Computers and Education Open 9 (2025) 100306

Contents lists available at ScienceDirect

Computers and Education Open

journal homepage: www.sciencedirect.com/journal/computers-and-education-open

Intelligent-TPACK in practice: design and evidence from a three-week teacher preparation module

Tugce Aldemir [a,*], Selcuk Kilinc [b], Ali Bicer [a], Patricia Grant [c], Trina Davis [a], Noelle Wall Sweany [a]

[a] Department of Teaching Learning and Culture, College of Education and Human Development, Texas A&M University, College Station, TX, USA
[b] Department of Teaching Learning and Culture, College of Education and Human Development, Texas A&M University, College Station, TX USA & College of Education, Middle East Technical University (METU), Ankara, Turkey
[c] Department of Educational Psychology, College of Education and Human Development, Texas A&M University, College Station, TX, USA

ARTICLE INFO

Keywords:
Intelligent-TPACK
AI-TPACK
AI literacy
Ethics
Teacher education

ABSTRACT

Rapid advances in generative AI sharpen the need for teachers to develop pedagogical and ethical capacities for AI-integrated instruction. While Technological Pedagogical Content Knowledge (TPACK) provides a valuable framework for technology integration, it does not fully capture AI's unique complexities. This study presents an integrated i-TPACK approach that extends Intelligent-TPACK by adding AI-as-content (i-CK) and AI-for-professional development (i-PD) and by threading a five-stage AI-literacy progression (Know→ Use→ Evaluate→ Ethics→ Create) within each domain, treating ethics as distributed and iterative. We designed and examined a three-week professional development module for preservice teachers using a convergent mixed-methods design. Pre–post surveys ($n = 25$ matched pairs) with a six-subscale Integrated i-TPACK instrument showed statistically significant gains across all domains (Wilcoxon, Holm-adjusted; medium-to-large effects). Qualitative analyses of lesson artifacts, decision logs, reflections, and micro-teaching documented instances of layered ethical decision-making (privacy/data governance, bias/fairness, transparency/provenance/accountability), progression along the AI literacy stages, and discipline-aligned pedagogical designs. Embedding an ethical decision-making checkpoint across performance-based activities made ethics visible in teacher work and coincided with more explicit safeguards and verification steps in lesson artifacts and micro-teaching within the module. By detailing this empirically grounded model, our study offers theoretical and practical insights for teacher educators seeking to cultivate principled GenAI-supported instruction.

1. Introduction

In recent years, research and practice on generative artificial intelligence (GenAI; hereafter, "AI") in education have accelerated, with systematic reviews noting sharp growth in classroom pilots, design studies, and policy debates [1–3]. This rapid uptake brings both opportunities and risks. While large language models and multimodal assistants can expand what teachers and students can do, scholars warn that, without explicit pedagogical and ethical guidance, classroom use may drift toward academic dishonesty, privacy violations, and inequitable practices [4–6]. Institutional responses have ranged from bans to managed-use policies ([7]; Petricini et al., 2023); emerging evidence suggests that blanket bans are difficult to enforce and may be counterproductive, whereas regulated integration can reduce misconduct and spur innovation in assessment [8]. Preparing future teachers who can exercise pedagogical creativity with ethical discernment is therefore essential.

Preservice teachers (PSTs), however, often feel underprepared to design or facilitate AI-mediated instruction, especially on ethical dimensions [9–12]. Many enter with a superficial understanding of AI concepts and limited experience in embedding safeguards [13,14]. Classic TPACK and AI-specific variants, together with AI literacy models, provide starting points [15–20]. Yet, ethics is too often appended rather than woven through design and enactment, and two adjacent domains remain underspecified for teacher preparation: i-CK (teaching about AI across subjects) and i-PD (teachers' sustained learning with AI).

* Corresponding author.
E-mail addresses: taldemir@tamu.edu (T. Aldemir), selcuklnc@tamu.edu (S. Kilinc), alibicer@tamu.edu (A. Bicer), pat.grant@tamu.edu (P. Grant), trinadavis@tamu.edu (T. Davis), nsweany@tamu.edu (N.W. Sweany).

https://doi.org/10.1016/j.caeo.2025.100306

Available online 8 November 2025

Responding to this need, we adopt an integrated i-TPACK approach that (a) extends Intelligent-TPACK [15] to six domains (i-TK, i-TPK, i-TCK, i-CK, i-PD, and integrated i-TPACK), and (b) threads a five-stage AI-literacy progression (Know → Use → Evaluate → Ethics → Create) within each domain, embedding ethics as a distributed, iterative practice across planning, enactment, and reflection. Building on this framing, we designed and implemented a three-week professional development (PD) module. This small-scale mixed-methods study examines PSTs' beginning- and end-of-module states and how integrated i-TPACK is evidenced in artifacts and enactment across this PD module. We draw on pre- and post-self-reports, as well as process-focused artifacts, to capture change and trace how ethics-infused integration appears in practice and self-reports.

2. Theoretical framework and previous work

2.1. From TPACK to i-TPACK

The classic TPACK framework [21] explains how effective technology-integrated teaching coordinates technological, pedagogical, and content knowledge. GenAI (hereafter AI), however, complicates this coordination with its capabilities shift rapidly, its reasoning is opaque, and its outputs are unstable. These qualities require teachers to verify results, audit for bias and accessibility, and maintain human accountability—needs that a technology-agnostic TPACK only gestures toward [22].

Thus, scholars have proposed several AI-TPACK variants, all recognizing that classic TPACK does not fully address AI's ethical, technical, and pedagogical complexities (e.g., [20,23]). Among these, Celik's [15] Intelligent-TPACK (i-TPACK) model is notable for its ethics focus. In this model, *i-TK* refers to knowledge of AI tools and systems, *i-TPK* to knowledge of orchestrating AI in teaching and assessment, and *i-TCK* to knowledge of how AI represents and supports disciplinary ideas. *Ethics* is treated as a co-equal domain that foregrounds transparency, accountability, inclusiveness, and fairness in instructional decision-making. Celik's validation indicates that while i-TK provides a necessary base, i-TPK, i-TCK, and Ethics strongly predict cumulative i-TPACK competence. Converging evidence from other AI-TPACK variants likewise shows that AI-TPK and to a lesser extent AI-TCK are the strongest drivers of integrated competence, with AI-TK foundational but insufficient on its own [20,23].

While i-TPACK and related frameworks provide a strong foundation for teaching with AI (via TK, TPK, TCK), two adjacent domains remain underspecified in teacher preparation yet are essential for responsible integration: content knowledge (CK) and professional development (PD). In classic TPACK, CK refers to teachers' disciplinary content knowledge [21]. In AI-integrated contexts, however, AI knowledge has typically been folded into TK (e.g., i-TK; [15]), where it is treated primarily as technical understanding and tool use. This framing positions AI as something to use rather than as subject matter to teach, leaving CK underdeveloped in most i-TPACK variants. Recent scholarship, however, argues that effective integration requires coupling teaching with AI and teaching about AI (e.g., [24,25]). Put differently, it is not enough for teachers to embed AI tools into instruction if students lack opportunities to understand what AI is and how to engage with it responsibly. From this perspective, AI is no longer confined to computer science but is knowledge to be addressed across all subject areas [26]. Accordingly, following Ng et al. [27], we added i-CK as a distinct domain, defining it as teachers' knowledge for teaching about AI, introducing core concepts and working mechanisms and guiding students' meaningful, ethical use of AI across subjects (e.g., a science teacher demonstrating dataset bias in a classifier, or an ELA teacher modeling provenance checks with a chatbot).

Second, i-PD (AI for professional development) is necessary because tools and norms shift rapidly: teachers need to learn with AI in order to keep teaching with it. i-PD aligns with UNESCO's AI Competency Framework for Teachers, which identifies AI-supported professional learning as a core strand. It includes competencies that help teachers sustain their own learning, for example, using chat assistants to support lesson study and planning, maintaining prompt/verification logs to guide professional reflection, and applying lightweight analytics for self-assessment as tools and norms evolve [28]. In short, without i-CK, teachers' use of AI risks missing its mark because students lack the concepts and norms to participate responsibly; and without i-PD, the pace of change makes early gains brittle.

With these two additions, we extended Celik's [15] i-TPACK to include six knowledge domains: i-TK, i-TPK, i-TCK, i-CK, i-TPACK, and i-PD. We then drew on AI literacy framings to operationalize each domain.

2.2. AI literacy as developmental mechanism

As AI tools enter K–12 classrooms, teachers are increasingly expected to be AI literate, that is, to understand core AI concepts and systems, use tools productively, evaluate outputs and affordances, and recognize and navigate ethical and social implications (e.g., [29,18,19,14]). Although definitions and assessments vary [29], the literature increasingly converges on five recurring dimensions: (1) Know: a foundational grasp of how AI works (concepts, common methods/algorithms, capabilities, limits) that supports confidence and readiness and relates to gains across other dimensions [9,30,19,31]; (2) Use: purposeful application of AI in authentic teaching (e.g., lesson/task design, content generation, assessment, differentiation), bridging "knowing about" AI to "teaching with" it and strongly predicting perceived AI literacy [9,19,31]; (3) Evaluate: critical appraisal of relevance, appropriateness, effectiveness, and fit by interrogating outputs, checking accuracy, attending to bias, and assessing pedagogical appropriateness for specific learners and contexts [29,18,19,31]; (4) Ethics: recognizing and responding to societal and moral implications of AI (e.g., fairness, transparency/explicability, safety, and data privacy) through concrete mitigation strategies [29,19,31]; and (5) Create: treating instructional needs as design problems by prototyping, testing, and refining AI enabled artifacts (e.g., feedback agents, multilingual supports) to meet pedagogical and ethical goals, a maker stance linked to self-efficacy, adoption intentions, and closer alignment between AI capabilities and curriculum aims [9,18,19].

These dimensions complement the i-TPACK knowledge domains discussed earlier. Whereas i-TPACK specifies the knowledge domains teachers draw on to teach with AI (i-TK, i-TPK, i-TCK, i-TPACK), to teach about AI (i-CK), and to sustain learning through AI (i-PD), AI literacy dimensions describe how competence develops and is enacted. Following Ng et al. [19], we therefore treat AI literacy as the process scaffold (the "how") that animates the i-TPACK domain structure (the "what"), turning a static knowledge map into an actionable developmental model for teacher preparation. Recent assessments likewise distinguish Know, Use, Evaluate, Ethics, and Create in teachers' AI literacy, revealing systematic gaps, especially in Ethics and Create, even when basic Use is present [32,19]. Reviews of teacher-facing resources similarly note strong activity design but limited assessment scaffolds and K–12 guidance [33]. Aligning AI literacy dimensions with i-TPACK domains can address some of these challenges by enabling targeted PD (e.g., focusing deliberately on i-TCK.Evaluate or i-TPK.Ethics rather than generic tool exposure) and providing a consistent analytic lens for studying and assessing i-TPACK.

Accordingly, we introduce an integrated i-TPACK approach that delineates six knowledge domains (i-TK, i-TPK, i-TCK, i-TPACK, i-PD, i-CK; [15,28,27,20]), with each domain articulated across five nonlinear AI literacy stages (Know → Use → Evaluate → Ethics → Create; e.g., i-TCK.Use; [9,19]). This integrated framing underpinned our module design (Appendix B) and analytical lens (survey for Appendix A; codebook for Appendix C).

2.3. Ethics as distributed, iterative decision-making process

As AI enters routine classroom work, teachers face situated choices with ethical stakes, including privacy, bias, provenance, accessibility, and learner agency. Studies routinely report that general "dos and don'ts" or one-off lectures leave teachers unsure how to enact ethics in real lessons, especially when tools, policies, and contexts vary (e.g., [25,34]). What is needed is not only knowledge of principles but practice-ready routines for deciding whether and how to use AI responsibly in specific lessons.

Emerging PD approaches show promise but remain uneven. Curriculum-linked PD with classroom enactment builds teacher preparedness through stakeholder analysis, ethical trade-offs, and bias investigations, though evidence is largely qualitative and emphasizes feasibility rather than competence [34,35]. Technical AI methods PD introduces fairness/bias alongside AI content, but ethics typically appears as topic exposure rather than an assessed teacher competency [36]. Co-design and microcredential efforts productively address ethical concerns by producing teacher-authored, standards-aligned materials, surface facilitation concerns, and emphasize just-in-time supports and tool monitoring, yet seldom measure teacher ethics competencies or classroom enactment [37–39]. Large, blended AI literacy offerings report pre-post gains, including "ethics awareness," but rely mainly on self-report and treat ethics generically [40,41,25].

Across these strands, three patterns emerge. First, ethics content is narrowly weighted toward bias/fairness and occasional stakeholder analysis; other core domains like privacy/data governance, transparency/provenance, accountability/governance, accessibility, and IP/authorship are underrepresented as explicit teacher competencies (e.g., [37,25,39]). Second, assessment is thin: studies lean on self-report or general literacy tests; performance tasks (e.g., rubric-scored audits, disclosure/provenance statements, datasheets/model cards), classroom observation of ethical facilitation, transfer to novel tools, and fidelity/durability reporting are rare (e.g., [5,42,43]). Finally, while reviews increasingly commend i-TPACK variants as a design spine, frameworks are seldom operationalized into objectives, tasks, and rubrics that make ethics observable in teacher work [5,44].

We therefore frame ethics as distributed and iterative. Distributed means layered attention to (a) technical routines (data minimization/no PII, model/version provenance, verification for accuracy and subgroup harms, accessibility), (b) pedagogical routines (maintaining learner agency, avoiding over-reliance, equitable differentiation, human-in-the-loop oversight), and (c) disciplinary routines (content integrity, representation and cultural/linguistic responsiveness, discipline-specific disclosures). Iterative means ethical work recurs across plan → enact → reflect/iterate, aligned with our AI-literacy progression where planning, enactment, reflection/iteration activate Know/Use/Evaluate/Ethics/Create, threading ethics through every i-TPACK domain (i-TK, i-TPK, i-TCK, i-CK, i-PD, integrated i-TPACK) rather than appending it after tool selection.

To operationalize this, we embedded an ethical decision-making checkpoint (see Appendix D Ethical decision-making checkpoint) before each performance task: (1) context & purpose; (2) risks & benefits; (3) policies & stakeholders; (4) a documented adopt/adapt/reject decision with layered safeguards (e.g., no student logins; verification routines); and (5) reflect & document (a decision log). This routine was used across module activities to ensure that ethical decision-making would be enacted and recorded, not merely discussed. Our codebook (see Appendix C Integrated i-TPACK codebook) then makes this work analyzable: we code ethics and evaluate (if ethics-focused) within each domain (e.g., i-TK.Evaluate for data/privacy safeguards; i-TPK.Ethics for classroom norms) and at the integrated level when safeguards are coordinated across a lesson flow. In this way, ethics becomes visible in lesson plans, decision logs, reflections, and micro-teaching.

2.4. Purpose of the current study

Prior work leaves four gaps: (i) i-TPACK variants do not specify i-CK and i-PD and often append ethics rather than weave it through design and enactment; (ii) interventions seldom align a five-stage AI-literacy progression within i-TPACK to organize learning and make growth observable; (iii) evaluation rarely moves beyond self-report to performance- and artifact-based evidence; and (iv) routines teachers can carry into planning and enactment are scarce. We address these gaps by extending Intelligent-TPACK to six domains (i-TK, i-TPK, i-TCK, i-CK, i-PD, integrated i-TPACK), threading Know → Use → Evaluate → Ethics → Create through each domain, and operationalizing ethics as a distributed, iterative practice via a brief ethical decision-making checkpoint and a domain-by-stage codebook.

This small-scale mixed-methods study examines whether, and if so how, PSTs' integrated i-TPACK changed during a three-week PD module designed around our integrated i-TPACK framing. We combined pre-post self-reports with process-focused artifacts to capture beginning and end states and to trace any changes throughout the module activities. Accordingly, we asked:

RQ1 (overarching): How did the beginning- and end-of-module states of PSTs' integrated i-TPACK compare?
RQ1a: To what extent did PSTs' self-ratings differ across the six integrated i-TPACK domains from pre- to post-module?
RQ1b: How did PSTs' conceptualizations of integrated i-TPACK practice change, if at all, between pre- and post-module reflections?
RQ2: In what ways, if any, did integrated i-TPACK emerge or shift in PSTs' instructional artifacts, reflections, and teaching demonstrations?

3. Methods

3.1. Research design, context, participants

We employed a convergent mixed-methods design to examine PSTs' integrated i-TPACK competencies, reflected in their pre-post self-reported measures as well as their activities during a three-week PD module [45,46]. We collected quantitative data from pre-post survey self-ratings (RQ1a) and qualitative data from pre-post survey reflections (RQ1b) and module artifacts across three-week activities (RQ2). Each dataset was analyzed separately, with integration occurring at the interpretation stage through joint displays and cross-source memoing [45].

The intervention was embedded in a required undergraduate course, "Technology in the Classroom," offered face-to-face at a large public university in the south-central United States. The PD module was integrated into the course's objective of supporting PSTs' digital literacy. It was implemented during the final three weeks of the semester, spanning nine 50-minute sessions, with each lesson addressing a specific facet of our integrated i-TPACK conceptualization. Twenty-nine of the 30 PSTs enrolled in the Spring 2025 course consented to participate in the study. The sample was predominantly female ($n = 28$) and composed mainly of sophomores ($n = 20$), with the remainder juniors ($n = 9$). Most participants were pursuing EC-6 Core Subjects/Generalist certification ($n = 20$, 69 %), with smaller groups enrolled in Bilingual Education ($n = 3$, 10.3 %), Middle Grades 4–8 Math/Science ($n = 3$, 10.3 %), and Special Education ($n = 3$, 10.3 %). Prior to the intervention, 69 % of PSTs ($n = 20$) reported having used AI for educational purposes, but fewer than half ($n = 13$, 45 %) described themselves as comfortable using AI tools in general. This suggests that prior exposure in classroom settings did not necessarily translate into broader self-reported confidence with AI tools.

This study was approved by the Institutional Review Board at [anonymized institution]. All participants received an information sheet and indicated written consent prior to participation. Data we deidentified and stored securely.

3.2. Integrated i-TPACK professional development module

We used backward design (goals→ evidence→ learning experiences) with constructive alignment [47,48] to ensure that assessments and activities targeted the same domain × stage outcomes. The integrated i-TPACK × AI literacy matrix operationalizes outcomes (domains) and progression (stages), with ethics embedded through the Evaluate/Ethics stages mapped across domains (see Appendix B Module activities mapping matrix). The three-week module, therefore, aimed to scaffold early competencies in PSTs' integrated i-TPACK for responsible AI uses by engaging them in repeated, structured activities across (a) teaching with AI (i-TK/i-TPK/i-TCK), (b) teaching about AI (i-CK), and (c) using AI for professional learning (i-PD), with ethics threaded throughout via a 5-step decision-making checkpoint. The intent was to seed routines and generate formative evidence of progress toward (1) evaluating tool fit and limits in content and pedagogical contexts and specifying safeguards; (2) designing a standards-aligned AI-integrated lesson with explicit teacher- and student-facing ethical practices (e.g., disclosure/verification) and micro-plans for when and how to teach about AI; and (3) enacting classroom-ready instructional practices and PD strategies for responsible AI use.

Guided by the i-TPACK × AI-literacy map, we designed nine lessons that moved PSTs from conceptual grounding to applied design and finally to peer-reviewed enactment (micro-teaching). Each lesson foregrounded one i-TPACK domain, while often and inevitably engaging others. For example, Lesson 2 foregrounded i-TK (foundational knowledge of AI systems/tools) but also drew on i-TPK through classroom-use reflections. As the sequence progressed (Weeks 2–3), the emphasis shifted toward integrated i-TPACK, where tool–pedagogy–content decisions and ethics were coordinated in full-lesson designs and enactments. Activities within a lesson were selected to advance PSTs nonlinearly through the AI-literacy stages, typically Know → Use → Evaluate/Ethics, and occasionally Create when time and task allowed. Not every lesson targeted every stage, but ethics was embedded in each one.

In practice, the lessons unfolded as follows: Week 1 laid the conceptual and ethical groundwork by (a) establishing a common language about how ML systems work and typical risks (i-TK); (b) building familiarity with different AI tools (i-TK/TPK); (c) practicing evaluation and ethical reasoning for pedagogical and content use (i-TPK/TCK); (d) beginning to connect AI mechanics to disciplinary aims (i-TCK); and (e) exploring initial strategies for teaching about AI within subject domains (i-CK). Week 2 moved toward classroom-ready practices and early i-PD routines by (a) designing a pedagogically aligned, student-facing activity with explicit safeguards (i-TK/TPK); (b) applying the integrated i-TPACK framework to analyze classroom scenarios with explicit ethical reasoning (i-TPACK with i-PD); (c) systematically evaluating a selected tool for classroom use to inform design choices (i-TK/TPK/CK); (d) exploring AI-for-PD opportunities and associated ethical considerations (i-PD); and (e) designing a complete, standards-aligned AI-integrated lesson plan that incorporated teacher- and student-facing safeguards and a built-in self-assessment activity for reflective practices (i-TPACK with i-PD). Finally, Week 3 focused on integrated enactment, where each team delivered a 15-minute micro-teaching lesson while classmates, role-playing target learners, submitted peer feedback. This live rehearsal supported real-time i-TPACK integration and introduced a professional development feedback loop. Table B.1 (Appendix B) details the learning objectives, activities, artifacts, focal domains, and targeted stages for each lesson.

3.3. Data sources and data collection

3.3.1. Survey instrument

Building on our integrated i-TPACK approach, we designed a survey to capture six theoretically grounded domains of PSTs' integrated i-TPACK competencies. A 37-item pool was developed from emerging AI literacy frameworks (e.g., [28,27]) and scholarship that extends TPACK to AI contexts (e.g., [15]). Two U.S. cohorts completed the draft survey prior to the present study (Sample 1 = 50; Sample 2 = 25). Given the small sample size and the exploratory nature of the study, we did not perform an EFA to test whether all items together formed a unified AI literacy construct. Instead, we followed a theoretically grounded approach: We defined the construct based on the literature and then examined whether each identified factor, as reflected in the item sets, consistently represented a unidimensional structure. We first confirmed the full use of the 5-point Likert range to verify whether item difficulty levels were appropriate. We then ran principal-axis EFAs with oblique rotation for each construct. Adequacy indices met accepted thresholds (KMO = .77–.84; [49]; Bartlett's χ^2 ps $< .001$; [50]), suggesting item interrelationships are suitable for factor analysis. Eigenvalues > 1, scree plots, and parallel analysis [51] all pointed to one-factor solutions. Finally, we conducted reliability analyses to assess the internal consistency of the items within each construct. Cronbach's alpha was used as a measure of internal reliability, with higher values indicating greater consistency and lower measurement error. In this context, internal consistency refers to how items within each factor perform similarly under comparable conditions. One cross-loading ethics item was removed from i-TK, resulting in a 36-item instrument. Single-factor models explained 60–85 % of the variance within each facet, and internal consistency was strong (α = .88–.96; see Table 1). Detailed loading matrices, communalities, scree plots, and the expert content-validity matrix are provided in Appendix A Detailed exploratory factor analysis results for the Integrated i-TPACK scale. Consistent with the Standards for Educational and Psychological Testing [52], a multi-factor confirmatory analysis will follow with a larger sample. The validated 36-item scale was administered twice: as a pre-survey before the 3-week module and a post-survey immediately afterward. Participants rated each item on a 1 (Strongly disagree) to 5 (Strongly agree) scale. Subscale scores were computed as the mean of their respective items; higher scores indicate stronger self-perceived competence. Pre/post reliability coefficients for the current sample are reported in the Data Analysis section.

Alongside the 36 Likert-scale items, pre- and post-surveys included five to seven open-ended questions. The pre-survey's open responses gauged participants' prior AI exposure, hopes and concerns (about ethics, equity, and instructional value), and their initial views on AI's role in teaching. The post-survey revisited these areas to capture their AI use, what helped or hindered their learning, how their ethical reasoning evolved, and shifts in their views of AI's role and plans for future integration.

3.3.2. Qualitative corpus

During the three-week module, we archived every artifact that required PSTs to design, enact, or critique AI-supported instruction. Nine distinct artefact sets were generated. First, each PST completed an *Ethical Scenario Analysis Worksheet* (A1, $n = 30$), in which they selected one or two AI tools encountered in class and worked through the structured five-step ethical decision-making checkpoint. Typical prompts included "Does this tool collect personal information? If yes, where is the data stored?" and "If the risks outweigh the benefits, what

Table 1
Summary Psychometrics for the Integrated i-TPACK Scale (pilot samples, n = 75).

Subscale	K	KMO	Variance explained (%)	Loading range	Cronbach's α
i-TK	6	.79	60.3	.72 – .85	.96
i-CK	5	.79	67.8	.75 – .89	.88
i-TCK	6	.77	73.5	.78 – .94	.93
i-TPK	6	.81	77.7	.69 – .96	.94
i-TPACK	7	.84	73.3	.75 – .94	.94
i-PD	6	.84	84.9	.82 – .96	.96

Note. k = number of items retained.

alternative would you adopt?"

Second, small groups built a *Teachable Machine mini-classifier and submitted a brief design sheet, along with screenshots* (A2, $n = 29$). Instructions asked them to "upload 5–10 images per category, test accuracy on a new image, and note any surprising mis-classifications or biases." *Individual Teachable-Machine reflections* (A3, $n = 30$) followed; one guiding question read, "Which step of the ethical cycle was most challenging while you trained the model and why?".

Next, groups created a *Teaching-AI design brief with reflection* (A4, $n = 29$), outlining a lesson that explicitly teaches students to "evaluate AI outcomes, use AI responsibly, and stay safe," while documenting where they "consulted policies or peers" during planning. At the end of Week 1, every PST wrote a *Comprehensive end-of-week reflection* (A5, $n = 30$) responding to items such as "What classroom tasks might benefit from AI and why?" and "When is human creativity indispensable? Provide a concrete example."

In Week 2, PSTs drafted a *Pedagogical AI-Integration Reflection* (A6, $n = 29$) that probed tool choice, learning goals, foreseeable risks, and next steps (e.g., "Which higher-order skill does this tool enhance and how might it also limit creativity?"). They also completed an *AI-Tool Rubric Evaluation* (A7, $n = 28$) covering bias, critical consumption, instructional value, and privacy; one item asked them to describe "a concrete instance where you noticed bias and how you would mitigate it."

Culminating group artifacts were an *AI-Integrated Lesson Plan* (A8, $n = 28$) created in a shared template with thirteen headings (standards alignment, assessment, differentiation, privacy safeguards, etc.) and a micro-teaching session in which each group enacted the lesson while the others role played target learners. *Slide decks* (A9, $n = 28$) and hand-outs were retained, and every PST submitted written *peer feedback* (A9) to two other groups, yielding sixty feedback forms. All materials were submitted through the university's learning-management system at the end of each activity, automatically time-stamped, and later downloaded by the instructor, first author, after the term ended. To protect confidentiality, the third author, uninvolved in teaching, replaced all personal identifiers with numeric codes (e.g., S3, G7) before passing the materials to the analysis team. The university's Institutional Review Board approved the study, and all PSTs signed written consent allowing their de-identified work to be used for research.

3.4. Data analysis

3.4.1. Quantitative analysis

The quantitative analysis addressed RQ1a and examined the pre-post differences in PSTs' integrated i-TPACK competencies. Quantitative survey data were analyzed in four steps.

First, we assessed the internal consistency of the six scale aggregates (i-TK, i-CK, i-TCK, i-TPK, i-TPACK, i-PD) using Cronbach's α at both administrations. Values of $\alpha \geq 0.80$ were considered satisfactory [53]. We also sought evidence that the scale remained stable with the focal cohort, consistent with the item-levels established in the earlier pilot [54]. Internal consistency remained high across administrations ($\alpha = 0.96$ pre-survey, $\alpha = 0.89$ post-survey), indicating that the six sub-scales functioned coherently within the study cohort.

Second, inspection of box plots, skewness-kurtosis indices, and Shapiro-Wilk statistics on the pre-post scores showed that all post-scores departed from normality ($p < .05$). Some also exhibited absolute skewness values > 1 (e.g., AI-TK-post skew $= -1.91$).

Third, because normality was violated, we examined pre/post changes with the non-parametric Wilcoxon signed-rank test, an approach recommended for matched samples with non-normal distributions (Field, 2018). To maintain the family-wise error rate at 0.05, p-values were Holm-Bonferroni adjusted ($m = 6$; [55]). Effect size was expressed as $r = Z/\sqrt{N}$, with 0.10/0.30/0.50 denoting small/-medium/large effects respectively ([56]; see also [57]).

Finally, we examined post-test inter-facet associations. For that, we examined Spearman correlations as a descriptive screen for overlap among domains, flagging coefficients ≥ 0.85 as potential conflation in this novice cohort rather than as a pass/fail test of discriminant validity [58]. We did not estimate confirmatory models due to sample size; a hierarchical CFA is planned with a larger cohort.

Of the 29 consented PSTs, four lacked complete pairs; the analytic sample was $n = 25$ with 0 % item-level missingness. Analyses were conducted in IBM SPSS Statistics 29 [59].

3.4.2. Qualitative analysis

To examine RQ1b and RQ2, we used Saldaña's (2s016) two-cycle coding with a hybrid deductive-inductive approach. For RQ1b, the unit of analysis was each complete pre/post reflection, capturing participants' conceptualizations of integrated i-TPACK practices (e.g., role of AI, ethical considerations, and pedagogical strategies) before and after the module. For RQ2, we analyzed a structured corpus of artifacts submitted across nine learning activities (e.g., ethical scenario analyses, teachable machine reports, full sAI-integrated lesson plans). The unit of analysis was typically each entire artifact or coherent segments (e.g., "AI Considerations" in a lesson plan) within each artifact, preserving contextual and conceptual coherence across epistemic, pedagogical, and ethical decision making.

In cycle 1 coding, we applied a codebook developed from our integrated i-TPACK approach (see Appendix C Integrated i-TPACK codebook). This included six knowledge domains (i-TK, i-TPK, i-TCK, i-CK, i-TPACK, i-PD) and five developmental stages (Know → Use → Evaluate → Ethics → Create). Two authors independently coded a stratified 20 % sample of responses and artifacts, achieving substantial agreement ($\kappa = .84$; [60]). Through constant comparison and collaborative memoing [61], we refined definitions and added emergent subcodes. We maintained a time-stamped audit trail for our analytic decisions to support dependability and transparency [62]. Once the codebook was finalized, the first author coded the remaining corpus. Then, both coders reviewed all codes and resolved discrepancies through line-by-line negotiation.

In cycle 2 coding [63,64], we focused on theme generation and interpretive synthesis. We clustered codes temporally (pre/post) and conceptually (e.g., stance, ethics, design agency) to capture the shifts in conceptualizations (RQ1a) and demonstrations of integrated i-TPACK practices (RQ2). For RQ2, we also organized the codes by artifact types and module's sinstructional sequence. This process yielded four themes: (1) Building the foundation, (2) Developing a critical stance, (3) From theory to practice, and (4) Fostering a professional identity. Each theme was explicitly mapped back to (a) most implicated domains of integrated i-TPACK and (b) predominantly enacted AI literacy stages, preserving a tight analytic link to the theoretical frame.

3.4.3. Integration, rigor and trustworthiness

Following a convergent mixed-methods design [45], we analyzed qualitative and quantitative strands in parallel and integrated them at the interpretation stage. We adopted an interpretive integration strategy [65] and used joint analytic memos and cross-source comparisons to realize complementarity [66].

Qualitative findings provided depth and explanatory power for interpreting quantitative trends. For example, although i-TCK showed the smallest self-reported gain (+0.39), participants demonstrated rich content-specific integration in lesson designs (e.g., using AI to scaffold science comparisons or personalize math problem sequences). In contrast, gains in i-TPK ($r = 0.63$) were mirrored by proactive ethical planning (e.g., privacy protocols and bias-checking strategies embedded in their instructional designs). In a few cases, qualitative data revealed competencies not fully reflected in survey scores (e.g., plans for ongoing professional development despite modest pre/post i-PD gains).

Rigor followed Lincoln and Guba's [62] four-criterion framework. For credibility, we used method triangulation with two survey administrations and nine artifact types, and analyst triangulation with substantial intercoder reliability ($\kappa = 0.84$). The three-week in-situ data-collection window provided prolonged engagement, further

supporting credibility [67]. Transferability was supported through thick description of the course context, participants, and learning tasks, and representative quotations from the participants. Dependability was addressed through a transparent audit trail, including raw files, codebooks, reflexive memos, and SPSS syntax [68]. Finally, we addressed confirmability with reflexive positionality memos, peer debriefs, and negative-case analysis that intentionally retain dissenting voices.

4. Results

4.1. RQ1. pre-post differences in PSTs' integrated i-TPACK

RQ1 examines how beginning- and end-of-module states of PSTs' integrated i-TPACK compare. We addressed this in two parts: RQ1a focuses on pre-post differences in self-rated competencies across the six domains, while RQ1b focuses on pre-post differences in PSTs' conceptualizations of integrated i-TPACK practice, as expressed in paired reflections.

4.1.1. RQ1a. pre-post differences in self-rated competencies across the six domains

To examine any differences in PSTs' pre- and post-module self-rated competencies, we analyzed matched pairs of integrated i-TPACK survey responses ($n = 25$). Descriptive statistics (Table 2) indicated higher post-module means for every integrated i-TPACK domain, with mean post-pre differences ranging from +0.39 (i-TCK) to +0.76 (i-CK). Because normality assumptions were violated, we employed Wilcoxon signed-rank tests with Holm-Bonferroni adjustment ($m = 6$). Results showed statistically reliable pre-post differences across all six domains (adjusted $p = .005$–0.044).

Effect sizes, expressed as standardized r ($Z/\sqrt{N}$), were large for four domains ($r = 0.63$–$.68$) and medium to large for i-TCK ($r = .46$) and i-PD ($r = .43$), indicating practically meaningful growth. Post-test correlations revealed two notable clusters: a tightly coupled technical–content triad (i-TK ↔ i-CK ↔ i-TCK, $\rho \geq .88$) and a pedagogical–integration pair (i-TPK ↔ i-TPACK, $\rho = .84$), with i-PD showing weaker associations ($\rho \leq .78$). Because TK/CK/TCK correlations exceeded our heuristic discriminant-validity guideline, we refrain from claiming discriminant validity for that cluster and interpret the pattern as novice conflation among closely related domains.

4.1.2. RQ1b. pre-post differences in conceptualizations of integrated i-TPACK practice

To examine differences in PSTs' pre- and post-module conceptualizations of integrated i-TPACK practice, we analyzed paired open-ended reflections from the surveys ($n = 25$). The analysis revealed three interconnected areas of conceptual change: (a) deeper integration of pedagogy and content, (b) movement toward creator-level enactment of AI-supported practices, and (c) greater attention to ethics routines.

The first pattern was a shift from viewing AI as a generic productivity aid to employing it as a content-specific pedagogical tool. Several respondents (9 of 25; 36 %) described integrated scenarios in which an AI tool was explicitly tied to both a content goal and a pedagogical move. For example, one PST noted, "*During my micro-teaching lesson for a 5th-grade math class, I guided students to use ChatGPT to generate custom multiplication problems… After solving, they could recheck their answers by asking ChatGPT to confirm the correct solutions*" (S9). Another described grammar practice where students would "*feed ChatGPT grammatically wrong sentences, critique its corrections, then reverse roles*" (S16). Such narratives signal a shift from viewing AI primarily as a productivity aid to leveraging it for content-aligned, intentionally scaffolded instruction, which is consistent with the RQ1a gains.

Second, a small subset of participants shifted from simply using AI to actively creating with it. Five PSTs (20 %) reported experimenting with creator-level practices absent in their pre-survey responses, such as tweaking model parameters or training mini models. One PST explained, "*I used Teachable Machine to show students how AI 'learns' by recognizing patterns, and they thought it was super cool seeing a model work live*" (S3). This reflection highlights an intent to teach AI mechanics (i-CK) through hands-on model training. The emergence of this 'create' stance, while limited to a few participants, points to early signs of technical knowledge and confidence to work with more intricate tools than generative systems like ChatGPT (i-TK.*Know/Use*), along with tentative efforts to design activities in which students engage with model training to explore AI functionalities (i-CK.*Create*).

Finally, the ethics conversation showed signs of maturation between pre-/post-module reflections. Early comments tended to highlight hazards in abstract terms (e.g., "*AI could just do the work for students*," S12, Pre-Survey). By post, however, many respondents described proactive management strategies, such as posting classroom privacy rules, staging a mini debate on bias before prompting, or specifying verification steps alongside outputs. One PST explained, "*We posted privacy rules on the slides and staged a mini debate on bias before letting students prompt the chatbot*" (S9, Post-Survey). While almost every respondent continued to voice at least one concern about classroom AI, these were now more often paired with concrete mitigation plans (i-TPK.*Ethics*), suggesting a shift from abstract caution to ethics routines.

Taken together, the findings suggest that PSTs' integrated i-TPACK grew across the module in both measured competencies and articulated practices. The survey data (RQ1a) showed gains across all six domains. The reflection data (RQ1b) complemented these results by highlighting how such growth was expressed in practice. RQ2 extends this analysis by examining how integrated i-TPACK was evidenced in PSTs' artifacts across the module, providing a process-oriented view of its emergence and evolution.

4.2. RQ2. evolution of PSTs' integrated i-TPACK across module activities

RQ2 focused on how PSTs' integrated i-TPACK emerged and evolved, if at all, in their artifacts across the module activities. Analysis revealed four major themes that illustrate the trajectory of PSTs' development from foundational knowledge to sophisticated, ethically informed practice: (1) Building the foundation: Acquiring conceptual and technical knowledge; (2) Developing a critical stance: Integrating evaluation and ethical deliberation; (3) From theory to practice: Enacting competencies through pedagogical design; and (4) Fostering a professional

Table 2
Pre-Post Differences in Self-Rated Competencies across Six Knowledge Domains of Integrated i-TPACK.

Domain	*M* ± *SD* Pre	*M* ± *SD* Post	Δ (Post–Pre)	*Z*	*R* ($Z/\sqrt{N}$)	*padj* (Holm)	Effect size
i-CK	3.24 ± 0.97	4.00 ± 0.89	+0.76	–3.40	.68	.005	Large ↑
i-TPACK	3.71 ± 0.96	4.34 ± 0.70	+0.63	–3.39	.68	.005	Large ↑
i-TPK	3.80 ± 1.07	4.41 ± 0.60	+0.61	–3.16	.63	.006	Large ↑
i-TK	3.59 ± 0.79	4.13 ± 0.88	+0.5	–3.15	.63	.008	Large ↑
i-TCK	3.80 ± 0.97	4.19 ± 0.89	+0.39	–2.30	.46	.044	Med–large ↑
i-PD	3.81 ± 1.05	4.30 ± 0.65	+0.49	–2.16	.43	.031	Med–large ↑

Note. $N = 25$ matched pairs. Wilcoxon signed-rank tests with Holm–Bonferroni adjustment ($m = 6$). Effect size r computed as $Z/\sqrt{N}$; interpretive cut-points: 0.10 = small, 0.30 = medium, 0.50 = large [56]. Positive Δ indicates Post > Pre.

identity: Leveraging AI for self-directed growth.

4.2.1. Theme 1: building the foundation: acquiring conceptual and technical knowledge

This theme illuminates the crucial first stage of participants' development: acquiring foundational declarative knowledge. It traces their evolution from initial uncertainty to confidently articulating technical and pedagogical AI concepts. This foundational stage provided the vocabulary and mental models needed for developing more advanced critical and creative competencies as demonstrated in the increasing sophistication of their reflections and design tasks from Week 1 to Week 3.

The module began by challenging PSTs' initial, often vague, understandings of AI. The first formal activity, *Week 1 Ethical Scenario Analysis Worksheet (A1)*, served as a catalyst for knowledge articulation. Coding of this activity revealed a strong cognitive focus on knowledge-building, with 61 instances of i-TK.*Know* (defining the technology) and 26 instances of i-TPK.*Know (*explaining its pedagogical purpose). This structured reflection prompted a shift from simple statements like, *"AI is a tool that skims the web to give answers"* (S1, Comprehensive End-of-Week Reflection) to more sophisticated definitions. The evolution is illustrated in the detailed explanation provided by one participant: *"AI…refers to computer systems that can perform tasks typically requiring human intelligence—like analyzing data, recognizing patterns, or generating language… AI follows programmed algorithms and learns from large datasets"* (S19, Comprehensive End-of-Week Reflection).

This initial definitional knowledge was deepened and contextualized through the hands-on design tasks later in Week 1 and in Week 2. Following the *Teachable Machine Design Report (A2)*, participants' knowledge was expressed in more applied ways. In their *Week 2 AI-Integrated Lesson Plan (A8),* their conceptual knowledge was solidified and integrated. The lesson plan template required them to move beyond simple definitions to formalize their understanding in sections like "Learning Objectives" and "i-TPACK Considerations". Evidence for this can be seen in the 18 instances of i-TCK.*Know* and 11 instances of i-TPK.*Know* within this single group artifact, demonstrating that their foundational knowledge was now being explicitly embedded within a comprehensive pedagogical design. For example, one group justified their instructional choices with this i-TPACK.*Know* statement: *"The use of AI enhances this lesson because it gives students more examples of what the different seasons look like and it encourages them to think critically through compare and contrast work. This aligns with the science TEKS for first grade"* (G6, AI-Integrated Lesson Plan).

During Week 3 Micro-Teaching Presentations (A9), participants confidently explained complex AI concepts in age-appropriate ways to their "students." For instance, one group introduced a complex tool with this i-CK.*Know* explanation: "*ChatGPT is a computer helper that gives us ideas to improve our writing!"* (G5, Micro-Teaching Presentations). This developmental arc, from simple definitions (Week 1) to design integration (Week 2), and confident articulation (Week 3), answers RQ2, demonstrating how sequenced activities systematically built foundational knowledge for advanced i-TPACK competencies.

4.2.2. Theme 2: developing a critical stance: integrating evaluation and ethical deliberation

This second theme captures the most significant development in the participants' journey: the shift from foundational knowledge to a sophisticated, critical mindset where technical evaluation and ethical deliberation became inseparable. This theme illustrates how PSTs evolved from a state of uncritical consumption to forging a critical partnership with the technology. Initially, as evidenced in the *Week 1 Ethical Scenario Analysis Worksheet (A1)*, participants tended to view AI in purely functional terms, defining it as *"essentially a tool that is engineered/ programmed to do certain tasks"* (S2, Ethical Scenario Analysis Worksheet). However, by the end of the module, their reflections in the *Week 2 AI-Tool Rubric Evaluation (A7)* demonstrated a shift towards actively questioning AI's outputs, fairness, and pedagogical appropriateness.

This initial, uncritical stance was characterized by a focus on AI's basic mechanics rather than its implications. For instance, one PST's early definition framed AI as merely *"a digital tool that works as a computer where you can input a command or something into the system and create digital outputs from it"* (S5, Ethical Scenario Analysis Worksheet). This perspective lacks the deeper inquiry that emerged later. As PSTs engaged in hands-on tasks, their critiques evolved to address the pedagogical and ethical limitations of AI. This development is powerfully illustrated by one participant who not only identified a key weakness but also articulated a sophisticated pedagogical strategy to manage it, stating that *"The tool may generate text that sounds polished but lacks depth. I'd need to model how to use AI as a starting point, not an endpoint, and emphasize revision. Students should know how to identify when content lacks nuance or voice"* (S19, AI-Tool Rubric Evaluation). The same participant further clarified the ultimate objective of this new, critical approach, explaining, *"My goal is to ensure students are not just using AI to 'complete' tasks but are developing a critical partnership with technology that enhances—not replaces—their own thinking"* (S19, AI-Tool Rubric Evaluation).

This critical lens sharpened when focused on issues of equity and bias. After building and testing their own models, PSTs began to identify specific, harmful biases that could impact students from marginalized groups. One PST illustrated this deeper analysis after her group's hands-on work with the Teachable Machine: *"it worked better with light skin tones… It didn't respond as well to students with darker skin tones, which could make some kids feel like the tech doesn't recognize or work for them"* (S8, AI-Tool Rubric Evaluation). Another participant, as shown in Fig. 1, demonstrated a sophisticated critique of AI's potential pedagogical misinterpretations, noting that ChatGPT might frame a student's *"'quietness' or 'lack of eye contact' as problems, rather than culturally appropriate norms or trauma responses,"* which could lead to unfair labeling (S20, AI-Tool Rubric Evaluation).

By the *Week 2 AI-Integrated Lesson Plan (A8),* participants had moved from simply identifying problems to designing pedagogical solutions infused with ethical considerations and practices. One group planned an activity where students *"ask an AI chatbot some science questions… Then they will read the AI's answer and identify any errors… The students will compare & contrast the AI's to teacher-approved materials/websites"* (G8, Teaching AI Design Report). Another group used the AI model's limitations as a central part of the lesson: *"We will discuss why the AI could only recognize English words better than the Spanish ones, which would emphasize the bias in AI training"* (G7, Teachable Machine Design Report). This proactive, communicative approach, where a PST plans to *"talk to students and families before using the tool… explain what it does… and make sure they feel okay with it"* (S4, AI-Tool Rubric Evaluation), contrasts with earlier uncritical consumptions.

4.2.3. Theme 3: from theory to practice: enacting competencies through pedagogical design

This theme provides the most tangible evidence of the participants' developing competencies by tracing their journey from conceptual knowledge and critical evaluation to the active doing and creating. It captures their shift from simple, scaffolded AI use to designing sophisticated, original, and ethically informed pedagogical experiences. This development is evident when contrasting their initial, guided interactions with AI in the *Week 1 Teachable Machine Design Report (A2)* with the complex, self-directed designs evident in *Week 2 AI-Tool Rubric Evaluation (A7), Week 2 AI-Integrated Lesson Plan (A8)* and final *Week 3 Micro-Teaching Presentations (A9)*.

The Teachable Machine Design Report in Week 1 provided their first opportunity to use and create with an AI tool, which was heavily scaffolded. Coding showed a balanced focus on both application and design, with 27 instances of i-TCK.*Use* and 14 instances of i-TCK.*Create.* This suggests that the hands-on task successfully prompted them to both

3. Example of bias and its impact:

In a prompt about behavior in the classroom, ChatGPT may frame “quietness” or “lack of eye contact” as problems, rather than culturally appropriate norms or trauma responses.
Impact: Teachers unfamiliar with these nuances might misinterpret student behavior, potentially labeling students unfairly.

Significance: Moderate – The AI can unintentionally reinforce mainstream expectations without acknowledging cultural or individual diversity.

Short vs. Long-term Implications:

- **Short-term**: Students may feel misunderstood or excluded.
- **Long-term**: May reinforce systemic bias if unchecked, particularly for multilingual or neurodiverse learners.

Fig. 1. AI Potential Pedagogical Misinterpretation Statement of S20 in AI-Tool Rubric Evaluation.
Note. This section is extracted from the AI-Tool Rubric Evaluation developed by *Student 20*.

5. Description of AI-Enhanced Activity

- **Activity Structure**
  - **i.** Students will type into Chat GPT, “Give me 5th grade long multiplication questions. Start easy and get progressively harder after each question I answer correctly.”
  - ii. Students will solve the question given by Chat GPT and type their answer into the answer box.
  - iii. Students will write down each question and show their solving to turn in to the teacher.
  - iv. Students will then answer questions given from Chat GPT until the 20 minute timer goes off.
  - **v.** We will introduce the AI tool by providing an in class example and doing some as a class.
  - vi. Any policy checks will be getting permission from parents and administrators to use Chat GPT in the classroom.

- **Interaction with the AI**
  - i. Students will engage with AI by asking it questions and responding to the questions.
  - **ii.** AI is enhancing this lesson plan by creating unique problems for each student and progressively getting more challenging as they go along.
  - iii. Teachers will walk around as students are doing the task and monitor the communication.
- **Accommodations & Differentiation**
  - i. Students with alternative needs have the ability to listen to the text aloud through chat gpt and can ask for any accommodations to the AI tec
- **Photos / Visuals**
  - **i.**

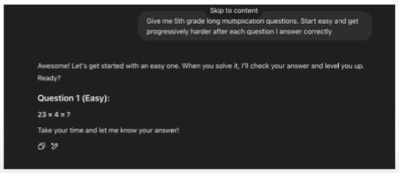

Fig. 2. Details of the Math Lesson of G7 in the Lesson Plan.
Note. This section is extracted from the Lesson Plan developed by *Group 7*.

operate the tool and begin thinking like designers. Notable at this early stage was participants' ability to transform the tool's limitations into a pedagogical opportunity. One PST reflected, *"I was surprised by how my model was not super accurate… but then I realized that we can use it to teach students to identify mistakes to help them learn. We could turn it into a lesson of them teaching the machine and correcting the mistakes it makes"* (S10, Teachable Machine Activity Reflection).

As PSTs constructed their AI-Integrated Lesson Plan in Week 2, initial practical experience evolved into more advanced instructional design. Analysis of this artifact revealed an increase in application-focused codes, with 118 instances of i-TPK.*Use* and 42 instances of i-TCK.*Use*, indicating that they were now applying AI tools to a wide range of pedagogical and content-specific goals. Importantly, numerous *Create* codes emerged, indicating a progression towards innovation as PSTs moved from just using tools to creating novel learning experiences. This pedagogical advancement was evident not just in the frequency of tool use but in the sophistication of the learning activities they designed.

For example, one participant designed a compare-and-contrast activity where students would first write their own essay outline and then critique an AI-generated version: *"In high school English class, I think that a good activity could be having the students write out an outline for their essay or writing assignment. After this, they could ask the AI tool to make them an outline based on the instructions provided. The students could then compare and contrast the two outlines and make edits as they see fit"* (S16, AI-Tool Rubric Evaluation). Another group designed an interactive grammar exercise that involved a multi-step cycle of evaluation and creation: *"Students will first input a prompt provided by us into ChatGPT. Students will edit the output provided by ChatGPT. After correctly editing this sentence, students will then provide ChatGPT with their own sentence including errors. ChatGPT will edit the provided sentence"* (G5, AI-Integrated Lesson Plan). Furthermore, another group moved beyond just using the tool to actively teaching students how to use it more effectively, for instance by modeling specific prompts to elicit age-appropriate responses: *"As mentioned before, to help students understand the AI better we will demonstrate example prompts. What is evaporation? Respond for a 4th grader."* (G8, AI-Integrated Lesson Plan). As shown in Fig. 2, one group designed a detailed, personalized math lesson using an AI chatbot: *"Students will type into Chat GPT, 'Give me 5th grade long multiplication questions. Start easy and get progressively harder after each question I answer correctly'"* (G7, AI-Integrated Lesson Plan).

The pinnacle of this theme was the creation of student-led projects. One of the groups designed a lesson in which students became AI creators rather than just learners: *"Students will use Teachable Machine to train a model to classify images of recyclable* vs. *non-recyclable materials… They will test the model by taking photos around the classroom or at home"* (G5, Teaching AI Design Report). This marks a substantial progression toward advanced pedagogical design, positioning students as active agents. In Week 3 Micro-Teaching, another group, demonstrated creativity by developing a ChatGPT-generated artifact, *"Water Cycle Song (Tune: If You're Happy and You Know It)"* (Fig. 3, G8, Micro-Teaching Presentations).

The progression from using a tool's limitations as a teaching moment, to designing personalized chatbot activities, to creating student-led AI projects and original learning materials, supports the pre- and post-survey shifts in RQ1.

4.2.4. Theme 4: fostering a professional identity: leveraging AI for self-directed growth

The final theme reveals a more advanced outcome: participants forming a professional identity that embraces AI for lifelong learning and collaboration. Initially, their conception of AI focused solely on K-12 classroom use, as seen in early *Week 1 Ethical Scenario Analysis Worksheet (A1)* and *Teachable Machine Design Report (A2)* which lacked references to professional growth. Through more complex evaluation and design tasks, they began recognizing AI's potential for their own practice, evidenced in later reflections that included self-directed, professionally oriented goals, particularly in the *Week 2 Pedagogical AI-Integration Reflection (A6)* and *AI-Tool Rubric Evaluation (A7)* activities. For example, by Week 2, one PST stated, "*PD is only helpful if it's relevant to what I'm doing in the classroom. I want PD that shows me how to use tech in a way that actually helps my ELs"* (S3, AI-Tool Rubric Evaluation). This reflection shows a shift from passive training recipient to active professional with specific learning needs.

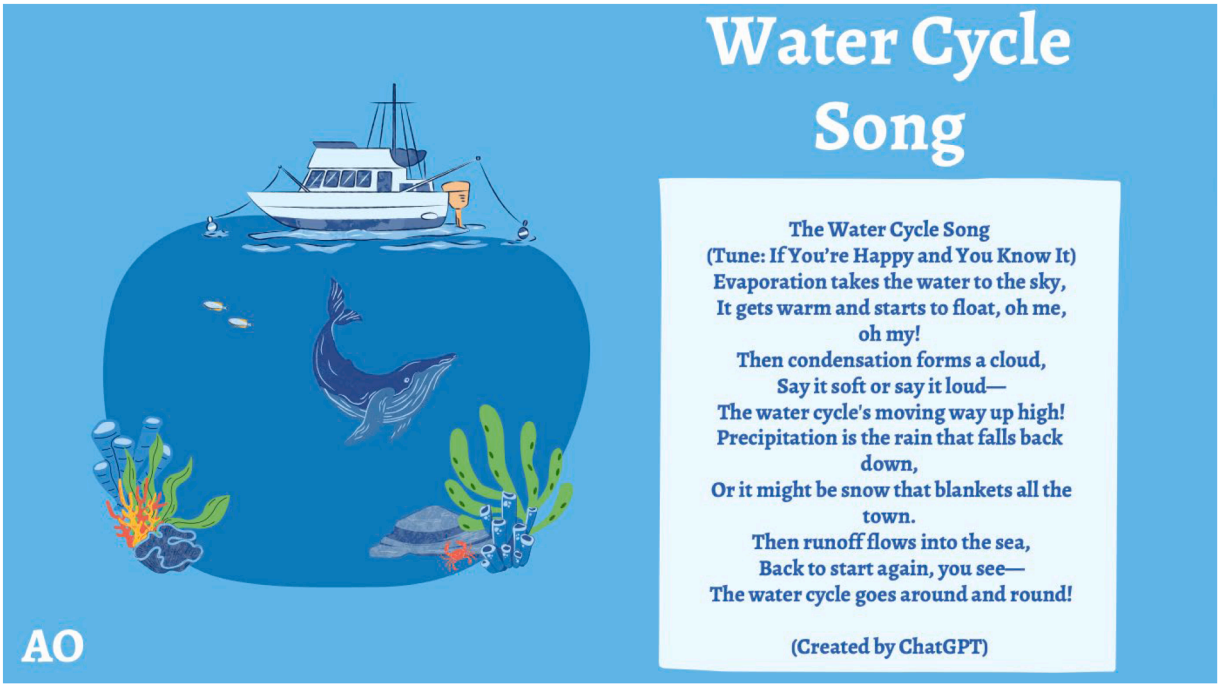

Fig. 3. Lyrics of the Song of G8 in Lesson Plan.
ote. This section is extracted from the Presentation developed by *Group 8*.

Participants' professional identities also became more collaborative and system-oriented, as they began to see themselves as part of a larger community of educators navigating AI's ethical complexities. One PST reflected, *"I also want to discuss ethical AI use with colleagues to build a shared understanding of its role in our classrooms"* (S19, Comprehensive End-of-Week Reflection). Another PST described consulting colleagues and standards when selecting tools, *"If I was unsure if the tool was good to use in my class, I would look at ISTE standards and consult my colleagues"* (S27, AI Tool Integration Reflection). This shift from an individualistic to a community-centered stance illustrates a meta-level competency that goes beyond immediate classroom practice. Ultimately, participants moved from viewing AI as a novel tool for children to recognizing it as integral to their own ongoing, critical, and collaborative professional learning.

5. Discussion

This study followed 29 PSTs through a three-week, integrated i-TPACK module that threaded a five-step ethical cycle across performance-based activities. Quantitative gains on all six domains of integrated i-TPACK, large for i-TK, i-TPK, i-TPACK, and i-CK, and medium-to-large for i-TCK and i-PD, paired with a four-stage qualitative trajectory, illuminate how technical insight, pedagogical integration, and ethical vigilance co-evolve in compressed time. The emergent arc, building foundational knowledge, adopting a critical stance, translating insight into design, and forging a professional identity, suggests that a tightly scaffolded "micro-cycle" can catalyze multidimensional growth while still revealing developmental pinch points.

First, building on prior attempts to "bolt-on" ethics to technology courses, our data suggest that treating ethical deliberation as the through-line, not the add-on, changes both the pace and the texture of growth. Participants were exposed to an ethical decision-making checkpoint in every activity; by the final micro-teaching, the majority spontaneously referenced privacy or bias mitigations without a rubric prompt, a behavior rarely reported in earlier i-TPACK interventions, where ethics was siloed in a single session [69,6]. This finding empirically builds upon Celik's [15] "Intelligent-TPACK" model, which positions ethics as a co-equal knowledge domain rather than a peripheral consideration. Moreover, the tight coupling we observed between i-TPK and i-TPACK post-test scores ($\rho = 0.84$) aligns with Ning et al.'s [20] structural equation work, which shows pedagogical AI knowledge as the primary driver of integrated competence. At the same time, very high correlations among i-TK, i-CK, and i-TCK ($\rho \geq .88$) potentially indicate that participants might have experienced tool knowledge, AI-as-content, and content–AI alignment as a single capability at course end. We therefore avoid claiming discriminant validity for that cluster in this cohort and treat the coupling as expected early-stage conflation that typically differentiates with extended practice and multi-method assessment. This interpretation aligns with our qualitative arc, where separation of roles emerged only after repeated design and enactment. Taken together, our mixed-methods results suggest that ethics could become actionable when PSTs can immediately test its consequences in lesson-design artefacts, supporting calls for performance-based AI education [70].

A second contribution concerns developmental sequencing. Mapping each activity onto the Know → Use → Evaluate → Ethics → Create ladder [19] appeared to compress learning curves: more than a quarter of the cohort reached the Create stage, i.e., training mini-models or scripting advanced prompts, within nine contact hours. Comparable studies that introduce AI tools absent a developmental scaffold report far lower rates of creator-level behaviors across entire semesters [71]. We hypothesize that situated redundancy, explicitly re-engaging the same ethical lens while the locus of control shifts from instructor demo to PST-led design, promotes adaptive expertise. This interpretation dovetails with Long and Magerko's [18] argument that AI literacy thrives when novices iteratively cycle between conceptual grounding and hands-on tinkering. Still, the comparatively modest gain in i-TCK indicates that novices struggle to see beyond cross-cutting skills (e.g., writing support) to deep, discipline-specific affordances. Future iterations should incorporate subject-specialist coaches and curated exemplar banks to accelerate i-TCK integration, echoing recommendations from recent scoping reviews [23].

The findings also underline identity formation as a driver of sustained engagement. By week three, reflective journals portrayed AI not as a productivity hack, but as a professional partner, an evolution consistent with identity-trajectory theory in teacher learning. This shift is noteworthy because previous surveys have documented persistent skepticism among PSTs about AI's pedagogical legitimacy [14]. We attribute the attitudinal change to two design choices: (a) foregrounding AI for PD use-cases (e.g., chatbot-mediated lesson study), and (b) embedding peer-feedback loops that cast PSTs as nascent members of an AI savvy professional community. Such design moves operationalize Floridi and Cowls' [72] principle of beneficence, promoting teacher well-being, while modelling autonomy-preserving practice.

Interpreted in the context of the literature, these findings contribute to three key conversations. First, they substantiate Celik's [15] and Ning et al.'s [20] claims that pedagogy-centered and ethics-aware knowledge elements are the strongest levers for integrated i-TPACK. By threading a five-step ethical decision checkpoint through every performance task, the module moved ethics from the "bolt-on" status criticized in prior reviews [73,6] to a catalytic practice that accelerated shifts from use to creation. Second, the uneven gain profile mirrors Sun et al.'s [26] observation that aligning AI affordances with disciplinary epistemologies (i-TCK) is slower than acquiring general technical skills; nonetheless, artifact data showed that even modest self-reported improvements translated into rich, subject-specific designs, indicating that performance evidence can outpace confidence. Third, the emergence of a creator-level stance, PSTs training mini-classifiers and crafting chatbot-generated materials, extends Long and Magerko's [18] AI literacy progression by showing that *Create* is feasible within a short, scaffolded window when ethics and critical evaluation are prerequisites rather than afterthoughts.

6. Implications, limitations and conclusion

Translating these insights into design guidance, four practical heuristics emerge. First, embed ethics early and often. When a brief ethical decision-making routine checkpoint preceded every learning task, novices transitioned ethical reflection from an abstract afterthought to a routine habit. This effect will only persist if courses continue to treat privacy, bias, and accountability as "must-pass" criteria at each design milestone. Second, rely on tightly sequenced performance tasks. Rapid cycling through the Know → Use → Evaluate → Ethics → Create ladder compressed learning curves; structuring a three-week "AI sprint" at the start of methods coursework can therefore bootstrap core competence before PSTs enter practicum. Third, supply discipline-specific examples. Persistent lags in i-TCK indicate that PSTs need concrete, subject-aligned cases, such as physics simulations, historical text-mining prompts, and genre-specific writing supports, co-taught or co-designed with content specialists, to recognize how AI can deepen conceptual learning rather than merely speed up task completion. Finally, cultivate AI-enabled reflective practice. Encouraging PSTs to use chatbots, analytics dashboards, or micro-journaling tools for lesson-study style reflection not only reinforces self-assessment skills but also establishes an i-PD habitus that can sustain growth after graduation.

In keeping a balanced perspective on these results, we acknowledge several limitations. First, the study involved a relatively small ($n = 29$), gender-skewed cohort from a single institution and course context, which may constrain both statistical power and the diversity of experiences represented. Second, the improvements we measured were largely self-reported and captured immediately after the three-week module, limiting our ability to assess long-term retention or transfer. Although

we triangulated these data with qualitative artifacts to enhance credibility, the evidence primarily reflects participants' intentions rather than their enacted classroom practices. Future research should examine how these competencies persist over time and manifest during student teaching. Finally, post-test correlations among i-TK/i-CK/i-TCK exceeded common heuristics for discriminant validity, consistent with novice conflation and single-method measurement. Larger, multi-method samples (e.g., objective knowledge tests and observation rubrics) and a hierarchical CFA will be needed to establish sharper construct boundaries. Nonetheless, by triangulating self-reports with nine distinct artifact sets, this study provides evidence at Kirkpatrick Level 2, supplemented by simulated practice via micro-teaching that potentially suggests readiness for Level 3. This is an empirical gap noted by Dogan et al. [44] and demonstrates that just nine contact hours could launch PSTs on an integrated i-TPACK trajectory. Scaling this work will require discipline-specific adaptations, the use of objective knowledge assessments to counteract potential optimism bias, and longitudinal tracking during practicum placements to examine whether create and ethics level behaviors persist when real classrooms and real stakes come into play.

In sum, a three-week, ethics-threaded PD module can launch PSTs from sandbox exploration to emerging i-TPACK integration while making visible the blind spots that warrant continued scaffolding. By demonstrating how technical fluency, pedagogical imagination, and ethical vigilance can co-evolve under time constraints, the study offers a replicable, time-efficient roadmap for preparing educators who can harness generative AI without compromising professional judgment or curricular rigor.

Data availability

Data generated or analyzed are available from the corresponding author by request.

Funding information

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

CRediT authorship contribution statement

Tugce Aldemir: Writing – review & editing, Writing – original draft, Methodology, Formal analysis, Data curation, Conceptualization. **Selcuk Kilinc:** Writing – review & editing, Writing – original draft, Conceptualization. **Ali Bicer:** Writing – review & editing, Writing – original draft, Methodology, Formal analysis. **Patricia Grant:** Writing – review & editing, Writing – original draft. **Trina Davis:** Writing – review & editing, Conceptualization. **Noelle Wall Sweany:** Writing – review & editing, Conceptualization.

Declaration of competing interest

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

Acknowledgement

The authors used OpenAI's ChatGPT (version 4o, accessed in June 2025) to support language editing and grammar correction. All content was originally written by the authors, and the tool was not used for generating ideas, data, or substantive content. After using this tool/service, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

Appendix A

Detailed Exploratory-Factor-Analysis Results for The Integrated i-TPACK Scale
Extraction = Principal-Axis Factoring, Rotation = Oblimin ($\delta = 0$).
All loadings < .40 suppressed for parsimony.

Table A.1
i-TK (Technological Knowledge) (6 Items).

Item	Loading
I can distinguish AI-driven tools (e.g., ChatGPT, image classifiers) from non-AI digital tools.	.809
I can explain generally how AI learns from data (e.g., pattern recognition).	.756
I can verify AI outputs, checking for bias or inaccuracies (e.g., reviewing AI-generated text for stereotypes) before using them in my tasks.	.848
I know how to operate basic AI software (e.g., lesson plan generators, chatbots) to produce outputs.	.728
I can compare multiple AI solutions (e.g., different chatbots) to identify which are most reliable for my professional tasks, checking user feedback or accuracy rates.	.790
I am comfortable customizing or adapting AI tools (e.g., adjusting prompts, small datasets) to suit my classrooms needs.	.724

Note. KMO = .79 Bartlett χ^2 (15) = 64.51, $p < .001$ Eigenvalue = 3.62 Variance = 60.3 %.

Table A.2
i-CK (AI as Content Knowledge) (5 Items).

Item	Loading
I can explain basic AI concepts (e.g., pattern recognition) in a clear and age-appropriate way for my students.	.745
I incorporate or plan to incorporate brief AI-related demos or discussions (e.g., What is an algorithm?) into my curriculum to build students' AI awareness.	.803
I can guide students to systematically evaluate AI outputs for potential biases, inaccuracies, and societal implications, exploring possible solutions to address these issues.	.877
I feel comfortable engaging students in hands-on AI activities (e.g., experimenting with a simple model) to deepen understanding of both AI's inner workings and ethical issues.	.891
I would encourage student-led AI projects (e.g., analyzing local data, exploring AI in social justice) to help them see real-world connections.	.792

Note. KMO = .79 Bartlett χ^2 (10) = 62.26, $p < .001$ Eigenvalue = 3.39 Variance = 67.8 %.

Table A.3
i-TCK (Technological Content Knowledge) (6 Items).

Items	Loading
I understand how AI tools (e.g., question generators, text analyzers) can align with my subject goals (e.g., math practice, language learning).	.873
I recognize that AI-generated materials (e.g., math problems, historical passages) may contain inaccuracies or reflect biases.	.780
I evaluate AI-generated resources for accuracy and cultural inclusivity before using them in my subject area.	.884
I can use AI-driven tasks (e.g., pattern recognition in science data, grammar checking in language arts) or tools to strengthen key content concepts.	.936
I can adapt or create AI-driven resources (e.g., customizing a dataset in biology, localizing texts for language learning) so that they enhance my subject instruction.	.829
I am comfortable inviting students to critique or improve AI-generated content to foster deeper engagement in my subject area.	.832

Note. KMO = .77 Bartlett χ^2 (15) = 123.46, $p < .001$ Eigenvalue = 4.41 Variance = 73.5 %.

Table A.4
i-TPK (Technological Pedagogical Knowledge) (6 Items).

Items	Loading
I can use basic AI features (e.g., auto-graded quizzes, chatbots) in teaching or lesson planning.	.955
I understand how basic AI features (e.g., auto-graded quizzes, chatbots) can support my teaching and lesson planning.	.947
I can adapt AI tasks (e.g., generating reading passages at different levels) and tools to support diverse learners (ELL students, varied reading levels, etc.).	.915
I can gather feedback (from peers or students) on AI-based strategies, adjusting if students seem confused or under-challenged.	.948
I can ensure that AI does not overshadow genuine student-teacher interaction, modifying tasks so students remain actively engaged and not passively reliant on AI.	.803
I can design advanced AI projects (e.g., problem-based tasks) that nurture creativity and critical thinking in the classroom	.685

Note. KMO = .81 Bartlett χ^2 (15) = 170.29, $p < .001$ Eigenvalue = 4.66 Variance = 77.7 %.

Table A.5
i-TPACK (Integrated) (7 Items).

Items	Loading
I recognize that effectively uniting AI, subject content, and pedagogy requires aligning all three from the planning stage.	.807
I am aware that student autonomy and teacher authority must be preserved even if AI suggests or automates certain tasks.	.752
I am prepared to regularly assess whether AI integration benefits both academic and socio-emotional outcomes (e.g., fosters critical thinking, respects diverse learners).	.916
I can create lessons in which AI tools, content goals, and teaching strategies complement each other for deeper learning.	.939
I can integrate AI, content, pedagogy, and ethical considerations (e.g., data privacy) from start to finish in my planning.	.879
I can design multi-day units or extended lessons where AI, content, pedagogy, and ethical concerns are fully integrated (e.g., from planning to assessment).	.832
I feel capable of creating basic AI-integrated lessons, checking at least minimal ethical concerns (e.g., data privacy) along the way.	.852

Note. KMO = .84 Bartlett χ^2 (21) = 145.63, $p < .001$ Eigenvalue = 5.13 Variance = 73.3 %.

Table A.6
i-PD (AI for Professional Development) (6 Items).

Items	Loading
I know my rights and responsibilities (e.g., data privacy, accountability) when using AI for my own professional growth (e.g., training modules, etc.).	.823
I can (or intend to) use AI tools or online communities to reflect on my teaching practice or gather new lesson ideas.	.958
I can (or intend to) use AI tools (like AI coaching bots or lesson-idea generators) to reflect on my teaching practice and gather new strategies.	.964
I can (or intend to) check whether these AI-based PD resources are valid, well-researched, or if they might be incomplete, biased, or manipulative.	.935
I can (or intend to) collaborate with peers to address the ethical aspects of AI-based PD, ensuring we do not rely on unscrupulous or privacy-violating platforms.	.955
I can (or intend to) co-create AI-based PD modules or policy guidelines that help other teachers ethically and effectively integrate AI in teaching.	.887

Note. KMO = .84 Bartlett χ^2 (15) = 194.84, $p < .001$ Eigenvalue = 5.09 Variance = 84.9 %.

Tables A.1-A.6

Appendix B Module activities mapping matrix

Table B.1
Details of the activities in the 3-week professional development module.

Lesson	Activity & Artifact	Key Tasks & Learning Objectives	Integrated i-TPACK Domain Foregrounded	AI Literacy Stage
Week 1				
W1 L1	Mini-Lecture: AI Mechanics & Ethics	• Review fundamentals of AI and Ethics. • Identify common ethical issues in AI (bias, privacy).	i-TK	Know → Ethics
W1 L1	AI Tool Exploration & Individual Reflection	• Use AI tools with varying functionalities (Craiyon, QuillBot, Claude). • Evaluate their benefits, risks, and limitations. • Propose initial classroom usage ethical guidelines.	i-TK, i-TPK	Know→ Use → Evaluate → Ethics
W1 L1	Ethical Scenario Analysis (*Artifact 1*)	• Analyze classroom scenarios and design mitigation Strategies (*with ethical decision-making checkpoint*).	i-TPK	Evaluate → Ethics
W1 L2	AI for Oceans Interactive Module	• Train a simple AI model to differentiate 'fish' vs. 'trash'. • Observe the effects of biased training data on outputs.	i-TK	Use → Evaluate→ Ethics

(*continued on next page*)

Table B.1 (*continued*)

Lesson	Activity & Artifact	Key Tasks & Learning Objectives	Integrated i-TPACK Domain Foregrounded	AI Literacy Stage
W1 L2	Teachable Machine Model Build & Reflection (*Artifacts 2 & 3*)	• Build and train a simple image classifier. • Design a mini-lesson concept integrating the model. • Reflect on content alignment, ethics, teaching AI (*with ethical decision-making checkpoint*).	i-TCK, i-TPK, i-CK (partial)	Use → Evaluate → Ethics→ Create
W1 L3	Teaching AI Design Report (*Artifacts 4)*	• Design a brief lesson to teach students how to evaluate AI, use it responsibly, and stay safe (*with ethical decision-making checkpoint*).	i-CK, i-TPK	Use → Evaluate → Ethics
W1 L3	Comprehensive End-of-Week Reflection (*Artifact 5*)	• Recap AI foundations and ethics (from • Week 1 activities. • Evaluate personal learning and plan next steps.	i-TK, i-TCK, i-TPK	Know → Use → Evaluate → Ethics
Week 2				
W2 L4	Pedagogical AI Integration Design & Reflection (*Artifact 6*)	• Select an AI tool for a specific pedagogical goal. • Design a student-facing activity, reflect on risks, and plan refinement (*with ethical decision making checkpoint*).	i-TPK, i-TK	Use → Evaluate → Ethics
W2 L5	Lecture: i-TPACK Framework & Scenario Evaluation	• Understand i-TPACK framework. • Apply the framework to evaluate classroom scenarios (*with ethical decision-making checkpoint*).	i-TPACK (with i-PD)	Know → Evaluate→ Ethics
W2 L5	AI Tool Rubric Evaluation (*Artifact 7*)	• Apply a four-dimensional rubric (bias, critical consumption, role in education, privacy) to systematically evaluate a chosen AI tool.	i-TK, i-TPK, i-CK	Evaluate → Ethics
W2 L6	Mini-Lecture: AI for Professional Development	• Discuss the use of AI for professional growth (e.g., AI coaching, reflective analytics).	i-PD	Know→ Ethics
W2 L6	AI-Integrated Lesson Plan (*Artifact 8*)	• Design a complete, standards-aligned lesson. • Plan for student interactions, assessment, and ethics. • Document and mitigate privacy concerns. • Evaluate i-TPACK integration of the lesson plan. • Apply ethical decision-making checkpoint in each phase.	i-TPACK (with i-PD)	Use → Evaluate → Ethics → (possibly) Create
Week 3				
W3 L7–9	Micro-Teaching & Peer Feedback (*Artifact 9*)	• Implement the AI-integrated lesson in a simulated setting. • Provide and receive formative peer feedback.	i-TPACK	Use → Evaluate → Ethics → (possibly) Create

Note. Each activity inevitably includes multiple knowledge domain × literacy stages. In this table, we highlight the most prominent emphases.

Appendix C. Integrated i-TPACK codebook

Table C.1
Integrated i-TPACK codebook.

Facet	Stage	Operational definition
i-TK (Tool/Systems Knowledge)	Know	Demonstrates conceptual understanding of the AI tool/system, i.e., what it is, how it works at a high level (data/algorithms), typical capabilities/limits, and conditions of use.
	Use	Operates the AI tool appropriately for the intended (non-pedagogical) purpose.
	Evaluate	Assesses tool behavior/outputs for reliability, quality (e.g., checks accuracy, latency, stability, hallucinations), and functional fit for the targeted task.
	Ethics	Identifies technical features that raise privacy, bias, accessibility, or safety risks and/or applies corresponding safeguards (privacy/data minimization, no PII, provenance/model version logging, consent/age-appropriateness, accessibility accommodations, human-in-the-loop for any high-stakes use).
	Create	Adapts or extends tool settings/workflows to better fit a use case (e.g., custom prompts/templates, lightweight parameter tuning, dataset curation/cleanup, simple pipelines/macros).
i-TPK (Pedagogical Integration)	Know	Identifies pedagogical affordances/constraints of AI for specific strategies (e.g., feedback cycles, differentiation, discussion scaffolds) and learner/context needs.
	Use	Selects and/or uses AI tools to meet a pedagogical objective (e.g., task flow, timing, grouping, scaffolds, monitoring).
	Evaluate	Specifies and/or applies verification/quality checks to judge alignment of AI outputs/affordances with learning goals, cognitive demand, and classroom context; and/or uses resulting evidence to adjust instruction.
	Ethics	Establishes and/or applies classroom norms/safeguards (e.g., acceptable use, disclosure/provenance statements, mitigation for bias/inequity, accessibility accommodations) that surface/address ethical risks in pedagogy.
	Create	Authors and/or substantially revises AI-supported pedagogical routines/materials (e.g., critique-revise cycles, peer-review workflows, prompt libraries) and/or iterates based on learner evidence.
i-TCK (Content–AI Alignment)	Know	Explains how AI can represent or support disciplinary ideas and/or where it commonly misleads in that subject.
	Use	Selects and/or uses AI tools to support subject learning (e.g., generating or transforming content-specific representations, examples, datasets, or visualizations that serve the stated learning goal).
	Evaluate	Assesses AI outputs against disciplinary criteria/standards (concept accuracy, concept integrity, representation coherence, unit consistency), may/may not include specifying/applying needed corrections.
	Ethics	Identifies and/or mitigates discipline-specific ethical risks (e.g., representation bias, cultural/linguistic exclusion, accessibility of representations), may/may not include disclosing known limitations to learners.
	Create	Designs and/or curates content-aligned AI artifacts/datasets/tasks that support subject learning (e.g., making target concepts visible, supporting misconception repair, supporting deeper disciplinary reasoning).
i-CK (Teaching About AI)	Know	Explains, in student-friendly terms, how AI systems work (e.g., learning from data, producing outputs, typical capabilities, limits, and sources of error/uncertainty).
	Use	Selects and/or facilitates activities (e.g., simple model demos, simulations, sandbox datasets, prompt comparisons) that let students experience core AI mechanics.
	Evaluate	Guides students in critiquing AI systems' technical limitations, ethical/societal impacts and/or developing strategies to navigate them.
	Ethics	Selects and/or facilitates explicit instruction and routines on fairness, accountability, privacy/data minimization, transparency/provenance, and accessibility within AI-integrated lessons.

(*continued on next page*)

Table C.1 (*continued*)

Facet	Stage	Operational definition
	Create	Supports student making/tuning (e.g., prototyping, prompt-engineering portfolios, basic model training) with/without metacognitive reflection on trade-offs and design decisions.
i-PD (AI for Professional Development)	Know	Identifies AI options and PD needs for one's context (goals, gaps, policies/constraints) and/or understands risks/benefits of teacher-facing AI for planning, feedback, and reflection.
	Use	Employs AI to support professional learning tasks (e.g., lesson study with a chat assistant, drafting plans with verification notes, curating resources, maintaining prompt/verification logs).
	Evaluate	Reviews the quality, credibility, and contextual fit of AI-assisted PD outputs/resources (e.g., alignment to standards/learners, bias, provenance), may/may not act on findings to refine practice.
	Ethics	Identifies and/or applies PD-specific safeguards (e.g., protects student/colleague data; respects copyright/provenance; documents model/version; avoids automated high-stakes judgments about teachers).
	Create	Develops and shares PD artifacts/workflows (e.g., templates, exemplars, micro-modules, reflective protocols) or contributes to communities of practice to advance ethical, effective AI use.
Integrated i-TPACK (Coordination Across Facets)	Know	Articulates how AI technology, pedagogical strategy, content goals, and contextual constraints (e.g., policies, inclusion, resources) interact in a given lesson or unit.
	Use	Designs and/or coordinates a learning activity so that tool choice, instructional method, content target, and an explicit ethical check (e.g., provenance/verification) operate together in real time.
	Evaluate	Assesses and/or revises an AI-supported design by balancing learning efficacy, technical feasibility, content alignment, and ethical duty, with/without specifying trade-offs and adjustments.
	Ethics	Designs/communicates/facilitate layered safeguards (technical, pedagogical, disciplinary) that foreground responsible/ethical AI use across a full lesson flow.
	Create	Conducts an iterative design/facilitation cycle (prototype → test → refine) that produces an integrated outcome (e.g., teacher-built aid, or student co-creation) in which AI functionality, disciplinary aims, pedagogy, and explicit ethical routines are woven together.

Note. Codes apply when the behavior is enacted in artifacts (e.g., lesson plans, micro-teaching slides, design reports) or explicitly articulated in reflections. Code the most specific facet implicated; permit stage co-codes (e.g., Use & Evaluate) when both intents are clear; facets are single-label unless integrative coordination is the focus (integrated i-TPACK).

Table C.1

Appendix D. Ethical decision-making checkpoint

When to use. Before any AI-mediated activity (planning or enactment), complete this 5-step checkpoint; keep a brief record (decision log) with the lesson artifacts.

1. **Context & purpose.** What is the learning goal? Who are the learners (grade, language needs, IEP/504, cultural/linguistic considerations)? Which AI tool/feature will be used and why (i-TK/TPK/TCK/TPACK/CK/PD link)?
2. **Risks & benefits.** Benefits for this context (engagement, differentiation, access)? Risks to anticipate now (privacy/data collection, bias/inaccuracy, workload, dependency, safety).
3. **Policies & stakeholders.** What do district/school policies, COPPA/FERPA (or local equivalents), and platform terms imply? Is consent needed? Who did you consult (admin, IT, mentor, families)? (*Accountability in action*)
4. **Decision & safeguards.** Adopt/adapt/reject. If adopt/adapt, list concrete safeguards (e.g., data minimization/no PII, disclosure/provenance statements, verification steps/human-in-the-loop, accessibility & translation, low-tech fallback).
5. **Reflect & document.** What did you decide and why? What will you monitor? What will you refine next time? (Attach prompts/templates; *this supports transparency and i-PD.*)
6. **Worked micro-example (Grade 7 ELA, summary writing with a chatbot)**
7. **Context & purpose.** Students practice gist statements and text-structure cues on a current events article. Target learners include multilingual students. (*i-TPK, i-TCK*)
8. **Risks & benefits.**
   a. **Risks:** student logins required (privacy/data retention unknown), English-only interface (inclusiveness), no source citations (transparency), possible hallucinations (accuracy), risk of over-reliance (agency).
   b. **Potential benefits (if present)**: leveled language scaffolds and quick examples; otherwise, benefits may be marginal if the class already uses leveled texts and teacher-made models.
9. **Policies & stakeholders**. District policy bans minors from creating accounts; IT confirms no Data Processing Agreement on file. Colleagues suggest a vetted reading platform with leveled texts. (*Accountability*)
10. **Decision & safeguards.**
    a. **Path a:** Adopt with safeguards (benefits > risks). Teacher pre-generates examples with a teacher account; no personally identifiable information; adds disclosure and a 3-step verification routine against the article; plans prompts that require students to critique and revise the AI output; provides bilingual glossaries and a no-AI option. (*Transparency, fairness, inclusiveness; human-in-the-loop*)
    b. **Path b:** Reject/replace (risks ≥ benefits or no clear benefit). Because student logins violate policy and unique learning value is weak, do not use the chatbot. Replace with vetted leveled texts and teacher-written model summaries; keep the critical-consumption piece by having students analyze two human-written summaries for accuracy and bias.
11. **Reflect & document.** Log the decision, rationale, and what would need to change (e.g., citation features, multilingual user interface) to reconsider later; save the verify routine for reuse. (*i-PD, transparency/accountability*)

References

[1] Al Samman AM. Harnessing potential: meta-analysis of AI integration in higher education. In: 2024 ASU International Conference in Emerging Technologies for Sustainability and Intelligent Systems (ICETSIS). IEEE; 2024. p. 1–7.
[2] Padmaja CVR, Narayana SL, Anga GL, Bhansali PK. The rise of artificial intelligence: a concise review. IAES Int J Artif Intell 2024;13:2224–33.
[3] Wang N, Wang X, Su YS. Critical analysis of the technological affordances, challenges and future directions of Generative AI in education: a systematic review. Asia Pacific J Educ 2024;44(1):139–55. https://doi.org/10.1080/02188791.2024.2305156.
[4] Ali H, Aysan AF. Ethical dimensions of generative AI: a cross-domain analysis using machine learning structural topic modeling. Int J Ethics Syst 2025;41(1):3–34. https://doi.org/10.1108/IJOES-04-2024-0112.
[5] Brandão A, Pedro L, Zagalo N. Teacher professional development for a future with generative artificial intelligence-an integrative literature review. Digital Educ Rev 2024;(45):151–7. https://doi.org/10.1344/der.2024.45.151-157.
[6] Rane N, Shirke S, Choudhary SP, Rane J. Education strategies for promoting academic integrity in the era of artificial intelligence and ChatGPT: ethical considerations, challenges, policies, and future directions. J ELT Stud 2024;1(1):36–59. https://doi.org/10.48185/jes.v1i1.1314.
[7] Xiao, P., Chen, Y., & Bao, W. (2023). Waiting, banning, and embracing: an empirical analysis of adapting policies for generative AI in higher education. *arXiv preprint* arXiv:2305.18617. https://doi.org/10.48550/arXiv.2305.18617.
[8] Furze L, Perkins M, Roe J, MacVaugh J. The AI Assessment Scale (AIAS) in action: a pilot implementation of GenAI-supported assessment. Australasian J Educ Technol 2024;40(4):38–55. https://doi.org/10.14742/ajet.9434.
[9] Ayanwale MA, Adelana OP, Molefi RR, Adeeko O, Ishola AM. Examining artificial intelligence literacy among pre-service teachers for future classrooms. Comput Educ Open 2024;6:100179. https://doi.org/10.1016/j.caeo.2024.100179.
[10] Bautista A, Estrada C, Jaravata AM, Mangaser LM, Narag F, Soquila R, Asuncion RJ. Preservice teachers' readiness towards integrating AI-based tools in education: a TPACK approach. Educ Process 2024;13(3):40–68. https://doi.org/10.22521/edupij.2024.133.3.
[11] Chan KKW, Tang WKW. Evaluating English teacher' artificial intelligence readiness and training needs with a TPACK-based model. World J English Lang 2025;15(1):129.
[12] Cukurova M, Kralj L, Hertz B, Saltidou E. Professional development for teachers in the age of ai. European Schoolnet Academy; 2024.
[13] Gander T, Harris G. Understanding AI literacy for higher education students: implications for assessment. He Rourou 2024;1(1):8. https://doi.org/10.54474/herourou.v1i1.10579.
[14] Tenberga I, Daniela L. Artificial intelligence literacy competencies for teachers through self-assessment tools. Sustainability 2024;16(23):10386. https://doi.org/10.3390/su162310386.
[15] Celik I. Towards intelligent-TPACK: an empirical study on teachers' professional knowledge to ethically integrate artificial intelligence (AI)-based tools into education. Comput Human Behav 2023;138:107468. https://doi.org/10.1016/j.chb.2022.107468.
[16] Karatas F, Atac BA. When TPACK meets artificial intelligence: analyzing TPACK and AI-TPACK components through structural equation modelling. Educ Inf Technol (Dordr) 2024;30:8979–9004. https://doi.org/10.1007/s10639-024-13164-2.
[17] Lintner T. A systematic review of AI literacy scales. NPJ Sci Learn 2024;9(1):50. https://doi.org/10.1038/s41539-024-00264-4.
[18] Long D, Magerko B. What is AI literacy? Competencies and design considerations. In: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. ACM Digital Library; 2020. p. 1–16. https://doi.org/10.1145/3313831.3376727.
[19] Ng DTK, Leung JKL, Chu SKW, Qiao MS. Conceptualizing AI literacy: an exploratory review. Comput Educ 2021;2:100041. https://doi.org/10.1016/j.caeai.2021.100041.
[20] Ning Y, Zhang C, Xu B, Zhou Y, Wijaya TT. Teachers' AI-TPACK: exploring the relationship between knowledge elements. Sustainability 2024;16(3):978. https://doi.org/10.3390/su16030978.
[21] Mishra P, Koehler MJ. Technological pedagogical content knowledge: a framework for teacher knowledge. Teach Coll Rec 2006;108(6):1017–54. https://doi.org/10.1111/j.1467-9620.2006.00684.
[22] Mishra P, Warr M, Islam R. TPACK in the age of ChatGPT and generative AI. J Digital Learn Teacher Educ 2023;39(4):235–51. https://doi.org/10.1080/21532974.2023.2247480.
[23] Sperling K, Stenberg CJ, McGrath C, Åkerfeldt A, Heintz F, Stenliden L. In search of artificial intelligence (AI) literacy in teacher education: a scoping review. Comput Educ Open 2024:100169. https://doi.org/10.1016/j.caeo.2024.100169.
[24] Hur JW. Fostering AI literacy: overcoming concerns and nurturing confidence among preservice teachers. Inf Learn Sci 2025;126(1/2):56–74. https://doi.org/10.1108/ILS-11-2023-0170.
[25] Lademann, J., Henze, J., Honke, N., Wollny, C., & Becker-Genschow, S. (2025). Teacher training in the age of AI: impact on AI literacy and teachers' attitudes. *arXiv preprint* arXiv:2507.03011. https://doi.org/10.48550/arXiv.2507.03011.
[26] Sun F, Tian P, Sun D, Fan Y, Yang Y. Pre-service teachers' inclination to integrate AI into STEM education: analysis of influencing factors. British J Educ Technol 2024;55(6):2574–96. https://doi.org/10.1111/bjet.13469.
[27] Ng DTK, Lee M, Tan RJY, Hu X, Downie JS, Chu SKW. A review of AI teaching and learning from 2000 to 2020. Educ Inf Technol (Dordr) 2023;28(7):8445–501. https://doi.org/10.1007/s10639-022-11491-w.
[28] Cukurova M, Miao F. AI competency framework for teachers. UNESCO Publishing; 2024.
[29] Almatrafi O, Johri A, Lee H. A systematic review of AI literacy conceptualization, constructs, and implementation and assessment efforts (2019-2023). Comput Educ Open 2024;6:100173. https://doi.org/10.1016/j.caeo.2024.100173.
[30] Ding ACE, Shi L, Yang H, Choi I. Enhancing teacher AI literacy and integration through different types of cases in teacher professional development. Comput Educ Open 2024;6:100178. https://doi.org/10.1016/j.caeo.2024.100178.
[31] Zhao L, Wu X, Luo H. Developing AI literacy for primary and middle school teachers in China: based on a structural equation modeling analysis. Sustainability 2022;14(21):14549. https://doi.org/10.3390/su142114549.
[32] Casal-Otero L, Catala A, Fernández-Morante C, Taboada M, Cebreiro B, Barro S. AI literacy in K-12: a systematic literature review. Int J STEM Educ 2023;10(1):29. https://doi.org/10.1186/s40594-023-00418-7.
[33] Druga S, Otero N, Ko AJ. The landscape of teaching resources for ai education. In: Proceedings of the 27th ACM Conference on Innovation and Technology in Computer Science Education. Vol. 1; 2022. p. 96–102.
[34] Ravi, P., Broski, A., Stump, G., Abelson, H., Klopfer, E., & Breazeal, C. (2023). Understanding teacher perspectives and experiences after deployment of AI literacy curriculum in middle-school classrooms. *arXiv preprint* arXiv:2312.04839. https://doi.org/10.48550/arXiv.2312.04839.
[35] Williams R, Kaputsos SP, Breazeal C. Teacher perspectives on how to train your robot: a middle school AI and ethics curriculum. Proc AAAI Conf Artif Intell 2021;35(17):15678–86. https://doi.org/10.1609/aaai.v35i17.17847.
[36] Lee I, Perret B. Preparing high school teachers to integrate AI methods into STEM classrooms. Proc AAAI Conf Artif Intell 2022;36(11):12783–91. https://doi.org/10.1609/aaai.v36i11.21557.
[37] Hutchins NM, Zhang S, Barrett JR, Isreal M. Empowering educators in AI: insights from co-designing an AI microcredential with and for K-12 educators. Proc AAAI Conf Artif Intell 2025;39(28):29137–44. https://doi.org/10.1609/aaai.v39i28.35186.
[38] Lin P, Van Brummelen J. Engaging teachers to co-design integrated AI curriculum for K-12 classrooms. In: Proceedings of the 2021 CHI conference on human factors in computing systems; 2021. p. 1–12. https://doi.org/10.1145/3411764.3445377.
[39] Xie B, Sarin P, Wolf J, Garcia RC, Delaney V, Sieh I, Lee VR. Co-designing AI education curriculum with cross-disciplinary high school teachers. Proc AAAI Conf Artif Intell 2024;38(21):23146–54. https://doi.org/10.1609/aaai.v38i21.30360.
….
[40] Graefe B, Porter A, Indorf J, Ahn S, Chuan CH, Gaines M. From chalkboards to ChatGPT: navigating the future of teaching with generative AI. J Interact Learn Res 2025;36(2):137–56. https://doi.org/10.70725/029002jsmbne.
[41] Jaksic N, Ansaf B, Massey M. Engaging high school teachers with artificial intelligence concepts, applications, and developments. ASEE Comput Educ 2025;14(3):1. https://doi.org/10.18260/B2B3-8F-68210.
[42] Walsh B, Dalton B, Forsyth S, Yeh T. Literacy and STEM teachers adapt AI ethics curriculum. Proc AAAI Conf Artif Intell 2023;37(13):16048–55. https://doi.org/10.1609/aaai.v37i13.26906.
[43] Zhang H, Lee I, Moore K. An effectiveness study of teacher-led AI literacy curriculum in K-12 classrooms. Proc AAAI Conf Artif Intell 2024;38(21):23318–25. https://doi.org/10.1609/aaai.v38i21.30380.
[44] Dogan S, Nalbantoglu UY, Celik I, Agacli Dogan N. Artificial intelligence professional development: a systematic review of TPACK, designs, and effects for teacher learning. Professional Dev Educ 2025;51(3):1–28. https://doi.org/10.1080/19415257.2025.2454457.
[45] Creswell JW, Plano Clark VL. Designing and conducting mixed methods research. 3rd ed. SAGE Publications; 2018.
[46] Yin RK. Case study research and applications: design and methods. Sage publications; 2017.
[47] Biggs J. Enhancing teaching through constructive alignment. High Educ (Dordr) 1996;32(3):347–64. https://doi.org/10.1007/BF00138871.
[48] Wiggins GP, McTighe J. Understanding by design. 2nd ed. ASCD; 2005.
[49] Kaiser HF. A second generation little jiffy. Psychometrika 1970;35(4):401–15. https://doi.org/10.1007/BF02291817.
[50] Bartlett MS. The effect of standardization on a $\chi 2$ approximation in factor analysis. Biometrika 1951;38(3/4):337–44.
[51] Bandalos DL. Measurement theory and applications for the social sciences. Guilford Publications; 2018.
[52] American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. Standards for educational and psychological testing. American Educational Research Association; 2014.
[53] Nunnally J, Bernstein I. Psychometric theory. 3rd ed. MacGraw-Hill; 1994.
[54] Authors. (under review).
[55] Holm S. A simple sequentially rejective multiple test procedure. Scandinavian J Stat 1979:65–70.
[56] Cohen J. A power primer. Psychol Bull 1992;112(1):155–9. https://doi.org/10.1037/0033-2909.112.1.155.
[57] Wilson ML, Ritzhaupt AD, Cheng L. The impact of teacher education courses for technology integration on pre-service teacher knowledge: a meta-analysis study. Comput Educ 2020;156:103941. https://doi.org/10.1016/j.compedu.2020.103941.
[58] Kline RB. Principles and practice of structural equation modeling. 4th ed. Guilford Publications; 2016.
[59] Field A. Discovering statistics using IBM SPSS statistics. Sage publications; 2024.
[60] McHugh ML. Interrater reliability: the kappa statistic. Biochem Med (Zagreb) 2012;22(3):276–82.

[61] Glaser BG, Strauss AL, Strutzel E. The discovery of grounded theory; strategies for qualitative research. Nurs Res 1968;17(4):364.
[62] Lincoln YS, Guba EG. Naturalistic inquiry. Sage; 1985.
[63] Saldaña J. The coding manual for qualitative researchers. 3rd ed. Sage; 2016.
[64] Strauss A, Corbin J. Basics of qualitative research: techniques and procedures for developing grounded theory. Sage publications; 1998.
[65] Fetters MD, Curry LA, Creswell JW. Achieving integration in mixed methods designs—Principles and practices. Health Serv Res 2013;48(6pt2):2134–56. https://doi.org/10.1111/1475-6773.12117.
[66] Greene JC. Mixed methods in social inquiry. John Wiley & Sons; 2007.
[67] Patton MQ. Enhancing the quality and credibility of qualitative analysis. Health Serv Res 1999;34(5 Pt 2):1189–208.
[68] Miles MB, Huberman AM, Saldana J. Qualitative data analysis: a methods sourcebook. Sage; 2014.
[69] Akgun S, Greenhow C. Artificial intelligence in education: addressing ethical challenges in K-12 settings. AI Ethics 2022;2(3):431–40. https://doi.org/10.1007/s43681-021-00096-7.
[70] Antonio RP. Promoting technological pedagogical content knowledge (TPACK) in preservice science teacher education: a scoping review of instructional strategies, interventions, and programs. Int J Stud Educ 2025;7(1):157–71.
[71] Mo Q, Yu J. AI-TPACK theory perspective: development of a model for enhancing teacher trainees' intelligent education literacy. J Intell Knowl Eng 2024;2:54–63. https://doi.org/10.62517/jike.202404410.
[72] Floridi L, Cowls J. A unified framework of five principles for AI in society. Harv Data Sci Rev 2019;1(1). https://doi.org/10.1162/99608f92.8cd550d1.
[73] Li Z. The shadow of the algorithm: the ethical blind spot of artificial intelligence education. J Educ Humanities Social Sci 2024;26:1146–52. https://doi.org/10.54097/pk0ajq60.