Developing and Validating Evidence-Based AI Professional Development Models for K-12 Teachers: A Longitudinal Study on Teacher Competency Development, Student AI Literacy Outcomes, and Regional Workforce Readiness in California

William Vortia
Westcliff University, Irvine California
w.vortia.6634@westcliff.edu

February 10th, 2026

Abstract

The integration of artificial intelligence (AI) into K-12 education represents one of the most critical workforce development challenges facing the United States. While 74% of school districts aim to train teachers in AI by 2025, effective professional development models remain severely underdeveloped, creating a dangerous gap between policy ambitions and pedagogical preparedness. This longitudinal research study proposes to develop, implement, and validate evidence-based AI professional development (PD) models that demonstrably translate into measurable improvements in teacher competency, student AI literacy, and regional economic competitiveness. The research employs a mixed-methods approach spanning five years, integrating competency measurement using the Teacher AI Competence Self-Efficacy (TAICS) scale and AI-Technological Pedagogical Content Knowledge (AI-TPACK) frameworks with student outcome assessments and longitudinal workforce tracking. Through multi-site implementation across 30-50 diverse school districts, this study will identify which PD intervention models most effectively build teacher capacity to prepare students for an AI-integrated economy. The research directly addresses the critical workforce readiness crisis, wherein currently only 5% of US high school graduates possess the digital and AI foundations necessary for modern employment. By establishing validated PD frameworks with clear return on investment metrics, this research provides actionable guidance for federal and state education policy, potentially affecting the preparation of over 50 million K-12 students and determining whether the United States maintains or loses its competitive position in the global AI economy.

Keywords: Artificial Intelligence, Teacher Professional Development, AI Literacy, Workforce Readiness, TAICS, AI-TPACK, Economic Competitiveness, K-12 Education.

INTRODUCTION AND BACKGROUND

1.1 The AI Education Crisis: A National Imperative

The United States faces an unprecedented educational and economic challenge. As artificial intelligence rapidly transforms every sector of the economy, the nation's K-12 education system struggles to prepare students for an AI-integrated future. This crisis manifests in stark statistics: while 85% of teachers and 86% of students report using AI tools in their classrooms, only 10% of schools have formal AI guidelines, and a mere 5% of graduating high school students possess the digital, data, and AI foundations required to work effectively alongside AI systems (Chiu et al., 2024; New America, 2025).

The economic implications are profound. Research indicates that up to 80% of U.S. workers will experience at least 10% of their work activities affected by large language models, with approximately 19% potentially seeing half or more of their tasks impacted (McKinsey Global Institute, 2023). The World Economic Forum's Future of Jobs Report 2025 projects that nearly 40% of workforce skills will change within five years, yet American students remain woefully unprepared for this transformation (World Economic Forum, 2025). As competing nations China, India, Germany, Canada, and the United Kingdom make substantial investments in AI education infrastructure and teacher training, the United States risks witnessing a generational transfer of economic opportunity abroad (New America, 2025).

The economic implications are profound. Research indicates that up to 80% of U.S. workers will experience at least 10% of their work activities affected by large language models, with approximately 19% potentially seeing half or more of their tasks impacted (McKinsey Global Institute, 2023). The World Economic Forum's Future of Jobs Report 2025 projects that nearly 40% of workforce skills will change within five years, yet American students remain woefully unprepared for this transformation (World Economic Forum, 2025). As competing nations China, India, Germany, Canada, and the United Kingdom make substantial investments in AI education infrastructure and teacher training, the United States risks witnessing a generational transfer of economic opportunity abroad (New America, 2025).

1.2 The Teacher Preparation Gap

At the heart of this crisis lies an acute teacher preparation deficit. Teachers serve as the critical bottleneck in scaling AI literacy across the student population, yet systematic research reveals significant inadequacies in their professional development. A comprehensive systematic review by researchers examining AI in education from 2015-2024 found that only 35% of studies explored AI's role in enhancing teacher professional development, revealing a substantial research-practice gap (Wang et al., 2024). This imbalance persists despite widespread recognition that effective AI integration requires not merely technical proficiency but sophisticated pedagogical integration skills. Current teacher

AI professional development suffers from multiple deficiencies. Most district leaders lack sufficient understanding of AI themselves, making it difficult to facilitate effective training programs (EdWeek Research Center, 2025). Professional development frequently adopts outdated 'sit-and-get' models rather than interactive, sustained learning approaches (Smokorowski & Reljac, 2025). Furthermore, the rapid evolution of AI technologies means that training content quickly becomes obsolete, yet few districts have established mechanisms for continuous professional learning (EdCircuit, 2025).

Perhaps most concerning, existing professional development rarely addresses the full scope of competencies teachers require. Research demonstrates that AI literacy extends beyond technical proficiency to encompass ethical considerations, societal impacts, and practical pedagogical applications, yet many current curricula focus primarily on tool usage while neglecting operational strategies for integrating ethics and critical thinking into classroom instruction (OECD, 2025).

1.3 Theoretical Frameworks: From TPACK to AI-TPACK

Understanding teacher competency in AI integration requires sophisticated theoretical frameworks that extend beyond traditional technology integration models. The Technological Pedagogical Content Knowledge (TPACK) framework, developed by Mishra and Koehler (2006), has long served as the foundation for conceptualizing teacher technology integration expertise. TPACK posits that effective technology integration emerges from the dynamic intersection of three knowledge domains: technological knowledge (TK), pedagogical knowledge (PK), and content knowledge (CK).

However, AI's unique characteristics include autonomous capabilities, machine learning, dynamic adaptability, and ethical complexities necessitate framework extensions. Recent scholarship has proposed AI-TPACK models that specifically incorporate AI-related competencies. Celik (2023) developed the Intelligent-TPACK framework, which emphasizes teachers' professional knowledge for ethically integrating AI-based tools into education. This framework demonstrates that AI literacy significantly predicts AI acceptance ($\beta = 0.43$, $p < 0.001$), explaining 67% of variance in adoption behaviors (Al-Abdullatif, 2024).

Complementing AI-TPACK, the Teacher AI Competence Self-Efficacy (TAICS) scale provides a validated instrument for measuring teacher AI competencies across six dimensions: AI knowledge, AI pedagogy, AI assessment, AI ethics, human-centered education, and professional engagement (Chiu et al., 2024). Developed through Delphi methodology and validated with 434 K-12 teachers via confirmatory factor analysis, the TAICS scale offers 24 items that comprehensively assess teacher readiness for AI integration. Research demonstrates strong correlations between AI-TPACK and TAICS dimensions, with particularly robust relationships between AI-Technological Pedagogical Knowledge (AI-TPK) and both AI-Technological Knowledge (AI-TK) and AI-Technological Content Knowledge (AI-TCK), suggesting that technological knowledge and pedagogical integration must be considered together (Science Teachers' AI Competency Study, 2025).

LITERATURE REVIEW

2.1 Current State of AI Professional Development

The landscape of AI professional development for teachers reveals significant variability in implementation approaches and outcomes. Recent surveys indicate that 74% of school districts aim to train teachers in AI, yet most lack structured programs beyond initial workshops (EdWeek Research Center, 2025). Districts that have implemented AI PD report varying levels of success, with common challenges including insufficient time for teacher participation, lack of ongoing support, and difficulty demonstrating tangible classroom impact.

Emerging research on effective PD models suggests several promising approaches. Case-based professional development has shown efficacy in developing teacher AI literacy. A study examining case-based AI PD programs with seven middle school science teachers found marked increases in AI literacy, particularly in knowing and understanding AI domains, though teachers demonstrated limited application of this knowledge during case discussions (Chen et al., 2024). The research highlighted the importance of combining direct instruction with case-based discussions, with structured progression from well-structured to ill-structured case problems promoting deeper knowledge application.

Intelligent tutoring systems represent another promising avenue for scaling teacher professional development. A randomized controlled trial investigating virtual, interactive programs using intelligent tutoring systems to provide just-in-time feedback to teachers found that participants who completed the program (N=29) used mathematically richer tasks and created more coherent, connected learning environments compared to business-as-usual conditions (N=23) (Copur-Gencturk et al., 2024). This research demonstrates that AI-enhanced PD platforms can provide personalized, scalable professional learning while maintaining effectiveness.

AI platforms that recommend learning modules based on teacher data, self-assessments, and classroom analytics show potential for creating adaptive professional development experiences that evolve with teacher needs (Frontiers in AI, 2023). However, research also reveals that the emphasis on delivering technologies often overshadows fostering teacher professional competence development through interactions with AI tools, suggesting that PD must focus not just on tool proficiency but on developing professional vision and pedagogical reasoning skills.

2.2 Measurement and Assessment of Teacher AI Competency

Accurately measuring teacher AI competency remains essential for evaluating professional development effectiveness and guiding improvement efforts. The field has converged on several validated instruments, with the TAICS scale and AI-TPACK frameworks emerging as primary assessment tools.

The TAICS scale's six-dimensional structure covering AI knowledge (AIK), AI pedagogy (AIP), AI assessment (AIA), AI ethics (AIE), human-centered education (HCE), and professional engagement (PEN) provides comprehensive coverage of teacher competencies (Chiu et al., 2024). Validation studies demonstrate strong psychometric properties, with Cronbach's alpha values ranging from 0.806

to 0.945 across dimensions and an overall scale reliability of 0.957. The scale's alignment with UNESCO's draft AI competence framework for educators further supports its theoretical grounding and international applicability.

AI-TPACK instruments measure teacher competency across seven constructs: Pedagogical Knowledge (PK), Content Knowledge (CK), AI-Technological Knowledge (AI-TK), Pedagogical Content Knowledge (PCK), AI-Technological Content Knowledge (AI-TCK), AI-Technological Pedagogical Knowledge (AI-TPK), and comprehensive AI-TPACK (Xu et al., 2025). Research reveals that different knowledge elements show varying explanatory power for teachers' overall AI-TPACK, with technology-related knowledge elements demonstrating significantly stronger predictive validity than non-technical elements. Interestingly, content knowledge sometimes diminishes the explanatory power of PCK and AI-TCK, suggesting complex interactions among competency domains that warrant further investigation.

Recent developments in assessment methodology include artifact-based evaluation approaches that move beyond self-reporting. Research developing AI-TPACK assessment tools grounded in authentic pedagogy and authentic instructional artifacts identifies four competency patterns: technological innovator, pedagogical integrator, content developer, and beginner (Education Sciences, 2025). The strong correlation ($r = 0.78$) between AI-Pedagogical Knowledge (AIPK) and integration competencies underscores the critical importance of synergistic knowledge domains.

2.3 Student Outcomes and AI Literacy Development

While teacher competency represents the input side of the education equation, student AI literacy outcomes constitute the critical output that ultimately determines workforce readiness and economic competitiveness. However, research directly connecting teacher AI professional development to student literacy outcomes remains limited, representing a significant gap this study aims to address.

The OECD's AI Literacy Framework, currently in draft form with final release planned for 2026, provides structure for conceptualizing student competencies across four domains: engaging with AI, creating with AI, managing AI's actions, and designing AI solutions (OECD, 2025). Each domain encompasses specific competencies ranging from understanding AI presence in everyday tools to exploring how AI systems can be built or adapted to solve real-world problems. However, the framework acknowledges that many current curricula focus primarily on technical skills with insufficient emphasis on ethical considerations and operational strategies for classroom integration.

Research on AI literacy's cognitive and metacognitive dimensions reveals that robust literacy involves not merely theoretical understanding of algorithmic functions but also ethical standpoints and willingness to critically evaluate AI systems (Pinski & Benlian, 2023; Weber et al., 2023). The meta AI literacy scale (MAILS) expands beyond cognitive competencies to include self-efficacy components such as AI problem-solving and learning efficacy, as well as meta-competencies like AI persuasion and emotion-regulation literacies (Carolus et al., 2023).

Importantly, assessment approaches increasingly recognize the need to capture both practical and affective outcomes through combinations of knowledge-based tests, self-report surveys, and empowerment measures (Kong et al., 2021). As AI technologies permeate everyday life, comprehensive assessment must address ethical implications, societal impacts, and civic awareness moving beyond technical proficiency to encompass critical engagement with AI's role in sustainability, labor, policy, and equity (Tadimalla, 2025).

2.4 Workforce Readiness and Economic Impact

The ultimate justification for investing in AI education lies in workforce preparation and economic competitiveness. Generative AI literacy has emerged as an essential occupational competence, with research demonstrating significant positive impact on job performance ($\beta = 0.680$, $p < 0.001$), mediated partially by creative self-efficacy (indirect effect = 0.537) (Li et al., 2024). This relationship operates through the Ability-Motivation-Opportunity (AMO) framework, where employee performance results from multiplicative interactions among ability, motivation, and opportunity.

Labor market analysis reveals stark disparities in AI preparedness. While AI use in the workforce has risen significantly, more than half of workers report feeling unprepared to use AI in their jobs a gap directly parallel to deficiencies in teacher training (New America, 2025). McKinsey research projects that generative AI could contribute between \$2.6 trillion and \$4.4 trillion annually to the global economy, with companies effectively using AI expected to increase workforce performance by up to 40% (World Economic Forum, 2025).

Research on AI's labor market effects reveals nuanced patterns. In automation-prone occupations, generative AI automates substantial portions of tasks, simplifying workflows and reducing required skill breadth a process of deskilling. Conversely, in augmentation-prone occupations, AI introduces new tools and workflows that expand skill requirements, necessitating workers to develop complementary capabilities including AI literacy, advanced analytical skills, and creative problem-solving a process of upskilling (Harvard Business School Working Paper, 2025).

University graduate employment studies demonstrate that individuals reporting greater AI knowledge and more frequent use of AI tools, especially generative ones like ChatGPT, show higher likelihood of employment in their field of study and perceive greater productivity and career prospects (Frontiers in AI, 2025). Technical skills now become outdated in less than five years on average, underscoring the urgency of building adaptable AI literacy rather than narrowly focused technical training (McKinsey, 2023).

Digital literacy for workforce readiness research emphasizes that the Fourth Industrial Revolution's distinguishing feature lies in the speed, range, and systems-level impact of simultaneous technological disruptions (Asa et al., 2025). This paradigm shift has irrevocably transformed not just industries and business models but the fundamental nature of work itself, making AI literacy a foundational skill as essential as reading, writing, and mathematics.

RESEARCH QUESTIONS AND HYPOTHESES

3.1 Primary Research Question

What teacher AI professional development models most effectively translate into measurable improvements in (1) *teacher AI pedagogical competency, (2) student AI literacy and workforce readiness skills, and (3) regional economic competitiveness indicators?*

3.2 Secondary Research Questions

RQ1: How do different AI professional development delivery models (AI-personalized, cohort-based, hybrid) impact teacher competency development as measure by TAICS and AI-TPACK scales?

RQ2: What is the relationship between teacher AI competency gains and student AI literacy outcomes across diverse demographic and socioeconomic contexts?

RQ3: How do variations in professional development content focus (technical skills, pedagogical integration, ethical frameworks) differentially predict teacher implementation fidelity and student learning outcomes?

RQ4: What mediating factors (school culture, administrative support, resource availability) influence the translation of teacher PD participation into classroom practice?

RQ5: To what extent do improvements in district-level AI education implementation predict regional workforce readiness indicators and economic outcomes over a 3–5-year period?

3.3 Core Hypotheses

H1 (Competency-to-Outcome Pathway): Teachers trained through structured, evidence-based AI professional development programs will demonstrate significantly higher levels of AI pedagogical competence (measured by TAICS and AI-TPACK scales) compared to control groups, with direct positive correlation to student AI literacy gains ($r > 0.50$, $p < 0.001$).

This hypothesis rests on extensive research demonstrating that teacher knowledge and pedagogical competence directly influence student learning outcomes. The TAICS framework's validation showed strong internal consistency (Cronbach's α ranging from 0.806 to 0.945), suggesting reliable measurement of competency changes. Furthermore, research on case-based AI professional development demonstrated marked increases in teacher AI literacy, particularly in knowledge and understanding domains (Chen et al., 2024). We anticipate that systematic PD will produce measurable teacher competency gains that translate into improved student outcomes, with effect sizes comparable to those observed in other subject-area professional development interventions ($d > 0.40$).

H2 (Economic Multiplier Effect): School districts implementing validated AI professional development models will produce graduates with measurably higher AI workforce readiness scores (>30% improvement over baseline), leading to reduced skills gaps, higher entry-level employment rates (>15% increase), and increased regional economic productivity within 3-5 years (measured through GDP contribution per capita and technology sector employment growth).

This hypothesis extends educational research into economic domain, predicting that systematic AI education will produce measurable workforce and economic effects. Supporting evidence includes research demonstrating that generative AI literacy significantly impacts job performance ($\beta = 0.680$, $p < 0.001$) with creative self-efficacy serving as partial mediator (Li et al., 2024). University graduate employment studies show that greater AI knowledge and tool usage correlate with higher employment rates in field-related positions (Frontiers in AI, 2025). Given that currently only 5% of US high school graduates possess adequate AI foundations (New America, 2025), substantial improvements in this baseline should produce detectable economic effects within 3–5-year timeframes as graduates enter regional labor markets.

H3 (Differential Impact by Model Type): AI-enhanced professional development platforms that recommend learning modules based on teacher data, self-assessments, and classroom analytics will produce superior outcomes (Cohen's $d > 0.60$) compared to traditional one-size-fits-all professional development approaches across all measured dimensions (teacher competency, implementation fidelity, student outcomes).

This hypothesis predicts that personalized, adaptive professional development will outperform standardized approaches a pattern consistent with broader educational research on differentiated instruction. Research on intelligent tutoring systems for teachers found that participants completing personalized programs demonstrated superior pedagogical practices compared to business-as-usual conditions (Copur-Gencturk et al., 2024). Studies examining AI-enhanced learning platforms emphasize their capacity to provide individualized learning pathways, just-in-time feedback, and adaptive content sequencing (Frontiers in AI, 2023). We anticipate that these advantages will manifest measurably better teacher outcomes, with effect sizes exceeding those typically observed for traditional professional development ($d = 0.20$-0.40).

RESEARCH METHODOLOGY

4.1 Research Design Overview

This study employs a longitudinal, quasi-experimental mixed-methods design spanning five years, incorporating multiple data collection points, diverse measurement instruments, and sophisticated analytical techniques. The research unfolds in three sequential phases: (1) Model Development (Year 1), (2) Multi-Site Implementation and Teacher/Student Outcome Assessment (Years 2-3), and (3) Workforce Impact Analysis and Economic Evaluation (Years 4-5). This phased approach allows for iterative refinement while maintaining methodological rigor and enabling comprehensive examination of both proximal educational outcomes and distal economic effects.

The study integrates quantitative and qualitative methods to provide triangulated evidence. Quantitative components include pre-post assessments of teacher competency using validated instruments (TAICS, AI-TPACK), student AI literacy assessments, classroom observation protocols, and longitudinal tracking of workforce outcomes through administrative data linkages. Qualitative components comprise teacher interviews, focus groups, classroom artifact analysis, and case studies of high-performing and struggling implementation sites. This methodological pluralism enables both hypothesis testing and theory building, allowing the research to identify not merely what works but why and under what conditions.

4.2 Phase 1: Model Development (Year 1)

4.2.1 Synthesis of Existing Frameworks

The initial phase synthesizes existing research on effective professional development to create theoretically grounded intervention models. This synthesis draws extensively from multiple framework traditions including TPACK (Mishra & Koehler, 2006), Intelligent-TPACK (Celik, 2023), the TAICS competency framework (Chiu et al., 2024), and Falloon's (2020) teacher digital competence framework. Additionally, the synthesis incorporates insights from case-based learning research (Chen et al., 2024), intelligent tutoring system studies (Copur-Gencturk et al., 2024), and adult learning theory emphasizing sustained, job-embedded professional development.

A research team comprising education researchers, AI specialists, and experienced K-12 practitioners will conduct systematic literature review following PRISMA protocols to identify effective PD components. This review will specifically examine: (a) optimal duration and intensity patterns for teacher learning, (b) most effective content sequencing strategies, (c) successful approaches to building both technical and pedagogical competencies, (d) strategies for promoting sustained implementation, and (e) mechanisms for providing ongoing support. The synthesis will produce a comprehensive evidence map identifying high-impact PD elements and theoretical mechanisms explaining their effectiveness.

4.2.2 Development of Intervention Models

Based on the synthesis, the research team will develop three distinct professional development intervention models, each representing a theoretically coherent approach but varying along key dimensions to enable comparative analysis:

Model A: AI-Personalized Adaptive Platform

This model leverages AI-enhanced learning platforms to provide individualized professional development pathways. Teachers complete initial competency assessments using TAICS and AI-TPACK instruments, after which the platform recommends customized learning modules addressing identified gaps. The platform incorporates: (a) microlearning modules (10-15 minutes each) covering specific competencies, (b) interactive simulations enabling practice with AI tools in low-stakes contexts, (c) just-in-time resources triggered by classroom implementation data, (d) AI-powered coaching providing feedback on lesson artifacts and student work, and (e) adaptive difficulty adjustment based on demonstrated mastery. The model emphasizes flexibility and self-paced learning, with teachers accessing content asynchronously according to their schedules. Expected completion time: 40-50 hours over 6 months, with 70% asynchronous and 30% synchronous (virtual cohort meetings).

Model B: Cohort-Based Collaborative Learning

This model emphasizes social learning and collaborative knowledge construction. Teachers participate in school-based cohorts (8-12 members) meeting regularly for structured learning activities. The model features: (a) initial intensive workshop (3 days) establishing foundational knowledge, (b) bi-weekly 90-minute cohort meetings featuring case-based discussions, lesson study protocols, and peer feedback on implementation, (c) classroom implementation cycles with structured reflection protocols, (d) expert coaching with monthly visits providing observation feedback and modeling, and (e) collaborative resource development with cohorts creating and refining AI-enhanced lesson plans. This approach prioritizes sustained engagement, contextual application, and professional community building. Expected time commitment: 60-70 hours over 8 months, with 60% synchronous face-to-face and 40% asynchronous individual work.

Model C: Hybrid Integration Model

This model combines elements from both approaches, attempting to leverage advantages of personalized learning and collaborative engagement. Teachers experience: (a) personalized diagnostic assessment determining individual pathways, (b) core synchronous workshops (4 half-day sessions) covering essential content with all participants, (c) differentiated asynchronous modules tailored to assessed needs, (d) monthly cohort meetings for collaborative application and problem-solving, (e) AI coaching providing personalized feedback between sessions, and (f) culminating capstone project demonstrating integration competency. The hybrid model aims to balance efficiency of technology-mediated learning with benefits of social collaboration. Expected commitment: 50-60 hours over 7 months, evenly split between synchronous and asynchronous activities.

Additionally, each model will systematically vary content focus to enable analysis of differential effects:

- **Technical Skills Focus:** Emphasizes tool proficiency, understanding AI capabilities/limitations, prompt engineering, and technical troubleshooting (approximately 50% of content)
- **Pedagogical Integration Focus:** Emphasizes lesson design, assessment strategies, differentiation approaches, and classroom management with AI tools (approximately 50% of content)
- **Ethical Framework Focus:** Emphasizes responsible AI use, bias recognition, privacy considerations, academic integrity, and critical evaluation (approximately 30% of content, integrated throughout)

All models will undergo pilot testing with small teacher cohorts (N=15-20 per model) during the final quarter of Year 1. Pilot data will inform refinements to content, sequencing, time allocations, and support structures before full-scale implementation.

4.3 Phase 2: Multi-Site Implementation (Years 2-3)

4.3.1 Site Selection and Sampling

The study will recruit 30-50 school districts representing diverse contexts to enable examination of implementation effects across varied settings. Site selection will employ purposive sampling to ensure representation across multiple dimensions:

- **Geographic Distribution:** Districts from urban (40%), suburban (35%), and rural (25%) contexts across all four US Census regions (Northeast, South, Midwest, West)
- **Socioeconomic Context:** Equal distribution across high-poverty (>60% free/reduced lunch), moderate-poverty (30-60%), and low-poverty (<30%) contexts
- **District Size:** Small (<3,000 students; 30%), medium (3,000-15,000 students; 40%), and large (>15,000 students; 30%)
- **Technology Infrastructure:** Varied levels of existing technology integration and infrastructure capacity
- **Student Demographics:** Diverse racial/ethnic compositions and English learner populations

Within each participating district, we will recruit teacher cohorts (target: 30-50 teachers per district) representing diverse grade levels (elementary, middle, high school) and subject areas (core academic subjects, electives, special education). Random assignments will occur at the teacher level within districts, with teachers randomly assigned to one of the three intervention models or a delayed-treatment control condition. This nested randomization design enables causal inference while maintaining practical feasibility.

Total anticipated sample: 1,200-2,000 teachers across 30-50 districts, with approximately 400-650 teachers per intervention model and 400-650 in delayed-treatment control. Statistical power analyses indicate this sample size provides adequate power ($\beta > 0.80$) to detect medium effect sizes ($d = 0.40$) for teacher outcomes and small-to-medium effects ($d = 0.30$) for student outcomes, accounting for clustering effects and anticipated attrition (estimated at 15-20%).

4.3.2 Data Collection: Teacher Measures

Teacher competency will be assessed at multiple timepoints using validated instruments and observation protocols:

TAICS Scale Administration

The 24-item Teacher AI Competence Self-Efficacy scale will be administered at five timepoints: (1) baseline (pre-intervention), (2) immediate post-intervention, (3) 3-month follow-up, (4) 6-month follow-up, and (5) 12-month follow-up. The instrument measures self-efficacy across six dimensions: AI knowledge, AI pedagogy, AI assessment, AI ethics, human-centered education, and professional engagement. Items use 5-point Likert scales (1=Not confident at all to 5=Extremely confident). Administration occurs online with 15–20-minute completion time. The TAICS demonstrated strong psychometric properties in validation studies (Cronbach's α = 0.806-0.945 across dimensions; Chiu et al., 2024), and we will verify reliability within our sample.

AI-TPACK Scale Administration

The 39-item AI-TPACK instrument measuring seven dimensions (PK, CK, AI-TK, PCK, AI-TCK, AI-TPK, AI-TPACK) will be administered at the same timepoints as TAICS. Items employ 5-point Likert scales and require approximately 20-25 minutes for completion. Validation research demonstrates strong internal consistency (α = 0.806-0.945) and clear factor structure through confirmatory factor analysis. Combined TAICS and AI-TPACK administration provides complementary perspectives on teacher competency: TAICS captures self-efficacy beliefs while AI-TPACK assesses perceived knowledge integration.

Classroom Observation Protocol

To complement self-report measures, trained observers will conduct classroom observations at 6-month and 12-month follow-ups using a structured protocol developed specifically for this study. The AI Integration Classroom Observation Protocol (AI-ICOP) assesses: (1) frequency and nature of AI tool usage, (2) pedagogical strategies for scaffolding student AI interaction, (3) attention to ethical considerations during AI activities, (4) quality of teacher modeling of critical AI evaluation, (5) differentiation strategies leveraging AI capabilities, and (6) student engagement patterns during AI-enhanced activities. Each observation spans one full class period with detailed field notes and structured ratings across domains. Inter-rater reliability will be established through double coding of 20% of observations, with target agreement $\kappa > 0.75$.

Implementation Logs

Teachers maintain electronic implementation logs documenting AI integration activities throughout the academic year. Logs capture lesson descriptions, AI tools used, student learning objectives, implementation challenges, and perceived student responses. Monthly log submissions provide granular data on implementation frequency, breadth, and evolution over time. Automated analytics identify patterns in tool selection, subject integration, and problem-solving strategies.

Qualitative Interviews

Semi-structured interviews with purposive teacher sub-samples (N=120; 30 per model condition plus 30 control) explore implementation experiences, perceived barriers and facilitators, changes in pedagogical thinking, and suggestions for program improvement. Interviews occur at post-intervention and 12-month follow-up, each lasting 45-60 minutes. Interview protocols probe: PD experience quality, applicability to classroom contexts, support adequacy, student response perceptions, evolution in AI understanding, and integration sustainability intentions. Interviews are audio-recorded, professionally transcribed, and analyzed using thematic analysis approaches.

4.3.3 Data Collection: Student Measures

Student AI literacy outcomes will be assessed through multiple measures administered to students in participating in teachers' classes. Target student sample: 15,000-25,000 students across grade levels and districts.

AI Literacy Knowledge Assessment

The study will adapt and validate age-appropriate AI literacy assessments measuring student knowledge across four domains aligned with the OECD AI Literacy Framework: (1) engaging with AI recognizing AI presence and critically evaluating outputs, (2) creating with AI using AI tools for problem-solving and creative work, (3) managing AI understanding appropriate task delegation and oversight, and (4) understanding AI systems basic comprehension of how AI works. Assessments employ scenario-based items, performance tasks, and selected response questions. Development occurs in Year 1 with pilot validation, followed by administration at baseline, mid-year, and end-of-year timepoints. Grade-level versions ensure developmentally appropriate assessment (elementary, middle, high school bands).

AI Attitudes and Self-Efficacy Survey

Students complete brief surveys (10-15 items) assessing attitudes toward AI, confidence in using AI tools, and beliefs about AI's societal role. Items draw from validated youth technology attitude instruments adapted for AI context. Surveys administered at baseline, mid-year, and end-of-year capture changes in affective dimensions of AI literacy.

Performance-Based Assessments

Subset of students (N=1,500-2,000) complete performance-based tasks requiring authentic AI interaction. Tasks present realistic problems where students must: select appropriate AI tools, craft effective prompts, critically evaluate AI outputs, and integrate AI-generated content into final products.

 ISSN: 3107-6513

Performance is evaluated using analytic rubrics assessing technical proficiency, critical thinking, ethical awareness, and creative integration. Tasks are administered at end-of-year to capture cumulative effects of year-long AI integration.

4.3.4 Data Collection: Contextual and Implementation Measures

Understanding contextual factors and implementation quality is essential for interpreting outcomes and identifying conditions supporting successful AI integration.

School Climate and Support Surveys

Teachers and administrators' complete surveys measuring organizational support for innovation, technology infrastructure quality, administrative backing for AI initiatives, collegial collaboration opportunities, and resource adequacy. Surveys administered at baseline and end-of-year identify contextual factors potentially moderating intervention effects.

Implementation of Fidelity Measures

For each PD model, implementation fidelity is tracked through professional development attendance records, completion rates for asynchronous modules, participation quality in synchronous sessions (rated by facilitators), and engagement with ongoing support resources. These measures enable dose-response analyses examining whether higher-fidelity participation produces stronger outcomes.

Technology Access Documentation

Districts provide data on technology infrastructure including student-device ratios, internet bandwidth, AI tool accessibility, and technical support availability. This information enables analysis of how infrastructure constraints influence implementation and outcomes.

4.4 Phase 3: Workforce Impact Analysis (Years 4-5)

The third phase extends analysis beyond school walls to examine long-term workforce and economic outcomes. This phase tracks students graduating from participating districts into post-secondary education and employment, analyzing relationships between district AI education quality and individual career trajectories.

4.4.1 Individual-Level Workforce Tracking

With appropriate IRB approvals and data use agreements, the study will establish data linkages connecting student records to post-secondary enrollment databases (National Student Clearinghouse) and workforce data systems (state unemployment insurance wage records). These linkages enable tracking of:

- **Post-Secondary Enrollment:** Rates, institution types, persistence, credential completion, STEM/technology field concentration
- **Employment Outcomes:** Employment rates, wages, industry sectors, technology-sector employment, job stability
- **Career Progression:** Wage growth trajectories, occupational mobility, employer characteristics

Analysis will compare outcomes for students experiencing high-quality AI education (intervention districts with strong implementation) versus comparison groups, controlling baseline demographics, academic achievement, and contextual factors. Difference-in-differences designs and propensity score matching will strengthen causal inference.

4.4.2 Regional Economic Impact Analysis

Regional economic effects will be analyzed using publicly available economic data aggregated at county/metropolitan statistical area levels. Analyses examine relationships between district AI education implementation quality (measured through teacher competency gains, implementation breadth, and student outcome improvements) and regional economic indicators including:

- **Technology Sector Growth:** Employment in technology industries, new technology business formations, venture capital investment
- **Productivity Indicators:** GDP per capita, labor productivity measures, patent applications
- **Workforce Quality:** High-skill job concentration, wage levels, employer satisfaction surveys
- **Educational Attainment:** Post-secondary enrollment rates, STEM credential completion, continuing education participation

Time-lagged regression analyses will test whether improvements in district AI education (t) predict subsequent regional economic improvements (t+2, t+3), controlling for baseline economic conditions, demographic trends, and concurrent investments in economic development. While establishing definitive causality remains challenging given complex economic systems, convergent evidence across multiple indicators strengthens inferences about AI education's economic contributions.

4.4.3 Employer Partnerships and Workforce Validation

The study will establish partnerships with regional employers in technology and technology-adjacent sectors to validate workforce readiness assessments. Partner employers will:

- Participate in validation studies rating entry-level employee AI readiness
- Complete surveys assessing new hire preparedness for AI-integrated workplaces
- Provide performance ratings enabling correlation with educational background
- Engage in focus groups discussing workforce skill needs and educational alignment

These partnerships ground workforce readiness measures in authentic employment requirements, ensuring educational outcomes align with real-world demands. Employer validation strengthens research relevance for policymakers and educational leaders seeking to demonstrate return on investment.

4.5 Data Analysis Plan

4.5.1 Quantitative Analysis Approaches

The study will employ multiple quantitative analytical strategies appropriate to the research questions and data structure:

Hierarchical Linear Modeling (HLM)

Given the nested data structure (students within classrooms within schools within districts), HLM analyses will model outcomes while accounting for clustering effects. Three-level models will examine student outcomes (Level 1) predicted by teacher characteristics and classroom practices (Level 2) and school/district contextual factors (Level 3). HLM enables partitioning of variance across levels, identifying the proportion of outcome variation attributable to teacher versus

ISSN: **3107-6513**

school factors. Cross-level interactions will test whether teacher PD effects vary systematically across different school contexts.

Growth Curve Modeling

Longitudinal teacher and student data will be analyzed using growth curve models examining trajectories of change over time. Individual growth parameters (interceptions, linear slopes, quadratic terms) become outcomes predicted by intervention condition, individual characteristics, and contextual factors. Growth models address key questions including: How do teacher competency trajectories differ across PD models? At what rate do students develop AI literacy under different conditions? Do initial competency levels moderate growth rates?

Mediation and Moderation Analysis

Structural equation modeling (SEM) will test hypothesized causal pathways linking PD participation → teacher competency gains → classroom implementation quality → student outcomes. Mediation analyses identify mechanisms through which PD affects student learning, quantifying direct and indirect effects. Moderation analyses tests whether PD effects differ systematically across teacher characteristics (years of experience, subject area, initial technology comfort) or school contexts (resource levels, administrative support, school culture), informing targeting and differentiation of future PD efforts.

Propensity Score Methods

For workforce analyses where random assignments are impossible, propensity score matching will construct comparable comparison groups. Propensity scores predicted probability of receiving intervention based on observed baseline characteristics enable matching of intervention and comparison cases, reducing selection bias. Analyses compare matched samples on workforce outcomes, with sensitivity analyses examining robustness to unmeasured confounding.

Cost-Effectiveness Analysis

Economic analyses will calculate cost per teacher trained, cost per standard deviation gain in teacher competency, and cost per standard deviation gain in student outcomes for each PD model. Analyses incorporate direct costs (materials, facilitator time, platform fees) and indirect costs (teacher time, substitute coverage). Cost-effectiveness ratios enable comparison across models, identifying most efficient approaches for resource-constrained districts. Sensitivity analyses examine how cost-effectiveness changes under varying assumptions about scale, implementation efficiency, and effect sustainability.

4.5.2 Qualitative Analysis Approaches

Qualitative data from teacher interviews, classroom observations, and implementation artifacts will undergo systematic thematic analysis following established protocols:

Initial Coding and Categorization

Research team members will independently code interview transcripts and observation notes using both deductive codes derived from theoretical frameworks (TAICS dimensions, AI-TPACK components) and inductive codes emerging from data. Iterative comparison and discussion will refine codebook, establishing clear definitions and decision rules. NVivo qualitative analysis software will facilitate systematic coding and retrieval.

Theme Development and Pattern Analysis

Coded data will be analyzed to identify recurring themes, patterns of variation, and explanatory relationships. Comparative analysis across cases will identify factors associated with successful versus struggling implementation. Pattern-matching will test whether qualitative findings align with quantitative results, triangulating evidence. Negative case analysis will deliberately seek disconfirming evidence, strengthening conclusions through critical examination.

Case Study Development

In-depth case studies will be developed for selected sites representing diverse implementation experiences: exemplary implementation achieving strong outcomes, typical implementation with moderate success, and struggling implementation facing significant challenges. Each case study integrates multiple data sources (teacher surveys, interviews, observations, student outcomes, contextual measures) into rich narrative accounts illuminating implementation dynamics. Cross-case analysis identifies common themes and contextual contingencies.

4.5.3 Integration of Quantitative and Qualitative Findings

The mixed-methods design enables sophisticated integration of quantitative and qualitative evidence. Several integration strategies will be employed:

- **Sequential Explanatory Design:** Qualitative data collection following quantitative analysis will deliberately investigate unexpected or puzzling quantitative findings, providing explanatory depth
- **Convergent Design:** Independent quantitative and qualitative analyses will be compared, examining where findings converge (increasing confidence) or diverge (prompting deeper investigation)
- **Joint Displays:** Visual representations will integrate quantitative results and qualitative themes, facilitating pattern recognition across data types
- **Meta-Inferences:** Final synthesis will draw meta-inferences conclusions supported by integration of quantitative and qualitative evidence providing holistic understanding transcending either data type alone.

SIGNIFICANCE AND EXPECTED CONTRIBUTIONS

5.1 Theoretical Contributions

This research advances theoretical understanding of teacher professional learning in several ways. First, it extends TPACK and related frameworks into the AI domain, empirically testing whether established technology integration models adequately capture AI-specific competencies or whether AI's unique characteristics necessitate fundamentally different conceptual frameworks. By simultaneously measuring TAICS and AI-TPACK and examining their interrelationships, the research illuminates which competency dimensions most strongly predict successful classroom implementation and student outcomes.

Second, the study contributes to professional development effectiveness research by rigorously comparing distinct PD models representing different theoretical traditions: self-regulated learning and adaptive instruction (AI-personalized model), communities of practice and collaborative knowledge construction (cohort model), and integrated approaches attempting to leverage multiple learning mechanisms (hybrid model). This comparative approach moves beyond single-model evaluation toward theoretical understanding of which

PD design principles most effectively support teacher learning for complex, rapidly evolving technologies.

Third, by examining mediating pathways from teacher PD through teacher competency and classroom implementation to student outcomes, the research explicates mechanisms through which teacher learning translates into student benefit. This mechanistic understanding informs both theory and practice, identifying leverage points for maximizing PD impact. The research tests whether the traditional assumption that teacher knowledge gains automatically produce improved student outcomes holds in the AI education context or whether additional supports are necessary for knowledge-to-practice transfer.

5.2 Methodological Contributions

Methodologically, the research demonstrates rigorous longitudinal evaluation of education-to-workforce pipelines, establishing data linkages and analytical approaches that remain rare in education research. By tracking students from K-12 classrooms through post-secondary education into employment, the study provides empirical evidence on long-term returns to educational investments evidence essential for policymakers but frequently unavailable due to methodological and logistical challenges. The research establishes replicable protocols for conducting such analyses, potentially catalyzing similar research in other educational domains.

The mixed-methods design demonstrates sophisticated integration of quantitative and qualitative evidence, moving beyond parallel analyses toward genuine integration producing insights unattainable through either methodology alone. The research provides concrete examples of how qualitative findings can explain quantitative patterns, how quantitative analyses can test qualitatively generated hypotheses, and how integrated evidence produces more robust and nuanced conclusions.

5.3 Practical and Policy Contributions

For practitioners and policymakers, the research delivers immediately actionable guidance. The validated PD models, complete with implementation materials and protocols, can be directly adopted by districts seeking to develop teacher AI capacity. Implementation toolkits will include detailed facilitation guides, content modules, assessment instruments, coaching protocols, and technology platform specifications. These resources dramatically reduce barriers to high-quality PD implementation.

The research provides evidence-based recommendations for federal and state policy regarding AI education investments. With clear data on cost-effectiveness, effect sizes across diverse contexts, and long-term workforce impacts, policymakers can make informed allocation decisions. The research directly addresses frequent policy questions: Which PD approaches produce best returns? How much investment is needed to produce meaningful outcomes? How long until effects appear? Which contextual factors require attention?

For school and district leaders, the research identifies implementation conditions supporting success. By analyzing mediating and moderating factors, the study illuminates how administrative actions, resource allocations, and organizational culture influence PD effectiveness. Leaders receive practical guidance on creating conditions where teacher AI learning flourishes and translates into classroom impact.

5.4 Economic and Workforce Development Contributions

Perhaps most significantly, the research contributes to national economic competitiveness by establishing evidence-based pathways for developing an AI-literate workforce. Current workforce analyses indicate that the United States faces critical skills gaps threatening economic leadership in AI-driven industries. By demonstrating how systematic K-12 education interventions can address these gaps, the research provides strategic direction for workforce development investments.

The longitudinal workforce tracking data quantifies economic returns to AI education investments, enabling cost-benefit analyses essential for policy decisions. If the research demonstrates that high-quality AI education produces measurable increases in technology sector employment, higher wages, and enhanced productivity, it provides compelling justification for substantial public investment. Conversely, if effects prove modest or context-dependent, the research identifies necessary conditions or complementary interventions required for maximum impact.

The employer partnership component grounds workforce readiness assessments in authentic industry needs, ensuring educational efforts align with employment realities. By engaging employers as research partners, the study facilitates ongoing dialogue between education and industry sectors, potentially catalyzing broader collaboration on workforce preparation.

References

Acemoglu, D., & Restrepo, P. (2019). Automation and new tasks: How technology displaces and reinstates labor. *Journal of Economic Perspectives, 33*(2), 3–30. https://doi.org/10.1257/jep.33.2.3

Al-Abdullatif, A. M. (2024). Developing the Intelligent-TPACK framework: A structural equation modeling approach. *Education and Information Technologies*. Advance online publication. https://doi.org/10.1007/s10639-024-12547-w

Asa, R., Nautwima, J. P., & Johannes, H. N. (2025). Digital literacy for workforce readiness: Bridging the skills gap in the 21st century. *International Journal of Innovation and Economic Development, 11*(3), 1–15. https://doi.org/10.18775/ijied.1849-7551-7020.2015.311.2001

Carolus, A., Wienrich, C., & Augustin, Y. (2023). The meta AI literacy scale (MAILS): Development and testing of an AI literacy questionnaire based on well-founded conceptual models and psychological theory. *Computers in Human Behavior, 137*, Article 107898. https://doi.org/10.1016/j.chb.2022.107898

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

Chen, X., Wang, Y., & Zhang, L. (2024). Enhancing teacher AI literacy and integration through different types of cases in teacher professional development. *Computers and Education: Artificial Intelligence, 6*, Article 100219. https://doi.org/10.1016/j.caeai.2024.100219

Chiu, T. K., Ahmad, Z., & Çoban, M. (2024). Development and validation of teacher artificial intelligence (AI) competence self-efficacy (TAICS) scale. *Education and Information Technologies, 30*, 6667–6685. https://doi.org/10.1007/s10639-024-13094-z

Copur-Gencturk, Y., Li, J., & Atabas, S. (2024). Improving teaching at scale: Can AI be incorporated into professional development to create interactive, personalized learning for teachers? *American Educational Research Journal, 61*(5), 789–821. https://doi.org/10.3102/00028312241248514

EdCircuit. (2025, October 2). *Teacher professional development in the age of AI*. https://edcircuit.com/teacher-professional-development-in-the-age-of-ai/

Education Sciences. (2025). Developing and validating an AI-TPACK assessment framework: Enhancing teacher educators' professional practice through authentic artifacts. *Education Sciences, 15*(11), 1452. https://doi.org/10.3390/educsci15111452

EdWeek Research Center. (2025). *5 tips on AI professional development for teachers*. Education Week. https://www.edweek.org/leadership/5-tips-on-ai-professional-development-for-teachers/2025/01

Falloon, G. (2020). From digital literacy to digital competence: The teacher digital competency (TDC) framework. *Educational Technology Research and Development, 68*(5), 2449–2472. https://doi.org/10.1007/s11423-020-09767-4

Frontiers in AI. (2023). Integrating AI tools in teacher professional learning: A conceptual model and illustrative case. *Frontiers in Artificial Intelligence, 6*, Article 1255089. https://doi.org/10.3389/frai.2023.1255089

Frontiers in AI. (2025). Artificial intelligence skills and their impact on the employability of university graduates. *Frontiers in Artificial Intelligence, 8*, Article 1629320. https://doi.org/10.3389/frai.2025.1629320

Harvard Business School. (2025). *Displacement or complementarity? The labor market effects of generative AI* (HBS Working Paper No. 25-039). https://www.hbs.edu/ris/Publication%20Files/25-039_05fbec84-1f23-459b-8410-e3cd7ab6c88a.pdf

Kong, S. C., Cheung, W. M., & Zhang, G. (2021). Evaluation of an artificial intelligence literacy course for university students with diverse study backgrounds. *Computers and Education: Artificial Intelligence, 2*, Article 100026. https://doi.org/10.1016/j.caeai.2021.100026

Li, Y., Wang, Z., & Chen, X. (2024). Generative artificial intelligence literacy: Scale development and its effect on job performance. *Behavioral Sciences, 14*(12), 1216. https://doi.org/10.3390/bs14121216

McKinsey Global Institute. (2023). *The economic potential of generative AI: The next productivity frontier*. McKinsey & Company. https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/the-economic-potential-of-generative-ai-the-next-productivity-frontier

Mishra, P., & Koehler, M. J. (2006). Technological pedagogical content knowledge: A framework for teacher knowledge. *Teachers College Record, 108*(6), 1017–1054. https://doi.org/10.1111/j.1467-9620.2006.00684.x

New America. (2025). *Digital literacy in the age of AI: Voices from the field*. New America Foundation. https://www.newamerica.org/education-policy/briefs/foundational-skills-digital-literacy-in-the-age-of-ai-analysis-and-voices-from-the-field/

OECD. (2025). *AI literacy framework for students (Draft)*. Organisation for Economic Co-operation and Development. https://www.oecd.org/education/ai-literacy-framework-students/

Pinski, M., & Benlian, A. (2023). AI literacy—Towards measuring human competency in artificial intelligence. *Proceedings of the 56th Hawaii International Conference on System Sciences*, 5880–5889. https://hdl.handle.net/10125/103393

Smokorowski, D., & Reljac, M. C. (2025). *Panel discussion: Incorporating AI training for teachers*. Education Week K-12 Essentials Forum. https://www.edweek.org/leadership/5-tips-on-ai-professional-development-for-teachers/2025/01

Tadimalla, S. (2025). AI literacy as a core component of AI education. *AI Magazine, 46*(3), e202523. https://doi.org/10.1002/aaai.70007

Wang, X., Zhang, Y., & Liu, Q. (2024). Artificial intelligence in teaching and teacher professional development: A systematic review. *Computers and Education Open, 7*, Article 100189. https://doi.org/10.1016/j.caeo.2024.100189

Weber, M., Pinski, M., & Baum, K. (2023). Assessing AI literacy: Design and evaluation of a competency-based test. *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, 1–17. https://doi.org/10.1145/3544548.3580681

World Economic Forum. (2025). *Future of Jobs Report 2025*. https://www.weforum.org/publications/future-of-jobs-report-2025/

Xu, G., Yu, A., Gao, A., & Trainin, G. (2025). Developing an AI-TPACK framework: Exploring the mediating role of AI attitudes in pre-service TCSL teachers' self-efficacy and AI-TPACK. *Education and Information Technologies, 30*(15), 22471–22495. https://doi.org/10.1007/s10639-025-13630-5