



From Paper-Based Monitoring to AI-Assisted Alignment Feedback: Design and Development of LP Monitor for Lesson Plan Supervision

Fernando B. Enad¹

¹San Agustin National High School, Department of Education–Schools Division of Bohol, Sagbayan, Bohol, Philippines

ORCID: <https://orcid.org/0009-0000-9837-0293>

Corresponding author: Fernando B. Enad • fernando.enad@sanhs.edu.ph

Received	28 August 2026	Revised	30 August 2026
Accepted	31 August 2026	Published	01 September 2026
Volume/Issue	Vol. 1, No. 1 (2026)	DOI	Pending assignment
Section	Design and Development Research Article	License	CC BY 4.0

ABSTRACT

Lesson plans are intended to guide instruction and provide a basis for professional reflection and instructional supervision, and their instructional value depends on the coherence among the intended learning competency, objectives, learning experiences, and assessment rather than on the mere presence of these components in a document (Biggs, 1996; Martone & Sireci, 2009). However, the processes used to submit, monitor, and review lesson plans may themselves create administrative and instructional difficulties. Paper-based monitoring requires recurrent printing, physical document movement, and manual tracking, and it frequently delays supervisory feedback until after a lesson has already been taught, which reduces much of that feedback's formative value (Hattie & Timperley, 2007; Shute, 2008). High submission volumes further limit the depth of structural analysis a supervisor can perform, often reducing review to compliance checking rather than instructional coaching. This design and development research (DDR) study documents the problem analysis, co-design, and formative tryout of LP Monitor, a web-based lesson-plan supervisory system built to digitize submission and monitoring while providing automated, rubric-based feedback on instructional alignment, consistent with current Philippine guidance on lesson planning (Department of Education, 2026b) and on human-centered use of artificial intelligence in basic education (Department of Education, 2026a). Problem identification was conducted through a focus group discussion held during a joint Learning Action Cell session with 18 teachers from the Languages (English and Filipino) and Science departments at San Agustin National High School, Bohol, Philippines; the session was documented through minutes independently recorded by Master Teachers from the requesting departments, and thematic coding was performed by a single coder. The same 18 teachers, who were not under the author's direct supervisory authority at any stage of the study, took part in a subsequent co-design session that shaped the system's requirements and interface, including the integration of a large language model that evaluates three alignment relationships anchored on an author-developed Scaffolding Objectives construct (Enad, 2026) informed by scaffolding theory (van de Pol et al., 2010) and by recent work applying large language models to constructive-alignment tasks (Pereira et al., 2024): Competency-to-Scaffolding Objective, Scaffolding Objective-to-Learning Experience, and Scaffolding Objective-to-Assessment. During a four-week formative tryout in August 2026, the same cohort submitted 91 lesson plans through the system, of which 50 received an automated recommendation for revision and 41 an approval recommendation; teachers used the automated feedback to revise and resubmit plans at their own discretion. The study is exploratory and descriptive: no control group, pre-registered hypothesis test, or pre-tryout baseline was used, and the automated alignment engine and the underlying Scaffolding Objectives construct have not undergone external psychometric validation. The findings describe LP Monitor's developmental logic and preliminary operational patterns rather than definitive claims of effectiveness, and the paper closes with a research agenda for psychometric validation, controlled comparison, and longitudinal follow-up.

Keywords: lesson planning, instructional supervision, artificial intelligence, constructive alignment, Scaffolding Objectives, developmental feedback, educational technology, design and development research, focus group discussion, co-design

1. Introduction

Lesson planning is a foundational instructional practice that enables teachers to structure learning goals, align pedagogical strategies, and prepare assessments, and its instructional value rests on the coherence among these components rather than on their mere presence in a document (Biggs, 1996; Martone & Sireci, 2009). In many public school systems, including the Department of Education in the Philippines, regular submission and supervisory review of lesson plans serve as primary mechanisms for quality assurance and instructional support. However, conventional monitoring workflows—often reliant on physical paper submissions or fragmented file-sharing folders—present systemic operational friction.

From an administrative perspective, physical handling, manual routing, and paper record-keeping generate significant delays. Supervisory feedback is frequently returned after a planned lesson has already been taught, stripping the review process of its formative value, since the developmental power of feedback depends considerably on its timing relative to the action it is meant to improve (Hattie & Timperley, 2007; Shute, 2008). From an instructional perspective, high submission volumes limit the depth of structural analysis supervisors can perform, often reducing reviews to compliance checking rather than meaningful instructional coaching. Current Philippine policy reinforces the need for a remedy: DepEd Order No. 016, s. 2026 revised national guidelines on lesson planning to emphasize flexible, reflective, and instructionally meaningful planning over rigid template compliance (Department of Education, 2026b), placing renewed importance on monitoring systems that can return feedback before, rather than only after, a lesson is delivered.

To address these challenges, this study explores the application of educational technology to lesson-plan supervision. By leveraging Design and Development Research (DDR), this project details the creation and field testing of LP Monitor, a web-based platform designed to streamline monitoring workflows and provide real-time automated structural feedback, consistent with national guidance encouraging complementary, human-centered use of artificial intelligence in basic education (Department of Education, 2026a).

1.1 Statement of the Problem

This study aimed to design, develop, and formatively evaluate LP Monitor, a web-based lesson-plan supervisory system. Specifically, it sought to answer the following research questions:

- 1. Contextual Problem Identification:** What operational friction points and instructional alignment challenges do secondary school teachers in the Languages and Science departments experience within conventional, paper-based lesson-plan monitoring workflows?
- 2. System Architecture and Requirements:** What technical requirements, system architecture, and interface workflows emerge from practitioner co-design to address these identified operational bottlenecks?
- 3. Operational Tryout Dynamics:** How does the LP Monitor system perform during a four-week formative tryout in terms of submission volume, system workflow activity, feedback generation, and teacher resubmission patterns?

2. Method

2.1 Research Design

The study employed Design and Development Research (DDR) to guide the conceptualization, creation, and preliminary field testing of an educational software prototype (Richey & Klein, 2007, 2014). The design comprised four sequential phases: (1) contextual problem identification, (2) requirements formulation and co-design, (3) prototype development, and (4) an exploratory formative tryout.

The study was exploratory and descriptive. No control group, pre-registered hypothesis test, or pre-tryout quantitative baseline was utilized. Consequently, findings describe the developmental logic and preliminary operational patterns rather than definitive causal effects.

2.2 Participants and Setting

The study was conducted at San Agustin National High School, a public secondary school in Bohol, Philippines. A single, continuous cohort of eighteen ($n = 18$) teachers from the Languages (English and Filipino) and Science departments participated across all study phases:

Problem Identification Phase (June 24, 2026): The 18 teachers participated in an initial Focus Group Discussion (FGD) during a requested joint Learning Action Cell (LAC) session.

Co-Design and Formative Tryout Phases (July–August 2026): The same cohort of 18 Languages and Science teachers completed the co-design refinements and executed the four-week formative tryout in August 2026.

Phase 1 (June 24, 2026): Incidental LAC FGD → $n = 18$ Teachers (Languages & Science Departments)

Phases 2 & 4 (July–August 2026): Co-Design & Tryout → $n = 18$ Teachers (Languages & Science Departments)

Methodological Risk Note: The 18 participating teachers were not under the direct supervisory authority of the author at any point during the study. This non-supervisory relationship mitigated power dynamics and supervisory bias during both problem identification and system evaluation. However, because participation was voluntary and operationalized outside formal evaluative supervision, system engagement patterns may differ from mandatory administrative workflows.

2.3 Problem Identification Procedure

The initial FGD was conducted incidentally during a joint LAC session requested by the Languages and Science departments on June 24, 2026. The session brought together the 18 teachers to examine operational difficulties in conventional lesson-plan submission, status tracking, feedback turnaround, version control, and administrative overhead.

Standard Operating Procedure (SOP) & Documentation Notes: To eliminate facilitator recording bias, official minutes of the meeting were independently recorded by Master Teachers from the requesting Languages and Science departments. The author facilitated the interactive discussion. Written informed consent was gathered prior to the session.

2.4 Qualitative Data Analysis and Requirements Derivation

The official minutes recorded by the external Master Teachers were analyzed to extract recurring operational and instructional friction points. Analysis focused on mapping operational bottlenecks and instructional gaps.

SOP Limitation: Qualitative thematic extraction from the meeting minutes was performed exclusively by the author as a single coder. No inter-coder agreement metric was calculated. Requirements were systematically derived using a two-column traceability framework mapping reported operational friction directly to technical specifications.

2.5 Co-Design and Prototype Development

Following problem derivation, the author convened with the cohort of 18 Languages and Science teachers for a co-design session. The team reviewed the baseline requirements derived from their June FGD, evaluated the proposed digital workflow against their department-specific planning routines, and critiqued the placement of feedback indicators. Direct modifications suggested during this session were integrated into the functional prototype.

Following co-design consensus, the author completed the LP Monitor prototype prior to the August 2026 tryout.



2.6 Formative Tryout Procedure

A four-week formative tryout was conducted during August 2026 with the 18 Languages and Science teachers. Participants utilized LP Monitor during their regular weekly lesson-plan submission routines, uploading authentic instructional plans.

Upon uploading a plan, teachers received feedback from the system. System logs automatically captured submission timestamps, review statuses, system outputs, and resubmission events.

3. Results

3.1 Qualitative Problem Identification (SOP 1)

Thematic analysis of the June 24, 2026 LAC session minutes—documented by independent Languages and Science Master Teachers—identified seven primary operational and instructional friction points. These findings reflect the reported experiences of the 18 participating teachers across the Languages and Science departments.

Physical Submission Burden: Teachers reported taking time away from instruction or preparation to physically deliver hard-copy lesson plans to supervisory offices, often requiring multiple follow-up visits to confirm receipt.

Recurring Printing Expenditures: Participants cited continuous personal or departmental costs for paper, ink cartridges, printer maintenance, and re-prints necessitated by minor revisions.

Administrative Document Handling: Supervisory review was delayed by administrative management tasks, including sorting, logging, filing, and physically returning paper folders.

Delayed Supervisory Feedback: Feedback was frequently returned after the planned lesson had already been delivered in the classroom, stripping the supervisory review of its formative value.

Opaque Review Status: Neither teachers nor supervisors possessed direct visibility into where a document sat within the review pipeline, leading to confusion over pending versus reviewed files.

Version Control Confusion: Repeated revisions generated multiple physical paper drafts, making it difficult to verify whether feedback was applied to the correct version.

Supervisory Constraints on Deep Alignment Analysis: Manual review of every instructional relationship across high submission volumes limited the depth of structural alignment analysis supervisors could perform.

3.2 Traceability Framework: Requirements to System Architecture (SOP 2)

To address these friction points, technical requirements were mapped directly to system features:

Table 1

Traceability of Identified Problems to Functional Requirements and System Architecture

Identified Problem	Functional Requirement	LP Monitor Technical Implementation
Physical & Cost Burden	Digital Document Portal	Web-based upload supporting PDF/Docx formats, eliminating printing.
Opaque Status & Delays	Real-time Dashboard	Centralized pipeline displaying submission status (Pending, Approved, Needs Revision).
Delayed / Shallow Feedback	Automated Alignment Feedback Engine	Integration evaluating alignment across Competency, Objectives, Activities, and Assessment.
Version Confusion	Automated Versioning	Multi-draft history log preserving prior submissions and feedback threads.

3.3 Co-Design Output: Integration of Large Language Model (LLM) and Scaffolding Objectives

As a direct outcome of the co-design sessions with the 18 teachers, the system architecture was configured to integrate a Large Language Model (LLM) API to generate automated preliminary alignment feedback. The team

operationalized instructional alignment through the construct of Scaffolding Objectives (Enad, 2026), drawing on scaffolding theory's emphasis on intermediate supports that bridge a learner's current performance to a target competency (van de Pol et al., 2010), and established prompt rubrics that analyze whether intermediate target objectives bridge the prescribed Learning Competency to the planned Learning Activities and Assessments. This approach is consistent with emerging work applying large language models to constructive-alignment analysis in other educational contexts (Pereira et al., 2024).

The integrated LLM engine processes three specific instructional alignment relationships:

Competency-to-Scaffolding Objective Alignment (R_1): Verifies if the target objective accurately reflects the depth and scope of the DepEd Curriculum Guide competency.

Scaffolding Objective-to-Learning Experience Alignment (R_2): Assesses whether classroom activities directly facilitate the attainment of the scaffolding objective.

Scaffolding Objective-to-Assessment Alignment (R_3): Checks if the evaluation task directly measures the mastery of the scaffolding objective.

3.4 System Architecture and Data Payloads

LP Monitor was constructed as a modern, full-stack web application. The frontend interface provides role-based dashboards for teachers and supervisors, while the backend handles file storage, authentication, and communication with the automated analysis engine.

The feedback module passes lesson plan components into the LLM API using structured prompt rubrics, returning JSON payloads containing:

- Numerical or categorical alignment scores for R_1 , R_2 , and R_3 .
- Specific textual suggestions highlighting observed misalignments.
- A binary overall recommendation: Approved or Needs Revision.

3.5 Formative Tryout Operational and Revision Results (SOP 3)

During the August 2026 tryout, the 18 participating Languages and Science teachers submitted 91 lesson plans through LP Monitor.

91 Submissions → Recommended Revision ($n = 50$) → Resubmitted by Participating Teachers

91 Submissions → Recommended Approval ($n = 41$)

Initial System Analysis: 50 submissions received an automated recommendation for revision; 41 received an approval recommendation.

Resubmission Activity: Participating teachers in the Languages and Science cohort utilized immediate automated feedback to revise and resubmit plans following revision recommendations.

SOP Reporting Gap: System logging did not systematically capture the total number of sub-iteration loops per document or qualitative teacher reflections regarding system interactions. Quantitative pre-tryout baselines regarding turnaround hours or printing costs were not recorded prior to the tryout.

4. Discussion

The design and field implementation of LP Monitor demonstrate the feasibility of transitioning lesson-plan supervision from paper-based routines to an automated digital workflow. By embedding automated alignment checks into the submission process—an innovation conceptualized during co-design—teachers receive immediate feedback prior to formal supervisory review, transforming lesson-plan submission into an active learning process.

4.1 Pedagogical and Operational Implications

The system addresses both operational friction and instructional feedback depth. Digitization eliminates paper and printing overhead, while centralized tracking provides full visibility to administrators and teachers. Furthermore, the automated evaluation of Scaffolding Objectives (R_1 , R_2 , R_3) offers formative guidance that assists teachers in tightening the internal consistency of their instructional plans before classroom execution, consistent with the broader case for scaffolded, timely feedback in instructional design (van de Pol et al., 2010; Hattie & Timperley, 2007).

4.2 Study Limitations

Key Study Limitations

- Non-Supervisory Context
- Single-Coder Analysis
- Unvalidated Scoring
- No Pre-Tryout Baseline
- Single-School Context
- Automated Model Limitations

Non-Supervisory Implementation Dynamic: Because the author was not the direct supervisor of the 18 Languages and Science teachers, participation was structured around formative tool evaluation rather than formal administrative performance evaluation.

Single-Coder Qualitative Extraction: Although official minutes were recorded by independent Master Teachers from the requesting departments, thematic coding of those minutes was performed by a single coder without inter-coder agreement verification.

Unvalidated Construct & Automated Models: The Scaffolding Objectives theoretical framework (Enad, 2026) and the LLM prompt rubrics developed during co-design have not undergone external psychometric or criterion-related validation.

Absence of Quantitative Baseline: The lack of pre-tryout data on feedback turnaround, printing costs, and reviewer workload precludes definitive comparative efficacy claims.

5. Future Research Agenda

To move beyond the descriptive nature of this design and development study, future research should follow a structured, multi-phase validation agenda:

1. Psychometric and Prompt Validation: Formally validate the prompt rubrics across larger, multi-rater datasets to establish inter-rater reliability, sensitivity, and specificity of the automated alignment engine against expert Master Teacher evaluations.

2. Controlled Comparative Studies: Implement quasi-experimental designs comparing schools using LP Monitor against control schools using conventional submission methods to measure impact on turnaround time, feedback adoption, and lesson plan quality.

3. Longitudinal Adoption and Classroom Observation: Conduct longitudinal studies tracking whether improvements in written lesson-plan alignment translate into enhanced classroom instructional delivery and student learning outcomes.

6. Declarations

Ethics Statement

No formal research-ethics-board clearance was obtained for the focus group discussion, co-design session, or tryout described in this article; all three were conducted as ordinary school-based professional-development and supervisory activities rather than as a formally cleared research study. Written informed consent was gathered from participants prior to the June 24, 2026 focus group discussion; no separately documented consent procedure is reported for the co-design session or the tryout. Readers should weigh the reported results accordingly.

Conflict of Interest

The author is the developer of LP Monitor, which constitutes a conflict of interest in interpreting the system's design rationale and reported performance. The participating teachers were not, however, under the author's direct supervisory authority at any stage of the study; this non-supervisory relationship is intended to mitigate, though it does not eliminate, the risk of participant deference or social-desirability bias in the focus group and co-design sessions the author facilitated. No independent evaluator verified the tryout data reported here. Readers should independently verify the study's favorable framing of LP Monitor's design and preliminary results before relying on them for procurement, adoption, or policy decisions.

Funding

The author received no specific funding for this work.

Author Contributions

Fernando B. Enad is the sole author of this article and is responsible for its conceptualization, system design and development, data collection, analysis, and writing.

Data Availability

The June 24, 2026 focus group discussion is documented through minutes independently recorded by participating departments' Master Teachers rather than a verbatim transcript, and the August 2026 tryout's submission records contain identifiable information about participating teachers and their lesson plans. These underlying materials are therefore not publicly available; de-identified summaries may be made available by the corresponding author upon reasonable request.

References

- Biggs, J. (1996). Enhancing teaching through constructive alignment. *Higher Education*, 32(3), 347–364. <https://doi.org/10.1007/BF00138871>
- Department of Education. (2026a). *DepEd Order No. 003, s. 2026: Foundational guidelines on artificial intelligence (AI) in basic education*. Department of Education, Republic of the Philippines.
- Department of Education. (2026b). *DepEd Order No. 016, s. 2026: Guidelines on lesson planning and learning design*. Department of Education, Republic of the Philippines.
- Enad, F. B. (2026). *Alignment matters: A coherence framework for intentional, equitable, and sustainable basic education*. Alignment Matters PH.
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81–112. <https://doi.org/10.3102/003465430298487>
- Martone, A., & Sireci, S. G. (2009). Evaluating alignment between curriculum, assessment, and instruction. *Review of Educational Research*, 79(4), 1332–1361. <https://doi.org/10.3102/0034654309341375>



- Pereira, E., Nsair, S., Pereira, L. R., & Grant, K. (2024). Constructive alignment in a graduate-level project management course: An innovative framework using large language models. *International Journal of Educational Technology in Higher Education*, 21, Article 25. <https://doi.org/10.1186/s41239-024-00457-2>
- Richey, R. C., & Klein, J. D. (2007). *Design and development research: Methods, strategies, and issues*. Lawrence Erlbaum Associates.
- Richey, R. C., & Klein, J. D. (2014). Design and development research. In J. M. Spector, M. D. Merrill, J. Elen, & M. J. Bishop (Eds.), *Handbook of research on educational communications and technology* (pp. 141–150). Springer. https://doi.org/10.1007/978-1-4614-3185-5_12
- Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research*, 78(1), 153–189. <https://doi.org/10.3102/0034654307313795>
- van de Pol, J., Volman, M., & Beishuizen, J. (2010). Scaffolding in teacher–student interaction: A decade of research. *Educational Psychology Review*, 22(3), 271–296. <https://doi.org/10.1007/s10648-010-9127-6>