrations, or no live AI at all while still asking students to analyze alternatives, verify evidence, articulate judgment, and transfer learning.

4 From Adjacent Frameworks to PEDAL

A design-oriented model of future faculty development must be theoretically grounded but operationally usable. Table 2 positions PEDAL among the principal adjacent frameworks. The comparison is not intended to diminish those frameworks. It clarifies the level at which each operates and the specific integration problem PEDAL addresses.

Constructive alignment supplies PEDAL's starting discipline: outcomes, activities, and assessment must cohere. It does not, by itself, specify what makes an AI-mediated task epistemically resistant to generic

Table 2 Positioning PEDAL among adjacent frameworks

Framework or lens	Main contribution	Limitation for the present problem	What PEDAL integrates
Constructive alignment and teaching as a design science (Biggs & Tang, 2011; Laurillard, 2012)	Aligns outcomes, activities, representations, feedback, and assessment	Not AI-specific; does not by itself specify how AI-mediated tasks preserve judgment, dialogue, or accountability	Converts alignment into five linked questions for AI-rich higher-order learning
TPACK/Intelligent-TPACK (Celik, 2023; Mishra & Koehler, 2006)	Explains technological, pedagogical, content, and ethical professional knowledge	Stronger on what faculty should know than on how to sequence decisions within a particular task	Links professional knowledge to task construction, facilitation, verification, and evidence
AI literacy and competency frameworks (Long & Magerko, 2020; Miao & Cukurova, 2024; Miao et al., 2024; Ng et al., 2021; Zhou & Schofield, 2024)	Foreground understanding, use, evaluation, creation, ethics, and human-centered agency	Do not necessarily specify structures that make particular higher-order outcomes visible in coursework	Embeds literacy inside a design cycle where discernment informs challenge, dialogue, and evidence
Co-regulation and cognitive presence (Garrison et al., 2001; Lodge et al., 2023a)	Explains inquiry, reflection, interaction, and distributed regulation	Under-specifies outcome selection, disclosure, assessment redesign, and evidentiary validity	Connects dialogic learning to explicit purpose, accountability, and outcome-aligned evidence
AI assessment guidance (Lodge et al., 2023b; Perkins et al., 2024)	Clarifies assessment reform and permissible levels of AI use	Permission or reform principles alone do not show how an individual task remains epistemically demanding	Connects AI-use boundaries to challenge, judgment, metacognition, transfer, and feasible evidence

Notes. PEDAL: purpose, epistemic challenge, dialogue, accountability, and learning evidence; TPACK: technological pedagogical content knowledge.

completion, how dialogue should be structured, or what accountability should look like when AI assistance is permitted. TPACK and Intelligent-TPACK clarify the knowledge faculty need but are less explicit about the sequence of design decisions within an assignment. AI literacy and competency frameworks specify responsible understanding and use, yet they do not necessarily translate those competencies into evidence of analysis, evaluation, creation, or transfer. Co-regulation and cognitive presence explain inquiry and monitoring but under-specify disclosure and assessment-validity requirements. Assessment frameworks clarify acceptable AI use and trustworthy judgment but may say more about permission levels or system-level reform than about the internal architecture of a particular learning task.

PEDAL's distinctive contribution is therefore integrative and meso-level. It reorganizes established insights around five questions that can be used by a faculty member, program team, or academic developer while redesigning authentic course assignments. The model links macro-level principles, such as human agency and institutional governance, to micro-level decisions about prompts, sources, discussion, verification, and evidence. It also treats faculty development as design development. The object of professional learning is not merely the tool or policy; it is the capacity to create and evaluate learning arrangements in which human judgment remains educationally indispensable.

The five dimensions are related but non-interchangeable. Purpose without epistemic challenge can produce a well-worded outcome attached to an answerable task. Challenge without dialogue can turn uncertainty into an unsupported struggle. Dialogue without accountability can normalize extensive AI assistance without verification or authorship responsibility. Accountability without proportionate learning evidence can become a surveillance or administrative burden. Learning evidence without a defined purpose can multiply artifacts without increasing validity. The model's value lies in making these dependencies visible and using evidence to revise the preceding design decisions.

This positioning also moderates the originality claim. PEDAL does not introduce previously unknown concepts. Its originality lies in selecting, relating, and sequencing concepts that have usually been distributed across curriculum, assessment, AI literacy, dialogic learning, and professional-development literatures. It offers a shared vocabulary and workflow for faculty-development settings in which academics need to move from abstract guidance to a defensible task design. The claim is thus practical and conceptual rather than foundational: PEDAL is a design heuristic that makes a cross-framework synthesis usable and empirically testable.

5 PEDAL Model for AI-Mediated Higher-Order Learning

Figure 1 presents the components of the PEDAL model as an iterative cycle. The circular arrangement is important: Learning evidence is not an endpoint added after teaching, but a source of information for revising purpose, challenge, dialogue, and accountability in subsequent enactments.

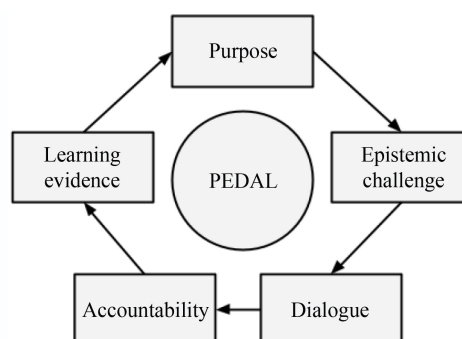


Figure 1 PEDAL model as an iterative design cycle for AI-mediated higher-order learning.

5.1 | Purpose

Purpose asks: What exact form of higher-order learning is intended, and what bounded role, if any, should AI play in relation to it? Institutional conversations often begin with tool affordances—summarization, feedback, role-play, image generation, or coding—and only later ask whether the activity advances a curriculum outcome. PEDAL reverses that sequence. Faculty first specify the intellectual operation and disciplinary object: for example, evaluate the evidential strength of competing causal explanations; create a design that meets specified safety and cost constraints; monitor how an interpretation changes after counterevidence; or transfer a diagnostic principle to an unfamiliar case.

Outcome specificity should include the criteria of quality. “Critically evaluate an AI response” is incomplete unless students know what counts as a strong evaluation in the discipline. In history, the criteria may include provenance, contextualization, and corroboration. In engineering, they may include technical validity, constraint satisfaction, uncertainty, and risk. In teacher education, they may include pedagogical alignment, inclusivity, language support, and feasibility. The role of AI is then assigned instrumentally: generator of alternative claims, simulated client, novice reasoner, source of misconceptions, debate partner, or producer of a flawed draft. AI is not the purpose of the

task; it is a designed condition through which the intended judgment is elicited.

Purpose also legitimizes non-use. Where the intended outcome is independent recall, unaided procedural fluency, or safety-critical performance, AI may be excluded during the evidentiary phase, even if it was used during preparation. Faculty development should therefore help academics distinguish between learning with AI, learning about AI, and demonstrating competence without AI. A coherent course may include all three.

5.2 | Epistemic Challenge

The epistemic challenge asks: Where must learners interpret ambiguity, manage uncertainty, weigh competing criteria, or make a context-sensitive judgment that cannot be replaced by fluent output? GenAI often produces plausible answers that smooth over disagreement and conceal evidential weakness. Tasks designed for higher-order learning must therefore contain features that make human deliberation necessary.

Such features include incomplete or conflicting evidence, stakeholder trade-offs, local constraints, multiple defensible solutions, disciplinary standards, and the possibility of error. Faculty can ask students to compare multiple AI-generated explanations, locate unsupported assumptions, verify claims against primary or scholarly sources, diagnose a persuasive hallucination, or recommend action when no option satisfies every criterion. The challenge should not be artificial complexity. It should resemble the uncertainty through which disciplinary judgment is formed.

An epistemically challenging task is intended to limit, and may reduce, the value of generic prompting. Contextual specificity is particularly powerful: local data, course-specific readings, an unfamiliar case, a requirement to justify criteria, or a changed constraint can make a stock response inadequate. Khlaif et al. (2025) similarly emphasized assessment redesign that preserves learner agency and human-AI collaboration rather than relying on resistance alone. Faculty development can operationalize this principle by stress-testing an assignment: Can a competent student obtain a high mark by submitting a lightly edited first response from a general-purpose model? If the answer is yes, the task probably needs a stronger epistemic core or richer evidence.

Epistemic challenges must be calibrated. Too little challenge permits substitution; too much unsupported ambiguity can disadvantage novices. Cognitive apprenticeship suggests that modeling, coaching, and scaffolding should precede independent exploration (Collins et al., 1989). Faculty can therefore provide

worked examples of claim verification, criteria application, or prompt interrogation before expecting students to perform these practices independently.

5.3 | Dialogue

Dialogue asks: How will learners question, probe, compare, and revise rather than accept the first output? GenAI can function as a conversational partner, yet interaction does not become educational merely because it is multiturn. A long prompt history may consist of repeated requests for polishing rather than inquiry. PEDAL therefore requires a protocol that structures movement between learner reasoning, AI output, disciplinary evidence, peer critique, and tutor questioning.

Dialogic protocols can require students to articulate an initial position before consulting AI; request contrasting explanations; identify what evidence would distinguish them; challenge the model with counterexamples; compare its response with course sources; and revise a claim in public or peer-reviewed form. In role play, AI can simulate a stakeholder, patient, client, novice, or skeptical reviewer, but the student must respond from a disciplinary position and explain the basis of the response. Co-regulation becomes visible when learners monitor not only the model but also their own and peers' reasoning (Lodge et al., 2023a).

Metacognition is explicit within dialogue rather than added as a sixth acronym element. Students should be asked what changed in their understanding, which suggestion they initially found persuasive, how they recognized an error, and what strategy they would use in a future task. Reflective prompts are strongest when tied to a concrete decision rather than inviting generic commentary. For example: "Identify one AI suggestion you rejected, state the disciplinary criterion that justified rejection, and explain how the episode changed your approach to the final recommendation." This creates evidence of regulation rather than a retrospective declaration that reflection occurred.

Dialogic design also resocializes AI use. Student-AI interaction can otherwise become an invisible, individual exchange. Connecting it to peer discussion, tutor questioning, collaborative comparison, and public justification may reduce the risk that the system will become the silent author of an unexamined product. It may also give faculty opportunities to observe misconceptions and reasoning strategies while learning is still in progress.

5.4 | Accountability

Accountability asks: How will students verify, disclose, and justify AI use while retaining responsibility for the final judgment? Institutional responses often reduce

accountability to detection and misconduct. PEDAL adopts a broader conception encompassing ethical, epistemic, authorship, privacy, and professional responsibilities.

Students should be accountable not only for whether they used AI but also for what role it played, which outputs they accepted or rejected, how factual claims were checked, what data were shared, and how responsibility for the final work was maintained. This is consistent with human-centered AI guidance (Miao & Holmes, 2023; UNESCO, 2021), competency frameworks that integrate ethics with use (Miao & Cukurova, 2024; Miao et al., 2024), and Intelligent-TPACK's inclusion of ethical professional knowledge (Celik, 2023). It also responds to empirical concerns about overreliance and reduced professional autonomy (Alhur et al., 2025).

Accountability can be operationalized through concise disclosure statements, source-verification tables, annotated excerpts of consequential prompts, bias or limitation checks, and rationales for accepted and rejected suggestions. Full prompt transcripts should not be a default requirement. They can expose personal data, create excessive marking burdens, and reward performative documentation rather than meaningful reasoning. Faculty should request only the evidence necessary to support the assessment inference.

Privacy and access belong inside accountability. Students should not be required to upload sensitive, copyrighted, clinical, proprietary, or identifiable material to external services. Where approved tools are unavailable or access is unequal, lecturers can supply pre-generated outputs or offer an equivalent non-AI route. Transparency expectations should apply to the institution as well as to the student: approved platforms, data-retention conditions, assessment criteria, and consequences of non-disclosure need to be clear.

5.5 | Learning Evidence

Learning evidence asks: What proportionate evidence will demonstrate thinking rather than only polished output? In AI-rich environments, the final products remain educationally relevant, but their interpretation requires additional evidence when the products could have been generated or extensively transformed by a model. The aim is not to document every cognitive event. It is to create a valid evidence set aligned with the intended outcome.

Evidence can include an initial position or concept map, a criteria table, source-verification notes, staged drafts, targeted prompt annotations, an error analysis, peer feedback, a short oral defense, an in-class application, a reflective memo, or a transfer task. Different forms support different inferences. An oral defense can establish ownership and responsiveness but may disadvantage students with particular communication needs if used without accommodation. A prompt log

can show interaction but not necessarily judgment. A transfer task can reveal durable understanding, but it should be brief enough to remain feasible. A design should therefore combine complementary forms rather than treat any single trace as definitive.

Metacognitive evidence is located here as well as in dialogue. Students can compare initial and revised reasoning, identify a turning point, explain the criteria used to reject an AI suggestion, and state how they would approach a new problem. This distinguishes evidence of learning from evidence of output quality. The final product may improve without the learner understanding why; conversely, a learner may demonstrate substantial conceptual development, even when the product remains imperfect.

Learning evidence must be assessed with validity and fairness in mind. Automated scoring should not be assumed to be equivalent to expert human judgment. Zacharis and Papadakis (2025), for example, found materially weaker AI-human agreement than human inter-rater agreement when complex university coursework was scored with a common rubric. AI may support administrative preparation or formative comparison, but high-stakes decisions require human oversight, moderation, and a clear account of what the evidence can support.

5.6 | Relationships and Boundary Conditions

PEDAL is an integrated architecture rather than a checklist of independent virtues. Purpose and epistemic challenge form the design core: they establish the intended outcome and protect the necessity of human judgment. Dialogue and accountability are enactment conditions; they shape how the learner engages with AI, peers, evidence, and ethical responsibilities. Learning evidence is proposed as an evidentiary check and source of feedback; it can indicate whether an enacted design appears to have produced the intended form of learning and can inform redesign.

Table 3 summarizes the relationship between each principle, particular higher-order outcomes, design moves, and evidence. The table deliberately avoids mapping every principle to every outcome. A design may target evaluation and metacognition without claiming creation or transfer. Such restraint improves conceptual precision and assessment validity.

PEDAL is proposed primarily for ill-structured, inquiry-rich, interpretive, problem-based, and design-oriented tasks in which ambiguity and justification are central. It is less central for initial recall, routine procedural practice, or time-pressured demonstration of independent safety-critical competence. In these contexts, AI may be limited to preparation or excluded. The model also does not override disciplinary ethics or

Table 3 PEDAL as a design heuristic for distinct higher-order outcomes

Principle	Primary outcome focus	Illustrative design move	Proportionate evidence
Purpose	Analysis, evaluation, creation, metacognition, or transfer, each specified rather than aggregated	State the intellectual operation, disciplinary object, quality criteria, and bounded AI role before choosing a tool	Outcome map; revised task brief; rubric criteria that name the intended judgment
Epistemic challenge	Analysis, evaluation, creation, and context-sensitive transfer	Introduce conflicting evidence, local constraints, plausible error, competing criteria, or an unfamiliar case	Annotated diagnosis; verified comparison; criteria-based recommendation; revised design under constraint
Dialogue	Analysis, evaluation, metacognition, and co-regulation	Require an initial position, multitransactional interrogation, peer challenge, evidence checking, and reasoned revision	Selected interaction annotation; peer critique; contrast between initial and revised reasoning
Accountability	Evaluative judgment and metacognitive responsibility	Disclose the AI role, verify consequential claims, justify accepted/rejected suggestions, and protect data	Concise AI-use statement; source-verification table; decision rationale; privacy confirmation
Learning evidence	Outcome-specific demonstration, especially metacognition and transfer	Combine final product with the smallest complementary set of process, oral, in-class, or novel-application evidence	Decision memo; staged artifact; brief structured oral defense (micro-viva); brief transfer task; evidence portfolio

professional regulation. A clinical, legal, engineering, or teacher education program may require additional constraints concerning confidentiality, public safety, licensure, or child protection.

6 Worked Redesigns and Proposed Feasibility Safeguards

The utility of a design heuristic depends on whether it can guide the transformation of actual assignments. [Table 4](#) presents two hypothetical worked assignment redesigns: a social science policy essay and an engineering design task. They are not empirical cases, and no claims are made about student outcomes. Their purpose is to show how the five principles can be enacted, from task brief to assessment evidence.

6.1 | Worked Case 1: From Conventional Essay to Evidence-Tested Policy Judgment

A conventional social science assignment might ask students to write a 2,000-word essay evaluating whether a government should adopt a contested education policy. The task nominally targets evaluation, but a general-purpose model can produce a plausible essay with balanced arguments, invented citations, and generic recommendations. The PEDAL redesign begins by specifying the outcome more precisely: Students must evaluate competing policy options using explicit criteria, verify evidence, explain a change in judgment, and transfer their criteria to a related but unfamiliar case.

The task opens with a short, unaided position statement and criteria map. AI is then assigned a

Table 4 Two worked assignment redesigns using PEDAL

Design element	Social sciences: contested policy judgment	Engineering: constrained system design
Conventional task	Write a 2,000-word essay evaluating whether a government should adopt a contested education policy	Select an optimal energy-system design and submit a technical report
Purpose and AI role	Evaluate alternatives, explain a change in judgment, and transfer criteria. AI generates three contrasting stakeholder memos; it does not write the final recommendation	Diagnose alternatives, justify trade-offs, create a revised design, and transfer the decision process. AI proposes alternatives and counterarguments; it does not validate calculations or standards
Epistemic challenge	Conflicting priorities, incomplete evidence, local context, and required distinction between empirical claims and value judgments	One superficially attractive but flawed option; uncertain assumptions; cost, safety, reliability, environmental, and regulatory trade-offs
Dialogue	Unaided initial position; structured AI interrogation; source checking; peer challenge; revised recommendation	Team diagnosis; AI-simulated client questions; tutor probe of one consequential decision; individual reflection on revisions
Accountability	Verify four consequential claims; disclose AI role; justify one accepted and one rejected suggestion; avoid uploading identifiable data	Verify calculations and standards; identify corrected AI content; disclose role; confirm no confidential or proprietary data were uploaded
Learning evidence and criteria	Initial criteria map, annotated claim check, final recommendation, one-page decision memo, metacognitive commentary, and brief transfer item. Criteria: evidence, judgment, assumptions, metacognition, and transparency	Alternative matrix, calculations, risk register, revised design rationale, short briefing, and individual changed-constraint task. Criteria: technical validity, uncertainty, trade-offs, creation, transfer, and accountability
Feasibility safeguards	One-page templates, guided peer review, lecturer-supplied outputs, selective micro-viva, and accessible non-AI route	Shared outputs, standard verification matrix, team artifacts plus individual transfer, sampled questioning, approved tools only

bounded role. It generates three stakeholder memos that contain different priorities, assumptions, and evidential weaknesses. Students interrogate one memo through a structured sequence, verify at least four consequential claims against course readings or primary sources, and annotate where the model has conflated evidence, value judgment, and speculation. A peer then challenges the criteria and requests one counterexample. Students submit a revised policy recommendation, a one-page evidence-and-decision memo, and a 250-word metacognitive commentary identifying one accepted and one rejected AI suggestion. A brief, in-class transfer item asks them to apply the same criteria to a new jurisdiction with different constraints.

The assessment criteria are aligned to the specified outcomes rather than to the volume of AI documentation. Disciplinary use of evidence and defensibility of judgment carry the greatest weight. The evidence memo assesses verification and accountability; the commentary assesses metacognitive regulation; and the transfer item tests whether the evaluative framework travels beyond the rehearsed case. The final prose remains important, but elegance alone cannot compensate for weak criteria or unverified claims.

6.2 | Worked Case 2: From Optimal Design Report to Constrained Engineering Judgment

A conventional engineering assignment might ask teams to select and justify an energy system design for a specified site. GenAI can produce candidate designs and polished rationales, but it cannot be trusted to satisfy technical assumptions, safety constraints, local standards, or trade-offs without verification. The redesigned outcome is therefore to diagnose design alternatives, justify trade-offs under uncertainty, create a defensible revised design, and transfer the decision process when a key constraint changes.

The lecturer supplies or approves three AI-generated alternatives, including one that is superficially attractive but technically flawed. Teams identify assumptions, run or reproduce necessary calculations, compare alternatives against cost, reliability, safety, and environmental criteria, and record where evidence is incomplete. AI may be used to propose counterarguments or simulate a client question, but calculations and standards must be verified through accepted disciplinary sources and tools. Each team submits an annotated alternative matrix, a risk-and-assumption register, and a revised design rationale. A short briefing allows the tutor to probe one consequential decision, while an individual timed transfer task introduces a changed load profile or regulatory constraint.

The design distinguishes collaborative creation from individual competence. The team artifacts

evidence distributed inquiry and design synthesis; the individual transfer item and targeted questioning evidence ownership. Accountability is focused on consequential decisions rather than every prompt. Students disclose the AI role, identify any generated calculations or claims that required correction, and confirm that no confidential or proprietary data were uploaded. Technical validity, uncertainty management, trade-off reasoning, and transfer are assessed separately so that a visually polished report cannot mask weak engineering judgment.

6.3 | Workload, Scalability, Privacy, Accessibility, and Fairness

Richer evidence may improve validity but may also increase burden. PEDAL therefore requires proportionality. Faculty should select the smallest evidence set that supports the intended inference rather than accumulating every possible trace. A one-page decision memo may be more informative than a full-prompt transcript. A three-minute micro-viva can be used for a stratified or rotating sample when a full oral defense is impractical, provided the sampling rule is declared and does not determine the entire grade. Structured templates, checklists, and analytic rubrics may help contain marking time while directing attention to consequential reasoning.

For large classes, evidence sampling and carefully bounded automation may help contain workload without outsourcing judgment. Students can use standardized verification tables; peer critique can be guided by two or three criteria; tutors can review selected decision points rather than every interaction; and moderation can focus on common exemplars. Group dialogue may be combined with an individual transfer task so that collaboration does not obscure individual learning. Where workload is especially constrained, the final product, one process artifact, and one brief transfer item may provide a defensible minimum.

Privacy should shape the task before students begin. No assessment should require the disclosure of sensitive personal information, clinical details, unpublished research, copyrighted examination material, or proprietary organizational data to an external system. Where institutional accounts and data agreements are unavailable, lecturers can provide static AI outputs for critique or permit an equivalent non-AI pathway. Students should be told whether prompts are retained, who can access them, and which information must not be entered. A disclosure requirement that forces students to reproduce private conversations or personal data is not ethically proportionate.

Accessibility and fairness require equal opportunities to demonstrate the same outcomes. Students

may differ in paid-tool access, bandwidth, disability-related communication needs, language proficiency, and prior AI experience. Institutions should provide an accessible baseline tool or lecturer-generated materials, allow reasonable format alternatives, and avoid making the speed of prompting a hidden criterion. Oral evidence should include accommodations and should not become a proxy test for accent, confidence, or spontaneous fluency. Multilingual students may be permitted to use AI for language support when the assessed outcome is disciplinary reasoning, provided the role is declared and the rubric does not falsely interpret surface language as independent conceptual mastery.

Fairness also requires procedural safeguards. AI detection scores should not be treated as proof of misconduct. Assessment briefs should state the permitted role of AI, the evidence required, and the process for clarifying uncertainty. Human review, moderation, and an avenue for student explanation or appeal are essential when authorship or integrity concerns arise. Accountability is educative when it clarifies responsibility and supports judgment; it becomes punitive when opaque tools are used as substitutes for evidence.

7 Translating PEDAL into Faculty Development and Institutional Support

If PEDAL is to function as a future faculty-development model rather than only an assignment checklist, each principle must be translated into professional capability, authentic learning activity, and institutional support. Faculty do not need a decontextualized inventory of prompts. They need structured opportunities to redesign real modules, test assumptions,

examine student work, and refine standards with colleagues. [Table 5](#) summarizes this translation.

7.1 | Developmental Progression

The first phase should establish foundational AI literacy. Faculty need a working understanding of how generative models produce outputs, why hallucination and bias occur, what privacy and copyright risks matter, and how automation can affect judgment. Functional prompting is included, but it is subordinate to evaluation, ethics, and disciplinary interpretation. The teacher competency framework ([Miao & Cukurova, 2024](#)), AI literacy models ([Long & Magerko, 2020](#); [Ng et al., 2021](#)), and Intelligent-TPACK ([Celik, 2023](#)) provide complementary resources for this baseline.

The second phase should use discipline-based design studios. Effective professional development is sustained, content-focused, collaborative, and connected to teachers' actual work ([Garet et al., 2001](#)). Faculty bring an existing task, identify its intended higher-order outcome, stress test its vulnerability to generic AI completion, assign a bounded AI role, and construct an evidence plan. Colleagues then critique the design from the perspectives of disciplinary validity, student access, workload, and policy. The output is a revised assessment brief and rubric, not merely workshop attendance.

The third phase should center on enactment, moderation, and interpretation of student work. Faculty trial redesigned tasks, collect proportionate evidence, and compare judgments using anonymized exemplars. This phase is crucial because design quality cannot be inferred from the brief alone. Evidence may reveal that students complied with prompt logging without engaging metacognitively, that an oral component assessed confidence rather than conceptual ownership, or that

Table 5 PEDAL as a faculty-development architecture

PEDAL principle	Faculty capability	Authentic professional-learning activity or output	Support continuum
Purpose	Articulate disciplinary higher-order outcomes and bounded AI roles	Outcome-mapping clinic using an existing assignment; revised brief and rubric	Minimum: shared outcome/AI-role template; coordinated: curriculum mapping, exemplars, and program-level assessment review; mature: cross-program repositories and longitudinal evaluation
Epistemic challenge	Design ambiguity, contestability, verification, and human judgment into tasks	Design studio that stress-tests answerability by AI; revised case, problem, or data set	Minimum: peer critique and redesign time; coordinated: facilitated studios and evaluation support; mature: disciplinary repositories and research support
Dialogue	Structure student–AI–peer–tutor interaction as inquiry and metacognitive regulation	Facilitation clinic using questioning protocols, peer critique, and reflective prompts	Minimum: approved or lecturer-supplied outputs; coordinated: accessible tools, mentoring, and multilingual guidance; mature: observation networks and cross-program exemplars
Accountability	Embed disclosure, verification, privacy, authorship, and ethical reasoning	Policy clinic; concise disclosure and source-verification templates; ethical case analysis	Minimum: agreed boundaries and a privacy-safe route; coordinated: policy alignment, due process, and student guidance; mature: data governance and procurement standards
Learning evidence	Assess judgment, revision, metacognition, and transfer without disproportionate burden	Rubric redesign and moderation with sample work; minimum-evidence planning	Minimum: shared moderation and one complementary evidence form; coordinated: workload recognition, accessibility support, and quality assurance; mature: longitudinal evaluation across programs

the AI role was so constrained that it added no educational value. Moderation turns PEDAL into an iterative professional-learning cycle.

The fourth phase can support scholarship and local evaluation. Design-based research, program evaluation, or structured reflective inquiry can examine implementation fidelity, student experience, workload, equity, and outcome evidence. Faculty communities and professional learning networks are useful because GenAI changes too rapidly for expertise to remain centralized in a single office (Trust et al., 2016). Networks also enable multilingual adaptation and the sharing of failures as well as successful exemplars.

7.2 | Minimum, Coordinated, and Mature Institutional Conditions

Institutional support should be differentiated rather than presented as an undifferentiated list of desirable resources. At a minimum, a program team may be able to pilot PEDAL without a large central infrastructure if four provisional conditions are in place: an agreed statement of permitted AI roles; a privacy-safe means of giving students access to the relevant AI outputs; time to redesign at least one assessment; and a shared moderation process. The program can use lecturer-prepared outputs, a concise disclosure template, and a common rubric, even where no enterprise AI platform exists. These conditions are proposed as a pragmatic starting point for a bounded pilot; their adequacy requires evaluation in context, and they do not resolve wider governance risks.

Coordinated institutional implementation requires central academic-development support, approved tools or alternatives, privacy and accessibility guidance, assessment-regulation alignment, and recognized workload. Design studios need facilitators and protected time. Policies need to distinguish acceptable assistance, disclosure, verification, and prohibited data use. Research on educators' trust similarly identifies professional support, training, and institutional conditions as central to adoption (Lelescu et al., 2025; Nazaretsky et al., 2022). Student-facing guidance should be consistent enough to avoid contradictory expectations across modules. Quality-assurance processes should recognize process and transfer evidence rather than requiring every module to retain product-dominant assessment.

A mature institutional architecture adds procurement standards, data-governance expertise, accessibility testing, longitudinal evaluation, cross-program repositories, student partnerships, and research support. It can monitor whether AI-related redesign disproportionately burdens particular staff or student groups, whether approved systems perform consistently across languages, and whether new evidence practices improve assessment validity. It can

also support cross-institutional networks and open exemplars, so that pedagogical innovation is not limited to universities with extensive commercial infrastructure.

7.3 | Resource-Diverse and Multilingual Implementation

The Palestinian studies incorporated into this review make clear that institutional readiness cannot be reduced to individual enthusiasm. Faculty may perceive GenAI as useful while simultaneously encountering policy ambiguity, unstable access, dependency concerns, and technostress (Alhur et al., 2025; Hamamra et al., 2026; Khlaif et al., 2024; Khlaif et al., 2026). Professional development in such settings should therefore avoid individualizing risk. Asking faculty to innovate without approved tools, privacy guidance, workload recognition, or assessment flexibility risks producing defensive compliance rather than thoughtful redesigns.

Low-resource implementation does not require low intellectual ambition. Many PEDAL activities can use one shared AI output projected in class, printed alternatives, or lecturer-generated examples. Students can compare, verify, debate, and revise without personal accounts or continuous connectivity. Multilingual guidance, local examples, and discipline-specific criteria may matter more than access to the newest model. Borderless faculty communities can pool templates and cases, but imported examples should be adapted to local assessment cultures, languages, and institutional constraints.

The distinction among minimum, coordinated, and mature support is therefore important for equity. A program should not delay all redesigns until enterprise procurement is complete, but institutions should also not use local ingenuity as an excuse to withhold governance and workload support. PEDAL is most credible when it makes all three levels visible: Faculty agency enables bounded pedagogical action, while institutional responsibility determines whether that action is safe, fair, and sustainable.

8 Discussion

8.1 | Principal Contributions

PEDAL makes four principal contributions to debates about GenAI and higher education. First, it explains pedagogical indeterminacy. The same system may support or undermine higher-order learning, depending on how purpose, challenge, dialogue, accountability, and evidence are related. This shifts attention away

from general claims about the technology and toward the design conditions under which students must reason, verify, revise, and transfer.

Second, PEDAL reframes faculty development as design development. Tool familiarity and policy awareness matter, but neither is sufficient. A faculty member may understand how to prompt a model and still lack a principled basis for deciding whether an assignment promotes evaluation rather than imitation. By sequencing five design questions, PEDAL specifies a more educationally defensible object of professional learning. Its meso-level position is central: It translates broad competency and ethical frameworks into task-level decisions while connecting those decisions to program and institutional support.

Third, the model foregrounds the evidentiary argument. In AI-rich environments, assessment cannot assume that polished output is a transparent proxy for understanding. PEDAL does not respond by requiring exhaustive surveillance. It proposes proportionate evidence ecologies in which product, process, metacognition, and transfer are combined according to the intended outcome. The worked examples show how this can be done without treating every prompt as assessable evidence or every student as a suspected offender.

Fourth, the model addresses recurrent weaknesses in institutional responses. Tool-first development begins with capability demonstrations rather than learning outcomes. Integrity-only framing identifies risks but supplies few design alternatives. Product-dominant assessment confuses fluency with ownership. Generic guidance ignores disciplinary epistemologies and unequal access. PEDAL responds by placing purpose first, embedding uncertainty in epistemic challenge, structuring inquiry through dialogue, broadening accountability beyond detection, and treating learning evidence as both validation and a mechanism for redesign.

8.2 | Limitations and Research Agenda

Several limitations remain. The review is conceptual and purposive, not systematic or exhaustive. The transparent search routes, selection functions, analytic strands, and derivation matrix reduce impressionism but do not eliminate the selectivity inherent in theory-building synthesis. The model also remains unvalidated. Its internal coherence and practical plausibility do not establish effectiveness for student learning, faculty capability, or institutional change.

The corpus is broad but cannot represent every region, language, discipline, or institutional type. Palestinian and Middle Eastern studies strengthen the analysis of resource diversity, yet they should not be used as proxies for the Global South as a whole. The

model will require contextual refinement in indigenous, multilingual, low-connectivity, transnational, and highly regulated professional environments. Industry and competency frameworks were included comparatively, but their educational value must be judged through independent research rather than authority or adoption alone.

PEDAL also assumes some capacity to redesign curriculum and assessment. Faculty working within fixed accreditation structures, central examinations, or precarious employment may have limited discretion. Richer evidence may improve validity but may also increase workload and create new inequities if oral, process, or digital components are poorly implemented. The feasibility principles offered here are design safeguards, not empirical proof that burden has been solved.

These limitations define a research agenda. Design-based studies should examine how faculty interpret each principle, where implementation diverges from intention, and how student reasoning changes over repeated cycles. Comparative studies should test the model across disciplines with different epistemic norms, including humanities, social sciences, engineering, health sciences, teacher education, and language education. Outcome measures should distinguish analysis, evaluation, creation, metacognition, and transfer rather than aggregating them under a single higher-order-thinking score.

Faculty-development research should compare formats such as sustained design studios, coaching, communities of practice, and webinar-based training. Relevant outcomes include design quality, faculty confidence, policy understanding, workload, and sustained enactment. Institutional studies should examine whether approved tools, policy clarity, data governance, and workload recognition moderate implementation. Equity research should investigate access, language, disability, and the differential effects of oral or process evidence.

Assessment research should evaluate the validity and reliability of evidence combinations. Studies could compare a final product alone with combinations, including decision memos, micro-vivas, source-verification tasks, and transfer items. They should also investigate the minimum evidence needed to support trustworthy judgment without multiplying burden. Cost, moderation time, student perceptions, due process, and unintended strategic behavior should be treated as outcomes rather than implementation details.

9 Conclusions

GenAI is forcing higher education to confront a long-standing pedagogical question: Is there evidence to

show that students have genuinely thought? The answer cannot lie in prohibition alone or in uncritical adoption. It lies in faculty capacity to design learning in ways that keep disciplinary analysis, evaluation, creation, metacognition, and transfer at the center of academic work.

This paper has proposed PEDAL as a meso-level design heuristic for that work. The model does not claim a new general theory of learning. Its contribution is to integrate established insights into a coherent sequence that faculty developers and program teams can use to redesign authentic tasks, preserve human judgment, and collect proportionate evidence. The worked cases and tiered institutional architecture illustrate a proposed approach in which responsible AI mediation is treated neither as better prompting nor as surveillance but as a question of educational design, professional capability, and institutional responsibility.

GenAI may have educational value when faculty design its use to prompt deeper questions, expose assumptions, sharpen judgment, and make evidence of learning more visible. Future faculty development should therefore help academics decide when AI is useful, when it is unnecessary, and how its presence might be designed as a context intended to encourage more thoughtful, rather than merely more efficient, learning.

Acknowledgments The authors did not receive support from any organization for the submitted work.

Declaration on the Use of AI The authors used GPT-5.3 Instant for language enhancement and formatting assistance. The authors reviewed the resulting text and take responsibility for the final manuscript. Both authors consented to the publication of this paper and its contents.

Author Contributions Both authors contributed to the conception of the article, development of the framework, interpretation of the literature, and revision of the manuscript.

Conflicts of Interest The authors declare no competing interests.

Data Availability Statements No datasets were generated or analyzed because this article is a conceptual review based on published literature.

References

- Alhur, A. A., Khlaif, Z. N., Hamamra, B., & Hussein, E. (2025). Paradox of AI in higher education: Qualitative inquiry into AI dependency among educators in Palestine. *JMIR Medical Education*, 11, e74947.
- Anderson, L. W., & Krathwohl, D. R. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. Longman.
- Azzam, A., & Charles, T. (2025). K-12 STEM teachers' experiences with artificial intelligence. *TEM Journal*, 14(1), 463–468.
- Biggs, J., & Tang, C. (2011). *Teaching for quality learning at university*. 4th ed. Open University Press.
- Borge, M., Smith, B. K., & Aldemir, T. (2024). Using generative AI as a simulation to support higher-order thinking. *International Journal of Computer-Supported Collaborative Learning*, 19(4), 479–532.
- Brandão, A., Pedro, L., & Zagalo, N. (2024). Teacher professional development for a future with generative artificial intelligence—An integrative literature review. *Digital Education Review*, (45), 151–157.
- Brookhart, S. M. (2010). *How to assess higher-order thinking skills in your classroom*. ASCD.
- Celik, I. (2023). Towards Intelligent-TPACK: An empirical study on teachers' professional knowledge to ethically integrate artificial intelligence (AI)-based tools into education. *Computers in Human Behavior*, 138, 107468.
- Chan, C. K. Y., & Hu, W. (2023). Students' voices on generative AI: Perceptions, benefits, and challenges in higher education. *International Journal of Educational Technology in Higher Education*, 20(1), 43.
- Charles, T., & Hill, C. (2026). *Integrity, assessment, and collaboration in the AI classroom*. IGI Global Scientific Publishing.
- Collins, A., Brown, J. S., & Newman, S. E. (1989). Cognitive apprenticeship: Teaching the crafts of reading, writing, and mathematics. In: Resnick, L. B., ed. *Knowing, learning, and instruction: Essays in honor of Robert Glaser*. Routledge, 453–494.
- Dakan, R., & Feller, J. (2025). *The AI fluency framework*. Available from Anthropic website.
- Facione, P. A. (1990). *Critical thinking: A statement of expert consensus for purposes of educational assessment and instruction: The Delphi report*. ERIC Doc. No. ED 315 423. The California Academic Press.
- Garet, M. S., Porter, A. C., Desimone, L., Birman, B. F., & Yoon, K. S. (2001). What makes professional development effective? Results from a national sample of teachers. *American Educational Research Journal*, 38(4), 915–945.
- Garrison, D. R., Anderson, T., & Archer, W. (2001). Critical thinking, cognitive presence, and computer conferencing in distance education. *American Journal of Distance Education*, 15(1), 7–23.
- Gordon, M., Daniel, M., Ajiboye, A., Uraiby, H., Xu, N. Y., Bartlett, R., Hanson, J., Haas, M., Spadafore, M., ... Thammasitboon, S. (2024). A scoping review of artificial intelligence in medical education: BEME Guide No. 84. *Medical Teacher*, 46(4), 446–470.
- Hague, C. (2024). *Fostering higher-order thinking skills online in higher education: A scoping review* [OECD Education Working Papers No. 306]. OECD.
- Hamamra, B., Khlaif, Z. N., & Mayaleh, A. (2026). Between promise and precarity: Palestinian university educators' experiences with AI integration in higher education. *Interactive*

- Learning Environments*, 34(4), 2574–2589.
- Hamamra, B., Mayaleh, A., & Khlaif, Z. N. (2024). Between tech and text: The use of generative AI in Palestinian universities—A ChatGPT case study. *Cogent Education*, 11(1), 2380622.
- Kasneci, E., Seßler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günemann, S., ... Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, 103, 102274.
- Khlaif, Z. N., Alkouk, W. A., Salama, N., & Abu Eideh, B. (2025). Redesigning assessments for AI-enhanced learning: A framework for educators in the generative AI era. *Education Sciences*, 15(2), 174.
- Khlaif, Z. N., Ayyoub, A., Hamamra, B., Bensalem, E., Mitwally, M. A. A., Ayyoub, A., Hattab, M. K., & Shadid, F. (2024). University teachers' views on the adoption and integration of generative AI tools for student assessment in higher education. *Education Sciences*, 14(10), 1090.
- Khlaif, Z. N., Hamamra, B., Bensalem, E., Mitwally, M. A. A., & Sanmugam, M. (2026). Factors influencing educators' technostress while using generative AI: A qualitative study. *Technology, Knowledge and Learning*, 31(3), 1391–1416.
- Laurillard, D. (2012). *Teaching as a design science: Building pedagogical patterns for learning and technology*. Routledge.
- Lee, H. P., Sarkar, A., Tankelevitch, L., Drosos, I., Rintel, S., Banks, R., & Wilson, N. (2025). The impact of generative AI on critical thinking: Self-reported reductions in cognitive effort and confidence effects from a survey of knowledge workers. In: *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*. Yokohama: Association for Computing Machinery, 1121.
- Lelescu, A., Sava, S., Grossecck, G., & Malita, L. (2025). Exploring trust in generative AI for higher education institutions: A systematic literature review focused on educators. *Humanities and Social Sciences Communications*, 12(1), 1961.
- Lodge, J. M., de Barba, P., & Broadbent, J. (2023a). Learning with generative artificial intelligence within a network of co-regulation. *Journal of University Teaching and Learning Practice*, 20(7), e1042.
- Lodge, J. M., Howard, S., Bearman, M., Dawson, P., Agostinho, S., Buckingham Shum, S., Deneen, C., Ellis, C., Fawns, T., ... Slade, C. (2023b). *Assessment reform for the age of artificial intelligence*. Tertiary Education Quality and Standards Agency.
- Long, D., & Magerko, B. (2020). What is AI literacy? Competencies and design considerations. In: *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. Honolulu: Association for Computing Machinery, 1–16.
- Miao, F. C., & Cukurova, M. (2024). *AI competency framework for teachers*. UNESCO.
- Miao, F. C., & Holmes, W. (2023). *Guidance for generative AI in education and research*. UNESCO.
- Miao, F. C., Shiohira, K., & Lao, N. (2024). *AI competency framework for students*. UNESCO.
- Mishra, P., & Koehler, M. J. (2006). Technological pedagogical content knowledge: A framework for teacher knowledge. *Teachers College Record: The Voice of Scholarship in Education*, 108(6), 1017–1054.
- Nazaretsky, T., Ariely, M., Cukurova, M., & Alexandron, G. (2022). Teachers' trust in AI-powered educational technology and a professional development program to improve it. *British Journal of Educational Technology*, 53(4), 914–931.
- Ng, D. T. K., Leung, J. K. L., Chu, S. K. W., & Qiao, M. S. (2021). Conceptualizing AI literacy: An exploratory review. *Computers & Education: Artificial Intelligence*, 2, 100041.
- Perkins, M., Furze, L., Roe, J., & MacVaugh, J. (2024). The artificial intelligence assessment scale (AIAS): A framework for ethical integration of generative AI in educational assessment. *Journal of University Teaching and Learning Practice*, 21(6), e810.
- Ruiz-Rojas, L. I., Salvador-Ullauri, L., & Acosta-Vargas, P. (2024). Collaborative working and critical thinking: Adoption of generative artificial intelligence tools in higher education. *Sustainability*, 16(13), 5367.
- Salido, A., Syarif, I., Sitepu, M. S., Suparjan, Wana, P. R., Taufika, R., & Melisa, R. (2025). Integrating critical thinking and artificial intelligence in higher education: A bibliometric and systematic review of skills and strategies. *Social Sciences & Humanities Open*, 12, 101924.
- Sharples, M. (2023). Towards social generative AI for education: Theory, practices and ethics. *Learning: Research and Practice*, 9(2), 159–167.
- Southworth, J., Migliaccio, K., Glover, J., Glover, J., Reed, D., McCarty, C., Brendemuhl, J., & Thomas, A. (2023). Developing a model for AI across the curriculum: Transforming the higher education landscape via innovation in AI literacy. *Computers & Education: Artificial Intelligence*, 4, 100127.
- Tan, X., Cheng, G., & Ling, M. H. (2025). Artificial intelligence in teaching and teacher professional development: A systematic review. *Computers & Education: Artificial Intelligence*, 8, 100355.
- Torraco, R. J. (2005). Writing integrative literature reviews: Guidelines and examples. *Human Resource Development Review*, 4(3), 356–367.
- Trust, T., Krutka, D. G., & Carpenter, J. P. (2016). “Together we are better”: Professional learning networks for teachers. *Computers & Education*, 102, 15–34.
- UNESCO. (2021). *Recommendation on the ethics of artificial intelligence*. UNESCO.
- Zacharis, G., & Papadakis, S. (2025). Can AI grade like a human? Validity, reliability, and fairness in university coursework assessment. *Educational Process: International Journal*, 19, e2025591.
- Zhou, X., & Schofield, L. (2024). Developing a conceptual framework for artificial intelligence (AI) literacy in higher education. *Journal of Learning Development in Higher Education*, (31), 1–14.