

Comprehensive Evaluation of The Support For Leadership Through Mentoring and Learning Program in Timor-Leste Elementary Schools

Mariano Dos Santos^{1*}, Edy Supriyadi², Aman³, Ahmad⁴, Maria Angelina Fransiska Mbari⁵, Petrus Redy Partus Jaya⁶

¹ Department of Educational Research and Evaluation, Graduate School, Universitas Negeri Yogyakarta, Indonesia

² Department of Electrical Engineering Education and Educational Research and Evaluation, Graduate School, Universitas Negeri Yogyakarta, Indonesia

³ Department of History Education, Faculty of Social Sciences, and Educational Research and Evaluation, Graduate School, Universitas Negeri Yogyakarta, Indonesia

^{4,5,6} Department of Educational Research and Evaluation, Graduate School, Universitas Negeri Yogyakarta, Indonesia

***Corresponding author:**

Email ID: mariano5001pasca.2023@student.uny.ac.id

Abstract:

The ALMA (*Apoiui Lideransa liu husi Mentoria no Aprendizajen*) or Support for Leadership through Mentoring and Learning Program was launched in 2016 as a bilateral collaboration between the Australian and Timorese governments with the objective of strengthening teacher professional competencies and supporting the implementation of Student-Centered Learning (SCL) in elementary schools across Timor-Leste. Despite its broad mandate, empirical evidence regarding the program's systemic effectiveness from resource allocation to long-term student learning outcomes remains limited. This study presents a comprehensive evaluation of ALMA using a Logic Model evaluation framework integrated with Partial Least Squares Structural Equation Modeling (SmartPLS) to assess the directional relationships among the program's structural components: Input, Process, Output, and Outcome. A descriptive-evaluative research design was adopted involving 100 participants from eight elementary schools, comprising 80 teachers, 6 mentors, 8 school principals, and 6 Ministry of Education officials. Data were collected through structured questionnaires, focus group discussions (FGD), and in-depth interviews, and analyzed using descriptive statistics and SmartPLS. The outer model results confirmed satisfactory convergent validity (all AVE > 0.50), discriminant validity (Fornell–Larcker criterion satisfied), and composite reliability (all ρ > 0.70). The inner model demonstrated that Input significantly predicted Process ($\beta = 0.716$, $t = 15.317$, $p < 0.001$) and Process significantly predicted Output ($\beta = 0.738$, $t = 17.881$, $p < 0.001$). The relationship between Output and Outcome was marginally significant and negative ($\beta = -0.224$, $t = 1.957$, $p = 0.050$), suggesting that while immediate training outputs are strong, their translation into long-term student learning outcomes is constrained by systemic factors. Goodness-of-Fit (GoF = 0.444) indicated a robust structural model. Qualitative findings corroborated quantitative results, highlighting workload concerns and the need for sustained institutional support. The findings underscore the importance of reinforcing post-training follow-up mechanisms and aligning program governance structures with the long-term capacity-building goals of educational reform in Timor-Leste.

Keywords: Support for Leadership through Mentoring and Learning (ALMA); Student-Centered Learning (SCL); Logic Model evaluation; SmartPLS; Timor-Leste; teacher professional development

1. Introduction

Quality education has long been recognized as a cornerstone of sustainable national development and a fundamental enabler of social equity. In the global policy landscape, Sustainable Development Goal 4 (SDG 4) underscores the imperative of ensuring inclusive, equitable, and quality education for all, with an explicit emphasis on improving learning outcomes and expanding access to continuous professional development opportunities for teachers as part of lifelong professional learning (Elmassah et al., 2022; United Nations, 2015). Research consistently demonstrates that investments in education quality particularly through strengthening teacher competencies translate directly into economic productivity, social cohesion, and innovation capacity, particularly in low- and middle-income countries where educational disparities are

most pronounced (Goczek et al., 2021; UNESCO, 2023).

Timor-Leste, as one of Asia's youngest nations, faces significant structural challenges in its education sector. The country's elementary school system is characterized by multilingual complexity, limited infrastructure, high rates of teacher absenteeism, and persistent deficiencies in foundational literacy and numeracy skills among students (Cassity et al., 2022; Quinn & Buchanan, 2022). These challenges are compounded by a historical legacy of institutional disruption and the ongoing transition toward a nationally coherent curriculum framework. Despite considerable investments in education reform, large-scale assessments have consistently revealed that a substantial proportion of Timorese elementary students remain below minimum proficiency benchmarks in reading and mathematics (Cassity & Chainey, 2020).

A paradigmatic response to these challenges has been the adoption of Student-Centered Learning (SCL) as the preferred pedagogical orientation for elementary education in Timor-Leste. SCL repositions the teacher from a transmitter of knowledge to a facilitator of active learning, placing student agency, critical inquiry, and collaborative problem-solving at the center of the instructional process (Abdigapbarova & Zhiyenbayeva, 2023; Tholibon et al., 2022). The theoretical and empirical case for SCL is compelling: meta-analytic evidence indicates that student-centered pedagogies are associated with significantly higher levels of student engagement, deeper conceptual understanding, and improved long-term retention compared to conventional teacher-led instruction (Freeman et al., 2014; Roca-Campos et al., 2021). However, successful SCL implementation demands a profound shift in teacher beliefs, classroom management strategies, and instructional design competencies—a shift that aligns closely with the principles of vocational pedagogy, where professional skills are developed through sustained practice, mentoring, and reflective learning rather than isolated theoretical instruction (Merino et al., 2025; Sahonero-Alvarez & Humberto, 2018).

It is within this context that the ALMA (*Apoiui Lideransa liu husi Mentoria no Aprendizajen*) was established in 2016 through a bilateral development partnership between the governments of Australia and Timor-Leste, administered under the auspices of the Australian Department of Foreign Affairs and Trade (DFAT) Education Analytics Service (Cassity et al., 2022). ALMA is a structured, multi-component professional development program targeting school principals and classroom teachers in elementary schools. Its four core components—school leader and teacher training, classroom observation, peer learning groups, and educational technology mentoring are designed to operate in a mutually reinforcing cycle that progressively builds teacher capacity to implement SCL with fidelity and consistency. In this sense, ALMA can be conceptualized as a form of in-service vocational education for teachers, where professional competencies are systematically developed through continuous practice-based learning and mentoring.

Despite ALMA's extensive operational scope and substantial resource investment, critiques have emerged regarding the program's tangible impact on student learning outcomes. National assessments using the Early Grade Reading Assessment (EGRA) and Early Grade Mathematics Assessment (EGMA) instruments have continued to report low literacy and numeracy scores among students in program schools, raising fundamental questions about the causal chain between professional development inputs and distal student outcomes (Cassity & Chainey, 2020; Quinn, 2023). Furthermore, implementation reports have documented significant teacher workload concerns associated with ALMA participation, suggesting that the program's design may inadvertently undermine the very classroom practices it seeks to improve (Cassity et al., 2022).

Evaluation of complex and multi-component education programs in resource-constrained contexts requires analytical frameworks capable of capturing both the sequential logic of program theory and the structural relationships among program components. In the context of teacher professional development, vocational education principles are particularly relevant because in-service teacher training represents a form of adult vocational learning that emphasizes skill mastery, workplace relevance, mentoring, and continuous competency improvement. Two complementary methodological approaches address this requirement. First, the Logic Model framework provides a systematic mapping of causal pathways linking program inputs, activities, outputs, and outcomes, enabling evaluators to assess whether observed program results are consistent with the underlying theory of change and the extent to which vocational-oriented professional learning principles are embedded across different stages of teacher development (Payton et al., 2022; Sentana Putra, 2025; W.K. Kellogg Foundation, 2004). Second, Partial Least Squares Structural Equation Modeling (PLS-SEM), implemented using SmartPLS, offers a variance-based approach for testing directional hypotheses among latent constructs with relatively small sample sizes, making it particularly suitable for evaluation research in developing country contexts, especially when examining structural relationships between professional development components and competency growth consistent with vocational pedagogy principles (Hair et al., 2019; Henseler et al., 2023).

The present study integrates both methodological approaches in a comprehensive evaluation of ALMA. Specifically, the study aims to assess stakeholder satisfaction with each component of the Logic Model (Input, Process/Activities, Output, Outcome-Impact), test the structural relationships among Logic Model components using SmartPLS; and derive evidence-based recommendations for program improvement and sustainability. By combining rigorous quantitative modeling with rich qualitative insights from school-level stakeholders, this study contributes to the empirical literature on teacher professional development evaluation in post-conflict, developing country contexts, particularly by framing teacher in-service training as a form of vocational and professional learning that strengthens pedagogical competence and instructional effectiveness.

LITERATURE REVIEW

Quality Education and the SDG 4 Framework in Developing Contexts

The SDG 4 framework represents the most ambitious global commitment to educational quality in history, encompassing targets related to early childhood education, primary and secondary school completion, tertiary access, and lifelong learning opportunities (UNESCO, 2023). For developing nations, achieving SDG 4 targets necessitates simultaneous action on supply-side constraints (school infrastructure, teacher supply, learning materials) and demand-side barriers (poverty, gender discrimination, geographic remoteness). Research across Sub-Saharan Africa, Southeast Asia, and the Pacific region has consistently identified teacher quality as the single most powerful school-level predictor of student learning outcomes, explaining more variance in achievement than class size, school resources, or curriculum design in isolation (Goczek et al., 2021; Nguyen et al., 2022).

In the Pacific Island context, small-state dynamics compound educational quality challenges. Limited economies of scale, geographic dispersion, and high dependence on bilateral aid partnerships shape the structure of educational governance and the design of teacher development programs. Timor-Leste exemplifies these dynamics: a population of approximately 1.3 million, significant internal migration, four national languages (Portuguese, Tetum, Indonesian, and English) co-existing in school settings, and an education system that is simultaneously transitioning from a colonial-influenced curriculum to a nationally authored framework (Quinn, 2023; Quinn & Buchanan, 2022). In this context, international assistance programs like ALMA are not merely supplementary they are constitutive of the national professional development infrastructure.

Student-Centered Learning: Theoretical Foundations and Implementation Challenges

Student-Centered Learning (SCL) is rooted in constructivist epistemology and draws on the work of Piaget, Vygotsky, and Dewey, among others, who argued that knowledge is actively constructed by learners through meaningful interaction with their environment and peers (Vygotsky, 1978). In contemporary educational practice, SCL encompasses a range of pedagogical strategies including inquiry-based learning, cooperative learning, problem-based learning, and differentiated instruction that share a common commitment to learner autonomy and active cognitive engagement (Abdigapbarova & Zhiyenbayeva, 2023).

The empirical evidence supporting SCL is robust. A landmark meta-analysis by (Nayir et al., 2023) found that students in active learning environments outperformed those in traditional lecture-based classrooms by 0.47 standard deviations in examination performance and were 1.5 times less likely to fail. More recent systematic reviews confirm these findings across diverse educational contexts, including primary education in Sub-Saharan Africa and Southeast Asia (Roca-Campos et al., 2021). However, the translation of SCL theory into practice is rarely straightforward, particularly in low-resource contexts. Teachers in Timor-Leste predominantly received their own schooling and initial teacher training through didactic, transmission-oriented pedagogies, creating deeply ingrained instructional habits that are resistant to change without sustained, context-sensitive professional support (Almansouri et al., 2022; Cassity et al., 2022).

Structural barriers to SCL implementation in Timor-Leste include inadequate classroom infrastructure (insufficient furniture for group work, lack of manipulatives and printed materials), large class sizes that complicate individualized attention, multilingual classroom dynamics that challenge the facilitation of student-led discussion, and assessment systems that continue to reward rote memorization over higher-order thinking (Almulla, 2023; Quinn, 2023). Addressing these barriers requires not only teacher-level interventions but also school-level leadership development and system-level policy alignment.

Teacher Professional Development and Mentoring in Elementary Education

The literature on effective teacher professional development converges on several design principles: content-focused learning connected to the curriculum teachers actually teach; active learning opportunities that allow teachers to engage with, discuss, and apply new knowledge; collaborative professional community structures that sustain learning beyond formal workshops; the use of models and modeling of effective practice; expert coaching and mentoring that supports transfer of learning to classroom contexts; and sufficient duration and intensity (Sancar et al., 2021; Sims & Fletcher-Wood, 2021). Critically, one-off training events however well-designed produce negligible lasting change in teacher practice without follow-up support (Merino et al., 2025).

Mentoring, as a specific form of professional support, has attracted considerable research attention. Peer mentoring and expert mentoring have been shown to improve teacher self-efficacy, reduce professional isolation, and accelerate the adoption of new pedagogical practices (De La Iglesia et al., 2024; Godfrey, 2020). Peer learning groups, in which teachers collaboratively examine student work, observe each other's instruction, and provide structured feedback, have emerged as particularly powerful professional learning structures (AKSU et al., 2023; Haelermans et al., 2022). The peer teaching component of ALMA where teachers prepare topic-specific lessons and receive structured feedback from peers aligns well with these evidence-based models.

Classroom observation with structured feedback is another high-leverage professional development strategy. (Acar et al., 2023) demonstrated that EFL teachers who received regular classroom observation with post-observation conferences reported significant gains in reflective practice and instructional quality. Similarly, (Bottazzi et al., 2023) and (Diwakar et al., 2023) emphasize that observation-based feedback creates accountability structures and supports the translation of training content into authentic classroom behavior. The combination of observation, feedback, and reflective discussion is a central design feature of ALMA and represents one of its key theoretical strengths.

The integration of educational technology into teacher development programs has gained increasing salience in the post-pandemic era. (Attard & Holmes, 2020) demonstrated that technology-mediated mentoring significantly enhanced teachers' adoption of SCL-aligned practices by providing access to exemplary lesson models and enabling flexible, asynchronous professional learning. In low-resource contexts, pre-loaded tablet devices as deployed in ALMA's educational technology mentoring component offer a pragmatic solution to connectivity constraints while expanding teachers' access to quality instructional resources (Merino et al., 2025).

Logic Model as an Evaluation Framework

The Logic Model is a widely used program evaluation tool that provides a visual and conceptual representation of the hypothesized causal chain linking a program's resources (inputs) to its activities, outputs, and intended outcomes (W.K. Kellogg Foundation, 2004). By making explicit the program's theory of change, the Logic Model enables evaluators to assess not only whether a program achieved its intended results but also how and why it did or did not succeed insights that are essential for program improvement and replication (Payton et al., 2022).

In educational program evaluation, Logic Models have been applied to assess professional development programs, curriculum reform initiatives, literacy interventions, and school improvement efforts across diverse contexts (Yee et al., 2022). A key strength of the Logic Model framework is its capacity to integrate multiple data sources quantitative performance data, stakeholder satisfaction surveys, and qualitative accounts into a coherent evaluative narrative. The framework also explicitly acknowledges the role of external factors (contextual conditions, political dynamics, resource constraints) that mediate the relationship between program activities and outcomes, a feature of particular relevance in the complex, multi-actor environment of educational reform in Timor-Leste (Hull et al., 2019; Payton et al., 2022).

The four Logic Model components evaluated in this study Input (resources and stakeholder support), Process/Activities (program implementation quality), Output (direct and measurable program results), and Outcome-Impact (short-, medium-, and long-term changes in teacher and student performance) provide a comprehensive evaluative architecture that accommodates both the program's operational complexity and the temporal dimensions of its intended effects.

Partial Least Squares Structural Equation Modeling (PLS-SEM) in Educational Evaluation

Partial Least Squares Structural Equation Modeling (PLS-SEM) is a second-generation multivariate analytical technique that enables researchers to simultaneously estimate measurement models (relating observed indicators to latent constructs) and structural models (specifying directional relationships among

latent constructs). PLS-SEM is particularly well-suited for exploratory and prediction-oriented research, for studies involving small to moderate samples, and for models with complex structural configurations involving mediating variables characteristics that are common in educational program evaluation research (Hair et al., 2022; Ringle et al., 2023).

Compared to covariance-based SEM (CB-SEM), PLS-SEM imposes fewer distributional assumptions, performs well with non-normal data distributions, and provides superior statistical power in samples below $n = 200$ (Hair et al., 2022). These properties make it an appropriate analytical choice for the present study, which involves a total sample of 100 participants. The two-stage analytical approach first assessing measurement model quality (outer model) before examining structural relationships (inner model) ensures that conclusions about construct relationships are grounded in valid and reliable measurement (Henseler et al., 2023).

Recent applications of PLS-SEM in educational research include studies examining the determinants of student academic achievement (Ringle et al., 2023), the effectiveness of teacher professional development programs (Fandos-Herrera et al., 2023), and the adoption of technology in educational settings (Attard & Holmes, 2020). The application of SmartPLS to Logic Model-structured program evaluation, as in the present study, represents a methodological contribution that bridges the evaluative rigor of SEM with the interpretive richness of the Logic Model framework.

The ALMA in the Context of Timorese Education Reform

ALMA was designed as a systemic response to the documented gap between teacher training and classroom practice in Timor-Leste's elementary schools. The program operates across eight districts, targeting schools where SCL implementation was assessed to be weakest, and involves a cadre of national and educational mentors who provide regular, school-based professional support to principals and teachers (Cassity et al., 2022; Partnership for Human Development, 2016).

Evaluation reports commissioned by DFAT's Education Analytics Service have noted mixed results. While stakeholder satisfaction with training activities has generally been positive, the translation of professional learning gains into sustained changes in classroom practice and, ultimately, student learning outcomes has been inconsistent (Cassity & Chainey, 2020). Qualitative findings from ALMA implementation studies consistently identify three systemic constraints insufficient post-training follow-up and coaching, high teacher turnover in program schools; and misalignment between ALMA's SCL-oriented professional development activities and the accountability structures embedded in the national curriculum and assessment system (Fillah et al., 2021; Quinn, 2023).

These findings suggest that while ALMA has successfully built a platform for professional learning at the school level, the conditions required for deep pedagogical change and durable student outcome improvement have not yet been fully established. A rigorous, data-driven evaluation that examines the structural pathways through which program components interact is therefore both timely and necessary.

RESEARCH METHOD

Research Design

This study employed a descriptive-evaluative research design that integrates quantitative and qualitative methodologies within a mixed-methods framework. The evaluative dimension is structured according to the Logic Model, which maps the sequential relationships among program components (Input -> Process -> Output -> Outcome-Impact). The quantitative component uses descriptive statistics to characterize stakeholder satisfaction and SmartPLS to model structural relationships among the Logic Model constructs. The qualitative component employs content analysis of FGD transcripts and interview narratives to provide interpretive depth and contextual explanation for the quantitative findings.

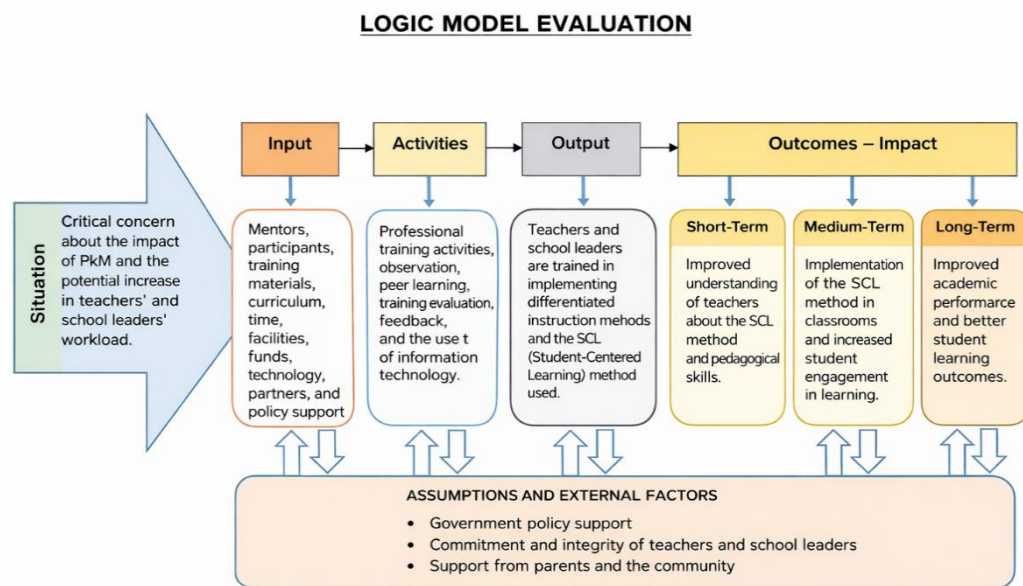


Figure 1. Logic Model Framework

Research Setting and Participants

The study was conducted in eight elementary schools in Timor-Leste participating in the ALMA during the 2022–2023 academic year. A census sampling technique was adopted, given the bounded and accessible nature of the total program population. The final sample comprised 100 participants: 80 classroom teachers, 6 national and educational mentors, 8 school principals, and 6 Ministry of Education district officials. This composition reflects the multi-stakeholder architecture of ALMA and ensures that evaluative judgments represent the perspectives of all key program actors.

Instruments and Data Collection

A structured questionnaire was developed to measure stakeholder perceptions of the four Logic Model components: Input (5 indicators), Process/Activities (5 indicators), Output (5 indicators), and Outcome-Impact (6 indicators). All items were scored on a five-point Likert scale ranging from 1 (very dissatisfied) to 5 (very satisfied). The questionnaire was reviewed by three independent measurement experts and pilot-tested with 20 non-study participants to establish face and content validity. Cronbach's alpha coefficients for the four subscales at the pilot stage ranged from 0.71 to 0.79.

In-depth interviews (30–45 minutes per participant) were conducted with representative school principals and district education officials ($n = 14$). FGD sessions (60–90 minutes each) were held with homogeneous groups of teachers ($n = 4$ groups) and mentors ($n = 2$ groups) to elicit collective perspectives on program implementation and impact. All interviews and FGD sessions were audio-recorded and transcribed verbatim. Member checking was performed with selected participants to verify the accuracy of transcriptions and interpretive summaries.

Student literacy and numeracy data were collected using standardized instruments: the Early Grade Reading Assessment (EGRA) for literacy and the Early Grade Mathematics Assessment (EGMA) for numeracy (RTI International, 2023). Assessments were administered at three time points before program commencement (T1), at mid-program (T2), and at program completion (T3) to enable longitudinal tracking of student learning gains.

Data Analysis

1. Quantitative Analysis

Descriptive statistics (mean, frequency, percentage distribution) were computed for each Logic Model component to characterize the distribution of stakeholder satisfaction levels. Structural relationships among Logic Model constructs were analyzed using SmartPLS version 4.0, following the two-stage analytical procedure recommended by Hair et al. (2022). The outer model was assessed for convergent validity (outer loadings > 0.50 ; Average Variance Extracted [AVE] > 0.50), discriminant validity (Fornell–Larcker criterion; HTMT ratio < 0.85), and reliability (composite reliability [ρ_c] > 0.70 ; Cronbach's alpha > 0.70). The inner model was evaluated using coefficient of determination (R^2), effect size (f^2), predictive relevance (Q^2), Goodness-of-Fit (GoF), and Variance Inflation Factor (VIF) for collinearity assessment. Path coefficient significance was

assessed using the bootstrapping procedure with 5,000 resamples at a 95% confidence level.

2. Qualitative Analysis

Qualitative data from interviews and FGD sessions were analyzed using conventional content analysis (Hsieh & Shannon, 2005). Transcripts were inductively coded by two independent researchers, with inter-rater agreement assessed using Cohen's kappa ($k = 0.82$, indicating strong agreement). Emergent themes were organized according to the Logic Model structure, enabling systematic triangulation with quantitative findings. Member checking and peer debriefing were employed to enhance the credibility and transferability of qualitative interpretations.

3. Ethical Considerations

This study was conducted in accordance with the ethical principles of the Declaration of Helsinki and the research ethics guidelines of Universitas Negeri Yogyakarta. Informed consent was obtained from all participants prior to data collection. Participants were assured of the confidentiality of their responses and their right to withdraw from the study at any time without consequence. Anonymization procedures were applied to all data at the point of analysis and reporting.

RESULTS AND DISCUSSION

Descriptive Analysis of Logic Model Components

1. Input Component

The Input component encompassed the foundational resources and enabling conditions of ALMA, including the quality and availability of mentors, participant engagement, training materials and curriculum relevance, time allocation, physical facilities, financial resources, technology provision, development partner support, and policy alignment. Descriptive analysis revealed that the majority of respondents expressed satisfaction with the Input component across all stakeholder groups: education officials (60% satisfied), teachers (55% satisfied), school principals and mentors (50% each). Notably, 20% of school principals reported being very satisfied, while 20% of teachers and approximately 5% of mentors and education officials reported dissatisfaction.

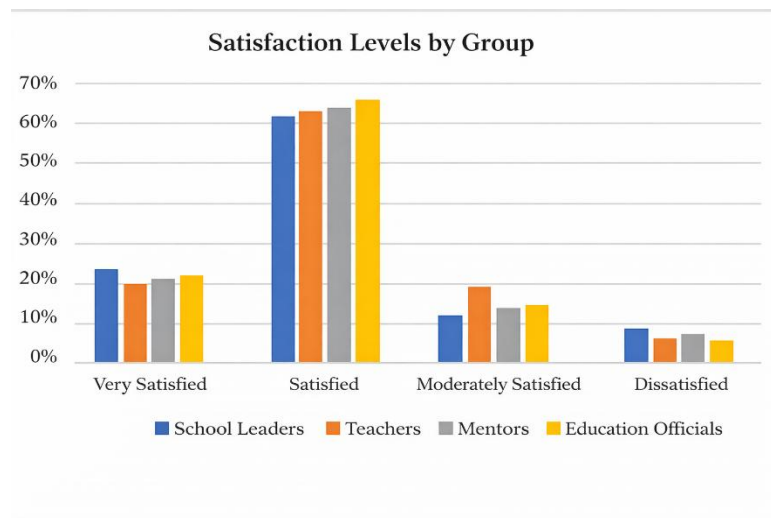


Figure 2. Results of the Input Component Analysis

These findings indicate that ALMA's input architecture has generally succeeded in meeting the baseline operational requirements for program implementation. The relatively high satisfaction rates among education officials likely reflect their proximity to policy-level decision-making and resource allocation processes. The lower though still majority satisfaction among teachers and mentors may reflect experiential concerns about material accessibility and the practical demands of program engagement. These observations align with Chambers et al. (2016) and Brown et al. (2020), who emphasize that stakeholder-differentiated resource needs must be addressed in program design to maximize implementation fidelity across actor groups.

2. Process/Activities Component

The Process component assessed the quality and perceived effectiveness of ALMA's four operational activities: professional training, classroom observation, peer learning groups, and educational technology

mentoring. Satisfaction levels were generally high across stakeholder groups: education officials and teachers (approximately 55% satisfied each), school principals (50%), and mentors (48%). Proportions reporting very high satisfaction were modest (school principals: 20%; teachers: 15%), while the proportion expressing moderate satisfaction was notable among teachers (20%) and school principals (15%).^t

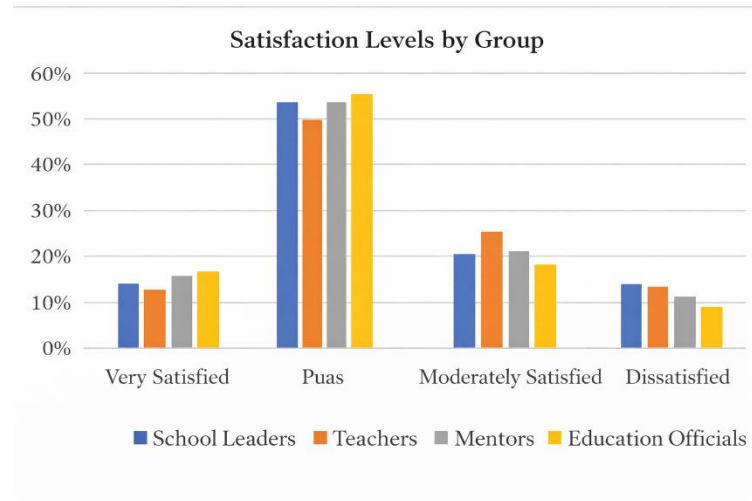


Figure 3. Results of the Activity Component Analysis

The high satisfaction scores for the Process component reflect the programmatic coherence and relevance of ALMA's four activities to teachers' actual professional needs. Qualitative data corroborated these findings: school principals characterized the program activities as "well-designed and school-appropriate," while teachers identified training and peer coaching sessions as genuinely informative, though occasionally difficult to integrate into existing workload demands. These results are consistent with Fandos-Herrera et al. (2023) and Roca-Campos et al. (2021), who demonstrate that professional development activities grounded in pedagogical practice rather than decontextualized theory generate higher participant engagement and satisfaction.

3. Output Component

The Output component captured the direct and observable results of ALMA, including the number and proportion of trained teachers and school principals, the updating and dissemination of teaching materials, and the measurable increase in SCL-aligned classroom practices. Satisfaction was distributed as follows: approximately 50% satisfied among school principals and education officials, with somewhat lower but still majority-level satisfaction among teachers and mentors. Proportions reporting very high satisfaction were below 20% across all groups, while 20% of teachers and mentors expressed only moderate satisfaction. A small but non-trivial proportion of respondents particularly among teachers and mentors reported dissatisfaction with output achievements.

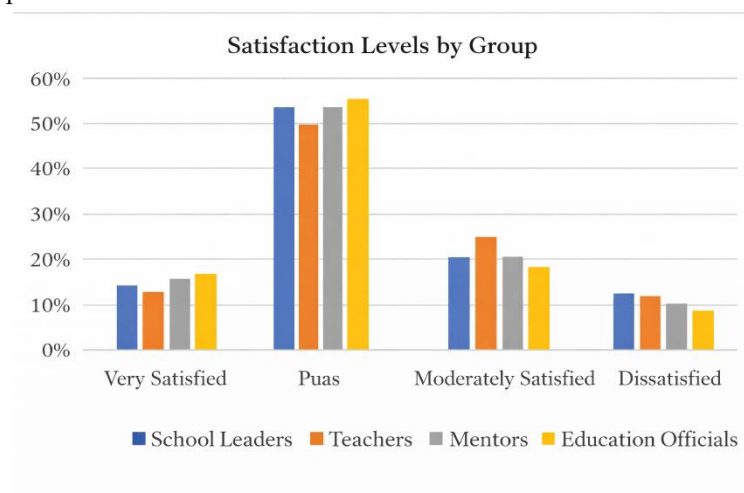


Figure 4. Results of the Output Component Analysis

The pattern of output satisfaction suggests that while ALMA has succeeded in delivering quantifiable

training outcomes, the quality and depth of those outcomes particularly as experienced by frontline practitioners does not fully meet expectations. School principals and education officials, whose assessment of program success may be more closely tied to administrative metrics (number of teachers trained, materials distributed), registered higher satisfaction than teachers and mentors, whose judgments reflect direct engagement with the instructional challenges of SCL implementation. This divergence highlights the importance of supplementing administrative output indicators with practitioner-centered measures of training quality and applicability (Lavilla-Gracia et al., 2023).

4. Outcome-Impact Component

The Outcome-Impact component was evaluated at three temporal levels: short-term, medium-term, and long-term.

a. Short-term outcomes

Teacher understanding of SCL methods increased markedly following program participation, rising from 37% to 67% a gain of 30 percentage points. Similarly, self-reported pedagogical skill competency increased from 35% to 55%. These gains confirm that ALMA's training activities successfully achieved their immediate knowledge and skill-building objectives, consistent with the findings of (Poulou et al., 2023) and (Sims & Fletcher-Wood, 2021)

b. Medium-term outcomes

Classroom implementation of SCL methods increased from approximately 30% to 50%, while measurable student engagement in active learning rose from approximately 35% to 45%. These gains, while positive, indicate that not all trained teachers consistently applied SCL principles in their classrooms, pointing to the existence of implementation barriers between professional learning and instructional change. (Ongori et al., 2020) and (Nguyen et al., 2022) similarly document this translation gap in comparable low-resource contexts, attributing it to insufficient in-classroom coaching support and persistence of habitual teaching behaviors.

c. Long-term outcomes (student literacy and numeracy)

Longitudinal data from EGRA and EGMA assessments across three measurement points (T1, T2, T3) revealed consistent improvement in both literacy and numeracy scores. Literacy scores were consistently higher than numeracy scores across all measurement points, reflecting the comparatively greater emphasis on reading instruction in ALMA's SCL training content. The upward trajectory across all three phases provides encouraging evidence of ALMA's capacity to improve foundational student learning outcomes over time, though the magnitude of improvement and the proportion of students reaching minimum proficiency benchmarks warrants continued monitoring.

Measurement Model Assessment

1. Convergent Validity-Outer Loadings

Initial measurement model testing revealed that four indicators (INP1, OTCM1, OTCM5, OUT5) had outer loading values below the 0.50 threshold recommended by Hair et al. (2022) and were therefore removed from the model in a sequential refinement procedure. Following indicator removal, all remaining 17 indicators demonstrated outer loadings ranging from 0.619 to 0.872, satisfying the convergent validity criterion. Table 1 presents the final outer loading values.

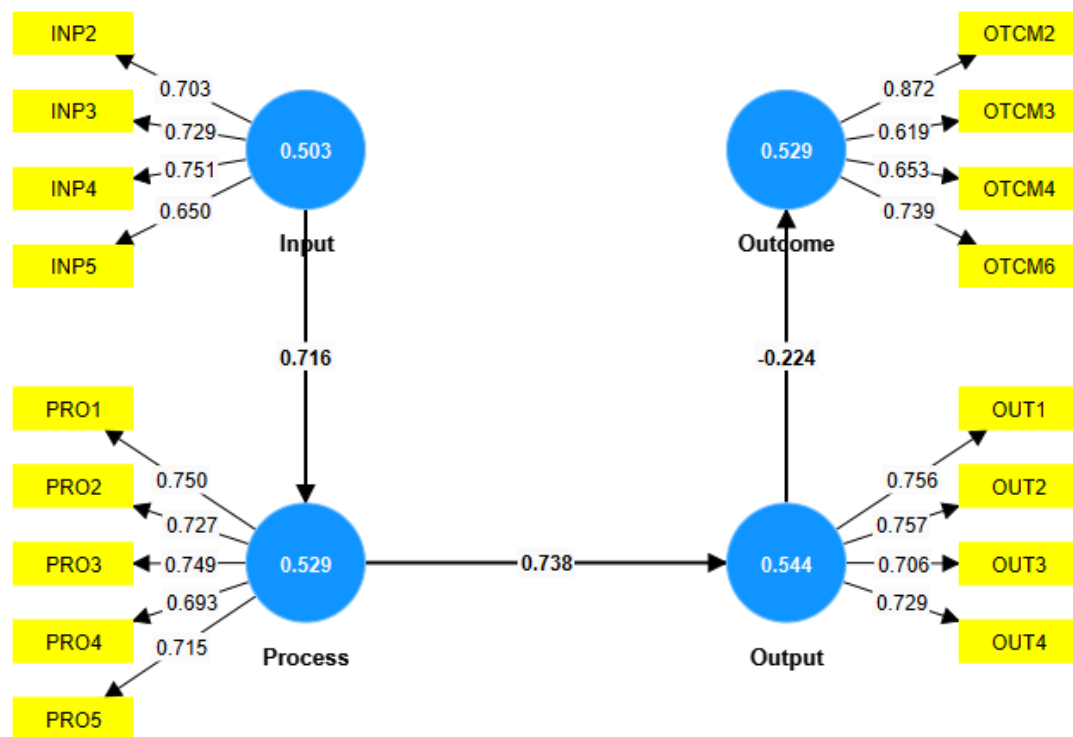


Figure 5. Outer Loading Validity

Table 1. Outer Loadings of Measurement Model Indicators (Final Model)

Indicator	Construct	Outer Loading
INP2	Input	0.703
INP3	Input	0.729
INP4	Input	0.751
INP5	Input	0.650
OTCM2	Outcome	0.872
OTCM3	Outcome	0.619
OTCM4	Outcome	0.653
OTCM6	Outcome	0.739
OUT1	Output	0.756
OUT2	Output	0.757
OUT3	Output	0.706
OUT4	Output	0.729
PRO1	Process	0.750
PRO2	Process	0.727
PRO3	Process	0.749

Indicator	Construct	Outer Loading
PRO4	Process	0.693
PRO5	Process	0.715

2. Discriminant Validity and Construct Reliability

AVE values for all four constructs in the final model exceeded the 0.50 threshold (Input = 0.503; Outcome = 0.529; Output = 0.544; Process = 0.529), confirming convergent validity at the construct level. Discriminant validity was assessed using the Fornell–Larcker criterion. Table 2 presents the AVE square roots (diagonal) alongside inter-construct correlations. The square root of each construct's AVE exceeded its highest correlation with any other construct, confirming that each construct shares more variance with its own indicators than with any other construct in the model.

Table 2. Discriminant Validity-Fornell–Larcker Criterion

Construct	Input	Outcome	Output	Process
Input	0.779 (√AVE)			
Outcome	−0.167	0.727 (√AVE)		
Output	0.752	−0.224	0.757 (√AVE)	
Process	0.716	−0.158	0.738	0.727 (√AVE)

Composite reliability (pc) values for all constructs were well above the 0.70 threshold: Input (pc = 0.801), Outcome (pc = 0.815), Output (pc = 0.827), and Process (pc = 0.849). These results confirm that the measurement instruments demonstrate satisfactory internal consistency and construct reliability, providing a sound psychometric foundation for structural model estimation.

Table 3. AVE and Composite Reliability Summary

Construct	AVE	√AVE	Composite Reliability (qc)
Input	0.503	0.779	0.801
Outcome	0.529	0.727	0.815
Output	0.544	0.757	0.827
Process	0.529	0.727	0.849

Structural Model Assessment

1. Coefficient of Determination (R²)

The coefficient of determination (R²) quantifies the proportion of variance in each endogenous construct explained by its predictors. As shown in Table 4, Input explained 51.2% of the variance in Process (R² = 0.512), indicating a substantial predictive effect. Process explained 54.4% of the variance in Output (R² = 0.544), confirming a strong explanatory relationship. In contrast, Output explained only 5.0% of the

variance in Outcome ($R^2 = 0.050$), indicating a weak direct predictive relationship between program outputs and long-term student outcomes a finding with important theoretical and practical implications that is elaborated in the discussion below.

Table 4. Coefficient of Determination (R^2) and Adjusted R^2

Endogenous Construct	R^2	R^2 Adjusted	Interpretation
Process	0.512	0.507	Substantial
Output	0.544	0.539	Substantial
Outcome	0.050	0.041	Weak

2. Effect Size (f^2)

Effect sizes (f^2) were computed to assess the relative contribution of each predictor path. Input→Process yielded $f^2 = 1.051$ (large effect), and Process→Output yielded $f^2 = 1.193$ (large effect), indicating that both relationships contribute strongly to the explanatory power of the model. Output→Outcome yielded $f^2 = 0.053$ (small effect), consistent with the low R^2 value for Outcome and further supporting the conclusion that Output's contribution to Outcome explanation is limited.

3. Predictive Relevance (Q^2)

Predictive relevance was assessed using the blindfolding procedure. Q^2 values for Process (0.260) and Output (0.285) indicate medium-to-large predictive relevance, confirming that the model has adequate out-of-sample predictive power for these constructs. The Q^2 value for Outcome (0.013) is positive but small, indicating limited but non-zero predictive relevance for long-term outcome variables.

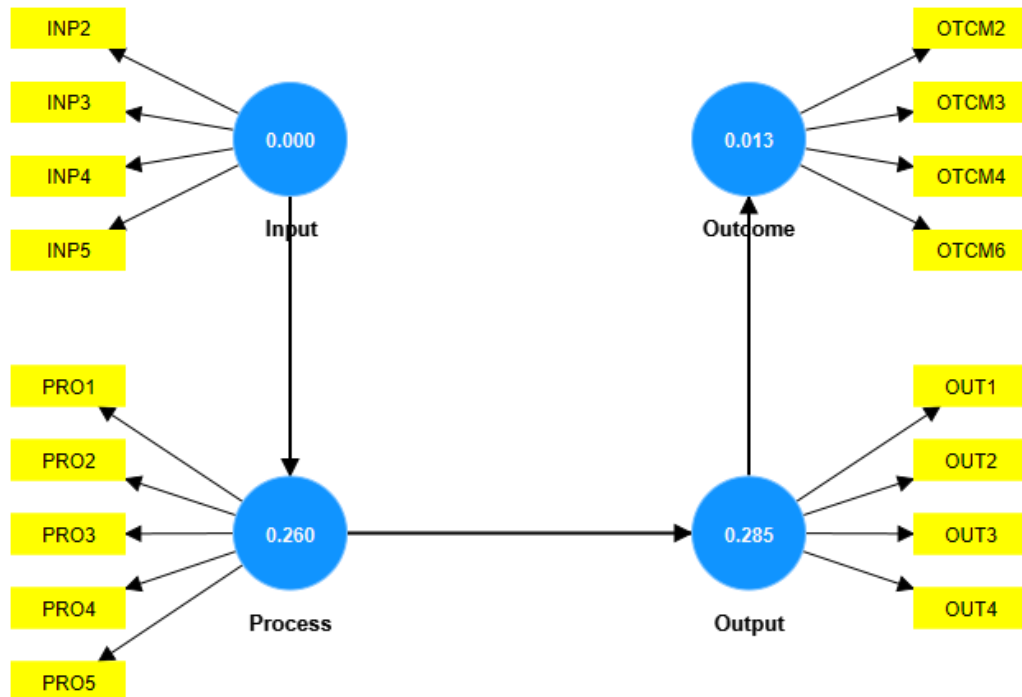


Figure 7. Q-square test results

4. Goodness-of-Fit (GoF)

Model fit was assessed using the GoF index, calculated as the geometric mean of the average AVE and average R^2 (excluding the Input construct, which has no antecedents in the model). The computed GoF value of 0.444 exceeds the threshold of 0.36 for large model fit, indicating that the overall structural model demonstrates strong explanatory power relative to the data (Henseler et al., 2023).

6. Collinearity Assessment (VIF)

Variance Inflation Factor (VIF) values were examined to detect multicollinearity among indicators. All VIF values ranged from 1.212 to 1.580, well below the threshold of 10 recommended by Hair et al. (2022), confirming the absence of multicollinearity in the measurement model.

7. Hypothesis Testing – Path Coefficients

Table 5 presents the results of the bootstrapping-based hypothesis tests for the structural paths.

Table 5. Hypothesis Testing Results – Direct and Indirect Effects

Path	β (O)	Mean (M)	SD	t-statistic	p-value	Decision
Input → Process	0.716	0.723	0.047	15.317	< 0.001	Supported
Process → Output	0.738	0.744	0.041	17.881	< 0.001	Supported
Output → Outcome	-0.224	-0.262	0.115	1.957	0.050	Marginal
Input→Process→Output→Outcome	-0.118	-0.141	0.063	1.869	0.062	Not Supported
Process→Output→Outcome	-0.165	-0.195	0.086	1.919	0.055	Not Supported

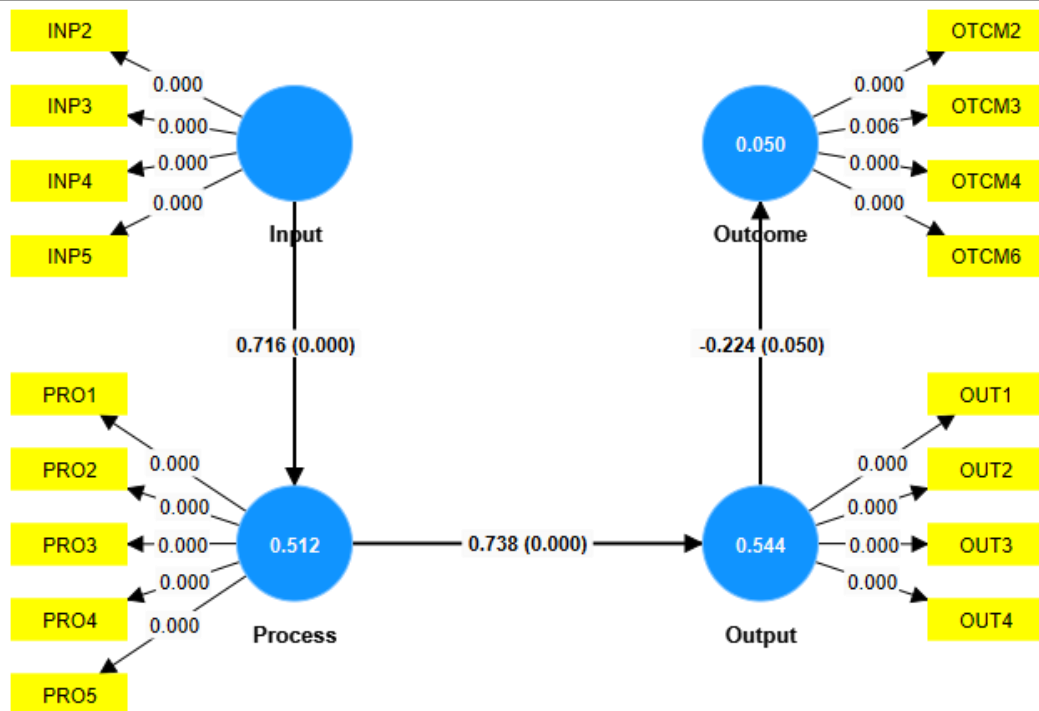


Figure 8. Results of the Significance Hypothesis Test

a. Hypothesis 1 (H1: Input → Process)

A significant positive direct effect of Input on Process was confirmed ($\beta = 0.716$, $t = 15.317$, $p < 0.001$). This finding indicates that a one-unit increase in Input quality is associated with a 71.6% increase in Process effectiveness, constituting a large and highly significant effect. This result underscores the importance of strong program foundations supported by adequate resources, including sufficient mentors, high-quality training materials, sufficient time allocation, and supportive policy environments. The availability of practice-oriented learning tools and appropriate technology further contributes to the quality of program implementation. This is consistent with (Hennessy et al., 2022) and (Brown et al., 2020), who emphasize that resource adequacy is a necessary, though not sufficient, condition for program implementation fidelity.

b. Hypothesis 2 (H2: Process → Output)

A significant positive direct effect of Process on Output was confirmed ($\beta = 0.738$, $t = 17.881$, $p < 0.001$), representing the largest effect in the model. This finding suggests that improvements in process quality—including effective training delivery, quality classroom observations, peer-learning engagement, and technology-supported mentoring—along with an emphasis on competency-based and practice-oriented approaches, explain 73.8% of the variance in Output achievement. This finding affirms that ALMA's operational activities constitute the primary engine of program performance: when the program's four core activities are implemented with quality and fidelity, tangible outputs (trained teachers, updated materials, increased SCL uptake) follow reliably.

c. Hypothesis 3 (H3: Output → Outcome)

The relationship between Output and Outcome was statistically marginally significant and negative ($\beta = -0.224$, $t = 1.957$, $p = 0.050$). This counterintuitive finding that higher output achievement is associated with marginally lower outcome attainment merits careful interpretation. Several explanatory mechanisms may be operating simultaneously. First, as noted by (Lavilla-Gracia et al., 2023), the temporal lag between professional development outputs and observable student learning outcomes means that cross-sectional or short-term evaluations systematically underestimate the impact of training on distal outcomes. Second, the negative coefficient may reflect a measurement paradox: teachers who have received more intensive training (high output) are more accurately able to diagnose their students' learning gaps, resulting in lower reported outcome satisfaction in the short term a form of the Dunning-Kruger effect in reverse. Third, as emphasized in qualitative findings, teachers who have received ALMA training consistently report increased workload pressures that may temporarily displace classroom instructional time, creating a short-term trade-off between professional development engagement and instructional delivery. Fourth, the low R^2 for Outcome (5%) suggests that the variance in long-term student learning outcomes is substantially explained by factors external to the ALMA program itself including family socioeconomic status, language background, classroom resource adequacy, and policy consistency that are not captured in the current model (Quinn, 2023), as well as the extent to which Technical & Vocational Education and Training (TVET) principles are integrated into classroom practice to support the development of practical and transferable student skills.

d. Indirect Effects

Neither the full mediation chain (Input→Process→Output→Outcome: $\beta = -0.118$, $t = 1.869$, $p = 0.062$) nor the partial mediation chain (Process→Output→Outcome: $\beta = -0.165$, $t = 1.919$, $p = 0.055$) reached statistical significance at the conventional $\alpha = 0.05$ threshold. These null indirect effect findings are consistent with the weak Output→Outcome direct path and reinforce the conclusion that ALMA's causal chain does not yet operate with sufficient strength to produce statistically detectable indirect effects on long-term student outcomes through the measured intermediate variables. This is not necessarily a program failure; rather, it signals the need for structural program enhancements that strengthen the connection between professional learning investments and sustained changes in student achievement. In this context, strengthening the practical, competency-oriented dimensions of teacher professional learning may serve as a potential mechanism to bridge the gap between professional development processes and long-term student learning outcomes.

Qualitative Synthesis and Triangulation

Qualitative findings from interviews and FGD sessions provided important contextual texture to the SmartPLS results across all four Logic Model components.

Input: School principals and education officials generally affirmed the adequacy of program facilities and resources, with one principal noting that "the resources provided truly helped us run the program smoothly." Teachers appreciated the relevance of the training materials to their classroom realities, although some requested more structured and systematically organized access to reference materials. Mentors highlighted the importance of technological tools in supporting their mentoring activities. In line with this, the implementation of Technical and Vocational Education and Training (TVET) principles in the provision of facilities, technology, and training materials was also regarded as an important component in strengthening teachers' readiness to develop skill-based and practice-oriented learning.

Process/Activities: The structured design and sequencing of ALMA activities received positive responses from all stakeholder groups. However, time management emerged as a cross-cutting concern: teachers reported that the demands of weekly peer learning sessions and classroom observation schedules were difficult to reconcile with existing instructional commitments. One teacher remarked that "the training is very supportive and gives us new insights, but sometimes we struggle to arrange the schedule." This

finding corroborates the documented workload tensions identified in DFAT evaluation reports (Cassity et al., 2022) and underscores the importance of designing professional development activities that are sensitive to teachers' existing temporal constraints.

Output: School principals and education officials observed tangible improvements in teacher competencies and classroom practice quality, with one district official stating, "we can see real improvement in teaching ability, although teachers still need continued mentoring." Teachers confirmed that program participation enhanced their instructional repertoire, while mentors emphasized the importance of teacher commitment and institutional accountability in consolidating training gains. The recurring qualification that improvement is evident but incomplete reflects the incremental and non-linear nature of pedagogical change in complex educational systems.

Outcome-Impact: The most cautiously optimistic assessments emerged regarding long-term student outcomes. Teachers and principals observed increased student participation in reading and mathematics activities, with one teacher reporting "students are more active in reading and counting after this program." One mentor offered a nuanced assessment: "the benefits are clearly visible, especially in student literacy, but for the long term, we need consistent support from central and regional government." These insights underscore the importance of systemic political will and sustained institutional investment as necessary complements to school-level professional development interventions a theme that resonates strongly with the recent literature on educational reform sustainability in fragile states (Quinn, 2023; UNESCO, 2023).

CONCLUSION

This study has presented a comprehensive evaluation of the Professional Learning and Mentoring Program (ALMA) in Timor-Leste elementary schools, employing an integrated Logic Model and SmartPLS analytical framework. The findings yield several significant conclusions. ALMA's Input and Process components have broadly achieved their operational objectives. High stakeholder satisfaction with program resources, training activities, classroom observation protocols, peer learning sessions, and technology mentoring reflects the program's strong foundational architecture. In this context, the emphasis on competency-based and practice-oriented professional learning further strengthens the relevance of program inputs and processes in supporting effective teacher development. SmartPLS analysis confirmed that Input exerts a large, significant positive effect on Process effectiveness ($\beta = 0.716$, $p < 0.001$), and that Process, in turn, strongly predicts Output quality ($\beta = 0.738$, $p < 0.001$). These findings validate the program's operational theory of change and affirm the value of its multi-component, school-based professional development model.

The relationship between program Outputs and Outcomes is complex, marginally negative, and not statistically propagated through the full mediation chain. This finding does not negate the program's value but rather illuminates the structural conditions under which professional development investments fail to translate automatically into student learning gains. The weak Output→Outcome pathway is attributable to a combination of temporal lags, workload displacement effects, the influence of unmeasured external factors on student achievement, and the inherently protracted timeline of deep pedagogical change. These insights call for targeted program enhancements: specifically, strengthening post-training coaching and follow-up support mechanisms, reducing administrative burdens on teachers, and aligning ALMA's SCL orientation with the accountability frameworks of the national curriculum and assessment system.

Longitudinal EGRA and EGMA data demonstrate consistent improvements in student literacy and numeracy across three assessment phases, providing encouraging evidence that ALMA's professional development investments are ultimately contributing to improved foundational learning outcomes, even if the statistical pathway between training outputs and student outcomes is not yet sufficiently robust to achieve significance in cross-sectional SEM analysis. The observed improvements also highlight the importance of sustained, practice-based professional development in enhancing instructional effectiveness over time.

From a policy standpoint, these findings support continued investment in ALMA as a structural component of Timor-Leste's educational reform agenda, subject to the following recommendations: (1) intensifying post-training mentoring and in-classroom coaching to bridge the implementation gap between professional learning and sustained instructional change; (2) reducing the administrative and scheduling burden of program participation to mitigate workload displacement effects on classroom teaching time; (3) strengthening data systems for tracking long-term student outcomes at program schools; and (4) embedding ALMA's SCL development framework within a broader national teacher certification and professional recognition system to create sustainable institutional incentives for continuous pedagogical improvement.

This study contributes to the evaluative literature on teacher professional development in post-conflict, developing country contexts by demonstrating the utility of integrating Logic Model evaluation with SmartPLS path modeling. Future research should employ longitudinal designs with larger samples and include school- and student-level control variables to more precisely isolate the causal effects of ALMA on long-term student learning outcomes.

Conflict of Interest

The authors declare no conflict of interest in the publication of this research. All contributions were made with full academic independence, without any financial or personal relationships that could have influenced the results or interpretations presented.

Acknowledgements

The authors sincerely thank all respondents, including school principals, classroom teachers, national and educational mentors, and district education officials, for their generous participation in this research. The authors also extend their gratitude to the supervisor and co-supervisor from Yogyakarta State University, as well as to the ICFP Campus in Baucau, Timor-Leste, for their valuable support

REFERENCES

- Abdigapbarova, U., & Zhiyenbayeva, N. (2023). Organization of student-centered learning within the professional training of a future teacher in a digital environment. *Education and Information Technologies*, 28(1), 647–661. <https://doi.org/10.1007/s10639-022-11159-5>
- Acar, S. P., Akgün Özpolat, E., & Çomoğlu, İ. (2023). Teacher-tailored classroom observation for professional growth of {EFL} instructors. *Journal on Efficiency and Responsibility in Education and Science*, 16(1), 26–35. <https://doi.org/10.7160/eriesj.2023.160103>
- AKSU, D., ÇALIŞIR, M., & SELLÜM, F. S. (2023). Reflective practices of pre-service teachers: Self-reflections on micro teaching. *Sosyal Bilimler ve Eğitim Dergisi*, 6(2), 254–274. <https://doi.org/10.53047/josse.1362951>
- Almansouri, M., Pange, J., & Kallery, M. (2022). Impacts on student learning and skills and implementation challenges of two student-centered learning methods applied in online education. *Sustainability*, 14(15), 9625. <https://doi.org/10.3390/su14159625>
- Almulla, M. A. (2023). Constructivism learning theory: {A} paradigm for students' critical thinking, creativity, and problem solving to affect academic performance in higher education. *Cogent Education*, 10(1), 2172929. <https://doi.org/10.1080/2331186X.2023.2172929>
- Attard, C., & Holmes, K. (2020). "It gives you that sense of hope": An exploration of technology use to mediate student engagement with mathematics. *Heliyon*, 6(1), e02945. <https://doi.org/10.1016/j.heliyon.2019.e02945>
- Bottazzi, P., Seck, S. M., Niang, M., & Moser, S. (2023). Beyond motivations: A framework unraveling the systemic barriers to organic farming adoption in northern Senegal. *Journal of Rural Studies*, 104, 103158. <https://doi.org/10.1016/j.jrurstud.2023.103158>
- Brown, C., Daly, A., & Liou, Y. H. (2020). Improving trust, improving schools: Findings from a social network analysis of 43 primary schools in England. *Journal of Professional Capital and Community*, 1(1), 1–19. <https://doi.org/10.1108/JPC-08-2015-0004>
- Cassity, E., & Chainey, J. (2020). *Education Analytics Service (EAS) Teacher Development Multi-Year Study Series: Timor-Leste*.
- Cassity, E., Chainey, J., & Wong, D. (2022). *Education Analytics Service: Teacher Development Multi-Year Studies (Timor-Leste: Final Report)*.
- De La Iglesia, B., Forteza, D., & Duma, L. (2024). Peer feedback in teacher professional development: A systematic review. *Revista Española de Pedagogía*, 82(288), 335–358. <https://doi.org/10.22550/2174-0909.4044>
- Diwakar, S., Kolil, V. K., Francis, S. P., & Achuthan, K. (2023). Intrinsic and extrinsic motivation among students for laboratory courses -- Assessing the impact of virtual laboratories. *Computers & Education*, 198, 104758. <https://doi.org/10.1016/j.compedu.2023.104758>
- Elmassah, S., Biltagy, M., & Gamal, D. (2022). Framing the role of higher education in sustainable development: A case study analysis. *International Journal of Sustainability in Higher Education*, 23(2), 320–355. <https://doi.org/10.1108/IJSHE-05-2020-0164>
- Fandos-Herrera, C., Jiménez-Martínez, J., Orús, C., Pérez-Rueda, A., & Pina, J. M. (2023). The influence of personality on learning outcomes and attitudes: The case of discussants in the classroom. *The International Journal*

of *Management Education*, 21(1), 100754. <https://doi.org/10.1016/j.ijme.2022.100754>

Fillah, A., Roesminingsih, E., & Sondang Sumbawati, M. (2021). The leadership of school principal in the education era 4.0. *Nidhomul Haq: Jurnal Manajemen Pendidikan Islam*, 6(2), 244–250.

Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences*, 111(23), 8410–8415. <https://doi.org/10.1073/pnas.1319030111>

Goczek, Ł., Witkowska, E., & Witkowski, B. (2021). How does education quality affect economic growth? *Sustainability*, 13(11), 6437. <https://doi.org/10.3390/su13116437>

Godfrey, D. (2020). From peer review to collaborative peer enquiry: Action research for school improvement and leadership development. *London Review of Education*, 18(3), 373–389. <https://doi.org/10.14324/LRE.18.3.04>

Haelermans, C., Aarts, B., Smeets, C., Kirezi, J. C., & Cabus, S. J. (2022). The effectiveness of peer-led and trainer-led e-learning leadership training for school leaders: A randomized experiment in Rwanda. *International Journal of Educational Management and Development Studies*, 18–26. <https://doi.org/10.53378/352901>

Hair, J. F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2022). *A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM)* (3rd ed.). SAGE Publications. <https://uk.sagepub.com/en-gb/eur/a-primer-on-partial-least-squares-structural-equation-modeling-pls-sem/book270548>

Hair, J. F., Risher, J. J., Sarstedt, M., & Ringle, C. M. (2019). When to use and how to report the results of PLS-SEM. *European Business Review*, 31(1), 2–24. <https://doi.org/10.1108/EBR-11-2018-0203>

Hennessy, S., D'Angelo, S., McIntyre, N., Koomar, S., Kreimeia, A., Cao, L., Brugha, M., & Zubairi, A. (2022). Technology use for teacher professional development in low- and middle-income countries: {A} systematic review. *Computers and Education Open*, 3, 100080. <https://doi.org/10.1016/j.caeo.2022.100080>

Henseler, J., Ringle, C. M., & Sarstedt, M. (2023). Using partial least squares path modeling in international advertising research: Basic concepts and recent issues. In S. Okazaki (Ed.), *Handbook of Research on International Advertising* (pp. 252–276). Edward Elgar. <https://doi.org/10.4337/9781781001042.00023>

Hull, L., Mandy, W., Lai, M. C., Baron-Cohen, S., Allison, C., Smith, P., & Petrides, K. V. (2019). Development and validation of the Camouflaging Autistic Traits Questionnaire (CAT-Q). *Journal of Autism and Developmental Disorders*, 49(3), 819–833. <https://doi.org/10.1007/s10803-018-3792-6>

Lavilla-Gracia, M., Pueyo-Garrigues, M., Fotopoulou, M., Canga-Armayor, A., Esandi, N., Alfaro-Díaz, C., & Canga-Armayor, N. (2023). Training nursing students in motivational interviewing for alcohol misuse: A mixed method study. *Nurse Education Today*, 125, 105775. <https://doi.org/10.1016/j.nedt.2023.105775>

Merino, C., Pacheco, G., Arenas-Martija, A., Becerra, R., & Solís-Pinilla, J. (2025). Continuing professional development in teachers: Insights for designing a formative trajectory in scientific education. *Frontiers in Education*, 10. <https://doi.org/10.3389/feduc.2025.1537502>

Nayir, F., Karaman, A., Ozturk, M., & Cinkir, S. (2023). The impact of active learning on academic achievement in {K-12} settings: {A} meta-analytic review. *Journal on School Educational Technology*, 18(3), 1–14. <https://doi.org/10.26634/jsch.18.3.19556>

Nguyen, Y. T. X., Ha, X. Van, & Tran, N. H. (2022). Vietnamese primary school teachers' needs for professional development in response to curriculum reform. *Education Research International*. <https://doi.org/10.1155/2022/4585376>

Ongori, J. G., Kitainge, K., & Kipkoech, L. (2020). Kenya Education Management Institute's capacity building programme and public primary school head teachers' competencies in teacher supervision in Uasin Gishu County, Kenya. *Asian Journal of Education and Social Studies*, 46–53. <https://doi.org/10.9734/ajess/2020/v11i130283>

Partnership for Human Development. (2016). *Partnership for Human Development first six-monthly report: Vol. I*.

Payton, C., Kumar, G. S., Kimball, S., Clarke, S. K., AlMasri, I., & Karaki, F. M. (2022). A logic model framework for planning an international refugee health research, evaluation, and ethics committee. *Health Promotion Practice*, 23(5), 852–860. <https://doi.org/10.1177/15248399211035703>

Poulou, M. S., Reddy, L. A., & Dudek, C. M. (2023). Teachers and school administrators' experiences with professional development feedback. *Frontiers in Psychology*, 14, 1074278. <https://doi.org/10.3389/fpsyg.2023.1074278>

Quinn, M. (2023). Implementing curriculum “more or less” in Timor-Leste. *International Journal on Studies in Education*, 5(4), 476–495. <https://doi.org/10.46328/ijonse.148>

- Quinn, M., & Buchanan, J. (2022). "A contribution to my country": Professional lives of teachers in Timor-Leste. *Asia Pacific Journal of Education*, 42(3), 497–512. <https://doi.org/10.1080/02188791.2021.1880370>
- Ringle, C. M., Sarstedt, M., Mitchell, R., & Gudergan, S. P. (2023). Partial least squares structural equation modeling in HRM research. *International Journal of Human Resource Management*, 31(12), 1617–1643. <https://doi.org/10.1080/09585192.2017.1416655>
- Roca-Campos, E., Renta-Davids, A. I., Marhuenda-Fluixá, F., & Flecha, R. (2021). Educational impact evaluation of professional development of in-service teachers: The case of the dialogic pedagogical gatherings at Valencia "On Giants' Shoulders". *Sustainability*, 13(8), 4275. <https://doi.org/10.3390/su13084275>
- RTI International. (2023). *Early Grade Reading Assessment (EGRA): Manual for Adaptation and Data Use*. <https://www.rti.org/egra>
- Sahonero-Alvarez, G., & Humberto, C. (2018). Implementation issues of student-centered learning based engineering education in developing countries universities. *2018 World Engineering Education Forum -- Global Engineering Deans Council (WEEF-GEDC)*. <https://doi.org/10.1109/WEEF-GEDC.2018.8629701>
- Sancar, R., Atal, D., & Deryakulu, D. (2021). A new framework for teachers' professional development. *Teaching and Teacher Education*, 101, 103305. <https://doi.org/10.1016/j.tate.2021.103305>
- Sentana Putra, I. G. N. (2025). Integrasi RFM, K-Means, dan XGBoost untuk Optimalisasi Retensi Pelanggan pada Online Retail. *KONSTELASI: Konvergensi Teknologi Dan Sistem Informasi*, 5(1), 1–10. <https://doi.org/10.24002/konstelasi.v5i1.11409>
- Sims, S., & Fletcher-Wood, H. (2021). Identifying the characteristics of effective teacher professional development: a critical review. *School Effectiveness and School Improvement*, 32(1), 47–63. <https://doi.org/10.1080/09243453.2020.1772841>
- Tholibon, D. A., Nujid, M. M., Mokhtar, H., Rahim, J. A., Rashid, S. S., Saadon, A., Tholibon, D., & Salam, R. (2022). The factors of students' involvement on student-centered learning method. *International Journal of Evaluation and Research in Education*, 11(4), 1637–1646. <https://doi.org/10.11591/ijere.v11i4.22314>
- UNESCO. (2023). *Global Education Monitoring Report 2023: Technology in Education -- A Tool on Whose Terms?* UNESCO Publishing. <https://doi.org/10.54676/UZQV8501>
- United Nations. (2015). *Transforming our world: The 2030 Agenda for Sustainable Development*.
- Vygotsky, L. S. (1978). *Mind in Society: The Development of Higher Psychological Processes*. Harvard University Press.
- W.K. Kellogg Foundation. (2004). *Logic Model Development Guide*. W.K. Kellogg Foundation.
- Yee, S., Deshler, J., Rogers, K. C., Petrulis, R., Potvin, C. D., & Sweeney, J. (2022). Bridging the gap between observation protocols and formative feedback. *Journal of Mathematics Teacher Education*, 25(2), 217–245. <https://doi.org/10.1007/s10857-020-09485-x>