

University of Tartu
Faculty of Social Sciences
Institute of Education
MA in Educational Technology programme

Kendrah Walker

**INTERPRETING AND ENACTING AI PROFESSIONAL DEVELOPMENT: AN
AUTOETHNOGRAPHIC STUDY OF TEACHER EXPERIENCES OF A
MANDATORY PROGRAM**

Master's thesis

Supervisor: Dr Emanuele Bardone

Tartu 2026

Abstract

Schools are adopting artificial intelligence (AI) professional development quickly, often through external vendors and with limited teacher input. This autoethnographic study examines how instructional staff at an American-curriculum preparatory school in Latin America interpreted and enacted a mandatory, 14-month AI professional development program. Data comprised the researcher's observation logs, seven interviews, a questionnaire ($n = 33$), and a program communications log, analyzed through reflexive thematic analysis. Results show teachers interpreted the program through a pervasive coherence failure and enacted it largely as structural compliance with cognitive withdrawal, while outcomes varied sharply by role. The study argues that teachers enacted, rather than implemented, the program, producing a distribution of strategic responses shaped by interpretive coherence, cognitive load, institutional voice, and role alignment.

Keywords: artificial intelligence, professional development, teacher experience, autoethnography, teacher agency

Contents

Abstract	2
1. Introduction	6
2. Theoretical Overview	8
2.1 The Role of AI in Education: Opportunities, Risks, and Policy Responses	8
2.2 Teacher Professional Learning for Digital and AI Integration.....	9
2.3 Characteristics of Effective Teacher Professional Development	10
2.4 AI Literacy, Teacher Agency, and Teacher Professional Identity	11
2.5 Conceptual Confusion and Implementation Overload	12
2.6 Autoethnography as a Lens for Teacher Experience.....	13
2.7 Gaps in the Literature and Contributions of the Present Study	14
3. Method	15
3.1 Research Design.....	15
3.2 Research Context and Sample.....	17
3.3 Data Collection.....	19
3.3.1 Autoethnographic Journal and Reflection.....	19
3.3.2 PD Session Observation Logs	19
3.3.3 Questionnaire	20
3.3.4 Semi-Structured Interviews.....	20
3.3.5 Program Communications Log	21
3.3.6 Additional Documentary and Artifactual Evidence	21
3.4 Data Analysis	22
3.5 Trustworthiness and Quality	23
3.6 Ethical Considerations.....	24
3.6.1 Disclosure of Generative AI Use	25
4. Results and Discussion.....	25
4.1 Chapter Overview	25

4.2 Status of the Research Questions	26
4.3 Sample Summary	26
4.4 Sub-RQ 1: How Conceptual Confusion Shaped Teachers' Interpretation	27
4.4.1 Coherence Failure: "Not Designed With Us in Mind"	27
4.4.2 Deflected Questions and the Breakdown of Trust	29
4.5 Sub-RQ 2: How Implementation Overload Shaped Teachers' Enactment.....	30
4.5.1 Disengagement-Within-Attendance	30
4.5.2 A Refinement: Load Was Not Uniform	32
4.5.3 Format and Cognitive Load.....	33
4.5.4 Productive Outcomes Despite Misalignment	34
4.6 Sub-RQ 3: Institutional and Identity Context	35
4.6.1 Voice and Choice Absence.....	35
4.6.2 Vendor Opacity and Institutional Communication Failure	37
4.6.3 Facilitator Quality as a Mediating Variable	38
4.6.4 Professional Identity Effects	39
5. Conclusion.....	40
5.1 Synthesis: Interpretation and Enactment as Mutually Constituting	40
5.2 Limitations	41
5.3 Applications and Implications.....	43
5.4 Recommendations for Further Research	44
Acknowledgments	46
Author's Declaration.....	46
References	47
Appendices	50
Appendix 1. PD Session Observation Log Template.....	50
Appendix 2. PD Observation Reference Card.....	57
Appendix 3. Interview Protocol	58

Appendix 4. Participant Online Questionnaire	64
Non-exclusive licence to reproduce the thesis and make the thesis public.....	71

1. Introduction

Artificial intelligence (AI) technologies have entered educational settings quickly, presenting opportunities and challenges for teachers and instructional staff. From adaptive tutoring platforms to generative text tools, AI applications have become more common in classrooms, changing expectations for instruction, assessment, and curriculum design (Holmes & Tuomi, 2022; Hwang & Chen, 2023). In response, schools and educational systems have moved quickly, and often without clear structure, to prepare teaching staff through professional learning initiatives aimed at building AI competencies. Research suggests that many teachers find the volume, pace, or conceptual complexity of these programs difficult to manage, often with little input into the design, delivery, or relevance of the professional learning they receive (Heine & König, 2025; Roshan et al., 2024).

Institutional pressure for AI integration, combined with shifting definitions of effective pedagogical practice, has produced two related conditions that form the analytical core of this study: conceptual confusion and implementation overload. Conceptual confusion refers to the ambiguity that arises when key terms, frameworks, and expectations shift, leaving practitioners uncertain about what they are being asked to learn, why, and how it connects to their existing knowledge and practice (Fullan, 2007). In a professional development setting, conceptual confusion appears as uncertainty about what a tool is for, how a recommended practice differs from current practice, or why a given activity matters for one's own teaching. Implementation overload describes the cumulative burden placed on teachers when multiple new initiatives, technologies, and policy mandates are introduced at once or in close succession without adequate support, time, or resources for integration (Fullan & Quinn, 2016). It appears as the expectation to adopt several tools or frameworks within a short period while maintaining existing teaching responsibilities. These two conditions are distinct: the first concerns how teachers make sense of what they encounter, and the second concerns what teachers are able to do with it. This study treats them as operating on two different aspects of the professional development experience—interpretation and enactment, respectively.

This study examines how teachers at an American-curriculum preparatory school in Latin America interpreted and enacted a mandatory AI professional development program delivered by an external vendor between January 2025 and March 2026. The distinction between interpretation and enactment is deliberate. Interpretation refers to how teachers understood the program's content—what they took it to mean and how they connected it, or failed to connect it, to their own knowledge and context. Enactment refers to what teachers actually did in

response—how they participated, what they adopted or set aside, and how they carried the program into or kept it out of their practice. Drawing on ecological accounts of teacher agency, which hold that teachers do not simply implement mandates but achieve a response to them through the interaction of their values, histories, and working conditions (Biesta et al., 2015; Priestley et al., 2015), the study treats teachers' responses as enactments rather than as degrees of compliance with an intended design.

Through reflexive narrative, supplemented by an anonymous staff questionnaire, voluntary participant interviews, and a documentary record of program communications, the study investigates how conceptual confusion shaped teachers' interpretation of the program and how implementation overload shaped their enactment of it, within the institutional and identity-related conditions that contextualized both. The following research questions guide the investigation.

The primary research question is: How do teachers at an American preparatory school interpret and enact mandatory, externally delivered AI professional development?

Three sub-questions structure the inquiry:

1. How does conceptual confusion shape teachers' interpretation of AI professional development content?
2. How does implementation overload shape teachers' enactment of AI professional development in everyday practice?
3. How do institutional culture, agency, and professional identity contextualize that interpretation and enactment?

The significance of this study lies in its contribution to a developing literature on how teachers experience externally delivered AI professional development, and on how they do so in a regional context that has received limited research attention. Much existing research on teacher training and AI integration has been conducted in well-resourced educational systems in North America, Europe, and East Asia (Holmes & Tuomi, 2022). Less is known about how teachers in Latin American school settings respond to the combined demands of institutional compliance and professional growth when working with unfamiliar technologies and limited information about the training itself. The school examined here occupies a position common to many internationally oriented private schools in the region: it adopts curricula and technologies developed elsewhere and employs a staff that includes both local citizens and expatriates. By documenting the experiences of the researcher and fellow staff members in this setting, the study extends the geographic scope of research on AI-related professional

development and examines the tensions, negotiations, and adaptive strategies that arise when professional learning does not account for teachers' existing expertise, context, and pedagogical commitments.

The following sections present a theoretical overview situating this research within the scholarship on AI in education, teacher professional development, professional identity, and teacher agency, and concluding with the gaps the study addresses.

2. Theoretical Overview

2.1 The Role of AI in Education: Opportunities, Risks, and Policy Responses

Artificial intelligence has been part of formal education for decades, though its presence has grown over the last several years. Early work in the 1960s and 1970s produced intelligent tutoring systems focused on diagnosing errors and providing step-by-step feedback. Recent developments in generative AI, particularly large language models capable of producing text, code, and multimodal content, have moved AI from a specialized instructional tool to a broader set of applications influencing assessment, administrative workflow, accessibility, and content creation (Holmes & Tuomi, 2022; Hwang & Chen, 2023). Rather than extending existing educational technologies, these systems introduce new ways of analyzing learner data and automating decisions that previously required professional judgment.

International policy documents reflect this shift. The U.S. Department of Education (2023) framed AI as a form of automation based on associations, emphasizing two transitions relevant to schooling: the shift from collecting information to detecting patterns within it, and the move from resource access to partially automated pedagogical decision-making. This framing identifies a central tension educators face: AI tools increasingly influence decisions that teachers once made themselves. Alongside potential benefits in personalization and reduced administrative load, the report identified unresolved challenges, including data governance, bias, the accuracy of AI outputs, and the risk of over-reliance on automation.

Policy discourse consistently emphasizes keeping educators involved at every stage of AI adoption, positioning teachers as both guides for implementation and shapers of expectations for how AI supports learning. This priority is relevant to the present study, in which participants were required to complete training that was designed and selected externally, with input from the administrative leadership team but with limited transparency about content, the evidence base, or the organizational rationale. The concerns raised in the international literature—vendor influence, limited empirical validation, and insufficient communication—

correspond to conditions reported by staff in the school examined here (Holmes & Tuomi, 2022). In addition, the school did not communicate a clear AI policy statement to instructional staff, which left uncertainty about acceptable use, ethical application, and alignment with institutional expectations. These policy-level tensions set the context for examining how professional development structures have, and have not, prepared teachers to address them.

2.2 Teacher Professional Learning for Digital and AI Integration

Research on professional development emphasizes that it must move beyond single workshops toward models that are sustained, collaborative, and connected to teachers' daily instructional work, particularly when new or unfamiliar technologies are involved. A review by Heine and König (2025), synthesizing research published between 2020 and 2024, found that programs grounded in teachers' existing knowledge, supported over time, and incorporating hands-on experimentation tended to produce improvements in confidence and instructional integration. Their analysis also identified persistent structural barriers: many educators reported lacking time, resources, or adequate training opportunities. Heine and König further noted that compressed or externally standardized training formats—formats in which content is delivered uniformly across contexts rather than adapted to local conditions—frequently failed to produce the intended outcomes. This concern corresponds to the format of the mandatory AI training in the present study, though whether the conditions are comparable is a question the present analysis examines rather than assumes.

The gap between institutional expectations and teacher readiness is not only a training problem; it reflects a structural mismatch between professional development design and classroom realities. Roshan et al. (2024) documented this mismatch in a survey of educators, finding limited familiarity with AI tools and a smaller proportion of teachers who felt confident applying them instructionally. They argued that many schools promote AI integration without establishing professional learning systems capable of sustaining long-term, practice-oriented learning, and recommended thorough, iterative models that combine technical understanding with pedagogical application.

The Technological Pedagogical Content Knowledge (TPACK) framework (Koehler & Mishra, 2009) offers one way to interpret why these gaps matter. Because effective technology use requires combining content knowledge, teaching methods, and technological affordances, professional development focused narrowly on technical skills is unlikely to prepare teachers for actual implementation. This disconnect is particularly relevant to AI-

related professional learning, where the novelty and perceived complexity of the technology can overshadow the pedagogical reasoning that should guide its use.

2.3 Characteristics of Effective Teacher Professional Development

Because professional development quality is central to this study's research questions, it is useful to establish what the research literature identifies as effective professional learning. The most widely cited framework comes from Darling-Hammond et al. (2017), whose review of 35 methodologically rigorous studies identified seven characteristics of professional development linked to improvements in teaching practice and student outcomes. Effective professional development, they found, is content focused, incorporates active learning grounded in adult learning theory, supports collaboration in job-embedded contexts, uses models and modeling of effective practice, provides coaching and expert support, offers opportunities for feedback and reflection, and is of sustained duration.

These characteristics offer a useful reference point against which the design and delivery of a professional development program can be considered. Several reviews reinforce them.

Desimone (2009) emphasized content focus, active learning, coherence, sustained duration, and collective participation as core features of professional development that produces measurable change in teacher knowledge, beliefs, and practice. Timperley et al. (2007) reported comparable findings in their best-evidence synthesis, noting that effective professional learning engages teachers in examining the relationship between their instructional decisions and student outcomes. More recently, Sims et al. (2025) proposed a mechanism-based framework identifying instruction, practical social support, modeling, feedback, and rehearsal as the techniques through which practice change occurs.

These frameworks matter to the present study in two ways. First, they provide a lens for examining the AI professional development program under investigation, which was characterized by mixed cohorts, compressed timelines, externally selected content, limited reflection opportunities, and minimal coaching or expert follow-up. Considered in relation to these criteria, several of the program's structural features appear to differ from the conditions the literature associates with meaningful teacher growth—a relationship the present study examines rather than assumes. Second, the frameworks help explain why such misalignment matters: professional development that lacks content relevance, active engagement, collaborative support, and sustained duration is less likely to produce durable learning and more likely to generate the conditions—conceptual confusion and implementation overload—

that this study examines. Desimone's (2009) five features were used to structure the observation instrument described in the Method chapter.

2.4 AI Literacy, Teacher Agency, and Teacher Professional Identity

Recent literature on AI in education has moved beyond technical training to examine the professional and identity-related implications of working with AI tools. Tripathi et al. (2025), drawing on AI literacy frameworks (Long & Magerko, 2020) and ecological accounts of teacher agency (Biesta et al., 2015; Priestley et al., 2015), found that teachers often hold inconsistent or ambivalent beliefs about AI. Their participants recognized AI's instructional potential while also expressing concern about student dependence and the possible erosion of teacher expertise. This tension reflects broader questions about authority, purpose, and professional identity that arise when automation enters pedagogical spaces.

Tripathi et al., referencing Long and Magerko (2020), also observed uneven development across four components of AI literacy: what teachers know, what skills they possess, how confidently they assess AI systems, and how they reason about ethical considerations.

Teachers tended to acquire operational skills—how to use AI for basic tasks—more quickly than skills related to critical evaluation or ethical reasoning. This imbalance appears across studies and is relevant for professional development programs that prioritize tool operation over engagement with implications, risks, and the reasoning behind AI outputs (Holmes & Tuomi, 2022; U.S. Department of Education, 2023).

The concept of teacher agency helps explain these dynamics. Ecological agency theory positions agency not as a fixed personal attribute but as something teachers achieve through the interaction of their values, professional histories, and working conditions (Biesta et al., 2015; Priestley et al., 2015). When educators respond to top-down mandates, particularly those involving complex technologies, their opportunities to act intentionally may narrow. In such conditions, teachers may comply without integrating new practices, or may disengage. This account of agency also clarifies the study's use of the term enactment: what teachers do with a mandate is an achieved response shaped by conditions, not a simple execution of an external plan. These patterns correspond to responses documented in the present study, where the structure of the AI professional development program shaped the kinds of learning staff felt able, or unable, to pursue.

Teacher identity research adds further detail and is relevant given the autoethnographic nature of this study. Beauchamp and Thomas (2009) describe professional identity as an ongoing

process of interpretation and reinterpretation shaped by institutional expectations, professional interactions, and teachers' evolving self-understandings. Kelchtermans (2009) extends this view by emphasizing the vulnerability that accompanies identity work, arguing that teachers' sense of professional self is closely tied to how they teach and what they value. When new institutional initiatives, particularly those involving unfamiliar technologies, conflict with teachers' existing self-understandings, teachers may experience a destabilization of professional identity that appears as resistance, withdrawal, or dissonance. Because autoethnography is concerned with the intersection of personal experience and cultural context, it is suited to surfacing these identity-based dimensions of professional learning. The tension between institutional demands for AI adoption and teachers' established pedagogical commitments is the kind of identity negotiation that autoethnographic inquiry can document.

2.5 Conceptual Confusion and Implementation Overload

The two conditions that anchor this study were defined in the introduction: conceptual confusion as the ambiguity arising when terms, frameworks, and expectations shift (Fullan, 2007), and implementation overload as the cumulative burden of multiple initiatives introduced without adequate support (Fullan & Quinn, 2016). Both are documented in the literature on educational reform and technology integration, though they have received less explicit attention in AI-related professional development. In AI contexts, conceptual confusion is compounded by the breadth and opacity of the technology: educators may struggle to distinguish machine learning, generative AI, and other forms of automation, and to judge the pedagogical soundness of specific tools (Holmes & Tuomi, 2022; Tripathi et al., 2025). Implementation overload, in turn, appears as the expectation to master an expanding set of tools while meeting existing teaching responsibilities, a demand intensified when training lacks coherence, transparency, or relevance. Research on educational change identifies reform overload as among the most significant barriers to sustainable improvement, contributing to shallow adoption and skepticism toward institutional initiatives (Hargreaves & Fullan, 2012). This study examines conceptual confusion as a condition affecting interpretation—how teachers understood what the program asked of them—and implementation overload as a condition affecting enactment—what teachers were able to do with it in practice.

The intersection of these two conditions has implications for how teachers process and retain new learning. Cognitive load theory (Sweller, 1988) provides the mechanism linking confusion and overload to impaired learning. When the volume and complexity of

information exceed a learner's processing capacity, learning becomes shallow and transfer to practice is reduced. Applied to professional development, the theory suggests that programs presenting too many unfamiliar concepts, tools, or frameworks in too short a period risk producing compliance without understanding—teachers who can perform prescribed tasks during sessions but cannot independently apply, adapt, or evaluate what they have learned in their own classrooms (Clark et al., 2012). This mechanism is relevant to the experiences documented in the present study, where the compressed format and breadth of the AI training may have exceeded participants' capacity for sustained engagement.

Several lines of evidence in the reviewed literature support these concerns. Time constraints and resource limitations are among the most significant barriers to AI integration that teachers report, and compressed or standardized training formats consistently fail to produce intended outcomes (Heine & König, 2025; Roshan et al., 2024). Tripathi et al. (2025) offered a more detailed view of how confusion and overload interact with teacher agency, showing that when educators lacked sufficient support to develop critical AI literacy, they were more likely to either reject AI tools or adopt them uncritically—neither of which constitutes the reflective, purposeful integration that effective professional development should support.

2.6 Autoethnography as a Lens for Teacher Experience

Autoethnography occupies a position at the intersection of autobiography and ethnography, allowing researchers to use their own lived experiences as a primary data source while situating those experiences within broader cultural, social, and institutional contexts (Ellis et al., 2011). In education, autoethnography has been used to examine the subjective dimensions of teaching—the emotions, tensions, identity negotiations, and commitments that shape how educators experience their professional lives (Chang, 2008; Pinnegar & Hamilton, 2009).

The choice of autoethnography for this study is deliberate. Approaches to evaluating professional development often rely on surveys, pre-post assessments, or observation protocols that measure changes in knowledge, skills, or behavior. While these methods produce useful data, they are less suited to capturing the experiential dimension of professional learning—the points of confusion, the adaptations, and the negotiations that unfold over time as teachers encounter and respond to new demands (Clandinin & Connelly, 2000). Autoethnography's emphasis on reflexive narrative allows the researcher to document not only what happened during training but how it was understood, what it meant, and how it interacted with existing beliefs, values, and commitments.

In this study, autoethnographic data are complemented by an anonymous staff questionnaire, voluntary participant interviews, and a documentary communications record. This design balances the depth of the researcher's first-person account with the range of perspectives offered by colleagues and with documentary evidence independent of any participant's recollection. Examining the researcher's experience alongside these other sources allows points of convergence and divergence to be identified, which strengthens the analysis (Denzin & Lincoln, 2018). The treatment of these multiple sources is described in the Method chapter, consistent with reflexive thematic analysis (Braun & Clarke, 2019, 2022).

Using autoethnography to examine teacher experiences with AI professional development is timely given the recent emergence of such programs. Because institutional AI integration efforts are relatively new, there is limited longitudinal or large-scale research on how teachers experience them from the inside. Autoethnography offers a means of generating contextualized accounts of a recent development that can inform future research designs, policy decisions, and professional development structures. By documenting the researcher's own uncertainties and moments of insight, the approach may also be useful to other educators navigating similar experiences (Adams et al., 2015).

2.7 Gaps in the Literature and Contributions of the Present Study

The literature reviewed above indicates several gaps this study addresses. First, externally selected, vendor-delivered AI training—in which teachers have limited knowledge of or input into the design and rationale of their professional development—has received relatively little scholarly attention, since most research assumes some institutional transparency about the purpose of training. This study examines how limited transparency and reduced teacher input interact with conceptual confusion and implementation overload to shape interpretation and enactment.

Second, much research on AI in education has focused on the design and evaluation of AI tools rather than on the experiences of the teachers expected to adopt them, and the teacher-focused studies that exist tend to rely on survey or interview methods that cannot capture the longitudinal, contextual, and personal dimensions autoethnography makes visible (Holmes & Tuomi, 2022; Tripathi et al., 2025).

Third, most empirical research on teacher AI professional development has been conducted in the Global North. Internationally oriented private schools in Latin America occupy a distinctive position, adopting curricula and technologies developed elsewhere while

employing staff with varied training backgrounds and professional cultures (UNESCO, 2023). This study extends the geographic scope of the field by documenting staff experiences in such a setting.

Finally, autoethnographies focused specifically on AI professional development remain rare; a search of ERIC, Education Source, and Google Scholar returned no published autoethnographic studies of teacher experiences with AI professional development as of early 2026. By combining autoethnographic reflexivity with questionnaire, interview, and documentary data, this study offers a multidimensional account of a professional challenge likely to grow as AI becomes more common in education.

3. Method

This chapter describes the research design, participant selection, data sources, collection procedures, and analytical approach used in this study. The study investigates how teachers at an American-curriculum preparatory school in Latin America interpreted and enacted a mandatory, externally delivered professional development program on artificial intelligence. Because the research examines how a mandated initiative was understood and carried out by the practitioners responsible for it, the design draws on analytic autoethnography supplemented by participant-generated and documentary data. Each subsection below explains the decisions made at each stage of the inquiry and addresses how the quality and ethical integrity of the research were maintained throughout.

3.1 Research Design

The study employs analytic autoethnography (Anderson, 2006) as its primary methodology. Anderson defines analytic autoethnography through five features: complete member researcher status, analytic reflexivity, visible narrative presence of the researcher, dialogue with informants beyond the self, and commitment to theoretical analysis. This approach was selected because the researcher is a practicing kindergarten teacher at the study site—a complete member of the group under investigation—and because the research questions center on how a professional development program was interpreted and enacted from the inside, not solely evaluated from the outside.

Unlike evocative autoethnography (Ellis & Bochner, 2000), which prioritizes emotional and aesthetic expression, analytic autoethnography uses the researcher's insider position to develop theoretical explanations of broader social phenomena. In this study, the researcher's firsthand participation in the AI professional development program provided access to

experiential data—moments of confusion, decisions about how to engage, shifts in professional self-understanding—that would be difficult to capture through external observation alone. Supplementary data from colleagues, gathered through a questionnaire and follow-up interviews, provided perspectives beyond the researcher's own and enabled dialogue between the researcher's experience and the experiences of other participants. A documentary record of program communications provided a third vantage point independent of any participant's recollection. This combination of self-narrative, participant data, and documentary evidence distinguishes the study from purely introspective autoethnography and aligns it with Anderson's (2006) call for autoethnographic work that contributes to broader analytic understanding.

The distinction between interpretation and enactment is central to the study's framing. Following the conceptual distinction between implementation and enactment in educational settings—implementation referring to how an initiative is understood by those who mandate it, and enactment referring to what practitioners actually do when a plan meets practice—this study treats teachers' responses to the program not as degrees of compliance with an intended design but as active enactments shaped by how the program was interpreted. The research questions operationalize this framing.

The primary research question is: *How do teachers at an American preparatory school interpret and enact mandatory, externally delivered AI professional development?*

Three sub-questions structure the inquiry:

1. How does conceptual confusion shape teachers' interpretation of AI professional development content?
2. How does implementation overload shape teachers' enactment of AI professional development in everyday practice?
3. How do institutional culture, agency, and professional identity contextualize that interpretation and enactment?

These questions operationalize the theoretical constructs established in the theoretical overview—conceptual confusion (Fullan, 2007), implementation overload (Fullan & Quinn, 2016), cognitive load (Sweller, 1988), teacher agency (Biesta et al., 2015; Priestley et al., 2015), and professional identity (Beauchamp & Thomas, 2009; Kelchtermans, 2009). The data collection instruments described below were designed around these constructs, and the research questions specify the relationships among them that the analysis examines.

3.2 Research Context and Sample

The study site is a private, bilingual (English–Spanish) American-curriculum preparatory school in Latin America that serves a diverse student population and employs approximately 250 instructional staff. The staff includes both local citizens and expatriates from English-speaking countries, working across divisions from early childhood through high school. Beginning in January 2025, the school mandated participation in an externally delivered AI professional development program for all instructional staff. The program was designed and facilitated by an outside vendor whose selection process was not disclosed to participants.

Content was delivered through a combination of synchronous Zoom sessions and asynchronous modules hosted on the vendor's learning management system (LMS), a platform separate from the school's existing instructional infrastructure. The vendor maintained two versions of the LMS—one English-forward and one Spanish-forward—though this was not disclosed to participants or to the researcher during the program; the existence of the dual-platform structure was communicated to the researcher by the school's Center for Teaching and Learning only after data collection was underway.

The program comprised five sequential courses: AI for Learning Design, AI for Assessment, AI for Differentiation, AI for Data Analysis, and AI for Routine Tasks. Delivery formats varied across courses. Some followed a weekly structure with approximately 1.5-hour synchronous sessions supplemented by asynchronous tasks; AI for Data Analysis departed from the weekly structure, delivered instead as a five-day intensive running Monday through Friday, with a 90-minute synchronous session each day. Participants were grouped into mixed cohorts that included teachers from early childhood through high school, specialists, and administrators, meaning that a kindergarten teacher and a secondary subject teacher received the same content in the same session. The mixing of cohorts in the Data Analysis intensive was not communicated in advance by the Curriculum, Teaching and Learning department or the vendor, and represented a departure from the cohort arrangement used in the other courses. All synchronous sessions attended by the researcher were conducted in English, and the LMS platform to which the researcher's cohort was assigned was English-forward. The researcher's observation logs and autoethnographic reflections therefore document the English-language experience of the program. The scope and content of the Spanish-forward platform—including whether its materials, sequencing, or facilitation differed from the English version—remain unknown to the researcher, as the dual-platform arrangement was not disclosed until after observation logging had concluded. This undisclosed structural

feature is itself relevant to the study's analysis of program transparency. The rationale behind the program's curriculum design, sequencing, and vendor selection was not communicated to participants at any point during the program.

Participant sampling for the supplementary data followed a purposive, convenience approach. All instructional staff who participated in the AI professional development program were eligible to complete the anonymous online questionnaire. Thirty-three staff members responded. Of those respondents, a subset indicated willingness to participate in a follow-up interview, and seven interviews were conducted. The seven interview participants worked across roles and divisions: early childhood and kindergarten, elementary, middle and high school subject areas, and administrative or support roles. Both English-language and Spanish-language participants are represented in the interview sample. The questionnaire respondents ($n = 33$) distributed across roles as follows: Elementary (24%), High School (21%), Early Childhood/Kindergarten (21%), Administrative/Support (18%), Middle School (12%), and Specialist (3%). Teaching experience clustered in mid- to late-career, with 42% reporting 6–10 years and 51% reporting 11 or more years. Prior AI familiarity was modest, with 64% describing themselves as only slightly familiar with AI tools before the program began.

The researcher is a kindergarten teacher at the study site with fewer than two years of employment at the school, eight years of classroom teaching experience, and 13 years of professional experience in education more broadly. This positionality is significant: the researcher participated in the same mandatory AI professional development as all other instructional staff, attended sessions, completed tasks, and navigated the same institutional expectations. The researcher's relatively recent arrival at the school provides a degree of outsider perspective within the insider role, which Anderson (2006) identifies as analytically productive for autoethnographic inquiry. The researcher's dual role as participant and investigator was formally communicated to the school's Center for Teaching and Learning during the program, and is documented in the program communications record described below. Notably, the researcher's position as a kindergarten teacher places her within the same early-childhood group whose experiences of grade-level misalignment feature prominently in the findings; this shared positionality is treated as a source of analytic insight and is examined reflexively rather than treated as a neutral vantage point.

3.3 Data Collection

Data collection occurred in two phases. An informal phase began in January 2025, coinciding with the start of the AI professional development program, during which the researcher began recording observations and reflections. In this early phase these took the form of scattered, unstructured notes—jottings made during or shortly after sessions, without a consistent template—which were later consolidated and which informed the design of the structured instruments used in the formal phase. A systematic, formal phase commenced in August 2025 and continued through the program's conclusion and the completion of interviews in May 2026. Multiple data sources were used to enable analysis across the researcher's own experience, participants' reported experiences, and documentary evidence of the program itself.

3.3.1 Autoethnographic Journal and Reflection

The researcher maintained a standalone reflective journal throughout the duration of the professional development program. These records captured firsthand accounts of sessions attended, including responses to content, moments of confusion or clarity, shifts in professional self-understanding, and informal interactions with colleagues before, during, and after sessions. The journal was the researcher's initial record; the structured observation logs described below were subsequently constructed from it, with the journal's reflective material informing and feeding into the logs' autoethnographic reflection sections. Following Chang's (2008) guidance on autoethnographic data management, the reflections distinguish between descriptive accounts of events and the researcher's interpretive reflections on those events.

3.3.2 PD Session Observation Logs

The researcher completed structured observation logs for synchronous professional development sessions across the five courses. The final corpus comprises observation logs for twelve synchronous sessions, supplemented by course-level observation summaries for each of the five courses, an LMS-completion observation log series documenting the asynchronous module experience, and a slide-deck content analysis for the AI for Differentiation course. Each log was guided by a custom-designed observation template organized around Desimone's (2009) five core features of effective professional development: content focus, active learning, coherence, duration, and collective participation. The template also included optional expansion sections aligned with Darling-Hammond et al.'s (2017) criteria for effective professional development and the U.S. Department of Education Office of

Educational Technology's (2023) policy recommendations for AI in education. For each Desimone feature, the log recorded evidence observed, evidence absent or contradicted, and a rating on a four-point scale (0 = not observed, 1 = minimal, 2 = partial, 3 = strong). Each log concluded with an autoethnographic reflection section that captured the researcher's lived experience of the session separately from the analytical framework ratings, maintaining methodological clarity between structural evaluation and narrative data. A non-attendance log was also completed for one Data Analysis session the researcher did not attend, documenting the experience of falling behind in a mandated sequence. The observation log template is presented in Appendix [X], and a condensed reference card developed to support consistent logging is presented in Appendix [X].

3.3.3 Questionnaire

A bilingual (English–Spanish) questionnaire gathered anonymous data on participants' experiences with the program. It was administered through Microsoft Forms with name recording disabled to preserve anonymity. The instrument comprised 23 items across six sections covering background information, general program experience, course-level perceptions, agency and professional identity, institutional context, and closing reflections, with an optional invitation to volunteer for a follow-up interview. It combined closed-ended items (Likert-type scales, multiple-choice, and ranking) with open-ended responses; Likert items used labeled rather than numeric options to reduce ambiguity across language groups. A consent statement at the outset informed participants that completion constituted voluntary consent. The full instrument is presented in Appendix [X]. The final response count was 33.

3.3.4 Semi-Structured Interviews

Seven semi-structured interviews were conducted with participants who volunteered through the questionnaire. Interviews were guided by a protocol organized around eight domains: opening narrative, conceptual clarity, relevance and implementation, pace and volume, agency and voice, professional identity and self-understanding, institutional context, and reflective closure. The protocol was designed following three principles: non-leading phrasing that allowed participants to name their own register; experience-before-evaluation sequencing consistent with autoethnographic methodology (Adams et al., 2015; Chang, 2008); and theoretical grounding without theoretical language—each domain maps to a construct from the literature (cognitive load, conceptual confusion, teacher agency, professional identity) but uses everyday language. The protocol comprised a pool of questions with follow-up probes,

from which questions were selected for each participant based on questionnaire responses. Interviews were conducted in the participant's preferred language (English or Spanish) and were recorded with consent, then transcribed for analysis. The full interview protocol is presented in Appendix [X].

The interview protocol was developed around the theoretical constructs that the research questions operationalize, rather than around any single phrasing of the research questions. Because the constructs—conceptual confusion, cognitive load, agency, and professional identity—remained stable across the refinement of the research questions, the protocol's domains align directly with the three sub-questions: conceptual clarity with interpretation (Sub-RQ 1); relevance, implementation, pace, and volume with enactment (Sub-RQ 2); and agency, identity, and institutional context with the contextualizing conditions (Sub-RQ 3).

3.3.5 Program Communications Log

A structured communications log was compiled as a distinct documentary data source. The log comprises 27 program-related communications spanning the 14-month duration of the program (January 2025–March 2026), including emails from the vendor, from the school's Center for Teaching and Learning, and from internal instructional support staff, together with associated calendar invitations, newsletters, and announcements. Each communication was logged with its date, sender, recipient group, and content, and was annotated by the researcher with notes mapping its content to the study's analytic constructs (transparency, coherence, agency, overload, and collective participation). The communications log functions as documentary evidence independent of participant recollection and the researcher's experiential record, providing a third source against which patterns identified in the autoethnographic and participant data could be examined. Because the log documents the program's design and delivery as institutionally enacted—what was communicated, when, by whom, and with what omissions—it is treated as a primary data source in its own right rather than as background material.

3.3.6 Additional Documentary and Artifactual Evidence

Beyond the communications log, additional documentary sources included content from the vendor's learning management system: Google Doc templates, vendor-created applications, externally linked AI tools, participant-submitted work, discussion-thread contributions and their engagement metrics, and tutorial recordings created by the vendor. The AI for Differentiation slide decks were analyzed separately as an artifact of program content. These

materials provide documentary evidence of the program's design, delivery, and content, supplementing the researcher's firsthand observations and participants' self-reported experiences.

3.4 Data Analysis

Data analysis followed the six-phase reflexive thematic analysis framework described by Braun and Clarke (2006, 2019, 2022). This approach was selected for its flexibility across data types, its compatibility with autoethnographic methodology, and its explicit positioning of the researcher's interpretive engagement as a resource for analysis rather than a source of bias to be eliminated.

In the first phase (familiarization), the researcher read and re-read all data sources—observation logs and their autoethnographic reflections, questionnaire responses, interview transcripts, the communications log, and documentary materials—to develop familiarity with the data's content and character. In the second phase (generating initial codes), the data were coded systematically, with segments tagged in relation to the study's theoretical constructs (conceptual confusion, implementation overload, cognitive load, teacher agency, and professional identity) and to patterns arising inductively from the data. Coding proceeded source by source, so that each data type generated its own initial codes before cross-source comparison. In the third phase (generating themes), codes were grouped into candidate themes capturing patterns of meaning across the dataset. Candidate themes were developed to hold across more than one source type where possible, with source-specific nuances retained as sub-themes. In the fourth phase (reviewing themes), candidate themes were checked against the coded data and the full dataset for coherence and adequacy of support, and a cross-source matrix was constructed to map where themes converged across sources, where one source extended or specified another, and where sources diverged. In the fifth phase (defining and naming themes), each theme was refined to clarify its scope and its relationship to the research questions. In the sixth phase (producing the report), the themes were organized into an analytical narrative addressing the research questions, integrating the researcher's autoethnographic reflections with participant data and documentary evidence.

Consistent with Braun and Clarke's (2019, 2022) account of reflexive thematic analysis, the use of multiple data sources is understood as multi-perspectival theme development rather than as positivist validity-checking. The aim of examining a pattern across the autoethnographic record, the interviews, the questionnaire, and the communications log was not to confirm a single objective account but to develop themes that are robust across vantage

points and to treat divergence among sources as analytically productive rather than as error to be resolved. Where the researcher's autoethnographic reading and participants' accounts diverged, the divergence was retained and examined.

Questionnaire data were handled according to the nature of each item. Closed-ended items (Likert-type scales, multiple-choice, and ranking questions) were analyzed using descriptive statistics to characterize the distribution of responses, and a cross-tabulation of overall program assessment by teaching role was computed to examine variation in experience across roles. Open-ended responses were incorporated into the thematic analysis alongside the other qualitative data.

The researcher used Claude (Anthropic), a generative AI assistant, as a tool at several stages of the analysis and writing process. Its use and limits are disclosed in full in the Ethical Considerations section below. The researcher retained interpretive authority over all coding and thematic decisions, reviewed all AI-assisted output against the underlying data, and is responsible for all analytical claims.

3.5 Trustworthiness and Quality

Several strategies support the credibility and trustworthiness of the study. First, the use of multiple data sources—autoethnographic observation logs and reflections, questionnaire data, interview transcripts, the communications log, and documentary artifacts—enabled multi-perspectival theme development. No single source carries the analysis alone; patterns identified in the researcher's experience were examined against participant reports and the documentary record, and the communications log in particular provided evidence independent of any participant's recollection.

Second, the structured observation log template, grounded in Desimone's (2009) framework, provided a consistent analytical lens across sessions. The explicit rating scale and guiding questions for each feature reduced the risk of selective or impressionistic observation. The template's separation of autoethnographic reflection from analytical observation helped maintain clarity between the researcher's lived experience and the structural evaluation of the program.

Third, the researcher maintained analytic reflexivity throughout, consistent with Anderson's (2006) criteria. This involved ongoing examination of how the researcher's positionality—as a kindergarten teacher, a relatively new staff member, and an expatriate working in a bilingual Latin American school—shaped what was observed, how it was interpreted, and what

meaning was made from the data. The researcher's membership in the early-childhood group whose experiences are prominent in the findings was treated as a position requiring particular reflexive scrutiny, given the risk that shared grievance could shape interpretation.

It should be noted that member checking was not conducted; interview participants did not review their transcripts or the researcher's interpretations prior to analysis. This decision, and its implications for the study's trustworthiness, is addressed among the limitations discussed in the Results and Discussion chapter. Reliance on a single coder, without independent inter-coder comparison, is likewise acknowledged there. These are reported transparently rather than presented as having been fully resolved.

3.6 Ethical Considerations

The study did not undergo formal institutional review board review, as the researcher's program does not require one. A research proposal was submitted to the school's administration, and permission to conduct the study at the site was granted. All questionnaire participation was voluntary, and responses were fully anonymous; the Microsoft Forms instrument was configured with name recording disabled, and participants were instructed not to include identifying information. The consent statement embedded in the questionnaire informed participants that completion constituted voluntary consent. Contact information provided by interview volunteers was stored separately from questionnaire data.

The researcher's dual role as colleague and investigator introduces particular ethical considerations. Participants may feel social pressure to participate or to moderate their responses, given that the researcher is a fellow staff member. To mitigate this, the questionnaire was anonymous and interview participation was entirely voluntary. The semi-structured protocol used non-leading questions to avoid channeling participants toward particular interpretations. The researcher was transparent about the study's purpose with participants and treats this dual positionality as a feature of the autoethnographic design rather than a limitation to be eliminated.

To protect confidentiality, interview participants are referred to by pseudonyms throughout the thesis, paired with general role descriptors. Questionnaire respondents are identified only by role and years of experience, never by name. The researcher recognizes that in a relatively small professional community, complete anonymity cannot be guaranteed even with these measures, and participants were made aware of this in the consent process. Where role descriptors are sufficiently specific to risk identification, they have been generalized.

A particular ethical matter surfaced during the analysis and is reported in the findings: in at least one course, participants were asked to input student data, including data concerning students with complex learner profiles, into vendor-hosted AI systems whose data-handling policies were not disclosed. This is treated in the analysis as a finding about the program rather than as a feature of the present study's own data handling; no student data was collected or used in this research.

3.6.1 Disclosure of Generative AI Use

In accordance with the University of Tartu guidelines for the use of AI applications in teaching and studies, the researcher's use of generative AI in the preparation of this thesis is disclosed here. The researcher used Claude (Anthropic) as an analytical and writing-support tool in four capacities: (a) to assist in identifying patterns across the data sources during coding; (b) to generate candidate themes for the researcher's review; (c) to expand and revise draft text for the Results and Discussion chapter, where the researcher produced an initial draft and supplied the data, evidence, and analytic direction, and the tool contributed additions and revisions; and (d) to revise chapters for clarity, cohesion, and clarity of research argument. In every case, the researcher supplied the underlying data and analytic framing, reviewed all generated and revised output against the source data, retained full interpretive authority over coding and thematic decisions, and edited and verified all text incorporated into the thesis. The researcher is solely responsible for the analytical claims, interpretations, and conclusions presented.

4. Results and Discussion

4.1 Chapter Overview

This chapter presents the results of the data analysis and interprets them in relation to the research questions and the theoretical framework, drawing on the four data sources described in the Method chapter. The seven interview participants are referred to by the pseudonyms Ana, Marcus, Sofia, Daniel, Helena, Beatriz, and Carmen; questionnaire respondents are identified by role and years of experience only. The chapter is organized around the study's three sub-questions, with themes presented within each. Following Anderson (2006), the researcher's first-person voice appears where the autoethnographic logs are the primary evidence, and third-person analytical voice predominates where interview and questionnaire

data are central. The concluding chapter then synthesizes the findings and addresses limitations, applications, and recommendations.

4.2 Status of the Research Questions

All three sub-questions were addressed by the data, and the primary question is answered through the synthesis presented in the concluding chapter: teachers did not implement the program as designed but enacted a distribution of strategic responses shaped by the interpretive coherence, cognitive demand, institutional voice, and role alignment available to each.

4.3 Sample Summary

The 33 questionnaire respondents spanned all divisions, with the largest groups in Elementary, High School, and Early Childhood/Kindergarten roles, and a substantial Administrative/Support group; experience clustered in mid- to late-career, and prior AI familiarity was modest (Table 1).

Table 1. *Sample Characteristics (Questionnaire, n = 33)*

Variable	Distribution
Teaching role	Elementary 24%; High School 21%; ECE/Kindergarten 21%; Admin/Support 18%; Middle School 12%; Specialist 3%
Years teaching	0–2 yrs 3%; 3–5 yrs 3%; 6–10 yrs 42%; 11–20 yrs 27%; >20 yrs 24%
Prior AI familiarity	Slightly familiar 64%; Moderately familiar 33%; Very familiar 3%
Preparedness for PD	Well prepared 6%; Somewhat prepared 49%; Somewhat unprepared 33%; Not at all prepared 12%

The seven interview participants, all of whom also completed the questionnaire, were: Ana (elementary, 6–10 years); Marcus (high school science, department lead, fewer than 2 years at the school); Sofia (kindergarten, early literacy and ESL); Daniel (high school humanities, 6–10 years); Helena (high school humanities, predominantly seniors); Beatriz (pre-kindergarten, English-language-learner classroom); and Carmen (administrative role, Spanish-first speaker). The sample spans early childhood through high school and includes administrative staff and both English- and Spanish-language participants. The researcher is a kindergarten teacher at the study site, placing her within the same early-childhood group whose accounts of grade-level misalignment are prominent in the findings; this shared position is examined reflexively rather than treated as neutral.

[Table 2 placeholder: course-level clarity and relevance ratings — insert final table.]

4.4 Sub-RQ 1: How Conceptual Confusion Shaped Teachers' Interpretation

The first sub-question concerns how conceptual confusion—ambiguity arising when key terms, frameworks, and expectations shift (Fullan, 2007)—shaped how teachers understood the program. Two themes emerged: a coherence failure in which content was misaligned with participants' grade level, subject, or context; and the deflection of teachers' substantive conceptual and ethical questions, which produced a breakdown of trust that itself shaped interpretation.

4.4.1 Coherence Failure: "Not Designed With Us in Mind"

The dominant interpretive frame teachers used was that the program had been built without regard for who would receive it. This frame appeared in every data source and across every role, most acutely among early childhood, elementary, and senior-level teachers—those at the margins of a curriculum apparently pitched to a generic middle range.

The questionnaire data show the pattern directly. Course-level clarity ratings were generally favorable (60–70% rated each course mostly or very clear, except Data Analysis at 33%), but relevance ratings were substantially lower. For four of the five courses the modal relevance response was "slightly relevant": Learning Design (42%), Assessment (49%), Differentiation (42%), and Data Analysis (36%). Only AI for Routine Tasks broke the pattern, with 36% rating it very relevant. Teachers could follow what they were shown; they could not connect it to their work.

The interview data specify the misalignment across the full grade span. Sofia, a kindergarten teacher, identified the first course as built without grade-level differentiation, asking how a single course could serve a pre-K teacher and an advanced secondary teacher at once, and concluding that the program had skipped the backward planning instructional design requires. Helena, who teaches predominantly seniors, reached the same judgment from the opposite end: she found the Data Analysis course inapplicable to students whose academic data predictably declines late in the college-admissions year, and identified mixed grouping as the core failure—she shared too little common ground with a lower-school teacher in a breakout room to engage the task. Asked to rate her sense of input and fit, she offered "a two" out of ten. Beatriz, a pre-kindergarten teacher whose English-language-learner class used tablets

twice in a year, described the content as "totally unrelated" to her work and the presenters as "very disassociated from what teachers are actually dealing with."

These accounts converge on a structural rather than stylistic objection: teachers did not object to AI as a topic but to the absence of differentiation in a program ostensibly about differentiation. The convergence across kindergarten, pre-K, and senior-level teachers is analytically important: the coherence failure was not confined to one division but was a property of the mixed-cohort design, producing misalignment at both ends of the grade span at once.

The vendor communications log documents institutional recognition of this gap. In an email dated October 7, 2025—eight months into the program—the school's Center for Teaching and Learning conceded to faculty that the program had not felt relevant for Pre-K, Kindergarten, and Grade 1 teachers. A vendor email two weeks later acknowledged the early-childhood mismatch and offered supplementary materials. The institutional response was a workaround offered late, not a redesign; the decision to use mixed cohorts had been made before any such feedback could shape it.

This finding aligns with Desimone's (2009) identification of coherence as a core feature of effective professional development: the alignment of content with teachers' knowledge, beliefs, and context. The data document a systematic failure of this feature. Fullan and Quinn's (2016) account of coherence as organizational alignment, and Darling-Hammond et al.'s (2017) emphasis on content focus grounded in teachers' own contexts, make the same point. The program's reliance on standardized content delivered to mixed-grade cohorts placed it at odds with the conditions the literature associates with durable teacher learning. The coherence failure was not incidental; it was the lens through which most participants interpreted everything that followed.

A second, finer dimension appears in the autoethnographic logs: breadth without prioritization. The Learning Design log records optional resources ranging from institutionally central material (AI policy templates) to speculative items (a piece on whether AI is conscious), with no scaffolding to indicate relative importance; the Differentiation course introduced several vendor tools in a single session at a pace that precluded evaluation or retention. The interpretive problem was not that individual elements were unintelligible but that teachers could not assemble them into a coherent trajectory. Helena named this directly: she valued a Learning Design exercise comparing outputs across Gemini, Claude, and ChatGPT, but wished for "a textbook, or a resource handbook" to return to, since instruction

was "verbal only, or within their PowerPoint slides only"—when she briefly left a session, she returned unable to find where the group was. The absence of durable, referenceable materials compounded the breadth problem and is discussed further under format and cognitive load.

4.4.2 Deflected Questions and the Breakdown of Trust

A second pattern within interpretation concerns the questions teachers brought and the program's response to them. Several participants entered with substantive concerns—about student over-reliance on AI, environmental impact, equity, and the displacement of teacher judgment—and reported that these were deflected rather than engaged. The consequence was not merely dissatisfaction; for some participants it was an early interpretive judgment that the program could not address what mattered to them, which then shaped all subsequent engagement.

Helena's account is the clearest instance of this pathway. Her central concern as a secondary humanities teacher was that students were outsourcing critical thinking to AI, and she reported that when she and colleagues raised this early in the program, the concern was dismissed rather than treated with what she called "the gravity that we actually really genuinely feel." She described the result plainly: she "wrote off this program very early on," locating the failure in trust. Once she judged that the program would not engage her primary professional worry, she disengaged from its content as a whole. This is an interpretation-driven enactment: an early interpretive verdict (the program will not address what concerns me) produced a durable behavioral stance (withdrawal). Her later observation that an instructor cited a social-media video as a source—after which, as a trained historian, she "checked out"—illustrates how individual moments were read through an interpretive frame already set against the program's credibility.

Ana framed her experience around the same deflection in a different register, reporting that she had hoped for substantive treatment of ethical AI use and found the topic "brushed over," leaving her using AI with "a guilty conscience" about its environmental and equity implications. Sofia raised the question of whether AI should be used with five-year-olds at all, using the phrase "moral justification" to name what she felt was absent. Beatriz, teaching pre-K, judged the ethics discussion "wasted breath" for her context on the grounds that the tools would be used by whoever could access them regardless of classroom guidance—a different but related interpretation, in which the ethical content was not so much deflected as mismatched to her students' developmental stage.

The questionnaire data echo these accounts across a wider sample. A high school humanities respondent reported that the program met resistance with the message that AI "is here to stay, so you better get used to it." Another stated, "I will not outsource my relationship with students to an algorithm." A reading specialist in the early grades reported that her concern about AI's environmental consequences received no action planning, and tied her professional role to "monitoring AI for errors and bias, and helping students do the same"—an ethical position the program did not engage.

This pattern maps onto the AI literacy framework of Long and Magerko (2020), operationalized for teacher PD by Tripathi et al. (2025), which distinguishes knowing about AI, using AI, evaluating AI, and reasoning ethically about it. Tripathi et al. found that teacher PD tends to privilege the operational dimension over the evaluative and ethical. The present study confirms this. Participants developed prompt-writing and tool-use competence (operational literacy) while reporting that critical and ethical literacy—the dimensions that would have addressed their substantive concerns—were deflected or treated as topical. The interpretive consequence is the more important finding: when professional questions are deflected, teachers interpret the program as either unable or unwilling to engage them, and that interpretation, once formed early, is resistant to later revision. The breakdown of trust functioned as an interpretive filter through which the rest of the program was read.

4.5 Sub-RQ 2: How Implementation Overload Shaped Teachers' Enactment

The second sub-question concerns enactment—what teachers actually did with the program. Implementation overload (Fullan & Quinn, 2016) shaped enactment through three patterns: disengagement-within-attendance; format-driven cognitive load; and uneven productive outcomes. The analysis also found that the load itself was not uniform across the program, a refinement of the overload construct that the data require.

4.5.1 Disengagement-Within-Attendance

The most distinctive enactment pattern is the gap between what the institution could observe—attendance, learning-management-system (LMS) completion, certificates—and what teachers actually did with the time. This gap is what makes enactment, rather than implementation, the appropriate frame: teachers did not fail to carry out a plan so much as enact their own response to it while the structural shell of compliance remained intact. Marcus, a high school science teacher, described keeping his camera on during a session whose pace he could not follow while completing other work in a separate window; his

attendance was recorded, his participation was not. Daniel described a more deliberate version of the same response—zoning out, leaving assignments unfinished, not pursuing the certificate—while remaining nominally present. Helena described the on-campus version: when she attended in person she would move to a colleague's classroom, where the two would fall into side conversations about the session's shortcomings, finding it "much easier to disengage" than when a live instructor was physically present. Her observation that teachers "behave differently" with a live instructor is a precise statement of how the delivery format shaped enactment. Ana reported the inverse vulnerability—remaining willing but losing motivation when her cohort visibly disengaged, contrasting an administrator-present cohort with high engagement against a later cohort in which "everyone had their cameras off and their mics off."

The questionnaire data show the pattern at scale, with participants converging on a shared vocabulary: an elementary teacher "found myself disengaging with the content"; a senior-level teacher reported the program "lost me within the first three sessions and never got me back"; a long-tenured middle school teacher "ended up checking out and not really trying, which is sad because I wanted to learn." Carmen, whose overall experience was positive, corroborated the pattern from the inside when she observed that the program worked "for the people that was engaged," while noting that the school still needed "to get everyone on board"—an acknowledgment, from a participant favorably disposed to the program, that engagement was unevenly distributed across the cohort.

The autoethnographic record documents the pattern from the inside: the Routine Tasks session-4 log notes the researcher's own minimal time-on-task while present, the Differentiation logs record extended breakout-room silences, and the Learning Design log provides the most direct evidence—discussion-board posts with zero likes, comments, and replies despite an explicit requirement to respond to colleagues. The institutional response is itself evidence of scale: a January 2026 faculty newsletter issued explicit reminders about cameras, names, and norms—issued only after compliance had visibly broken down, in the program's final months.

Cognitive load theory (Sweller, 1988) and the change literature on overload (Fullan & Quinn, 2016; Hargreaves & Fullan, 2012) help explain why this particular pattern emerged.

Cognitive load theory holds that when the demands of a task exceed a learner's working-memory capacity, the learner cannot process the material deeply, and learning and transfer are impaired (Clark et al., 2012). The overload literature adds that when teachers face cumulative

initiative demands without adequate time or support, they tend to adopt surface compliance rather than substantive change. Taken together, the two accounts anticipate the response observed here: a teacher who cannot productively process the session content, and who has no sanctioned way to opt out of a mandatory program, is left with attendance as the only available form of participation. The predictable result is not open refusal or visible struggle but the quiet decoupling of presence from engagement—remaining in the session while withdrawing cognitive investment. The data show this decoupling at the individual, cohort, and institutional levels. The pattern is therefore not a failure of teacher will but the rational adaptive response of professionals required to remain present in a program they could not productively engage.

4.5.2 A Refinement: Load Was Not Uniform

The interview data require a refinement of the overload construct. Implementation overload, as theorized, describes excessive cumulative demand. Several participants experienced precisely this in the Data Analysis course, which Helena described as "filled to the brim" and which she could not follow because instruction proceeded verbally, without written steps, through technical content—coding references that, she said, registered as undifferentiated noise. But the same participants reported the opposite condition in other courses. Helena characterized Learning Design and Assessment as thinly resourced: an instructor would talk briefly, assign a thirty-minute breakout, and the group would exhaust the task and sit idle, having, in her words, secured their attendance credit with time to spare. Beatriz made the parallel point that her sessions "could have been an email"—a two-hour synchronous meeting for a task she could have completed in twenty minutes with a written guide.

This is not a contradiction in the data but a more precise account of how the program mismanaged cognitive load in both directions. Some sessions exceeded capacity (excessive intrinsic and extraneous load, in cognitive-load terms), while others underused the synchronous time (low germane load, high opportunity cost). Both conditions produced disengagement, but by opposite mechanisms: overload produced withdrawal through saturation, while underload produced withdrawal through irrelevance and the perception of wasted time. The overload construct (Fullan & Quinn, 2016) thus captures part of the enactment story; the fuller account is one of poor load calibration, in which the program failed to match cognitive demand to the synchronous format and to participants' varying prior knowledge. This refinement strengthens rather than weakens the study's framing: it specifies

that what impaired enactment was not demand as such but the program's inability to calibrate demand to learners and format.

4.5.3 Format and Cognitive Load

The virtual synchronous format itself was a structural impediment that multiple participants named. Daniel described the difficulty of watching a facilitator demonstrate a task on one screen while performing it on another, at a pace faster than his own work, and proposed a flipped structure—recorded pre-work that could be paused, followed by synchronous application. This proposal corresponds directly to the split-attention effect documented in cognitive load research (Sweller, 1988): when learners must divide attention between simultaneous sources, working memory is overloaded. Beatriz independently described the same toggling problem and the same preference for a written guide she could work through at her own pace. Helena's account of being unable to recover her place after a brief absence, given verbal-only instruction, is a further instance: the format provided no persistent external representation to offload working memory.

The questionnaire data corroborate the format finding (instruction "moved too quickly," difficulty following demonstrations across windows), as does the after-hours scheduling concern raised by several participants, including the cost of childcare to attend. Beatriz noted gratitude that the program had not begun in August but identified the after-contract-hours timing as a shared grievance. The autoethnographic logs locate the format problem within a larger infrastructure failure: the Assessment session-2 log records a facilitator's divided attention and a 20-minute breakout silence, and the Learning Design log records reading materials embedded in documents that returned access-denied screens to participants' school credentials. The format constraints were compounded by content infrastructure that did not reliably function for the participants it was meant to serve.

The contrast participants drew with in-person or team-based delivery strengthens the interpretation. Daniel cited an internal, in-person workshop led by a colleague as more productive than equivalent synchronous online time. Ana and Marcus preferred in-person delivery with their actual teams. Helena, while noting that she personally focused better at home (living alone, free of domestic distraction), nonetheless argued that on-campus group PD should be conducted with a live instructor, precisely because physical presence changes participant behavior. These preferences map onto Desimone's (2009) collective participation feature: professional learning is more productive when teachers work with colleagues who

share content, grade level, or context—conditions the mixed-cohort virtual format did not provide.

4.5.4 Productive Outcomes Despite Misalignment

Despite the structural critiques, participants reported concrete productive outcomes, and an accurate account must report them. These outcomes were uneven, clustering by course content and by role in ways that themselves carry analytic significance, and they coexisted with the disengagement documented above rather than displacing it. The framing here is deliberate: teachers extracted use value from a misaligned program, often through their own adaptive enactment and outside the structures the program had designed, rather than the program straightforwardly producing positive outcomes.

Marcus presents the clearest productive arc. Entering resentful of the demand on his time, he ended by describing how he had used the program's tools to rebuild ecology and tectonics units faster and more thoroughly than he could have unaided, through a tool the school already licensed. Others reported narrower gains: Ana automated routine tasks while keeping her ethical reservations; Sofia found administrative uses (syllabus and IEP-goal drafting) if not instructional ones; Daniel used AI in a poetry unit to have students rewrite their own work across forms. Carmen, in an administrative role, described the most unambiguously positive arc—initial overwhelm resolving into confidence and time saved—and her case is notable for tracing overload and a productive outcome in the same person: "behind all the time," she finished the work "in one week," yet emerged confident.

The questionnaire data show similar outcomes concentrated by content and role. Productive responses clustered around the Routine Tasks course—the only course rated modally "very relevant"—and among administrative and support staff. The cross-tabulation of overall assessment by role makes this concentration explicit: Administrative/Support staff rated the program positively in 5 of 6 cases (83%), with none negative; Elementary teachers showed the opposite distribution (1 of 8 positive, 6 of 8 negative); Early Childhood teachers were mixed and predominantly neutral-to-negative; High School teachers were the most polarized; and Middle School teachers were predominantly positive. Carmen's positive interview account and Beatriz's and Helena's critical accounts are not in tension with one another; they are the interview-level expression of this role distribution.

The role effect is not a peripheral result. It complicates any unified verdict on the program. The conditions that produced positive outcomes for administrative staff—content (routine-

task automation) aligned with their actual work, and proximity to the institutional rationale—were the same conditions whose absence produced disengagement for early-grade classroom teachers. A program is not effective in the aggregate when its effectiveness is conditional on which segment of a mandated cohort one occupies. The productive outcomes are real; they are also distributed in a way that tracks the coherence and voice conditions analyzed elsewhere in this chapter.

Two qualifications emerged. First, several participants reported that their gains were achievable by other means: Daniel argued that one general-purpose model met his classroom needs and the specialized tools added complexity without proportionate benefit, and a questionnaire respondent noted that the one useful skill she retained took about thirty minutes to learn. Second, the productive outcomes had a ceiling the program's framing obscured. Attempting the culminating project—building a tool to arrange any song into two-part harmony for a hypothetical choir—Beatriz was assured by a general-purpose model that it could do so, and after five hours received an unusable output. Her conclusion, that the tool would "flat out lie" rather than disclose its limits and that she could have written the arrangements herself in the time spent, names the gap between implied and delivered capability. The program's optimistic framing set expectations the tools could not meet; at least one early-childhood participant reported emerging more AI-averse than she began. The productive outcomes were genuine but bounded, uneven by role, and frequently achieved in spite of the program's design rather than because of it.

4.6 Sub-RQ 3: Institutional and Identity Context

The third sub-question concerns the institutional culture, agency, and professional identity within which interpretation and enactment occurred. Four themes are presented: the absence of teacher voice and choice; vendor opacity and institutional communication failure; facilitator quality as a mediating variable; and professional identity effects.

4.6.1 Voice and Choice Absence

On the questionnaire's voice item, 64% of respondents (21 of 33) reported very little or no voice in how they engaged with the program; only four reported significant voice. This is among the strongest convergent patterns in the dataset, and it appears with consistent language across sources.

Interview participants articulated it directly. Daniel described having "very little say." Sofia framed the absence as institutional inconsistency: a school that promotes student voice and

choice as pedagogy denied teachers an analogous voice, and she observed that the institution did not "practice what they preach." Helena rated her sense of input "a two" out of ten and located the core failure in the decision to group all staff together regardless of role and need. Beatriz provided the most concrete evidence of voice denied: her pre-K department of seventeen teachers approached administration collectively with an alternative professional development of equivalent time and commitment and was told to remain with the AI program. The same after-hours scheduling that recurred across accounts compounded the sense that the program was something done to staff rather than with them.

The communications log corroborates this at the documentary level. Cohort assignments were delivered without consultation, with faculty instructed to "verify placement" but given no mechanism to influence grouping; schedule changes were announced without faculty involvement; and by the program's final months the communications had shifted to compliance framing, invoking deadlines and certificates rather than relevance. This pattern aligns with the ecological account of agency (Biesta et al., 2015; Priestley et al., 2015), in which agency is achieved or constrained through working conditions rather than possessed as a trait. When teachers cannot influence what, when, with whom, or whether they learn, their capacity for intentional action narrows—and, as Beatriz's account shows, so does their willingness to invest effort.

Beatriz stated the link between voice and enactment more directly than any other participant: if the institution does not care about her individual growth, she asked, "then why should I put any effort into it?" This is the agency–enactment connection in its clearest form. The denial of voice was interpreted as institutional indifference to her professional development, and that interpretation licensed a withdrawal of effort. The voice finding (Sub-RQ 3) and the disengagement finding (Sub-RQ 2) are, in this light, two descriptions of one process: the institutional condition (absent voice) and the enacted response (withdrawn effort) joined by an interpretive judgment (the institution is indifferent to my growth).

A counter-pattern qualifies the theme. Carmen, in an administrative role, did not experience the program as imposed in the same way; she dealt directly with her trainer, did not feel the need to involve school leadership, and valued the option to attend from school or home. Several Spanish-language administrative respondents likewise reported positive experiences emphasizing capable facilitation and useful tools. Voice and choice, in other words, were distributed unevenly, and the distribution tracked the same role pattern visible in overall

assessment. The absence of voice was not absolute; it was concentrated among classroom teachers, and most acutely among those at the margins of the grade span.

4.6.2 Vendor Opacity and Institutional Communication Failure

A second theme concerns the program's structural communication conditions. The autoethnographic logs and the communications log document sustained vendor opacity—about both the program's design rationale and the commercial relationships embedded in its content—alongside institutional communication failures that shifted logistical and contextual labor onto participants.

The Learning Design log documents that the vendor included its own AI assistant among the course tools, listed its own products alongside third-party tools in exploration activities, offered a discount code for a partner product, and administered feedback through its own platform—without disclosing these commercial relationships. The Differentiation course extended the pattern, building the design workflow around vendor-proprietary tools in a way that created dependency participants could not reproduce independently. A more serious transparency concern also appears in the Differentiation course: participants were asked to input student data—including data on students with complex learner profiles—into vendor-hosted AI systems whose data-handling policies were not disclosed, and one underlying platform displayed its own warning against entering sensitive information, which the activity required participants to disregard. This data-privacy concern is identified here as a finding about the program and is discussed further in the applications section; no student data was used in the present research.

The institutional communication failures are documented across the 27 logged communications, which included welcome emails without session dates or access links, calendar invitations without explanatory content, the vendor's name misspelled in official newsletters, peer-to-peer forwarding of meeting links when official channels failed, and a structured peer-sharing protocol introduced only in the program's final session; two questionnaire respondents reported not receiving invitations for entire courses. Participants read these failures through the school's relationship with the vendor. Daniel observed that the administration appeared to have placed substantial trust in the vendor while lacking the capacity to course-correct, describing the school as "a little bit blind." Helena extended this into a critique of institutional decision-making, locating the program within a pattern of the school adopting initiatives "without ever giving any of them really the time to mature," and suspecting that the program as pitched differed from the program as delivered. Her reading

matches a documentary reality visible in the log: internal school staff, not the vendor, performed much of the contextualization labor required to make the content usable, particularly in the Data Analysis module.

This institutional-churn critique—the adoption of successive initiatives without sustained implementation—connects the present case to the broader change literature. Hargreaves and Fullan (2012) identify initiative overload and the failure to sustain reforms as primary barriers to durable improvement. Helena's account suggests that teachers read the AI program through the memory of prior under-sustained initiatives, which shaped their interpretation before the program's own merits could be assessed. Vendor opacity is treated here as a stand-alone theme because the documentary evidence is uniquely strong and because the commercial-relationship and data-privacy concerns raise questions the coherence framework alone does not address: not whether the content fit teachers' contexts, but who designed the program, who benefited from it, and what was withheld.

4.6.3 Facilitator Quality as a Mediating Variable

A third theme is finer-grained: the same content produced different interpretation and enactment outcomes depending on who delivered it. The program's multiple facilitators created a natural variation in delivery that participants noticed and that emerges most clearly when interview and questionnaire data are read together.

Marcus drew the sharpest contrast, describing one facilitator's slow, exploratory pace as productive for him and another's rapid movement through technically dense material as a trigger for disengagement. The questionnaire data both corroborate and complicate this: a middle school respondent identified the same fast-paced facilitator as a reason for her disengagement, while an administrative respondent praised that same facilitator as a reason for her focus. Carmen attributed part of her positive experience to her trainer specifically, valuing the rapport she built "with persons." A Spanish-language respondent identified facilitator selection as a strength of the program, noting that capable facilitation made the content feel achievable. Helena offered the converse, recalling an instructor whose appeal to a social-media source forfeited her credibility.

The contradiction between participants' assessments of the same facilitator is analytically productive. Facilitator effectiveness, in these data, is not an intrinsic property; it is a function of the interaction between facilitator pace and style, participant prior knowledge, cohort composition, and content type. A facilitator who is too fast for a participant unfamiliar with

the technology may be well paced for one who is not; a facilitator effective with an administrative cohort on an automation task may falter with a mixed classroom-teacher cohort on a pedagogical one. The autoethnographic logs sharpen this: the Assessment session-2 log records that when facilitator attention was sustained on a small, content-aligned group using authentic student artifacts, the Desimone features rated higher (partial rather than minimal), while the same facilitator under different conditions produced weaker results. Facilitator quality is therefore best understood as a configuration of conditions rather than a personal attribute—a framing with direct implications for how schools structure external PD partnerships.

4.6.4 Professional Identity Effects

A final, cross-cutting theme concerns the relationship between the program and teachers' professional identities. The theoretical framework drew on Beauchamp and Thomas (2009) and Kelchtermans (2009) to argue that identity is an ongoing process of interpretation shaped by institutional expectations, and that initiatives conflicting with teachers' self-understandings can destabilize identity. The data show identity effects in three directions.

For some, the program reinforced an existing identity. Marcus, who described himself as comfortable with technology and long paperless, found the program extended an identity he already held. For others, it produced a hardening or reversal. Beatriz's case is the clearest: she entered enthusiastic ("sell it to me, I'm ready") and emerged committed to avoiding AI "as much as I possibly can," having found the tools "glitchy and full of mistakes" and concluding that if she were to make mistakes in her classroom she "want[ed] them to be my own." The culminating-project experience described earlier—five hours yielding an unusable musical arrangement—crystallized this reversal. A questionnaire respondent reported the analogous effect of the program strengthening her confidence in her own pedagogical knowledge precisely because the AI outputs underperformed it. For a third group, including Daniel, Sofia, and Helena, the program produced identity maintenance under pressure. Daniel's statement that "nothing's really changed" is itself an identity claim: he engaged, remained unconvinced, and exited with his prior practice intact. Sofia framed her stance as curiosity without adoption. Helena, who described "no change," nonetheless credited AI generally—though not the program—with supporting her lesson design, maintaining a clear boundary between a tool she had chosen independently and a program she had not.

These three patterns—reinforcement, hardening, and maintenance—are all forms of identity work, and none necessarily moved in the direction the program intended. Tripathi et al.'s

(2025) finding that teacher AI integration is shaped by ambivalent beliefs, ecological agency, and professional identity is confirmed here. The program could not produce uniform identity outcomes because identity is not a deliverable; a time-bound external program does not override the longer arc of professional commitment teachers bring to and carry from any mandate. Kelchtermans's (2009) emphasis on vulnerability is relevant: where the program's framing implied that teachers were too busy to do their work without AI assistance—a framing Beatriz heard as the claim that the tools would let her "do everything"—it touched directly on professional self-understanding, and the response was defensive rather than receptive.

5. Conclusion

The preceding chapter presented the study's results and interpreted them in relation to the research questions and theoretical framework, drawing connections to the literature as each theme was developed. This concluding chapter takes stock of the study as a whole: it synthesizes the findings into an integrated account of interpretation and enactment, states the study's limitations, considers the applications of the results, and offers recommendations for further research.

5.1 Synthesis: Interpretation and Enactment as Mutually Constituting

The three sub-questions separated interpretation, enactment, and context for analytic clarity. The data show that these are not sequential stages but mutually constituting processes. How teachers interpreted the program shaped how they enacted it; how they enacted it revised their interpretation; and the institutional context conditioned both at once.

The coherence failure under Sub-RQ 1 created the interpretive conditions for the disengagement-within-attendance under Sub-RQ 2. When a kindergarten teacher interprets a course as not built for her, or a teacher of seniors interprets a differentiation course as inapplicable to students who have largely completed their academic trajectory, the rational enactment is to satisfy the structural requirements while withdrawing the cognitive investment the content does not reward. Helena's case shows the temporal compression of this process: an early interpretive judgment that the program would not address her central concern produced an enactment stance—writing the program off—that then filtered every subsequent session. Interpretation and enactment were not separated by a planning interval; they unfolded together, each revising the other in real time.

The institutional conditions of Sub-RQ 3 were the medium in which both occurred. The voice finding and the disengagement finding are two views of one process, joined by interpretation: the institutional condition (absent voice) was interpreted as indifference to teachers' growth (Beatriz's "why should I put any effort into it?"), and that interpretation licensed withdrawn effort. The role effect makes the same point structurally. Participants whose role aligned with the program's content and who stood closer to its institutional rationale—predominantly administrative staff such as Carmen—interpreted, enacted, and assessed the program differently from classroom teachers whose role did not align. The program did not have a single effect; it had a distribution of effects organized by the interpretive coherence, cognitive calibration, institutional voice, and role alignment available to each participant.

In answer to the primary research question: teachers at this school did not implement a mandatory AI professional development program. They enacted it. Enactment is what occurs when a plan meets the practitioners responsible for carrying it out, and what occurred here was a structured distribution of strategic responses. Conceptual confusion and implementation overload did not impair learning uniformly. Conceptual confusion shaped interpretation—most powerfully through a coherence failure that taught teachers to read the program as not designed for them, and through a deflection of substantive questions that broke trust early. Implementation overload shaped enactment—though the data refine the construct, showing that it was poor calibration of cognitive demand, in both excess and deficit, rather than demand alone, that drove withdrawal. For many participants, the most adaptive enactment was to maintain structural compliance while withholding the cognitive and professional commitments the program could not productively engage. For a minority—concentrated in administrative roles and around the program's most role-relevant content—the same program afforded genuine, if bounded, productive outcomes. The program's effect was not its design realized; it was the sum of these enactments.

5.2 Limitations

Several limitations should be acknowledged. First, although data collection is now complete—seven interviews, 33 questionnaire responses, observation logs across the five courses, and the communications log—the interview sample remains small and self-selected from questionnaire respondents who volunteered for follow-up. Volunteers may differ systematically from non-volunteers, for instance in the strength of their views about the program. The questionnaire's broader sample partially offsets this, but the interview findings should be read as illustrative depth rather than representative breadth.

Second, the interview sample's linguistic composition is uneven. Six interviews were conducted primarily in English; Carmen's interview included the Spanish-language administrative perspective the sample would otherwise lack, but the transcript also shows points where language difference constrained the exchange. Spanish-speaking participants are well represented in the questionnaire but thinly represented in interview depth. For a bilingual school, this is a meaningful limitation, and it bears on the transferability of the interpretation, particularly given that some evidence (the pre-K assessment bot's value to Spanish-speaking teachers) suggests the program may have been experienced differently across language groups.

Third, the researcher's position as a complete-member researcher (Anderson, 2006) affords contextual access and carries interpretive risk. The autoethnographic logs capture experiential dimensions external observation could not, and they reflect the researcher's frames and commitments. The researcher is also a member of the early-childhood group whose grievances feature prominently in the findings, which raises the risk that shared experience shaped interpretation. The use of four data sources—including the communications log, which is independent of the researcher's experience—is intended to address this by allowing patterns from the researcher's experience to be examined against participants' independent accounts and the documentary record. Following Braun and Clarke (2019, 2022), this multi-source design is understood as multi-perspectival theme development rather than positivist validity-checking; divergences among sources were retained and examined rather than resolved.

Fourth, member checking was not conducted: participants did not review transcripts or interpretations. Analysis relied on a single coder without independent inter-coder comparison. These choices constrain one form of trustworthiness and are reported transparently rather than presented as resolved.

Fifth, the study is a single case—one school, one vendor, one 14-month implementation. The findings do not generalize statistically. They generalize analytically, by illuminating mechanisms (how coherence failure shapes interpretation, how load miscalibration shapes enactment, how role mediates outcomes) that may operate in comparable settings, and by surfacing patterns—particularly the role effect—that warrant testing elsewhere.

Sixth, the program's dual-language LMS structure was disclosed to the researcher only after observation logging was underway, so the autoethnographic data document the English-forward platform. Whether the Spanish-forward platform differed in materials, sequencing, or

facilitation is unknown. The non-disclosure of this structure is itself reported as a finding under vendor opacity, but it also bounds what the autoethnographic data can claim.

5.3 Applications and Implications

The findings carry implications for schools, vendors, and the research field.

For schools, the role effect indicates that uniform, mandatory AI professional development produces uneven outcomes. The participants who benefited most were administrative and support staff, for whom routine-task automation was directly applicable; early-grade and senior-level classroom teachers encountered the sharpest coherence failures. Schools should be cautious about mandating uniform participation across heterogeneous roles and should consider role- or grade-band-specific cohorts from the outset. In the present case, institutional acknowledgment that the program did not serve early-childhood teachers came eight months in, after the mixed-cohort decision had already shaped the experience. The voice findings suggest that mandatory PD without structured teacher input is interpreted as inconsistent with a school's own stated commitments to voice and choice, and that this interpretation depresses engagement; consulting teachers on content, pacing, and grouping before commitment—as Carmen's recommendation of an introductory orientation and preparation time also implies—would address the disengagement pattern more effectively than mid-program reminders about compliance.

For vendors and for schools contracting them, the opacity findings suggest concrete questions to ask before commitment: What evidence base underlies the program's design? What commercial relationships connect the vendor to the tools it recommends? What data-handling policies govern any student data the activities require, and are those activities consistent with the platforms' own terms? What is the cost per participant relative to alternatives such as in-person workshops, internal coaching, or sustained team-based projects? In the present case the first three questions were not visibly asked or answered, and the fourth was raised independently by participants who questioned the program's cost relative to its value. The data-privacy concern—activities requiring sensitive student data to be entered into systems with undisclosed data handling, against a platform's own warning—is the most serious of these and should be treated as a threshold issue, not a detail.

For the research field, the study extends work on teacher experiences of AI professional development (Holmes & Tuomi, 2022; Tripathi et al., 2025) by foregrounding enactment as a distinct category. Research that measures teacher response through completion rates and tool-

adoption metrics will miss the substantive distribution of effects documented here: structural compliance with cognitive withdrawal, productive use alongside rejection of the program's framing, and identity hardening in directions the program did not intend. The refinement of the overload construct—from excessive demand to miscalibrated demand in both directions—offers a more precise analytic tool for examining why mandated PD produces disengagement. Methodologically, the study suggests that participant-researcher positioning, combined with a structured observation instrument that pairs an established PD framework (Desimone, 2009) with separate autoethnographic reflection, is productive for studying mandatory professional development from the inside, provided the researcher's experience is triangulated against independent participant and documentary data.

5.4 Recommendations for Further Research

Several directions follow from these findings. First, multi-site comparative work should test the role effect—the concentration of positive outcomes among administrative staff and of negative outcomes among classroom teachers, particularly at the margins of the grade span—to determine whether it is a general feature of mixed-cohort AI PD or specific to this program. A design varying cohort composition (homogeneous versus mixed by role) and delivery format (in-person versus virtual, sustained versus episodic) would isolate the conditions under which AI PD produces more equitable outcomes.

Second, longitudinal work should follow teachers' AI integration beyond a program's end. The present data capture the within-program experience; the persistence or decay of productive outcomes, and the trajectory of the reverse effect documented in at least one participant, would clarify the durable consequences of mandated AI PD.

Third, research should examine the configuration of conditions that produces disengagement-within-attendance. The present study found this pattern co-occurring with virtual delivery, after-hours scheduling, mixed cohorts, vendor opacity, absent voice, and miscalibrated cognitive load. Which of these conditions, alone or in combination, are most consequential remains open.

Fourth, comparative work on vendor-delivered versus internally led AI professional development would address a question the present findings raise but cannot resolve: several participants preferred internally led, team-based, or in-person alternatives, and one department proposed a specific substitute. Whether schools should contract external vendors for AI PD,

given the alternatives and the opacity risks documented here, warrants empirical comparison across schools that have made different choices.

Fifth, Spanish-language and Latin American teachers' experiences of AI PD merit dedicated study, ideally by Spanish-language researchers. The present study operates primarily in the researcher's working language and could only partially represent the Spanish-language perspective. Given preliminary evidence that the program was experienced differently across language groups, and that AI tools' language handling itself shaped their usefulness, this is a substantive gap rather than a marginal one.

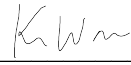
Acknowledgments

I want to acknowledge and thank my supervisor, Dr Emanuele Bardone, for his guidance, his questions, and the conversations that shaped this study. I thank my professional colleagues and the support staff at my school, who shared this experience and who gave their time and candid reflections through the interviews and questionnaire that made this research possible. I am grateful to my employing school for granting permission to conduct the study on site. I also thank my cohort peers in the MA in Educational Technology programme for their feedback and encouragement throughout. Finally, this work is offered in recognition of teachers everywhere, whose tenacity and ethical judgment sustain the work of education.

Author's Declaration

I hereby declare that I have written this thesis independently and that all contributions of other authors and supporters have been referenced. The thesis has been written in accordance with the requirements for graduation theses of the Institute of Education of the University of Tartu and is in compliance with good academic practices.

Name: Kendrah Walker

Signature: 

Date: May 27, 2026

References

- Adams, T. E., Holman Jones, S., & Ellis, C. (2015). *Autoethnography: Understanding qualitative research*. Oxford University Press.
- Anderson, L. (2006). Analytic autoethnography. *Journal of Contemporary Ethnography*, 35(4), 373–395. <https://doi.org/10.1177/0891241605280449>
- Beauchamp, C., & Thomas, L. (2009). Understanding teacher identity: An overview of issues in the literature and implications for teacher education. *Cambridge Journal of Education*, 39(2), 175–189. <https://doi.org/10.1080/03057640902902252>
- Biesta, G., Priestley, M., & Robinson, S. (2015). The role of beliefs in teacher agency. *Teachers and Teaching*, 21(6), 624–640. <https://doi.org/10.1080/13540602.2015.1044325>
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp063oa>
- Braun, V., & Clarke, V. (2019). Reflecting on reflexive thematic analysis. *Qualitative Research in Sport, Exercise and Health*, 11(4), 589–597. <https://doi.org/10.1080/2159676X.2019.1628806>
- Braun, V., & Clarke, V. (2022). *Thematic analysis: A practical guide*. SAGE.
- Chang, H. (2008). *Autoethnography as method*. Left Coast Press.
- Clandinin, D. J., & Connelly, F. M. (2000). *Narrative inquiry: Experience and story in qualitative research*. Jossey-Bass.
- Clark, R. E., Kirschner, P. A., & Sweller, J. (2012). Putting students on the path to learning: The case for fully guided instruction. *American Educator*, 36(1), 6–11.
- Darling-Hammond, L., Hyler, M. E., & Gardner, M. (2017). *Effective teacher professional development*. Learning Policy Institute. <https://doi.org/10.54300/122.311>
- Denzin, N. K., & Lincoln, Y. S. (Eds.). (2018). *The SAGE handbook of qualitative research* (5th ed.). SAGE.
- Desimone, L. M. (2009). Improving impact studies of teachers' professional development: Toward better conceptualizations and measures. *Educational Researcher*, 38(3), 181–199. <https://doi.org/10.3102/0013189X08331140>

- Ellis, C., & Bochner, A. P. (2000). Autoethnography, personal narrative, reflexivity: Researcher as subject. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (2nd ed., pp. 733–768). Sage.
- Ellis, C., Adams, T. E., & Bochner, A. P. (2011). Autoethnography: An overview. *Historical Social Research*, 36(4), 273–290. <https://doi.org/10.12759/hsr.36.2011.4.273-290>
- Fullan, M. (2007). *The new meaning of educational change* (4th ed.). Teachers College Press.
- Fullan, M., & Quinn, J. (2016). *Coherence: The right drivers in action for schools, districts, and systems*. Corwin Press.
- Hargreaves, A., & Fullan, M. (2012). *Professional capital: Transforming teaching in every school*. Teachers College Press.
- Heine, S., & König, L. (2025). Systematic review on the impact of teacher professional development on digital instructional integration and teaching practices. *European Journal of Education*. Advance online publication.
- Holmes, W., & Tuomi, I. (2022). State of the art and practice in AI in education. *European Journal of Education*, 57(4), 542–570. <https://doi.org/10.1111/ejed.12533>
- Hwang, G.-J., & Chen, N.-S. (2023). Editorial position paper: Exploring the potential of generative artificial intelligence in education: Applications, challenges, and future research directions. *Educational Technology & Society*, 26(2). [https://doi.org/10.30191/ETS.202304_26\(2\).0014](https://doi.org/10.30191/ETS.202304_26(2).0014)
- Kelchtermans, G. (2009). Who I am in how I teach is the message: Self-understanding, vulnerability and reflection. *Teachers and Teaching*, 15(2), 257–272. <https://doi.org/10.1080/13540600902875332>
- Koehler, M. J., & Mishra, P. (2009). What is technological pedagogical content knowledge? *Contemporary Issues in Technology and Teacher Education*, 9(1), 60–70.
- Long, D., & Magerko, B. (2020). What is AI literacy? Competencies and design considerations. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1–16). ACM. <https://doi.org/10.1145/3313831.3376727>
- Pinnegar, S., & Hamilton, M. L. (2009). *Self-study of practice as a genre of qualitative research: Theory, methodology, and practice*. Springer.

- Priestley, M., Biesta, G., & Robinson, S. (2015). *Teacher agency: An ecological approach*. Bloomsbury Academic.
- Roshan, M., Iqbal, M. Z., & Qing, E. (2024). Teacher training and professional development for implementing AI-based educational tools. *International Journal of Educational Research and Innovation*. Advance online publication.
- Sims, S., Fletcher-Wood, H., O'Mara-Eves, A., Cottingham, S., Stansfield, C., Goodrich, J., Van Herwegen, J., & Anders, J. (2025). Effective teacher professional development: New theory and a meta-analytic test. *Review of Educational Research*. Advance online publication. <https://doi.org/10.3102/00346543231217480>
- Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, 12(2), 257–285. https://doi.org/10.1207/s15516709cog1202_4
- Timperley, H., Wilson, A., Barrar, H., & Fung, I. (2007). *Teacher professional learning and development: Best evidence synthesis iteration*. New Zealand Ministry of Education.
- Tripathi, S., Kumar, A., Gupta, P., & Sharma, R. (2025). Teaching and learning with AI: A qualitative study on K–12 teachers' use and engagement with artificial intelligence. *Education and Information Technologies*. Advance online publication.
- UNESCO. (2023). *Global education monitoring report 2023: Technology in education – A tool on whose terms?* UNESCO Publishing. <https://doi.org/10.54676/UZQV8501>
- U.S. Department of Education, Office of Educational Technology. (2023). *Artificial intelligence and the future of teaching and learning: Insights and recommendations*. <https://tech.ed.gov>

Appendices

Appendix 1. PD Session Observation Log Template

PD Session Observation Log

Structured Analysis of Professional Development Training Sessions

Primary Framework: **Desimone (2009)** | Optional Expansions: **Darling-Hammond et al. (2017)** | **U.S. DoE OET (2023)**

The blue sections are required for the Desimone crosswalk. The purple sections (Darling-Hammond) and green sections (U.S. DoE policy) are optional expansions.

Rating Scale: 0 = Not observed or not applicable; 1 = Minimal (brief, superficial, or token evidence); 2 = Partial (present but incomplete or inconsistent); 3 = Strong (substantive, well-integrated, clearly intentional).

The Autoethnographic Reflection section at the end.

Course Reference

1. AI for Learning Design | 2. AI for Assessment | 3. AI for Differentiation | 4. AI for Data Analysis | 5. AI for Routine Tasks

SESSION METADATA			
Course Title:		Session #:	
Session Date:		Duration:	
Delivery Format:		Language(s):	
Facilitator(s):			
Materials Used:			

Feature 1: Content Focus

Definition: *PD is centered on subject matter content and how students learn that content, rather than on generic teaching strategies (Desimone, 2009).*

OBSERVATION LOG — CONTENT FOCUS	
Guiding Questions	• Was the session grounded in specific AI tools, concepts, or applications relevant to teaching? • Did the content address how students would encounter or use AI in classroom contexts? • Was pedagogical content knowledge (not just technical

Evidence Observed Evidence Absent or Contradicted Rating	operation) addressed? • Was the AI content contextualized for PreK–12 instructional practice, or was it generic/industry-oriented?
	Circle one: 0 — Not observed 1 — Minimal 2 — Partial 3 — Strong
OPTIONAL EXPANSION (<i>Darling-Hammond, Hyster, & Gardner, 2017</i>)	
Models of Effective Practice	
POLICY ALIGNMENT (<i>U.S. DoE, Office of Educational Technology, 2023</i>)	
Rec #2: Align AI to a Shared Vision for Education Rec #3: Design Using Modern Learning Principles	

Feature 2: Active Learning

Definition: *Teachers engage in meaningful analysis, discussion, practice, and feedback rather than passively receiving information (Desimone, 2009).*

OBSERVATION LOG — ACTIVE LEARNING

Guiding Questions

- Were participants given opportunities to practice with AI tools during the session?
- Was there structured discussion, peer interaction, or collaborative problem-solving?
- Did the format allow for questions, experimentation, or sense-making?
- What proportion of time was lecture/presentation vs. participant-driven activity?

Evidence Observed

Evidence Absent or Contradicted

Rating

Circle one: 0 — Not observed 1 — Minimal 2 — Partial 3 — Strong

OPTIONAL EXPANSION *(Darling-Hammond, Hylar, & Gardner, 2017)*

Coaching & Expert Support

Feedback & Reflection

POLICY ALIGNMENT *(U.S. DoE, Office of Educational Technology, 2023)*

Rec #5: Inform and Involve Educators

Feature 3: Coherence

Definition: *PD is consistent with teachers' knowledge, beliefs, and the policies and practices of their school and system (Desimone, 2009).*

OBSERVATION LOG — COHERENCE

Guiding Questions

- Did the session content connect to participants' existing teaching practices and curricular context?

Evidence Observed Evidence Absent or Contradicted	• Was there alignment between the PD content and the school's stated goals or strategic plan? • Did the content acknowledge participants' prior knowledge or varied entry points? • Were there tensions between what was presented and the institutional or cultural context of the school?
Rating	Circle one: 0 — Not observed 1 — Minimal 2 — Partial 3 — Strong
OPTIONAL EXPANSION (<i>Darling-Hammond, Hyler, & Gardner, 2017</i>)	
Sustained Coherence Across Sessions	
POLICY ALIGNMENT (<i>U.S. DoE, Office of Educational Technology, 2023</i>)	
Foundation 4: Promote Transparency Rec #1: Humans in the Loop	

Feature 4: Duration

Definition: *PD is sustained over time with sufficient contact hours to allow for intellectual engagement, practice, and integration (Desimone, 2009).*

OBSERVATION LOG — DURATION

Guiding Questions	• How long was this session? Was it sufficient for the complexity of the material? • Was there time for both input and processing, or did the session feel rushed? • Is this session part of a longer sequence with built-in follow-up or reinforcement? • Was there an expectation of between-session practice or application?
Evidence Observed	
Evidence Absent or Contradicted	
Rating	Circle one: 0 — Not observed 1 — Minimal 2 — Partial 3 — Strong
OPTIONAL EXPANSION (<i>Darling-Hammond, Hyle, & Gardner, 2017</i>)	
Sustained Duration & Follow-Up	

Feature 5: Collective Participation

Definition: *Groups of teachers from the same school, grade, or department learn together, building shared language and collaborative capacity (Desimone, 2009).*

OBSERVATION LOG — COLLECTIVE PARTICIPATION

Guiding Questions	• Were participants grouped by grade level, department, or shared context—or mixed across roles and levels? • Did the session design leverage participants' shared institutional knowledge? • Were there opportunities for collaborative planning or peer learning? • Did language differences (English/Spanish) affect participation
--------------------------	---

Evidence Observed Evidence Absent or Contradicted	dynamics?
Rating	Circle one: 0 — Not observed 1 — Minimal 2 — Partial 3 — Strong
POLICY ALIGNMENT (<i>U.S. DoE, Office of Educational Technology, 2023</i>)	
Foundation 1: Center People	
Foundation 2: Advance Equity	

Autoethnographic Reflection

AUTOETHNOGRAPHIC REFLECTION

Section to record first-person experience. Guiding Questions

What did I feel during this session? What moments of confusion, resistance, clarity, or engagement stood out?

How did this session affect my sense of professional identity or agency as a teacher?

Where did I notice conceptual confusion or implementation overload—in myself or in colleagues?

Session Summary

SESSION SUMMARY RATINGS

Desimone Feature	Rating (0–3)	D-H Expansion?	DoE Alignment?
1. Content Focus			
2. Active Learning			
3. Coherence			
4. Duration			
5. Collective Participation			
Overall Session Notes:			

Framework References

Desimone, L. M. (2009). Improving impact studies of teachers' professional development: Toward better conceptualizations and measures. *Educational Researcher*, 38(3), 181–199. <https://doi.org/10.3102/0013189X08331140>

Darling-Hammond, L., Hyler, M. E., & Gardner, M. (2017). Effective teacher professional development. Learning Policy Institute. <https://doi.org/10.54300/122.311>

U.S. Department of Education, Office of Educational Technology. (2023). Artificial intelligence and the future of teaching and learning: Insights and recommendations. Washington, DC. <https://tech.ed.gov/ai-future-of-teaching-and-learning/>

Appendix 2. PD Observation Reference Card

PD Observation Log — Quick Reference Card

Virtual/Zoom Sessions with Breakout Rooms | Co-located & Remote Participants

Rating: 0 = Not observed 1 = Minimal 2 = Partial 3 = Strong

DESIMONE FEATURE	POSITIVE EVIDENCE — LOOK FOR	RED FLAGS — DOCUMENT	OPTIONAL EXPANSIONS (D-H / DoE)
1. Content Focus	✓ Facilitator connects AI tool to how students learn, not just how it works ✓ Examples drawn from PreK–12 contexts, not corporate/industry ✓ Pedagogical content knowledge addressed alongside technical operation ✓ Screenshare shows a real lesson taught with the tool, not a marketing demo	- X Demo shows features/buttons with no curricular connection - X All examples are from generic or U.S. monolingual settings - X Tool presented as an end in itself, no link to learning goals - X Co-located teachers can't see screenshare clearly on shared laptop	D-H: Models of Effective Practice Did facilitator show an actual taught lesson vs. a vendor demo? Were student work samples included? DoE Rec #2: Shared Vision Any reference to your school's mission, strategic plan, or accreditation standards? Or content identical across all client schools? DoE Rec #3: Learning Principles Does facilitator use formative checks, scaffolding, differentiated breakout tracks? Or is it a slide deck with a tacked-on activity?
2. Active Learning	✓ Breakout task is meaningful: "redesign your assessment using this tool" ✓ Facilitator drops into breakout rooms, stays, gives targeted feedback ✓ Structured reflection: Padlet, poll, shared doc, report-out with prompts ✓ Practice task has clear instructions and success criteria	- X Breakout task is vague: "click around for five minutes" - X Facilitator joins breakout room for 30 seconds, says "how's it going?" and leaves - X "Any reflections?" into silence → "Great" → moves on - X Co-located group sits near each other but works independently on separate devices	D-H: Coaching & Expert Support Time-stamp facilitator's breakout room visits. How long? What was said? Was support substantive? D-H: Feedback & Reflection Did anyone review participants' practice output? Or did the activity just end when the timer closed rooms? DoE Rec #5: Inform & Involve Were teachers asked what they already know? Did the task invite professional judgment or just procedural compliance?
3. Coherence	✓ Facilitator references your school's grading system, LMS, or curriculum ✓ Session opens with connection to the previous session ✓ Materials from prior sessions are linked and accessible ✓ Facilitator acknowledges varied entry points and comfort levels	- X Session starts cold, no reference to what came before - X Recommends tools requiring infrastructure your school lacks - X Content feels designed for a generic audience, not your context - X Chat comments: "Wait, didn't we cover this?" or "How does this connect?"	D-H: Sustained Coherence Does this session build on the prior one, or does each module exist in isolation? DoE Found. 4: Transparency Is the rationale for vendor, curriculum, or sequencing ever explained? Or is it a black box? DoE Rec #1: Humans in the Loop "When you implement this" (compliance) vs. "If you determine this fits" (agency). Which framing?
4. Duration	✓ Pacing allows input AND processing time ✓ Between-session tasks assigned and actually debriefed next time ✓ Sufficient spacing between modules to allow classroom experimentation ✓ Session ends with clear bridge to what comes next	- X Session feels rushed at the end; final topic gets 3 minutes - X "Try this before next time" but next session never references it - X Five-day intensive: new AI application every day, zero practice time - X Too many tools/concepts crammed into 90 minutes	D-H: Sustained Duration & Follow-Up Are there structured follow-up activities, check-ins, or iterative practice cycles? Or is each session terminal?
5. Collective Participation	✓ Breakout groups organized by grade band, department, or shared context ✓ Tasks leverage shared institutional knowledge	- X K teacher and HS physics teacher receive identical advice - X Breakout groups are random with no shared context	DoE Found. 1: Center People Are different roles (teacher, specialist, admin) acknowledged? Or is everyone treated as interchangeable?

	✓ Language differences explicitly accommodated in grouping ✓ Collaborative planning time for participants who teach similar students	- X Spanish-dominant colleagues participate less or disengage on Zoom - X One bilingual teacher informally translates for colleagues (invisible labor)	DoE Found. 2: Advance Equity Materials in both languages? AI tools demoed in Spanish? Digital literacy gaps accommodated? Platform itself a barrier?
AUTOETHNOGRAPHIC CUES	• Identity: When was your expertise acknowledged or dismissed? Did you feel like a learner or like your experience was overlooked? • Confusion: Not tool confusion — purpose confusion. Why would I use this? When? What problem does it solve? Can I articulate the difference between three tools presented? • Overload: "This is the fourth new tool this week and I haven't tried the first one." Sighing, phone-checking, off-mic side conversations, mentally checking out. • Agency: Did the mandatory framing surface emotionally? Excitement vs. resignation? "Are we required to do this?" whispered off-camera. • Peer signals: Colleagues' body language, chat messages, informal debriefs after session ends — especially from co-located groups.		

Blue = Desimone (2009, required) | Purple = Darling-Hammond et al. (2017, optional) | Green = U.S. DoE OET (2023, optional)

Appendix 3. Interview Protocol

Semi-Structured Interview Protocol

Condensed for 30-Minute Interviews — 12 Core Questions

Protocolo condensado para entrevistas de 30 minutos — 12 preguntas principales

Interview Structure

30-minute semi-structured interview. It contains 12 core questions with optional follow-up probes. Each question is presented in English followed by its Spanish translation. Follow-up probes are listed beneath each question; use them as needed based on the participant's response.

Opening

Constructs: Autoethnographic method; lived experience (Chang, 2008; Ellis et al., 2011)

Q1. Think back to the beginning of the AI professional development program. What do you remember about your initial reaction, and how would you describe the experience overall?

Piense en el inicio del programa de desarrollo profesional en IA. ¿Qué recuerda de su reacción inicial y cómo describiría la experiencia en general?

→ Was there a particular moment early on that has stayed with you?

¿Hubo algún momento al inicio que se le haya quedado grabado?

→ If you were describing this experience to a colleague at another school, what would you want them to understand?

Si estuviera describiendo esta experiencia a un colega de otra escuela, ¿qué querría que entendiera?

AI for Learning Design — The Foundational Course

Constructs: AI literacy; TPACK; conceptual foundation; coherence (Long & Magerko, 2020; Koehler & Mishra, 2009; Fullan & Quinn, 2016)

Q2. The first course, AI for Learning Design, covered broad topics like evaluating AI tools, ethical use, and opportunities in education. What stands out to you when you think back to that course?

El primer curso, IA para Diseño de Aprendizaje, cubrió temas amplios como la evaluación de herramientas de IA, el uso ético y las oportunidades en educación. ¿Qué le viene a la mente cuando piensa en ese curso?

→ How well did you feel you understood the ideas presented in that course?

¿Qué tan bien sintió que comprendió las ideas presentadas en ese curso?

→ Looking back, how did that first course shape the way you approached the courses that followed?

Mirando hacia atrás, ¿cómo influyó ese primer curso en la forma en que abordó los cursos siguientes?

Q2b. The course included objectives around ethical and responsible AI adoption. How did you experience that part of the training?

El curso incluyó objetivos sobre la adopción ética y responsable de la IA. ¿Cómo vivió esa parte de la capacitación?

→ Did the ethical framing feel relevant to your teaching context?

¿Le pareció relevante el enfoque ético para su contexto de enseñanza?

→ Were there ethical considerations you felt were missing?

¿Hubo consideraciones éticas que sintió que faltaban?

Q2c. One objective of the course was to collaborate with colleagues to plan and reflect on AI tool integration. What was that collaborative element like for you?

Uno de los objetivos del curso era colaborar con colegas para planificar y reflexionar sobre la integración de herramientas de IA. ¿Cómo fue ese elemento colaborativo para usted?

→ How did working with colleagues from different grade levels or roles affect that process?

¿Cómo afectó trabajar con colegas de diferentes niveles o roles a ese proceso?

Conceptual Clarity and Understanding

Constructs: Conceptual confusion; cognitive load (Sweller, 1988; Clark et al., 2012; Long & Magerko, 2020)

Q3. Across the entire program, what stands out as something you felt you understood well? And were there ideas that were harder to make sense of?

A lo largo de todo el programa, ¿qué es lo que siente que comprendió bien? ¿Y hubo ideas que fueron más difíciles de asimilar?

→ What made the difference between content that was clear and content that wasn't?

¿Qué marcaba la diferencia entre el contenido que era claro y el que no lo era?

→ Did you see connections across the five courses, or did they feel more like separate topics?

¿Vio conexiones entre los cinco cursos, o se sintieron más como temas separados?

Relevance and Implementation

Constructs: Implementation overload; effective PD criteria (Darling-Hammond et al., 2017; Roshan et al., 2024; Tripathi et al., 2025)

Q4. How would you describe the relationship between what was covered in the training and the realities of your day-to-day work?

¿Cómo describiría la relación entre lo que se cubrió en la capacitación y la realidad de su trabajo diario?

→ Was there a specific tool or idea you tried to bring into your practice? What happened?

¿Hubo alguna herramienta o idea específica que intentó incorporar a su práctica? ¿Qué ocurrió?

→ Were there things you felt were missing, or things that felt well-suited to your context?

¿Hubo cosas que sintió que faltaban, o cosas que se adaptaban bien a su contexto?

Pace, Volume, and Workload

Constructs: Implementation overload; cognitive load; change fatigue (Sweller, 1988; Fullan, 2007; Heine & König, 2025)

Q5. The program ran from January 2025 through March 2026 and included five courses. How did you experience the pacing and the amount of content over that time?

El programa se desarrolló desde enero de 2025 hasta marzo de 2026 e incluyó cinco cursos.

¿Cómo vivió el ritmo y la cantidad de contenido durante ese tiempo?

→ Were there times when the pace felt about right, or times when it felt too fast?

¿Hubo momentos en que el ritmo le pareció adecuado, o momentos en que sintió que era demasiado acelerado?

→ How did you manage the workload alongside your regular responsibilities?

¿Cómo manejó la carga de trabajo junto con sus responsabilidades habituales?

Q5b. Did your experience with the program change as it progressed from course to course? In what way?

¿Su experiencia con el programa cambió a medida que avanzaba de curso en curso? ¿De qué manera?

→ Were there points when your energy or attention shifted?

¿Hubo momentos en que su energía o atención cambió?

Agency and Voice

Constructs: Teacher agency; ecological agency model (Biesta et al., 2015; Priestley et al., 2015)

Q6. The professional development was required by the school and delivered by Unconstrained.co. How much say did you feel you had in how you engaged with it?

El desarrollo profesional fue requerido por la escuela e impartido por Unconstrained.co. ¿Cuánta capacidad de decisión sintió que tuvo en cómo participaba?

→ Did the mandatory nature or the external delivery affect how you engaged with the content?

¿El carácter obligatorio o la entrega externa afectaron la forma en que se involucró con el contenido?

→ Were there ways you adapted the content to make it work for your particular context?

¿Hubo formas en que adaptó el contenido para que funcionara en su contexto particular?

Professional Identity

Constructs: Teacher identity; self-understanding; TPACK (Beauchamp & Thomas, 2009; Kelchtermans, 2009; Koehler & Mishra, 2009; Hargreaves & Fullan, 2012)

Q7. Has this experience affected how you see yourself as a professional — as a teacher, as a learner, or in your relationship to technology?

¿Ha afectado esta experiencia la forma en que se ve a sí mismo/a como profesional — como docente, como aprendiz, o en su relación con la tecnología?

→ Some teachers feel more confident after PD; others feel less certain. Where would you place your experience?

Algunos docentes se sienten más seguros después del desarrollo profesional; otros se sienten menos seguros. ¿Dónde ubicaría su experiencia?

→ Were there aspects of your existing practice that felt affirmed? Aspects that felt challenged?

¿Hubo aspectos de su práctica actual que se sintieron reafirmados? ¿Aspectos que se sintieron cuestionados?

Institutional and Social Context

Constructs: Coherence; institutional support; AI adoption context (Fullan & Quinn, 2016; Holmes & Tuomi, 2022; UNESCO, 2023)

Q8. How did you understand the school's reasons for adopting this program, and did that align with what you experienced in the training?

¿Cómo entendió las razones de la escuela para adoptar este programa, y coincidió eso con lo que experimentó en la capacitación?

→ How would you describe the support you received from school leadership during this process?

¿Cómo describiría el apoyo que recibió de parte de la dirección escolar durante este proceso?

→ What was the experience like among your colleagues — how did conversations about the program unfold?

¿Cómo fue la experiencia entre sus colegas — cómo se dieron las conversaciones sobre el programa?

Reflective Closure

Constructs: Meaning-making; participant authority (Pinnegar & Hamilton, 2009; Adams et al., 2015)

Q9. If the school were to do something like this again, what would you want to be different? What would you want to stay the same?

Si la escuela volviera a hacer algo así, ¿qué le gustaría que fuera diferente? ¿Qué le gustaría que se mantuviera igual?

→ What would make the biggest difference for you personally?

¿Qué marcaría la mayor diferencia para usted personalmente?

Q10. Is there anything about your experience with this program that we haven't talked about that feels important to you?

¿Hay algo sobre su experiencia con este programa de lo que no hayamos hablado y que sienta que es importante?

Reference Mapping: Questions to Constructs

Domain	Thesis Constructs	Key References	Q#
Opening	Autoethnographic method; lived experience	Chang (2008); Ellis et al. (2011)	Q1
AI for Learning Design	AI literacy; TPACK; coherence	Long & Magerko (2020); Koehler & Mishra (2009); Fullan & Quinn (2016)	Q2
Conceptual Clarity	Conceptual confusion; cognitive load	Sweller (1988); Clark et al. (2012)	Q3
Relevance	Implementation overload; effective PD	Darling-Hammond et al. (2017); Roshan et al. (2024); Tripathi et al. (2025)	Q4
Pace & Volume	Cognitive load; change fatigue	Sweller (1988); Fullan (2007); Heine & König (2025)	Q5
Agency & Voice	Teacher agency; ecological model	Biesta et al. (2015); Priestley et al. (2015)	Q6
Professional Identity	Identity; self-understanding; TPACK	Beauchamp & Thomas (2009); Kelchtermans (2009); Hargreaves & Fullan (2012)	Q7
Institutional Context	Coherence; institutional support	Fullan & Quinn (2016); Holmes & Tuomi (2022); UNESCO (2023)	Q8
Reflective Closure	Meaning-making; participant authority	Pinnegar & Hamilton (2009); Adams et al. (2015)	Q9–Q10

Appendix 4. Participant Online Questionnaire

Appendix 4. Online Questionnaire (English–Spanish)

Note. The questionnaire was administered through Microsoft Forms with name recording disabled. Both English and Spanish text were presented within each item. The instrument is reproduced below as administered. The name of the external vendor has been redacted and replaced with "[the vendor]" to preserve anonymity, consistent with the body of the thesis.

Consent / Information for Participants

This questionnaire is part of a master's research study exploring how teachers experience mandated professional development on artificial intelligence tools. Your participation is entirely voluntary, and all responses are completely anonymous. No identifying information will be collected. There are no right or wrong answers — your honest perspective is what matters most.

Este cuestionario forma parte de un estudio de investigación de maestría que explora cómo los docentes viven el desarrollo profesional obligatorio sobre herramientas de inteligencia artificial. Su participación es completamente voluntaria y todas las respuestas son anónimas. No se recopilará ninguna información identificable. No hay respuestas correctas ni incorrectas — lo que más importa es su perspectiva honesta.

By completing this questionnaire, you indicate your voluntary consent to participate in this study. / Al completar este cuestionario, usted indica su consentimiento voluntario para participar en este estudio.

This questionnaire refers specifically to the artificial intelligence professional development program delivered by [AI PD vendor] between January 2025 and March 2026. The program consisted of the following five courses: AI for Learning Design, AI for Assessment, AI for Differentiation, AI for Data Analysis, and AI for Routine Tasks. Your responses should reflect your experiences with this particular program. This questionnaire does not address your personal use of AI tools outside of this training, nor any other professional development you may have participated in during this period. // Este cuestionario se refiere específicamente al programa de desarrollo profesional en inteligencia artificial impartido por [AI PD vendor] entre enero de 2025 y marzo de 2026. El programa consistió en los siguientes cinco cursos: IA para Diseño de Aprendizaje, IA para Evaluación, IA para Diferenciación, IA para Análisis de Datos e IA para Tareas Rutinarias. Sus respuestas deben reflejar sus experiencias con este programa en particular. Este cuestionario no aborda su uso personal de herramientas de IA

fuera de esta capacitación, ni ningún otro desarrollo profesional en el que haya participado durante este período.

Section 1: Background Information / Información General

1. What is your primary teaching role or grade level? / ¿Cuál es su rol docente principal o nivel de grado?

- Early Childhood / Kindergarten // Educación Inicial / Kindergarten
- Elementary (Grades 1–5) // Primaria (1.º–5.º)
- Middle School (Grades 6–8) // Secundaria Baja (6.º–8.º)
- High School (Grades 9–12) // Secundaria Alta (9.º–12.º)
- Specialist / Elective // Especialista / Electiva
- Administrative / Support // Administrativo / Apoyo

2. How many years have you been teaching? / ¿Cuántos años lleva enseñando?

- 0–2 years // 0–2 años
- 3–5 years // 3–5 años
- 6–10 years // 6–10 años
- 11–20 years // 11–20 años
- More than 20 years // Más de 20 años

3. Before this professional development program, how would you describe your familiarity with AI tools? / Antes de este programa de desarrollo profesional, ¿cómo describiría su familiaridad con las herramientas de IA?

- Very familiar // Muy familiarizado/a
- Moderately familiar // Moderadamente familiarizado/a
- Slightly familiar // Algo familiarizado/a
- Not at all familiar // Nada familiarizado/a

Section 2: Experiences with the AI PD Program / Experiencias con el Programa

4. To what extent did you feel prepared to engage with the content of this professional development program? / ¿En qué medida se sintió preparado/a para participar en los contenidos de este programa de desarrollo profesional?

- Well prepared // Bien preparado/a
- Somewhat prepared // Algo preparado/a
- Somewhat unprepared // Poco preparado/a
- Not at all prepared // Nada preparado/a

5. How relevant did you find the content of the professional development to your daily teaching practice? / ¿Qué tan relevante consideró el contenido del desarrollo profesional para su práctica docente diaria?

- Very relevant // Muy relevante
- Moderately relevant // Moderadamente relevante
- Slightly relevant // Poco relevante
- Not at all relevant // Nada relevante

6. What aspects of the professional development program did you find most meaningful or useful? Please explain. / ¿Qué aspectos del programa de desarrollo profesional encontró más significativos o útiles? Por favor explique. (*Open-ended response / Respuesta abierta*)

7. Were there aspects of the program that you found challenging, confusing, or difficult to connect to your work? Please describe. / ¿Hubo aspectos del programa que encontró desafiantes, confusos o difíciles de conectar con su trabajo? Por favor describa. (*Open-ended response / Respuesta abierta*)

Section 3: Course-Level Perceptions / Percepciones por Curso

This section asks about your experience with each of the five courses in the AI professional development program. To help you recall each course, brief descriptions are provided below. // Esta sección pregunta sobre su experiencia con cada uno de los cinco cursos del programa de desarrollo profesional en IA. Para ayudarlo a recordar cada curso, se proporcionan descripciones breves a continuación.

AI for Learning Design // IA para Diseño de Aprendizaje — AI opportunities, tool evaluation, and ethical adoption // Oportunidades de IA, evaluación de herramientas y adopción ética

AI for Assessment // IA para Evaluación — AI-assisted assessments and feedback // Evaluaciones y retroalimentación asistidas por IA

AI for Differentiation // IA para Diferenciación — AI tools for diverse learner needs // Herramientas de IA para necesidades diversas de estudiantes

AI for Data Analysis // IA para Análisis de Datos — Data collection, automation, and trend analysis with AI // Recopilación, automatización y análisis de tendencias con IA

AI for Routine Tasks // IA para Tareas Rutinarias — Automating recurring professional tasks with AI // Automatización de tareas profesionales recurrentes con IA

8. For each course, how clear was the content? / Para cada curso, ¿qué tan claro fue el contenido?

*Rate each of the five courses: Very clear // Bastante claro // Slightly clear // Not at all clear
(Muy claro / Bastante claro / Poco claro / Nada claro)*

- Artificial Intelligence for Learning Design // Inteligencia Artificial para Diseño de Aprendizaje
- Artificial Intelligence for Assessment // Inteligencia Artificial para Evaluación
- Artificial Intelligence for Differentiation // Inteligencia Artificial para Diferenciación
- Artificial Intelligence for Data Analysis // Inteligencia Artificial para Análisis de Datos
- Artificial Intelligence for Routine Tasks // Inteligencia Artificial para Tareas Rutinarias

9. For each course, how relevant was the content to your practice? / Para cada curso, ¿qué tan relevante fue el contenido para su práctica?

*Rate each of the five courses: Very relevant // Mostly relevant // Slightly relevant // Not at all relevant
(Muy relevante / Bastante relevante / Poco relevante / Nada relevante)*

- Artificial Intelligence for Learning Design // Inteligencia Artificial para Diseño de Aprendizaje
- Artificial Intelligence for Assessment // Inteligencia Artificial para Evaluación
- Artificial Intelligence for Differentiation // Inteligencia Artificial para Diferenciación
- Artificial Intelligence for Data Analysis // Inteligencia Artificial para Análisis de Datos
- Artificial Intelligence for Routine Tasks // Inteligencia Artificial para Tareas Rutinarias

10. Rank the five courses from most valuable to least valuable to your professional practice (1 = most valuable). / Ordene los cinco cursos del más valioso al menos valioso para su práctica profesional (1 = más valioso). (Ranking of the five courses listed above / Clasificación de los cinco cursos)

11. Was there a particular course that stood out to you — either positively or negatively? Please describe what made it stand out and how it affected your engagement with the program. / ¿Hubo algún curso que se destacara para usted, ya sea de forma positiva o negativa? Por favor describa qué lo hizo destacar y cómo afectó su participación en el programa. (Open-ended response / Respuesta abierta)

Section 4: Agency and Professional Identity / Agencia e Identidad Profesional

12. To what extent did you feel you had a voice or choice in how you engaged with this professional development? / ¿En qué medida sintió que tenía voz o elección en la forma en que participaba en este desarrollo profesional?

- Significant voice // Voz significativa
- Some voice // Algo de voz
- Very little voice // Muy poca voz
- No voice at all // Ninguna voz

13. How, if at all, has participating in this program affected the way you think about yourself as a teacher? / ¿De qué manera, si es que alguna, ha afectado su participación en este programa a la forma en que se percibe como docente? (*Open-ended response / Respuesta abierta*)

14. Which of the following best describes how you felt about integrating AI tools into your teaching as a result of this program? (Select all that apply) / ¿Cuál de las siguientes opciones describe mejor cómo se sintió respecto a integrar herramientas de IA en su enseñanza como resultado de este programa? (Seleccione todas las que correspondan)

- Enthusiastic and motivated // Entusiasmado/a y motivado/a
- Confident in applying new tools // Seguro/a para aplicar nuevas herramientas
- Curious but uncertain // Curioso/a pero inseguro/a
- Neutral / No significant change // Sin cambio significativo
- Unsupported or overwhelmed // Sin apoyo o abrumado/a

15. Did participating in this program change the way you use AI tools in your classroom? Please describe what changed, or explain why things stayed the same. / ¿Participar en este programa cambió la forma en que usa las herramientas de IA en su aula? Por favor describa qué cambió, o explique por qué las cosas permanecieron igual. (*Open-ended response / Respuesta abierta*)

Section 5: Institutional Context / Contexto Institucional

16. To what extent did you feel supported by school leadership in your engagement with this professional development? / ¿En qué medida se sintió apoyado/a por la dirección escolar en su participación en este desarrollo profesional?

- Strongly supported // Apoyo significativo
- Moderately supported // Apoyo moderado
- Minimally supported // Apoyo mínimo
- Not supported at all // Sin ningún apoyo

17. How would you describe the relationship between this professional development program and your sense of professional accountability? / ¿Cómo describiría la relación entre este programa de desarrollo profesional y su sentido de responsabilidad profesional?

- It increased my sense of accountability in a positive way // Incrementó mi sentido de responsabilidad de manera positive
- It felt like a compliance exercise with little connection to my professional values // Se sintió como un ejercicio de cumplimiento con poca conexión con mis valores profesionales
- It created tension between institutional expectations and my personal teaching values // Creó tensión entre las expectativas institucionales y mis valores personales
- It had little effect on my sense of professional accountability // Tuvo poco efecto en mi sentido de responsabilidad profesional
- Other // Otro

18. If other was selected, please describe // Otro (por favor, descríballo a continuación)
(Open-ended response / Respuesta abierta)

19. Is there anything about the institutional context of this program — such as how it was communicated, structured, or supported — that you would like to share? / ¿Hay algo sobre el contexto institucional de este programa — como la forma en que se comunicó, se estructuró o se apoyó — que le gustaría compartir? (Open-ended response / Respuesta abierta)

Section 6: Reflections and Looking Forward / Reflexiones y Perspectivas

20. Overall, how would you assess your experience with this professional development program? / En general, ¿cómo evaluaría su experiencia con este programa de desarrollo profesional?

- Very positive // Muy positiva
- Somewhat positive // Algo positiva
- Neutral // Neutral
- Somewhat negative // Algo negativa
- Very negative // Muy negativa

21. Is there anything else you would like to share about your experience with AI professional development that has not been addressed in this questionnaire? / ¿Hay algo

más que quisiera compartir sobre su experiencia con el desarrollo profesional en IA que no se haya abordado en este cuestionario? (*Open-ended response / Respuesta abierta*)

Optional: Follow-Up Interview / Opcional: Entrevista de Seguimiento

As part of this research study, I am inviting a small number of participants to take part in a voluntary, one-on-one conversation with me, the researcher, to explore your experiences in greater depth. The interview would last approximately 30–45 minutes and can be conducted in person at a location convenient to you, or virtually via a platform of your choice (e.g., Zoom or Google Meet). The interview would be conducted in English, but participants can answer in their first language. All interviews will be audio-recorded only with your explicit permission, and you may withdraw at any time without consequence. If you choose to provide your name and contact information below, that information will be stored separately from your questionnaire responses and used only to arrange the interview. // Como parte de este estudio de investigación, invito a un pequeño grupo de participantes a participar en una conversación individual y voluntaria conmigo, la investigadora, con el fin de profundizar en sus experiencias. La entrevista duraría aproximadamente entre 30 y 45 minutos y puede realizarse en persona en un lugar que le resulte conveniente, o de manera virtual a través de la plataforma que usted elija (por ejemplo, Zoom o Google Meet). La entrevista se llevará a cabo en inglés, pero los participantes pueden responder en su lengua materna. Todas las entrevistas se grabarán en audio solo con su permiso explícito, y usted puede retirarse en cualquier momento sin consecuencias. Si elige proporcionar su nombre e información de contacto a continuación, dicha información se almacenará por separado de sus respuestas al cuestionario y se utilizará únicamente para coordinar la entrevista.

22. Would you be willing to participate in a follow-up interview? / ¿Estaría dispuesto/a a participar en una entrevista de seguimiento?

- Yes, I volunteer to participate in an interview // Sí, me ofrezco como voluntario/a para participar en una entrevista
- No thank you // No gracias

23. If yes: Your name / Su nombre; Your email address to be contacted / Su dirección de correo electrónico para ser contactado/a (*Short answer / Respuesta breve*)

24. Which interview format would you prefer? / ¿Qué formato de entrevista preferiría?

- In-person interview / Entrevista presencial
- Online interview / Entrevista en línea
- I am open to either format / Estoy dispuesto/a a cualquiera de los dos formatos

Non-exclusive licence to reproduce the thesis and make the thesis public

I, Kendrah Walker,

1. grant the University of Tartu a free permit (non-exclusive licence) to reproduce, for the purpose of preservation, including for adding to the digital archives of the University of Tartu until the expiry of the term of copyright, my thesis *Interpreting and Enacting AI Professional Development: An Autoethnographic Study of Teacher Experiences of a Mandatory Program*, supervised by Dr Emanuele Bardone;
2. grant the University of Tartu a permit to make the thesis specified in point 1 available to the public via the web environment of the University of Tartu, including via the digital archives, under the Creative Commons licence CC BY-NC-ND 4.0, which allows, by giving appropriate credit to the author, to reproduce, distribute the work and communicate it to the public, and prohibits the creation of derivative works and any commercial use of the work until the expiry of the term of copyright;
3. am aware of the fact that the author retains the rights specified in points 1 and 2;
4. confirm that granting the non-exclusive licence does not infringe other persons' intellectual property rights or rights arising from the personal data protection legislation.

Kendrah Walker

May 27, 2026

