

Research Article

Evaluation of the Implementation of Pedagogic Competence Training Policy: Analysis of the Effectiveness of the Teacher Human Resource Development Program (Dili District, Timor-Leste)

Dulce Pereira Lopes ^{1*}, Suparno ², Aris Toening Winarni ³

1. Program Studi Magister Administrasi Publik Gak Ilmu Sosial dan Politik, Universitas 17 Agustus 1945 Semarang, Indonesia; email: lopesdulcepereira913@gmail.com
 2. Program Studi Magister Administrasi Publik Gak Ilmu Sosial dan Politik, Universitas 17 Agustus 1945 Semarang, Indonesia; email: suparno@untagsmg.ac.id
 3. Program Studi Magister Administrasi Publik Gak Ilmu Sosial dan Politik, Universitas 17 Agustus 1945 Semarang, Indonesia; email: aris-toening@untagsmg.ac.id
- * Corresponding Author: Dulce Pereira Lopes

Abstract. This study investigates the effectiveness of teacher training policy implementation in Dili Regency, Timor-Leste, utilizing the CIPP (Context, Input, Process, Product) evaluation model. A structured survey was administered to 150 teachers and education stakeholders to assess key determinants influencing implementation outcomes. Quantitative analysis employed Pearson correlation, multiple regression, and path analysis to examine relationships among variables. Findings revealed strong correlations between all independent variables and implementation effectiveness ($r > 0.60$, $p < 0.01$). The multiple regression model accounted for 76.3% of the variance in implementation effectiveness ($R^2 = 0.763$, $F = 123.45$, $p < 0.001$), with Policy Standards exerting the greatest influence ($\beta = 0.243$), followed by Resources ($\beta = 0.203$) and Implementer Disposition ($\beta = 0.201$). Path analysis confirmed the significance of all direct causal pathways, reinforcing the robustness of the model. These results underscore the critical role of clear policy standards, adequate resource allocation, and implementer commitment in enhancing the success of teacher training programs. The study contributes empirical evidence to the discourse on education reform in developing and post-conflict regions, highlighting strategic priorities for policy makers and practitioners. By aligning training initiatives with contextual realities and stakeholder capacities, the findings offer actionable insights for improving teacher professional development in Timor-Leste and similar settings.

Received: June 30, 2025;
Revised: July 30, 2025;
Accepted: August 26, 2025;
Published: August 28, 2025
Curr. Ver.: September 02, 2025

Keywords: CIPP Evaluation; Implementation Effectiveness; Teacher Professional; Teacher Training; Timor-Leste

1. Introduction

Human resource (HR) development is a fundamental challenge in public administration, particularly for developing countries facing limited institutional capacity and resources. In the context of modern governance, investment in civil servant capacity building is a key determinant of the successful implementation of public policies and the achievement of national development goals (Grindle, 2017). Timor-Leste, a newly independent nation in 2002, faces unique complexities in building an effective public administration system, including in the Education sector, a strategic priority for national development. The Education sector in Timor-Leste is not only a domain of social policy but also represents the state's ability to manage human resource development programs systematically. The Timor-Leste government, through Law No. 23/2010 of the Democratic Republic of Timor-Leste, has established a mandatory competency framework for teaching staff encompassing four dimensions: proficiency in the official language, technical-scientific knowledge, pedagogical techniques, and professional ethics. This regulation reflects state-building efforts through



Copyright: © 2025 by the authors.
Submitted for possible open
access publication under the
terms and conditions of the
Creative Commons Attribution
(CC BY SA) license
(<https://creativecommons.org/licenses/by-sa/4.0/>)

standardizing public sector HR qualifications, which aligns with the principles of New Public Management to improve bureaucratic accountability and performance.

The implementation of the teacher competency development policy through a pedagogical training program since 2020 is a concrete manifestation of the capacity building agenda mandated in the 2011-2030 Strategic Development Plan (PED). However, like many development programs in post-conflict countries, the implementation of this policy faces various systemic challenges related to organizational capacity, resource allocation, and institutional coordination. This situation demands a comprehensive evaluation to identify factors influencing the program's effectiveness and provide recommendations for improvement to policymakers. Public administration literature has long recognized the importance of capacity building as an instrument for improving public sector performance. Theorists such as Grindle (2007) and Morgan (2006) emphasize that capacity building is not simply the transfer of technical skills, but rather a complex process involving individual, organizational, and institutional change. In this context, pedagogical competency training programs can be understood as capacity-building interventions aimed at enhancing individual teacher abilities while strengthening the organizational capacity of schools and the Education system as a whole.

The theoretical framework for policy implementation developed by Pressman and Wildavsky (1973), and refined by Mazmanian and Sabatier (1983), provides an analytical perspective for understanding the complexities of implementing teacher training programs. This Theory identifies that successful implementation is influenced by factors such as clarity of policy objectives, resource adequacy, institutional support, and commitment of implementing agencies. In the context of Timor-Leste, these factors become even more crucial given the limited administrative capacity and challenges of inter-agency coordination. Program evaluation studies in developing countries indicate that the effectiveness of capacity-building programs is highly dependent on contextual factors and institutional arrangements (World Bank, 2012). Christensen and Lægveid's (2015) research on administrative reforms in small states indicates that country size and institutional characteristics influence policy implementation patterns. However, literature specifically analyzing teacher human resource development programs within the framework of public administration in small island developing states is still limited.

In the context of program evaluation, the framework of Rossi, Lipsey, and Freeman (2018) provides a methodological foundation for comprehensive process and outcome evaluation. This approach allows for identification not only of what works, but also of how and why interventions succeed or fail. A gap in the existing literature lies in the lack of studies integrating policy implementation Theory perspectives with evaluation research in the context of capacity building in the Education sector in post-conflict countries. Furthermore, literature on Public Sector Human Resource Management (PSHRM) in developing countries shows that training programs often face implementation challenges such as inadequate needs assessments, poor training design, and limited follow-up mechanisms (Berman et al., 2019). Larbi's (1999) research on public sector reforms in Africa suggests that the sustainability of training interventions depends on institutional embedding and organizational support systems. However, the application of these findings to the specific context of Timor-Leste, a country with unique post-independence characteristics, still requires in-depth empirical investigation.

Although the pedagogical competency training program has been implemented for four years, there has been no systematic evaluation analyzing the program's effectiveness from a public administration perspective. This situation creates a significant knowledge gap for policymakers in understanding the factors influencing the successful implementation of capacity-building programs in the Education sector. From a theoretical perspective, the limited research on policy implementation in the context of small post-conflict states requires empirical contributions to enrich the body of knowledge in public administration. Another relevant issue is the lack of adequate feedback mechanisms for ongoing program monitoring and evaluation. These limitations impact the system's ability to engage in adaptive management and continuous improvement, fundamental principles of good governance. Furthermore, coordination challenges between stakeholders and limited resource allocation add to the complexity of program implementation. Based on these identified problems, this study poses three main research questions: (1) How effective is the implementation of the pedagogical competency training policy from a policy implementation Theory perspective?

(2) What organizational and institutional factors Influence the success or failure of the teacher capacity building program in Dili Regency? (3) What program evaluation model is appropriate for the institutional context of Timor-Leste in improving public sector accountability and performance management?

This research makes significant contributions at three levels of analysis. First, from a theoretical perspective, this study enriches the policy implementation literature by exploring the application of the Theory in the context of post-conflict small states. The research findings will provide insights into how contextual factors Influence implementation processes and outcomes, thereby contributing to the development of more nuanced implementation Theory. Second, the practical contribution of this research lies in providing evidence-based recommendations for policymakers in Timor-Leste. The evaluation results will assist the government in redesigning more effective capacity-building programs, improving resource allocation, and strengthening institutional mechanisms for future development programs. The resulting recommendations can also serve as a reference for other sectors in implementing human resource development programs.

Third, the methodological contribution of this research is the development of an evaluation framework appropriate to the institutional context of Timor-Leste. This framework can be adapted for evaluating other public programs in countries with similar characteristics, thus contributing to the advancement of evaluation practices in developing countries. From a policy-relevant perspective, this research supports the Sustainable Development Goals (SDGs) agenda, particularly Goal 4 on quality Education and Goal 16 on strong institutions. The research findings will provide insights into how strengthening public sector capacity can contribute to the sustainable achievement of development objectives. This research uses a mixed-methods approach with a sequential explanatory design, where a quantitative phase is followed by a qualitative phase to provide explanatory insights. The geographic scope of the research is limited to Dili Regency, which represents an urban setting with a concentration of educational institutions and administrative capacity. The temporal scope covers the program implementation period from 2020 to 2024, with a focus on the analysis of implementation processes and immediate outcomes.

2. Literature Review

Pedagogical Competency Training Policy

The pedagogical competency training policy is a systematic government intervention to develop teachers' abilities to integrate content knowledge with effective pedagogical strategies. Pedagogical competency encompasses instructional planning, learning delivery, classroom management, and evaluation skills that enable the transformation of knowledge into meaningful learning (Padillo et al., 2021). The theoretical foundation of this concept is rooted in Shulman's (1986) Pedagogical Content Knowledge (PCK), which defines PCK as a teacher's understanding of the most useful forms of representation, powerful analogies, illustrations, examples, explanations, and demonstrations, to make subject matter understandable to others (McComas, 2014). PCK is the unique knowledge possessed by teachers to transform specific subject matter for students while considering potential misconceptions (Depaepe & König, 2018).

The development of the PCK concept gave rise to the TPACK Framework (Technology, Pedagogy, and Content Knowledge), which integrates the technological dimension. TPACK expands Shulman's PCK construct by incorporating technological knowledge as a third component that interacts complexly with content and pedagogical knowledge (Mishra & Koehler, 2006). Pedagogical competence in higher Education involves the capability to meet complex demands by utilizing psychological resources (International Labor Organization, 2008). The implementation of pedagogical competence training policies utilizes Continuous Professional Development (CPD) Theory, which emphasizes lifelong learning. CPD is vital for teachers to broaden and deepen their knowledge, stay abreast of new research, current tools, and practices, and respond to changing student needs (OECD, 2024).

Organizational Factors

Organizational factors are internal elements within the school structure or program-hosting organization that significantly Influence the success of capacity building implementation. School leadership support is a key determinant, with principals actively promoting teacher participation, providing adequate resources, and integrating training

outcomes into daily teaching practices, enhancing the overall sustainability of the program (Harris & Jones, 2011). Conversely, a lack of commitment from leadership can lead to training programs being perceived as merely an administrative burden.

A school organizational culture that supports collaboration, innovation, and continuous learning has been shown to promote the success of capacity-building programs. A work environment that encourages teachers to share knowledge and best practices and does not discourage experimentation with new teaching methods will maximize the positive impact of the training provided (Fullan, 2007). Furthermore, the appropriate allocation of resources, both financial and material, such as books, technological devices, and adequate classroom facilities, is a crucial prerequisite for program success. Capacity-building programs can potentially fail to achieve their goals if teachers lack access to the tools and resources necessary to apply the new skills they have acquired through training (UNESCO, 2014).

Timor-Leste's Public Sector Evaluation Model

Timor-Leste, as a young nation, faces complex challenges in developing its public sector institutions. Post-conflict countries like Timor-Leste generally experience limited institutional capacity, policy fragmentation, and high dependence on international aid (D'Arcy & Nasi, 2018). In this context, effective accountability systems and measurable performance management are crucial to ensuring optimal resource utilization for the public good (Brinkerhoff, 2011). Traditional evaluation models, which rely on a positivistic approach with an emphasis on quantitative data and statistical analysis (Patton, 2015), have significant limitations when applied to complex development programs in developing countries. These limitations include difficulty isolating variables, a lack of reliable historical data, and a failure to capture the socio-cultural nuances that influence program outcomes (Stern et al., 2012).

Contemporary literature recommends a more realistic and contextual approach. Participatory evaluation emphasizes the active involvement of local stakeholders throughout the evaluation cycle, not only enhancing the validity of findings through the integration of local perspectives but also serving as a capacity-building tool (Cousins & Whitmore, 1998). Meanwhile, a Theory of Change (ToC)-based evaluation provides a systematic framework for mapping a logical path from inputs to long-term impacts, explaining not only whether a program is successful but also why and how (Vogel, 2012). The integration of participatory approaches and ToC is considered the most appropriate evaluation model for Timor-Leste (Jackson, 2013). This integrated model creates a comprehensive framework for strengthening accountability through a transparent process and provides a structured roadmap for performance management, enabling regular evaluation and timely strategic adjustments (Stern et al., 2012).

Previous Research and Theoretical Framework

Several previous studies have identified critical factors in the implementation of teacher training programs. Nuraini and Supriyadi (2021) found that training is effective when supported by adequate resources, communication between stakeholders, and post-training follow-up. Lemos et al. (2020) identified that strengthening teacher competencies in Timor-Leste is influenced by organizational culture, government role, and participatory training design, supporting the importance of implementer disposition and institutional context. Widodo and Jannah (2022) highlighted factors contributing to policy implementation failure, including unclear vertical communication and a lack of monitoring. Andrade and Almeida (2019) identified obstacles to training implementation due to limited resources and weak ongoing evaluation, supporting the need for a CIPP-based evaluation model for the Timor-Leste context. Lopes and Soares (2022) emphasized that training programs need to adopt a contextual and participatory evaluation approach to align with local realities.

The theoretical framework of this research is based on Law No. 14 of 2005 concerning Teachers and Lecturers and UNESCO standards (2018) define pedagogical competence as encompassing student understanding, learning planning and implementation, and evaluation of learning outcomes. Grindle (1997) and Morgan (2006) emphasize that implementing organizations require institutional capacity encompassing human resources, structures, regulations, and management systems to run effective programs. The CIPP evaluation model by Stufflebeam (2003) is used to assess the effectiveness of training implementation through the components of Context, Input, Process, and Product.

The conceptual framework is built on the integration of policy implementation Theory and the CIPP evaluation model as an analytical tool. The flow begins with the teacher pedagogical competency training policy, followed by policy implementation in the field using

the Mazmanian and Sabatier approach, which considers actors, structures, and context. Then, an evaluation of program effectiveness through the CIPP model is carried out. Finally, an analysis of the results and recommendations for evidence-based Education policy is concluded. The research hypotheses are formulated as follows: H1 - There is a positive relationship between the quality of training policy implementation and the improvement of teachers' pedagogical competence, and H2 - Contextual factors such as organizational support, resources, and the social environment significantly Influence training effectiveness. This approach allows for objective measurement of effectiveness and the identification of inhibiting/enabling factors based on statistical data to develop applicable recommendations in the Timor-Leste context.

3. Research Methods

This study employed a quantitative approach with explanatory research, aiming to test causal relationships between variables through a cross-sectional survey research design with individual units of analysis. The study population consisted of 1,250 elementary school teachers in Dili Regency who had participated in the pedagogical competence training program, 45 Education officials and administrators at the district and national levels, and 120 principals and Education supervisors, for a total population of 1,415. Using the Slovin formula with a 5% margin of error, a sample of 312 respondents was obtained, distributed among 280 teachers (90%), 12 officials/administrators (4%), and 20 principals/supervisors (6%). The data collection method used a structured questionnaire with a Likert scale of 1-5, consisting of 85 questions for all variables, supplemented by in-depth interviews with key informants and direct observation at the training site. The validity and Reliability of the instrument were tested through a pilot study. Data analysis techniques included descriptive analysis using descriptive statistics, frequency distribution, and crosstabulation, as well as inferential analysis using tests of normality, homoscedasticity, multicollinearity, and linearity. The primary analysis used Pearson Product-Moment correlation to test the strength and direction of relationships between variables, supplemented by multiple linear regression analysis to determine the effect of independent variables on the dependent variable simultaneously and partially.

4. Results And Discussion

Respondent Characteristics Based on Demographics

Based on the results of the study of 312 respondents, the demographic characteristics showed a fairly even distribution, with 170 female respondents (54.5 percent) and 142 male respondents (45.5 percent). In terms of age, the majority of respondents were in the 30-40 age range (134 or 42.9 percent), followed by the 41-50 age group (85 or 27.2 percent), the under-30 age group (78 or 25 percent), and the over-50 age group (15 or 4.8 percent). The respondents' educational level was dominated by bachelor's degree graduates (223 or 71.5 percent), followed by master's degree graduates (44 or 14.1 percent), and diploma graduates (45 or 14.4 percent).

Based on length of service, the majority of respondents (156 respondents, or 50 percent) had 5-15 years of work experience, followed by 89 respondents (28.5 percent) with less than 5 years of work experience, and 67 respondents (21.5 percent) with more than 15 years of work experience. Based on school level, the majority of respondents were elementary school graduates (168 respondents, or 53.8 percent), junior high school graduates (89 respondents, or 28.5 percent), and high school graduates (55 respondents, or 17.6 percent). These characteristics indicate that the respondents have a representative background and experience in Education.

Based on school location, the majority of respondents (187 respondents, or 59.9 percent) came from urban areas, while 125 respondents (40.1 percent) came from rural areas. In terms of school status, the majority of respondents (245 respondents, or 78.5 percent) taught at public schools, while 67 (21.5 percent) taught at private schools. Based on school accreditation, the majority of respondents (156, or 50 percent) came from B-accredited schools, followed by A-accredited schools (89, or 28.5 percent), and C-accredited schools (67, or 21.5 percent).

Descriptive Statistical Analysis

The table shows the results of the evaluation of the implementation of the teacher training policy using the CIPP (Context, Input, Process, Product) model. Data were obtained

from respondents using a Likert scale with the following categories: high (>3.40), moderate ($2.60-3.40$), and low (<2.60). The analysis includes the mean, standard deviation, and category of each indicator.

Table 1. Descriptive Statistical.

Indicator	Mean	Std. Deviation	Category
Program budget (X2.1)	3.67	0.89	High
Training facilities (X2.2)	3.45	0.95	Medium
Training staff (X2.3)	3.78	0.82	High
Learning materials (X2.4)	3.52	0.91	High
Implementation time (X2.5)	3.61	0.89	High
Mean Construct X2	3.23	1.02	Medium
Identification of training needs (Y1.1)	3.15	1.08	Medium
Analysis of educational problems (Y1.2)	3.67	0.85	High
Relevance to national priorities (Y1.3)	3.54	0.92	High
Stakeholder support (Y1.4)	3.29	0.98	Medium
Mean Y1	3.38	0.97	Medium
Program design quality (Y2.1)	3.72	0.86	High
Trainer qualifications (Y2.2)	3.58	0.91	High
Budget suitability (Y2.3)	3.81	0.79	High
Completeness of materials (Y2.4)	3.64	0.88	High
Process (Y3)	3.69	0.86	High
Quality of training implementation (Y3.1)	3.57	0.89	High
Participant participation (Y3.2)	3.74	0.82	High
Monitoring during the process (Y3.3)	3.21	1.04	Medium
Problem solving (Y3.4)	3.63	0.87	High
Mean Y3			
Product (Y4)	3.78	0.83	High
Improved teacher competency (Y4.1)	4.02	0.74	High
Changes in teaching practices (Y4.2)	3.45	0.94	Medium
Participant satisfaction (Y4.3)	3.52	0.92	High
Impact on students (Y4.4)	3.69	0.86	High
Program budget (X2.1)			
Training facilities (X2.2)	3.95	0.77	High
Training staff (X2.3)	3.68	0.85	High
Learning materials (X2.4)	3.87	0.79	High
Implementation time (X2.5)	3.54	0.91	High
Mean Y4	3.76	0.83	High

The evaluation results indicate that the implementation of the teacher training policy is categorized as good. The context evaluation (3.61) achieved a high rating, indicating clear and relevant program objectives. The input evaluation (3.38) was categorized as moderate, with weaknesses in budget and facilities. The process evaluation (3.69) was categorized as high, supported by excellent participant participation (4.02). The product evaluation (3.76) was categorized as high, indicating a significant increase in teacher competency. Overall, the program successfully achieved its objectives, although improvements are needed in input aspects, particularly the training budget and facilities.

Construct Validity Test

The table displays the results of the validity and Reliability test of the research constructs using Confirmatory Factor Analysis (CFA). The parameters measured included loading factor (>0.7), Cronbach's Alpha (CA >0.7), and Average Variance Extracted (AVE >0.5) to determine the validity and Reliability of each construct in the implementation research model.

Table 2. Measurement Test.

Construct	Item	Loading Factor	CA	AVE	Status
Policy Standards and Objectives (X1)			0.847	0.685	Reliabel
a). Clarity of program objectives	X1.1	0.823			Valid
b). Competency standard specifications	X1.2	0.789			Valid
c). Consistency of national policies	X1.3	0.856			Valid
d). Relevance to teacher needs					
e). Realistic achievement targets	X1.4	0.794			Valid
Resources (X2)			0.823	0.589	Reliabel
a). Program budget	X2.1	0.756			Valid
b). Training facilities	X2.2	0.742			Valid
c). Trainers	X2.3	0.824			Valid
d). Learning materials	X2.4	0.801			Valid
e). Implementation time	X2.5	0.778			Valid
Interorganizational Communication X3			0.639	0.639	Reliabel
a). Communication between the Ministry of Education and INFORDEPE	X3.1	0.812			
b). Timeliness of information	X3.2	0.798			Valid
c). Coordination between institutions	X3.3	0.734			Valid
d). Consistency of national policies	X3.4	0.845			Valid
e). Feedback mechanisms	X3.5	0.769			Valid
Characteristics of Implementing Institutions X4			0.871	0.664	Reliabel
a). INFORDEPE organizational capacity	X41	0.789			
b). Organizational structure	X42	0.823			
c). Human resource competency	X43	0.806			
d). Management system	X44	0.831			
e). Experience in implementing	X45	0.798			
Economic, Social, and Political Conditions X5			0.834	0.601	Reliabel
a). Economic conditions	X51	0.856			
b). Community support	X52	0.823			
c) Political stability	X53	0.812			
c). Local culture	X54	0.847			
d). International donor support	X55	0.834			
Disposition of Implementers X6			0.889	0.694	Reliabel
a). Full commitment	X61	0.856			
b). Understanding the importance of the program	X62	0.823			
c). Motivation to implement results	X63	0.812			
d). Policy support	X64	0.847			
e). Sense of responsibility		0.834			
Organizational Factors Z1			0.863	0.708	Reliabel
a). Principal support	Z11	0.834			

b). Culture of continuous learning	Z12	0.845			
c). Reward system	Z13	0.823			
d). Support for the implementation of results	Z14	0.867			
e). Collaboration among teachers	Z15	0.851			
Institutional Factors	Z2		0.841	0.677	Reliabel
a). Clarity of regulations	Z21	0.812			
b). Accountability system	Z22	0.798			
c). Monitoring and evaluation mechanisms	Z23	0.834			
d). Standard operating procedures	Z24	0.845			
e). Documentation reporting system	Z25	0.829			
Effectiveness of Policy Implementation	Y		0.934	0.708	Reliabel
a). Context Evaluation	Y1	0.798			
b). Input Evaluation	Y2	0.789			
c). Process Evaluation	Y3	0.812			
d). Product Evaluation	Y4	0.823			
e). Impact Evaluation	Y5	0.829			

The analysis results show that all constructs meet validity and Reliability criteria. Factor loadings ranged from 0.734 to 0.867 (>0.7), indicating that the indicators validly measure their constructs. Cronbach's Alpha values of 0.639 to 0.934 (>0.6) and AVE values of 0.589 to 0.708 (>0.5) indicate good Reliability and convergent validity. The construct Policy Implementation Effectiveness had the highest Reliability (CA=0.934), while Interorganizational Communication had the lowest but still acceptable Reliability (CA=0.639). Overall, the research instrument is suitable for measuring the effectiveness of teacher training policy implementation with high accuracy. The Fornell-Larcker Criterion table tests the discriminant validity of constructs by comparing the square root of AVE (diagonal values) with the correlation between constructs (off-diagonal values). Discriminant validity is met if the diagonal value is greater than the correlation value in the same row and column, indicating each construct is unique and distinct.

Table 3. Fornell-Larcker Criterion.

Construct	X1	X2	X3	X4	X5	X6	Y	Z1	Z2
X1	0.828								
X2	0.623	0.767							
X3	0.645	0.567	0.598	0.689	0.578	0.833			
Y	0.742	0.698	0.721	0.756	0.687	0.734	0.841		
Z1	0.598	0.567	0.589	0.634	0.612	0.645	0.678	0.841	
Z2	0.612	0.589	0.598	0.623	0.634	0.598	0.689	0.656	0.823

The test results showed that all constructs met the Fornell-Larcker discriminant validity criteria. The diagonal values (square root of AVE) ranged from 0.767 to 0.841, all greater than the inter-construct correlations, which ranged from 0.567 to 0.756. Construct Y (Policy Implementation Effectiveness) had the highest discriminant validity (0.841), while X3 had several values approaching the limit but still met the criteria. The highest correlation occurred between X4 and Y (0.756), indicating a strong relationship, but the constructs remained distinct. Overall, all constructs were proven valid and statistically distinguishable.

Teacher Human Resource Development in Dili Regency, Timor Leste

Teacher training is a key instrument in policies to improve Education quality, particularly in teacher human resource development. According to Darling-Hammond, Hyler, & Gardner (2017), effective training policies must be based on teacher needs, relevant to the

local context, and provide ongoing post-training support. They emphasized that teacher training cannot stand alone but must be part of a long-term policy of teacher professional development. In Timor-Leste, improving the quality of teacher human resources has been a priority in the National Education Strategic Plan 2011–2030, which underscores the importance of training to equip teachers with basic pedagogical skills. The training program, organized by INFORDEPE (Instituto Nacional de Formação de Docentes e Profissionais da Educação), in collaboration with UNICEF and UNESCO, focused on improving the pedagogical capacity of primary school teachers (UNESCO, 2018).

However, the World Bank's Education Sector Analysis report in Timor-Leste (2020) noted that although training has reached many teachers, its effectiveness remains limited by weak implementation evaluation, a lack of professional trainers, and a mismatch between training materials and classroom practices. The teacher training policy is one form of strategic government intervention aimed at improving the quality of Education. Training is a crucial component of human resources (HR) development, particularly in the context of the teaching profession, which demands continuous competency development. According to UNESCO (2015), effective teacher training policies must align with national Education goals, local school needs, and current pedagogical developments. However, as noted by the World Bank (2020), the implementation of teacher training policies in Timor-Leste still faces structural obstacles, such as limited budgets, a lack of competent trainers, and a weak monitoring and evaluation system. Therefore, evaluating the implementation of this policy is crucial to assessing the effectiveness of the training in improving teacher competency and the quality of learning.

Correlation Analysis

The Pearson Product-Moment correlation table measures the strength and direction of the linear relationship between the independent variables and the dependent variable, policy implementation effectiveness. The analysis uses a significance level of 0.01 (two-tailed), with interpretation of correlation strength based on Cohen's classification to determine the degree of relationship between the research variables.

Table 4. Pearson Product-Moment Correlation

Variable	Effectiveness	Corelation	Interpretasi
Policy Standards	0.742**	Strong	Very Significant
Resources	0.698**	Strong	Very Significant
Communication	0.721**	Strong	Very Significant
Implementing Agencies	0.756**	Strong	Very Significant
Contextual Conditions	0.687**	Strong	Very Significant
Implementer Disposition	0.734**	Strong	Very Significant
Organizational Factors	0.678**	Strong	Very Significant
Institutional Factors	0.689**	Strong	Very Significant

The results of the Pearson correlation analysis showed a strong and highly significant positive relationship between all independent variables and the effectiveness of policy implementation ($p < 0.01$). The characteristics of the implementing agency had the highest correlation ($r = 0.756$), followed by policy standards and objectives ($r = 0.742$), and implementer disposition ($r = 0.734$). Meanwhile, contextual conditions showed the lowest but still strong correlation ($r = 0.687$). According to Cohen (1988), all correlation values were in the strong category (0.60–0.799), indicating that each factor contributed significantly to the successful implementation of the teacher training policy, with the capacity of the implementing organization as the strongest predictor.

Multiple Linear Regression Analysis

The Model Summary table displays the results of the multiple linear regression analysis, which measures the ability of the independent variables to explain variation in the dependent variable. Parameters displayed include the correlation coefficient (R), coefficient of determination (R^2), adjusted R^2 , standard error, and an F-test to evaluate the overall strength and significance of the regression model.

Table 5. Model Summary.

Model	R	R ²	Adjusted R ²	Std. Error	F	Sig.
1	0.874	0.763	0.757	0.428	123.45	0.000

The analysis results show that the regression model has excellent predictive power. An R value of 0.874 indicates a very strong correlation between the independent variables and the effectiveness of policy implementation. The coefficient of determination (R²) of 0.763 explains that 76.3% of the variation in implementation effectiveness can be predicted by factors within the model, with the remaining 23.7% influenced by other variables. An adjusted R² of 0.757 indicates model consistency even with the addition of new variables. An F test of 123.45 with a significance of $p < 0.001$ proves that the model is statistically highly significant and suitable for prediction, with a relatively small standard error (0.428). The Coefficients table displays the results of the regression test, which measures the individual Influence of each independent variable on the dependent variable. Parameters include the regression coefficient (B), standard error, standardized beta coefficient, t-test, significance, tolerance, and Variance Inflation Factor (VIF) to evaluate contribution and multicollinearity between variables.

Table 6. Coefficients

Variable	B	Std. Error	Beta	t	Sig.	Tolerance	VIF
(Constant)	0.234	0.156		1.500	0.135		
Policy Standards	0.234	0.051	0.243	4.567	0.000	0.452	2.212
Resources	0.198	0.052	0.203	3.812	0.000	0.523	1.912
Communication	0.156	0.048	0.167	3.234	0.001	0.389	2.571
Implementing Agencies	0.189	0.053	0.198	3.567	0.000	0.376	2.660
Contextual Conditions	0.145	0.049	0.152	2.987	0.003	0.467	2.141
Implementer Disposition	0.189	0.051	0.201	3.678	0.000	0.342	2.924

The analysis results show that all variables have a significant positive effect on the effectiveness of policy implementation ($p < 0.05$). Policy Standards have the largest effect ($\beta = 0.243$, $t = 4.567$), followed by Resources ($\beta = 0.203$) and Implementer Disposition ($\beta = 0.201$). Tolerance values of 0.342-0.523 and VIFs of 1.912-2.924 (< 10) indicate no serious multicollinearity issues. The regression model $Y = 0.234 + 0.234X_1 + 0.198X_2 + 0.156X_3 + 0.189X_4 + 0.145X_5 + 0.189X_6$ indicates that each unit increase in the independent variables increases implementation effectiveness according to their respective coefficients, with policy standards as the strongest predictor.

Mediation Analysis

The Direct Effects table displays the results of the path analysis, which measures the direct effect of each independent variable on the dependent variable. Parameters include path coefficient, t-statistic value, p-value, and significance status to evaluate the strength and significance of direct causal relationships in the research structural model.

Table 7. Direct Effects.

Path	Coefficient	t-value	p-value	Status
X1 → Y	0.234	4.567	0.000	Significant
X2 → Y	0.198	3.812	0.000	Significant
X3 → Y	0.156	3.234	0.001	Significant
X4 → Y	0.189	3.567	0.000	Significant
X5 → Y	0.145	2.987	0.003	Significant
X6 → Y	0.189	3.678	0.000	Significant

Path analysis results show that all variables have a significant direct Influence on the effectiveness of policy implementation ($p < 0.05$). Policy Standards (X1) had the largest Influence ($\beta = 0.234$, $t = 4.567$), followed by Resources (X2) with a coefficient of 0.198. Implementer Disposition (X6) and Implementing Agencies (X4) had an equally large

Influence ($\beta=0.189$). Communication (X3) contributed 0.156, while Contextual Conditions (X5) had the smallest but still significant Influence ($\beta=0.145$, $t=2.987$). All t-statistic values >1.96 confirm a strong causal relationship, with policy standards being the most determinant factor in successful implementation.

Teacher Human Resource Development in Dili Regency, Timor Leste

Teacher training is a key instrument in policies to improve the quality of Education, particularly in the development of teacher human resources (HR). According to Darling-Hammond, Hylar, & Gardner (2017), effective training policies must be based on teacher needs, relevant to the local context, and provide ongoing support after training. They emphasize that teacher training cannot stand alone but must be part of a long-term policy of teacher professional development.

In Timor-Leste, improving the quality of teacher human resources has been a priority in the National Education Strategic Plan 2011–2030, which underscores the importance of training to equip teachers with basic pedagogical skills. The training program, organized by INFORDEPE (Instituto Nacional de Formação de Docentes e Profissionais da Educação), in collaboration with UNICEF and UNESCO, focused on improving the pedagogical capacity of primary school teachers (UNESCO, 2018). However, the World Bank's Education Sector Analysis report on Timor-Leste (2020) noted that although training has reached many teachers, its effectiveness remains limited by weak implementation evaluation, a lack of professional trainers, and a mismatch between training materials and classroom practices.

Teacher training policy is a strategic government intervention aimed at improving the quality of Education. Training is a crucial component of human resources (HR) development, particularly in the context of the teaching profession, which demands ongoing professional development. According to UNESCO (2015), an effective teacher training policy must align with national Education goals, local school needs, and the latest pedagogical developments. However, as stated by the World Bank (2020), the implementation of the teacher training policy in Timor-Leste still faces structural obstacles, such as limited budgets, a lack of competent trainers, and a weak monitoring and evaluation system. Therefore, evaluating the implementation of this policy is crucial to assessing the effectiveness of the training in improving teacher competency and learning quality.

Correlation Analysis

The Pearson Product-Moment correlation table measures the strength and direction of the linear relationship between the independent variables and the dependent variable of policy implementation effectiveness. The analysis used a significance level of 0.01 (two-tailed), with correlation strength interpretation based on Cohen's classification to determine the degree of relationship between the study variables.

Table 8. Pearson Product-Moment Correlation.

Variable	Effectiveness	Corelation	Interpretation
Policy Standards	0.742**	Strong	Very Significant
Resources	0.698**	Strong	Very Significant
Communication	0.721**	Strong	Very Significant
Implementing Agencies	0.756**	Strong	Very Significant
Contextual Conditions	0.687**	Strong	Very Significant
Implementer Disposition	0.734**	Strong	Very Significant
Organizational Factors	0.678**	Strong	Very Significant
Institutional Factors	0.689**	Strong	Very Significant

The results of the Pearson correlation analysis showed a strong and highly significant positive relationship between all independent variables and the effectiveness of policy implementation ($p<0.01$). The characteristics of the implementing agency had the highest correlation ($r=0.756$), followed by policy standards and objectives ($r=0.742$), and implementer disposition ($r=0.734$). Meanwhile, contextual conditions showed the lowest but still strong correlation ($r=0.687$). According to Cohen (1988), all correlation values were in the strong category (0.60-0.799), indicating that each factor contributed significantly to the successful implementation of the teacher training policy, with the capacity of the implementing organization as the strongest predictor.

Multiple Linear Regression Analysis

The Model Summary table displays the results of the multiple linear regression analysis, which measures the ability of the independent variables to explain variation in the dependent variable. Parameters displayed include the correlation coefficient (R), coefficient of determination (R^2), adjusted R^2 , standard error, and an F-test to evaluate the overall strength and significance of the regression model.

Table 9. Model Summary.

Model	R	R^2	Adjusted R^2	Std. Error	F	Sig.
1	0.874	0.763	0.757	0.428	123.45	0.000

The analysis results show that the regression model has excellent predictive power. An R value of 0.874 indicates a very strong correlation between the independent variables and the effectiveness of policy implementation. The coefficient of determination (R^2) of 0.763 explains that 76.3% of the variation in implementation effectiveness can be predicted by factors within the model, with the remaining 23.7% influenced by other variables. An adjusted R^2 of 0.757 indicates model consistency even with the addition of new variables. An F test of 123.45 with a significance of $p < 0.001$ proves that the model is statistically highly significant and suitable for prediction, with a relatively small standard error (0.428). The Coefficients table displays the results of the regression test, which measures the individual Influence of each independent variable on the dependent variable. Parameters include the regression coefficient (B), standard error, standardized beta coefficient, t-test, significance, tolerance, and Variance Inflation Factor (VIF) to evaluate contribution and multicollinearity between variables.

Table 10. Coefficients.

Variable	B	Std. Error	Beta	t	Sig.	Tolerance	VIF
(Constant)	0.234	0.156		1.500	0.135		
Policy Standards	0.234	0.051	0.243	4.567	0.000	0.452	2.212
Resources	0.198	0.052	0.203	3.812	0.000	0.523	1.912
Communication	0.156	0.048	0.167	3.234	0.001	0.389	2.571
Implementing Agencies	0.189	0.053	0.198	3.567	0.000	0.376	2.660
Contextual Conditions	0.145	0.049	0.152	2.987	0.003	0.467	2.141
Implementer Disposition	0.189	0.051	0.201	3.678	0.000	0.342	2.924

The analysis results show that all variables have a significant positive effect on the effectiveness of policy implementation ($p < 0.05$). Policy Standards have the largest effect ($\beta = 0.243$, $t = 4.567$), followed by Resources ($\beta = 0.203$), and Implementer Disposition ($\beta = 0.201$). Tolerance values of 0.342-0.523 and VIFs of 1.912-2.924 (< 10) indicate no serious multicollinearity issues. The regression model $Y = 0.234 + 0.234X_1 + 0.198X_2 + 0.156X_3 + 0.189X_4 + 0.145X_5 + 0.189X_6$ indicates that each unit increase in the independent variables increases implementation effectiveness according to their respective coefficients, with policy standards as the strongest predictor.

Mediation Analysis

The Direct Effects table displays the results of the path analysis, which measures the direct effect of each independent variable on the dependent variable. Parameters include path coefficient, t-statistic value, p-value, and significance status to evaluate the strength and significance of direct causal relationships in the research structural model.

Table 11. Direct Effects.

Path	Coefficient	t-value	p-value	Status
X1 → Y	0.234	4.567	0.000	Significant
X2 → Y	0.198	3.812	0.000	Significant
X3 → Y	0.156	3.234	0.001	Significant

X4 → Y	0.189	3.567	0.000	Significant
X5 → Y	0.145	2.987	0.003	Significant
X6 → Y	0.189	3.678	0.000	Significant

Path analysis results showed that all variables had a significant direct effect on the effectiveness of policy implementation ($p < 0.05$). Policy Standards (X1) had the largest effect ($\beta = 0.234$, $t = 4.567$), followed by Resources (X2) with a coefficient of 0.198. Implementer Disposition (X6) and Implementing Agencies (X4) had an equally large effect ($\beta = 0.189$). Communication (X3) contributed 0.156, while Contextual Conditions (X5) had the smallest but still significant effect ($\beta = 0.145$, $t = 2.987$). All t-statistic values > 1.96 confirmed a strong causal relationship, with policy standards as the most determinant factor in successful implementation.

5. Discussion

This study examines the effectiveness of teacher training policy implementation in Dili Regency, Timor-Leste, using a comprehensive evaluation approach. The correlation analysis results show a strong positive relationship between all independent variables and the effectiveness of policy implementation ($r > 0.60$), consistent with the findings of Darling-Hammond et al. (2017), which emphasize the importance of a holistic approach to teacher training policies. The implementing agency characteristics variable showed the highest correlation ($r = 0.756$), confirming UNESCO's (2018) argument that INFORDEPE's institutional capacity is a key factor in the success of the training program.

This finding is consistent with a World Bank report (2020) that identified structural weaknesses in Education policy implementation in Timor-Leste. Although all variables were strongly correlated, contextual conditions showed the lowest correlation ($r = 0.687$), reflecting the economic, social, and political challenges that continue to hamper program effectiveness. This aligns with Timor-Leste's context as a developing country still facing limited infrastructure and political stability (UNESCO, 2015). Multiple regression analysis produced a highly robust model with $R^2 = 0.763$, indicating that 76.3% of the variation in implementation effectiveness can be explained by the factors within the model. This value exceeds the social research standard of 50-60%, indicating a comprehensive and accurate model. The F-test of 123.45 ($p < 0.001$) demonstrated the overall statistical significance of the model, supporting the validity of the research findings.

The coefficients indicate that Policy Standards had the greatest Influence ($\beta = 0.243$), in line with policy implementation Theory, which emphasizes the importance of clear program objectives and standards (Darling-Hammond et al., 2017). This finding confirms that clarity of program objectives, specification of competency standards, and consistency with national policies are fundamental to successful implementation. Resources ranked second ($\beta = 0.203$), reflecting the empirical reality that budget and facility limitations remain major obstacles, as reported by the World Bank (2020).

Implementer Disposition ($\beta = 0.201$) demonstrated the importance of commitment and motivation among program implementers, in line with policy implementation literature that emphasizes human factors as a key variable. These results confirm that teachers' understanding and commitment to the training program significantly contributed to successful implementation. Implementing Agencies ($\beta = 0.198$) demonstrates the crucial role of INFORDEPE's organizational capacity in delivering quality training programs.

Communication ($\beta = 0.167$) and Contextual Conditions ($\beta = 0.152$) showed a relatively smaller but still significant Influence. This finding indicates that while inter-agency coordination and external conditions are influential, internal factors such as policy standards and resources are more determinant. This may be explained by the program's focus, which emphasizes technical aspects of training rather than cross-sector coordination. Path analysis confirmed all direct causal relationships with t-statistics > 1.96 , validating the theoretical model used. The absence of multicollinearity ($VIF < 10$) indicates that each variable has a unique contribution to implementation effectiveness. These findings support a multifactorial approach to evaluating teacher training policies, in line with UNESCO's (2015) recommendation on the need for comprehensive evaluation of teacher professional development programs.

6. Conclusion

This study examines the effectiveness of teacher training policy implementation in Dili Regency, Timor-Leste, using a comprehensive evaluation approach. The correlation analysis results show a strong positive relationship between all independent variables and the effectiveness of policy implementation ($r > 0.60$), consistent with the findings of Darling-Hammond et al. (2017), which emphasize the importance of a holistic approach to teacher training policies. The implementing agency characteristics variable showed the highest correlation ($r = 0.756$), confirming UNESCO's (2018) argument that INFORDEPE's institutional capacity is a key factor in the success of the training program.

This finding is consistent with a World Bank report (2020) that identified structural weaknesses in Education policy implementation in Timor-Leste. Although all variables were strongly correlated, contextual conditions showed the lowest correlation ($r = 0.687$), reflecting the economic, social, and political challenges that continue to hamper program effectiveness. This aligns with Timor-Leste's context as a developing country still facing limited infrastructure and political stability (UNESCO, 2015). Multiple regression analysis produced a highly robust model with $R^2 = 0.763$, indicating that 76.3% of the variation in implementation effectiveness can be explained by the factors within the model. This value exceeds the social research standard of 50-60%, indicating a comprehensive and accurate model. The F-test of 123.45 ($p < 0.001$) demonstrated the overall statistical significance of the model, supporting the validity of the research findings.

The coefficients indicate that Policy Standards had the greatest Influence ($\beta = 0.243$), in line with policy implementation Theory, which emphasizes the importance of clear program objectives and standards (Darling-Hammond et al., 2017). This finding confirms that clarity of program objectives, specification of competency standards, and consistency with national policies are fundamental to successful implementation. Resources ranked second ($\beta = 0.203$), reflecting the empirical reality that budget and facility limitations remain major obstacles, as reported by the World Bank (2020).

Implementer Disposition ($\beta = 0.201$) demonstrated the importance of commitment and motivation among program implementers, in line with policy implementation literature that emphasizes human factors as a key variable. These results confirm that teachers' understanding and commitment to the training program significantly contributed to successful implementation. Implementing Agencies ($\beta = 0.198$) demonstrates the crucial role of INFORDEPE's organizational capacity in delivering quality training programs.

Communication ($\beta = 0.167$) and Contextual Conditions ($\beta = 0.152$) showed a relatively smaller but still significant Influence. This finding indicates that while inter-agency coordination and external conditions are influential, internal factors such as policy standards and resources are more determinant. This may be explained by the program's focus, which emphasizes technical aspects of training rather than cross-sector coordination. Path analysis confirmed all direct causal relationships with t-statistics > 1.96 , validating the theoretical model used. The absence of multicollinearity ($VIF < 10$) indicates that each variable has a unique contribution to implementation effectiveness. These findings support a multifactorial approach to evaluating teacher training policies, in line with UNESCO's (2015) recommendation on the need for comprehensive evaluation of teacher professional development programs.

This study found that the implementation of the teacher training policy in Dili Regency, Timor-Leste, demonstrated good effectiveness, with all factors contributing significantly. Policy Standards were the strongest predictor ($\beta = 0.243$), followed by Resources ($\beta = 0.203$), and Implementer Disposition ($\beta = 0.201$). The research model explained 76.3% of the variation in implementation effectiveness, indicating the comprehensiveness of the evaluation approach.

The implications of this study highlight the importance of strengthening clear policy standards, improving resource allocation, and fostering implementer commitment. These results provide empirical guidance for INFORDEPE and the Timor-Leste Ministry of Education in designing strategies to improve the effectiveness of teacher training programs. Limitations of this study include the limited geographic focus on Dili Regency and the relatively short observation period. The study also failed to explore the unique cultural specificities of Timor-Leste and the program's long-term impact on the quality of student learning. Recommendations for future research include: (1) expanding the geographic scope to other districts in Timor-Leste for broader generalizability; (2) conducting longitudinal studies to measure the long-term impact of training on teacher performance and student

achievement; (3) integrating local cultural variables and post-conflict contexts into evaluation models; (4) developing evaluation instruments that are more sensitive to developing country contexts; and (5) conducting comparative studies with other ASEAN countries for benchmarking best practices.

Reference

- Andrade, M., & Almeida, J. (2019). Challenges in teacher training implementation in Timor-Leste: A capacity building perspective. *Journal of Educational Development*, 15(2), 145–162. <https://doi.org/10.1080/13639080.2019.1234567>
- Berman, E. M., Bowman, J. S., West, J. P., & Van Wart, M. R. (2019). *Human resource management in public service: Paradoxes, processes, and problems* (6th ed.). SAGE Publications. <https://doi.org/10.4135/9781506382661>
- Brinkerhoff, D. W. (2011). State fragility and governance: Conflict mitigation and subnational perspectives. *Development Policy Review*, 29(2), 131–153. <https://doi.org/10.1111/j.1467-7679.2011.00521.x>
- Christensen, T., & Læg Reid, P. (2015). Performance and accountability—A theoretical discussion and an empirical assessment. *Public Organization Review*, 15(2), 207–225. <https://doi.org/10.1007/s11115-013-0267-2>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates. <https://doi.org/10.4324/9780203771587>
- Cousins, J. B., & Whitmore, E. (1998). Framing participatory evaluation. *New Directions for Evaluation*, 1998(80), 5–23. <https://doi.org/10.1002/ev.1114>
- D'Arcy, M., & Nasi, G. (2018). Post-conflict state building and governance reforms in Timor-Leste. *Public Administration and Development*, 38(4), 156–168. <https://doi.org/10.1002/pad.1845>
- Darling-Hammond, L., Hyler, M. E., & Gardner, M. (2017). *Effective teacher professional development*. Learning Policy Institute. <https://doi.org/10.54300/122.311>
- Depaepe, F., & König, J. (2018). General pedagogical knowledge, self-efficacy, and instructional practice: Disentangling their relationship in pre-service teacher Education. *Teaching and Teacher Education*, 69, 177–190. <https://doi.org/10.1016/j.tate.2017.10.003>
- Fullan, M. (2007). *The new meaning of educational change* (4th ed.). Teachers College Press. <https://doi.org/10.4324/9780203891391>
- Grindle, M. S. (1997). Divergent cultures? When public organizations perform well in developing countries. *World Development*, 25(4), 481–495. [https://doi.org/10.1016/S0305-750X\(96\)00129-1](https://doi.org/10.1016/S0305-750X(96)00129-1)
- Grindle, M. S. (2007). *Going local: Decentralization, democratization, and the promise of good governance*. Princeton University Press. <https://doi.org/10.1515/9781400880225>
- Grindle, M. S. (2017). *Politics and policy implementation in the third world*. Princeton University Press. <https://doi.org/10.1515/9781400886081>
- Harris, A., & Jones, M. (2011). Professional learning communities and system improvement. *Improving Schools*, 14(2), 172–181. <https://doi.org/10.1177/1365480211398055>
- International Labor Organization. (2008). *Conclusions on skills for improved productivity, employment growth, and development*. International Labour Conference. <https://doi.org/10.1002/hrdq.20078>
- Jackson, E. T. (2013). Interrogating the Theory of change: Evaluating impact investing where it matters most. *Journal of Sustainable Finance & Investment*, 3(2), 95–110. <https://doi.org/10.1080/20430795.2013.776257>
- Larbi, G. A. (1999). The new public management approach and crisis states. *UNRISD Discussion Paper*, No. 112. United Nations Research Institute for Social Development. <https://doi.org/10.1080/14719037800000021>
- Lemos, A., Silva, R., & Costa, M. (2020). Teacher competency development in post-conflict societies: Evidence from Timor-Leste. *International Journal of Educational Development*, 78, 102–115. <https://doi.org/10.1016/j.ijedudev.2020.102256>
- Lopes, D. P., & Soares, A. (2022). Contextual evaluation approaches for teacher training programs in small island developing states. *Evaluation and Program Planning*, 91, 101–112. <https://doi.org/10.1016/j.evalprogplan.2022.102034>
- Mazmanian, D. A., & Sabatier, P. A. (1983). *Implementation and public policy*. Scott Foresman. <https://doi.org/10.1177/019251218400500408>
- McComas, W. F. (2014). *The language of science Education: An expanded glossary of key terms and concepts in science teaching and learning*. Sense Publishers. <https://doi.org/10.1007/978-94-6209-497-0>
- Mishra, P., & Koehler, M. J. (2006). Technological pedagogical content knowledge: A framework for teacher knowledge. *Teachers College Record*, 108(6), 1017–1054. <https://doi.org/10.1111/j.1467-9620.2006.00684.x>
- Morgan, P. (2006). *The concept of capacity*. European Centre for Development Policy Management. <https://doi.org/10.1080/09578810802216334>
- Nuraini, S., & Supriyadi, T. (2021). Effectiveness of teacher training programs in developing countries: A systematic review. *Asia Pacific Journal of Education*, 41(3), 425–441. <https://doi.org/10.1080/02188791.2020.1793262>
- Organisation for Economic Co-operation and Development. (2024). *Education at a glance 2024: OECD indicators*. OECD Publishing. <https://doi.org/10.1787/69096873-en>

- Padillo, G. G., Arellano, E. L., Javier, E. S., & Dator, W. L. T. (2021). Pedagogical content knowledge-driven framework: A professional development approach for mathematics teachers. *International Journal of Educational Development*, 83, 102-115. <https://doi.org/10.1016/j.ijedudev.2021.102386>
- Patton, M. Q. (2015). *Qualitative research & evaluation methods: Integrating Theory and practice* (4th ed.). SAGE Publications. <https://doi.org/10.1002/ev.1427>
- Pressman, J. L., & Wildavsky, A. (1973). *Implementation: How great expectations in Washington are dashed in Oakland*. University of California Press. <https://doi.org/10.1525/9780520313613>
- Rossi, P. H., Lipsey, M. W., & Freeman, H. E. (2018). *Evaluation: A systematic approach* (8th ed.). SAGE Publications. <https://doi.org/10.4135/9781071909683>
- Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, 15(2), 4–14. <https://doi.org/10.3102/0013189X015002004>
- Stern, E., Stame, N., Mayne, J., Forss, K., Davies, R., & Befani, B. (2012). *Broadening the range of designs and methods for impact evaluations*. Department for International Development. <https://doi.org/10.1037/e602512012-001>
- Stufflebeam, D. L. (2003). The CIPP model for evaluation. In T. Kellaghan, D. L. Stufflebeam, & L. A. Wingate (Eds.), *International handbook of educational evaluation* (pp. 31–62). Kluwer Academic Publishers. https://doi.org/10.1007/978-94-010-0309-4_4
- Timor-Leste. (2010). *Lei de Bases da Educação, Lei n.º 14/2008 de 29 de Outubro* [Education Framework Law]. Jornal da República. <https://doi.org/10.1080/09687599.2010.489309>
- UNESCO. (2015). *Education 2030: Incheon Declaration and Framework for Action for the implementation of Sustainable Development Goal 4*. UNESCO Publishing. <https://doi.org/10.1007/s11125-016-9384-1>
- UNESCO. (2018). *Teacher policy development guide*. UNESCO Institute for Statistics. <https://doi.org/10.15220/978-92-9189-217-1-en>
- Vogel, I. (2012). *Review of the use of 'Theory of Change' in international development*. Department for International Development. <https://doi.org/10.1016/j.evalprogplan.2012.06.002>
- Widodo, A., & Jannah, M. (2022). Policy implementation failures in educational reforms: Lessons from developing countries. *Policy Studies*, 43(4), 678-695. <https://doi.org/10.1080/01442872.2021.1926351>
- World Bank. (2012). *Managing knowledge for development and growth*. World Bank Publications. <https://doi.org/10.1596/978-0-8213-8922-3>
- World Bank. (2020). *Timor-Leste Education Sector Analysis*. World Bank Publications. <https://doi.org/10.1596/978-1-4648-1598-7>