Journal of Language Teaching and Learning Today
ISSN: 3083-256X
Volume 1, Issue 2, 2026
pp. 1-14

A Critical Review of the CIPP Model and Its Relevance to Language Education

Nguyen Doan Thao Chi
University of Foreign Language Studies, The University of Danang, Vietnam
Corresponding author email: ndtchi@ufl.udn.vn
https://doi.org/10.65956/ltlt.2026.136

Abstract

This critical review provides a synthesis of the Context, Input, Process, and Product (CIPP) evaluation model and its practical relevance to language education. While traditional evaluation frameworks, such as Tyler's objective-based model, often treat instructional delivery as an unexamined "black box" by focusing exclusively on summative outcomes, the CIPP model offers a proactive, systems-oriented approach designed for continuous programme improvement. Through a structured synthesis of foundational literature and recent empirical studies, this paper evaluates the methodological advantages of CIPP across diverse language education contexts, including reading interventions, tertiary speaking programmes, national curriculum reforms, and pre-service teacher training. The findings show that the CIPP framework can uncover instructional and psychological friction during active implementation, helping language programmes align with learners' actual needs. The review also highlights the model's implications for language materials evaluation by showing how its four dimensions closely parallel the standard phases of needs analysis, pre-use, while-use, and post-use assessment. The study concludes that, although full CIPP implementation requires substantial resources, a pragmatic and context-driven approach enables institutions to promote democratic, learner-centred, and sustainable language education.

Keywords CIPP model, language education, curriculum evaluation, educational evaluation, language materials evaluation

Article history Received: 06 Jul 2026 | Accepted: 04 Sep 2026 | Available: 06 Sep 2026

INTRODUCTION

Educational evaluation has a rich and complex history that predates the modern era by at least 150 years, evolving from the Age of Reform before 1900 through the Tylerian Age (1930-1945) and into the Age of Development (1958-1972), as outlined by Madaus and Stufflebeam (2000). During the Age of Development, evaluators began to recognise that purely objective-based evaluation models were insufficient for capturing the complexity of educational programmes. Instead of measuring only whether predetermined objectives had been achieved, evaluation needed to examine the wider context, available inputs, implementation process, and eventual outcomes of an educational programme.

In response to this crisis in educational evaluation, the CIPP Evaluation Model was conceptualised and developed by Stufflebeam and his colleagues in the late 1960s. Stufflebeam (2003) notes that the model was originally formulated to help improve and achieve accountability for U.S. public

school projects, especially those funded under federal education initiatives. The model was later refined as a decision-oriented framework that could support administrators, policymakers, teachers, and other stakeholders throughout an entire programme cycle.

The significance of the CIPP model lies in its proactive, systems-oriented philosophy. Stufflebeam (2003) emphasises that its fundamental principle is that "evaluation's most important purpose is not to prove, but to improve" (p. 31). Unlike older approaches that judged a programme only at the end, CIPP treats evaluation as a continuous process that begins before implementation and continues during and after the programme. This orientation makes the model relevant to language education, where curriculum design, materials selection, classroom implementation, and learner outcomes are closely connected.

1. What are the core theoretical foundations and dimensions of the CIPP model?
2. What are the methodological strengths and limitations of the CIPP model compared to traditional evaluation frameworks?
3. How can the full implementation of the CIPP model be feasibly applied and justified in resource-constrained language programmes?

LITERATURE REVIEW

Description of the CIPP Model

The CIPP model provides a systems-based framework that addresses four distinct but interrelated dimensions of programme evaluation, namely Context, Input, Process, and Product. House (2003) highlights that CIPP is one of the oldest and most thoroughly tested approaches, originally formulated to help educational decision-makers improve programmes rather than merely judge them. In this model, Context evaluation identifies needs and problems, Input evaluation examines strategies and resources, Process evaluation monitors implementation, and Product evaluation assesses outcomes.

Examining these aspects in greater detail, context evaluation aims to define the institutional or service context, identify the target population, and diagnose specific needs, problems, assets, and opportunities (Stufflebeam, 2000). This diagnostic phase provides the rationale for setting priorities and clarifying goals. Input evaluation then assesses alternative strategies, action plans, staffing, materials, and budgets to determine whether a programme is feasible and appropriate. Process evaluation monitors implementation and detects problems while the programme is still operating. Product evaluation examines intended and unintended outcomes, short-term and long-term effects, and the overall value of the programme.

Beyond its structural components, the theoretical foundations of the CIPP model are rooted in an objectivist orientation. This perspective asserts that moral good and educational value are objective, so evaluations must be grounded in explicit ethical principles and remain free from human bias as far as possible (Scriven, 1991; Stufflebeam, 2014). This objectivist stance also explains why CIPP is closely associated with accountability, democratic participation, and service to stakeholders. Rather than serving only administrators, the model requires evaluators to consider the needs and rights of all affected groups.

Theoretical Foundations and Dimensions

A review of the literature indicates that the CIPP model's core theoretical strength lies in its systems-oriented philosophy. Rather than functioning merely as a post-hoc judgement tool, CIPP approaches programme evaluation as an ongoing, systemic mechanism for continuous improvement. Recent systematic accounts describe the model as useful because it connects needs assessment, planning, implementation monitoring, and outcome evaluation within a single framework (Popham, 1995; Stufflebeam, 2014).

Methodological Strengths and Limitations Compared to Traditional Frameworks

The synthesised literature suggests that the CIPP model offers notable methodological strengths over traditional evaluation frameworks, particularly Tyler's objective-based model. Tyler's classical framework operates on the premise that curriculum evaluation is synonymous with measuring the extent to which stated objectives have been achieved. Although this approach is clear and practical, it treats actual instructional delivery as an unexamined "black box". Tyler's model focuses mainly on summative outputs, leaving daily classroom challenges, such as student anxiety, resource constraints, and socio-economic disparities, insufficiently examined (Madaus & Stufflebeam, 2000; Vo, 2018).

In contrast, the CIPP model addresses this limitation through its detailed Process and Context evaluations. Recent empirical applications demonstrate CIPP's methodological flexibility, as it successfully integrates mixed-methods designs such as concurrent triangulation to merge quantitative performance data with qualitative stakeholder insights (Abdulkareem & Hameed, 2024; Aybek & Oğuz, 2025; Solis, 2026). This flexibility allows the model to explain why programmes succeed or fail, not simply whether predetermined outcomes have been met.

Review of Previous Studies

A critical synthesis of the literature reveals broad agreement on the CIPP evaluation model's theoretical reliability and practical adaptability across diverse disciplines, although ongoing discussions highlight the need for continuous methodological refinement. Foundational scholars such as House (2003), Owston (2008), and Zhang et al. (2011) emphasise CIPP's value as a decision-oriented model, while recent language education studies show how it can be adapted to specific instructional contexts.

While the basic structure of the CIPP model is well established, researchers today often focus on how it continues to evolve and integrate with other methods, particularly through concurrent triangulation and mixed-methods designs (Abdulkareem & Hameed, 2024; Aybek & Oğuz, 2025; Solis, 2026). Accordingly, the model is no longer treated only as a broad evaluation framework but also as a practical tool for diagnosing specific instructional problems in complex language education settings.

This pragmatic integration allowed the model to be effectively implemented in complex public sectors. For instance, Madaus and Stufflebeam (2000) point out how scholars such as Webster (1975, as cited in Madaus & Stufflebeam, 2000) applied the framework to manage the diverse evaluation

needs of entire school systems. This example shows how CIPP can operate across multiple levels, from individual classrooms to large-scale educational programmes.

METHODOLOGY

To provide a transparent and analytical overview of the CIPP model and its relevance to language education, this critical review adopted a systematic literature search and synthesis strategy. The methodology was designed to capture recent empirical applications in language teaching and curriculum evaluation, while also drawing on foundational texts that explain the model's historical development and theoretical assumptions.

Literature Search Strategy

To build a theoretical and empirical foundation for this review, a systematic literature search was conducted across major academic databases, including Google Scholar, ERIC, and Scopus. The search used combinations of the following keywords and search strings, namely ("CIPP model" OR "Context, Input, Process, and Product") AND ("language education" OR "language teaching" OR "language curriculum" OR "materials evaluation"). Additional searches were conducted using the names of foundational theorists and recent empirical authors.

Inclusion and Exclusion Criteria

To maintain the quality and relevance of the reviewed literature, strict inclusion and exclusion criteria were established. The literature was divided into two tiers, with foundational theoretical papers having no publication date restrictions to capture the historical evolution of evaluation models, and recent empirical studies limited to 2023-2026 to capture current applications in language education. Studies were included if they directly applied the CIPP model or compared it with traditional models in educational contexts. Studies outside education were excluded unless they offered transferable insights for programme evaluation.

Table 1. *Inclusion and Exclusion Criteria for Literature Selection*

Criteria category	Inclusion criteria	Exclusion criteria
Publication Type	Peer-reviewed journal articles, seminal academic books, and published doctoral dissertations.	Non-peer-reviewed articles, opinion pieces, and blogs.
Timeframe	Theoretical texts: No date limit Empirical studies: 2023-2026	Empirical studies published before 2023 were excluded from the recent findings section.
Topic and Scope	Studies directly applying the CIPP model or comparing it with traditional models (e.g., Tyler's model) in language education.	Studies applying CIPP to non-educational sectors (e.g., healthcare, corporate management) unless used for general theory.

Language	Publications written in English.	Publications written in languages other than English.

Data Synthesis and Analysis

Following the screening process, the selected literature was subjected to a detailed thematic synthesis. This synthesis involved three steps. First, initial open coding was conducted to extract recurring concepts related to curriculum evaluation. Second, these initial codes were grouped into broader analytical categories, including theoretical foundations, methodological strengths, practical limitations, empirical applications, and relevance to language education. Third, the findings were synthesised in relation to the three research questions.

FINDINGS

To address the review questions formulated in the introduction, this section synthesises findings from foundational literature and recent open-access empirical studies conducted between 2023 and 2026. Moving beyond summary, this section critically analyses the CIPP model's specific theoretical foundations, methodological strengths, limitations, and applications in language education.

Recent Empirical Applications in Language Education

Recent empirical studies provide evidence for the CIPP model's practical utility across diverse language education domains. The following analysis of four recent case studies illustrates how the four dimensions of CIPP are used to diagnose, monitor, and improve language programmes. The detailed synthesis is presented in Table 2.

Table 2. *Synthesis of Recent Empirical Applications of the CIPP Model in Language Education*

Authors (Year)	Educational Context and Focus	Methodology and Participants	Core CIPP Findings	Critical Appraisal
Solis (2026)	Philippines Junior High School Reading Intervention (Project LIGHT)	Mixed methods 60 struggling readers and 4 English teachers	**Context:** Severe reading distress; >70% below independent level **Input:** Quality reading materials and trained teachers utilized **Process:** Effective scaffolding, but consistency dropped due to scheduling conflicts **Product:** Substantial reading gains, though younger cohorts progressed slower	Adequate purposive sampling. Utilized validated instruments (Phil-IRI) ensuring high reliability. CIPP genuinely structured the evaluation from its inception to diagnose needs and monitor interventions, rather than being applied as a post-hoc label.
Aybek & Oğuz (2025)	Turkey National Primary Turkish	Mixed methods	**Context:** Culturally aligned and flexible, but lacked inclusivity	Strong quantitative validity due to a large, representative national sample (n=112). CIPP served effectively as an

	Language Curriculum	112 first-grade primary school teachers	**Input:** Good visual materials, but severe infrastructure disparities **Process:** Constructivist activities hindered by overcrowded classes and time constraints **Product:** Improved literacy, but failed to address individual learning differences.	overarching macro-level framework, though the findings heavily relied on self-reported teacher perceptions.
Abdulkareem & Hameed (2024)	Iraq University English Speaking Programme	Mixed methods 140 undergraduate students and 10 lecturers	**Context:** High speaking anxiety and cultural friction identified **Input:** CLT materials claimed, but actual resources were outdated **Process:** Teachers reverted to rigid, teacher-controlled methods **Product:** Low student satisfaction; lacked real-world speaking readiness	Rigid concurrent triangulation design. High instrument validity achieved by cross-checking quantitative student surveys with qualitative field observations. CIPP was successfully used to expose discrepancies between claimed inputs and actual instructional processes.
Nurjanah et al. (2023)	Indonesia Pre-Service English Teacher Development Programme.	Qualitative field study Pre-service English teachers and mentors.	**Context:** Aimed to produce teachers with comprehensive competencies. **Input:** Thorough combination of theoretical and practical courses **Process:** Successful application of theory during school practicum **Product:** Strong pedagogical growth, but lacked spontaneous spoken English fluency.	Appropriate qualitative sampling for an in-depth field study. High ecological validity derived from direct, labor-intensive mentor observations. CIPP functioned well as an evaluative lens to assess practical teaching mastery beyond theoretical knowledge.

Evaluating reading interventions and scaffolding strategies (Philippines)

Solis (2026) applied the CIPP model to evaluate Project LIGHT, a targeted reading intervention for junior high school students at Eastern Bacoor National High School. The Context evaluation used the Philippine Informal Reading Inventory (Phil-IRI) to diagnose that more than 70% of the students were operating below independent reading levels and struggling with severe reading distress and emotional barriers. During the Process evaluation, the study monitored the implementation of scaffolding strategies, revealing that student-centred instruction was sound but that instructional quality dropped significantly when session regularity declined because of scheduling conflicts. Finally, the Product evaluation demonstrated substantial post-test gains in reading comprehension, vocabulary, and self-confidence, although it also highlighted a developmental limitation because younger Grade 7 cohorts showed slower progress than older students. This study illustrates how CIPP connects contextual needs with process friction and final reading outcomes.

Evaluating tertiary English speaking programmes (Iraq)

At the university level, Abdulkareem and Hameed (2024) used a concurrent triangulation mixed-methods CIPP design to evaluate the English speaking programme at Al-Maaqal University. The Input evaluation showed that lecturers claimed to use Communicative Language Teaching (CLT) materials consistently. However, the Process evaluation uncovered a severe psychological disconnect, as student satisfaction was low because the speaking activities remained highly rigid and teacher-controlled. Contextual data revealed that students suffered from debilitating speaking anxiety and cultural friction, preventing them from engaging in spontaneous peer communication unless directly prompted. This empirical finding shows a major advantage of the CIPP model. It showed that adopting a communicative syllabus on paper (Input) does not guarantee communicative instruction in practice (Process) if affective barriers are ignored.

Evaluating pre-service English teacher development (Indonesia)

Nurjanah et al. (2023) used the CIPP model to assess the English Language Pre-service Teacher Professional Development Programme (PPG Prajabatan) at Universitas Muhammadiyah Malang. The Input and Process evaluations confirmed that the programme's combination of theoretical coursework and school-based practicum (PPL) provided a strong foundation. The Product evaluation documented a significant enhancement in the participants' professional competencies, particularly in pedagogical content knowledge, ELT material creation, and school administrative management. Nevertheless, the evaluation identified an important product deficiency. Mentor teachers noted that pre-service teachers still experienced confusion regarding certain practical English language theories and lacked fluent, spontaneous spoken English during active teaching. This insight allowed the institution to recognise the need for more intensive practical language training within the curriculum.

Evaluating national primary language curriculum reforms (Turkey)

At a macro level, Aybek and Oğuz (2025) used the CIPP framework to evaluate the 1st Grade Turkish Language Curriculum newly restructured under the Century of Turkey Education Model. The Context evaluation revealed positive teacher perceptions, indicating that the curriculum's flexible structure successfully aligned with cultural values and students' developmental levels. However, Input and Process evaluations exposed severe structural inequalities. Teachers reported that, despite the curriculum's constructivist design, implementation was heavily hindered by a lack of technological infrastructure, disparities in physical resources across schools, and time management constraints in overcrowded classrooms. This large-scale study illustrates CIPP's important role in bridging the gap between ambitious national policy designs and the realistic operational capacities of local schools.

In summary, these recent empirical applications demonstrate that the CIPP model is a useful tool in language education. By systematically evaluating Context, Input, Process, and Product, the framework allows educators and policymakers to identify not only whether a language programme succeeds or fails, but also where and why instructional breakdowns occur.

DISCUSSION

Overcoming the Limitations of Traditional Evaluation Models

A critical review of the CIPP model, supported by recent empirical findings, suggests several methodological strengths over traditional frameworks such as Tyler's objective model. Traditional models focus heavily on summative outcomes and neglect the actual delivery of educational services. For instance, while Ergin and Kambur (2025) successfully used Tyler's model to prove statistical gains in students' speaking, listening, and writing skills, their objective-based approach could not explain the daily instructional friction or psychological barriers present in the classroom. The CIPP framework effectively addresses this limitation. As demonstrated by Abdulkareem and Hameed (2024), using the Input and Process dimensions allows evaluators to uncover severe instructional disconnects caused by affective barriers such as speaking anxiety. Similarly, the macro-evaluation by Aybek and Oğuz (2025) illustrated that, while a curriculum might be conceptually sound and culturally aligned, Process evaluations are important for exposing structural inequalities such as overcrowded classrooms and lack of technological resources that hinder policy implementation. Furthermore, when contrasted with Scriven's goal-free evaluation or Stake's responsive model (1975), CIPP offers a more organised and objective set of steps for official accountability. However, the empirical studies also highlight methodological challenges. Implementing the full CIPP framework requires significant time and resources, and detailed monitoring can overwhelm under-resourced schools if not adapted pragmatically (Aybek & Oğuz, 2025; Dang et al., 2026).

Evaluating Higher Education and Pre-Service Teacher Programmes

The empirical evidence supports the practical application of the CIPP model in higher education and teacher training programmes. CIPP is particularly effective at uncovering discrepancies between intended curriculum design and actual student outcomes. For instance, the evaluation of the pre-service English teacher programme by Nurjanah et al. (2023) demonstrated that, while the Input and Process phases, such as theoretical coursework and school-based practicums, were structurally effective, the Product evaluation identified an important competency gap. Specifically, student teachers lacked spontaneous spoken English fluency in real classroom settings. This empirical finding highlights CIPP's diagnostic power to ensure that teacher education goes beyond theoretical knowledge to verify practical pedagogical mastery.

Furthermore, the evaluation of the tertiary speaking programme by Abdulkareem and Hameed (2024) illustrates how the CIPP dimensions work together to diagnose instructional breakdowns. Although lecturers claimed to use Communicative Language Teaching materials during the Input phase, the Process evaluation revealed that high speaking anxiety and cultural friction forced teachers to return to rigid, teacher-controlled methods. Consequently, the Product evaluation reflected high student dissatisfaction. When discussed together, these empirical studies demonstrate that the CIPP framework provides educational administrators with precise and actionable data. It shows that adopting a modern communicative syllabus on paper does not guarantee instructional success unless both learners' psychological barriers and educators' practical fluency are continuously monitored and addressed.

Feasibility and Pragmatic Application in Resource-Constrained Contexts

While the empirical studies highlight the diagnostic power of the CIPP model, they also expose an important challenge, as full CIPP implementation is highly resource-intensive. As Aybek and Oğuz

(2025) and Nurjanah et al. (2023) demonstrated, detailed process and product evaluations demand significant time, labour, and logistical stability, which are often lacking in under-resourced language programmes.

How, then, can CIPP be justified and feasibly applied in such constrained settings? The synthesised literature and empirical evidence suggest a pragmatic, context-driven approach. First, full implementation is justified primarily in high-stakes, macro-level contexts such as the national curriculum reform evaluated by Aybek and Oğuz (2025), where the cost of programme failure far outweighs the cost of the evaluation itself. Second, for ordinary, under-resourced language programmes, feasibility is achieved by avoiding rigid, full-scale formal assessments in favour of integration. As Stufflebeam (2003) advises, institutions should use existing data such as student reading inventories, attendance records, and teacher logs rather than inventing costly new data collection tools. Furthermore, the CIPP framework can be feasibly scaled down into a curriculum-embedded evaluation or conducted as a self-evaluation by individual teachers. Rather than exhausting resources on all four dimensions simultaneously, resource-constrained programmes can pragmatically focus on specific areas of friction, such as using Process evaluation to uncover immediate psychological barriers such as speaking anxiety (Abdulkareem & Hameed, 2024), thereby making the model a sustainable tool for daily instructional improvement.

Limitations and Future Directions

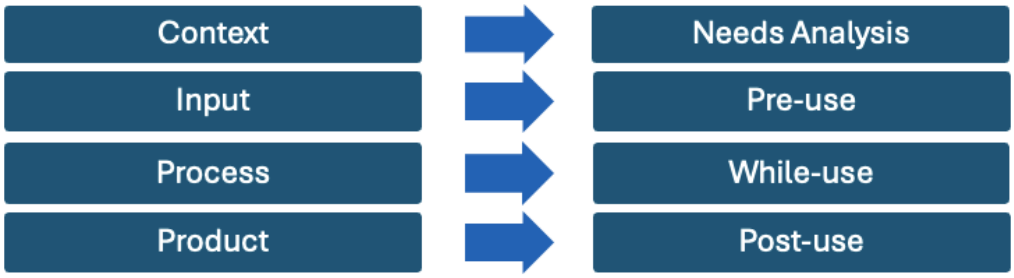
This critical review acknowledges several methodological limitations. First, the systematic search returned a relatively small corpus of only four recent empirical studies that directly apply the CIPP model to language education. Second, the exclusion of inaccessible paywalled papers and non-English publications introduced an access constraint, potentially biasing the reviewed sample towards open-access and English-only venues. Consequently, the findings may not fully capture the global or historical application of CIPP in language programmes.

To address these gaps, future research should expand the corpus by securing access to subscription-based databases and including multilingual studies. Moreover, future empirical studies should rigorously document the cost-benefit ratio and practical constraints of full CIPP implementation to determine its true feasibility and cost-effectiveness in ordinary, under-resourced language programmes.

IMPLICATIONS FOR LANGUAGE EDUCATION

Language education is uniquely complex because it involves not only the acquisition of linguistic knowledge but also the development of communicative competence and intercultural awareness. Applying the CIPP evaluation framework to language curriculum design offers a systematic and responsive way to manage this complexity. Conceptually, the model's dimensions closely correspond to standard phases of language materials evaluation (e.g., Tomlinson, 2013), providing educators with a structured pathway from initial planning to final assessment. As visually summarised in Figure 1, this one-to-one mapping provides a practical blueprint. It bridges the gap between macro-level educational evaluation and micro-level classroom practices, allowing language educators to integrate the CIPP framework into their familiar instructional cycles effectively.

Figure 1. *Mapping the CIPP Model to the Phases of Language Materials Evaluation*



Context Evaluation as Needs Analysis

In the initial phase, context evaluation closely aligns with what language teachers traditionally call needs analysis. Rather than relying on generic language goals, curriculum designers must use context evaluation to diagnose learners' specific target situations, current proficiency levels, and the broader sociolinguistic demands of the community (Brown, 2016; Dang & Le, 2022). The recent empirical case of Project LIGHT (Solis, 2026) illustrates this broader implication for language curriculum design. By using context evaluation to diagnose specific "reading distress" and emotional barriers before implementing an intervention, the study demonstrates that teachers must develop local criteria tailored to their students' psychological and linguistic realities rather than relying solely on universal commercial checklists.

Input Evaluation as Pre-Use Evaluation

In the next step, input evaluation functions as pre-use evaluation. Language curriculum design frequently requires teachers to choose among competing methodologies, such as Task-Based Language Teaching (TBLT) or Content and Language Integrated Learning (CLIL). However, language learning materials are often evaluated based on personal impressions rather than clear criteria. The CIPP model's Input phase addresses this problem by helping schools compare different materials carefully to determine whether they fit their budgets and teachers' actual pedagogical skills (Tomlinson, 2013, 2016). This ensures that the chosen language syllabus is theoretically grounded and practically feasible before institutions invest in trendy but potentially ineffective resources.

Process Evaluation as While-Use Evaluation

Once the language programme is operational, Process evaluation serves as an ongoing monitor of instructional delivery. In communicative language classrooms, this naturally functions as the whilst-use evaluation phase (Bachman & Damböck, 2018). Continuous classroom-based assessment allows teachers to monitor learner engagement, clarity of instructions, and the immediate effectiveness of communicative activities in real time. The empirical evidence from Abdulkareem and Hameed (2024) illustrates the urgency of this phase, as it is during Process evaluation that teachers can uncover hidden affective barriers, such as speaking anxiety, that hinder communicative activities. Furthermore, as Solis (2026) demonstrated, Process evaluation is essential for tracking instructional consistency and adjusting scaffolding strategies immediately when disruptions occur.

Product Evaluation as Post-Use Evaluation

Finally, Product evaluation assesses the ultimate communicative proficiency of language learners. Tomlinson (2013) asserts that post-use evaluation is arguably the most valuable yet least administered stage in language education. Going beyond traditional, discrete-point grammar and vocabulary tests, CIPP's Product phase uses authentic performance assessments to determine whether students can use the language effectively in real-life contexts. By identifying specific developmental limitations, such as the finding that younger cohorts progressed more slowly than older ones (Solis, 2026), the Product phase provides teachers with evidence-based justifications for whether specific coursebooks, methodologies, or language modules should be retained, adapted, or completely replaced in subsequent academic years (Nation & Macalister, 2020).

CONCLUSION

In conclusion, the CIPP model stands as an effective, comprehensive, and practical framework for evaluating educational programmes. Rooted in the fundamental belief that evaluation's most important purpose is "not to prove, but to improve", it reconciles ongoing formative development with final summative judgement. By systematically examining Context, Input, Process, and Product, the framework equips educators with a clear guide to diagnose needs, select appropriate strategies, monitor active implementation, and measure outcomes. Furthermore, its strong ethical foundation in fair and democratic values ensures that the evaluation process remains objective and includes the diverse voices of all stakeholders. For language education, the CIPP framework is relevant because its four dimensions closely parallel the standard phases of language materials evaluation, namely needs analysis, pre-use, while-use, and post-use assessment, offering a systematic way to manage the complexity of communicative language teaching.

To enhance the utility of the CIPP model in real-world educational settings, schools and teachers should adopt a pragmatic approach. First, recognising that implementing the full framework requires significant time and financial resources, institutions can exercise the model's inherent flexibility by selectively prioritising the evaluation phases that best fit their current budgets and immediate needs. Second, language educators must actively use the Context and Input stages to develop specific, localised criteria. This ensures that language courses and instructional materials are not merely generic but are authentically tailored to match their students' cultural backgrounds and real-life communication needs. Third, institutions should place greater emphasis on the often-neglected Product, or post-use, stage to rigorously verify not only long-term language acquisition but also the durable effects of materials on learner motivation over time. Finally, to ensure that the evaluation process itself remains fair, accurate, and helpful, all CIPP applications should be subject to continuous metaevaluation so that accountability continues to serve educational improvement.

ACKNOWLEDGEMENTS

The author would like to express her sincere gratitude to the University of Foreign Language Studies, The University of Danang, for its administrative and academic support throughout this study. The author also extends her thanks to the anonymous reviewers for their constructive feedback, which significantly improved the quality of this manuscript.

DECLARATION

Ethics Approval and Informed Consent

Ethics approval and informed consent were not applicable because this article is a review study and involved no human participants or identifiable personal data.

AI Tool Use

During the preparation of this manuscript, the author(s) used Gemini Flash 3.5 to check grammar and improve language clarity and expression. The tool was used in the Introduction and Literature Review sections. The author(s) reviewed, edited, and verified all AI-assisted output and take full responsibility for the accuracy, originality, and integrity of the manuscript.

Funding

This research received no external funding.

Conflicts of Interest

The author declares no conflicts of interest.

Data Availability

The data supporting the findings of this study are available from the corresponding author upon reasonable request.

Author Contributions

The sole author was responsible for the conceptualization, research design, analysis, interpretation, and preparation of the manuscript and accepts full responsibility for the work.

REFERENCES

Abdulkareem, M. N., & Hameed, J. Q. (2024). A mixed method study adopted as an evaluative means by using CIPP model to evaluate the teaching process of speaking skill at Almaaqal University. *Journal of Sustainable Studies*, *6*(6), 2409-2430.

Aybek, B., & Oğuz, O. (2025). Sustainable and inclusive education reform in Türkiye: A CIPP evaluation of the primary Turkish language curriculum. *Sustainability*, *17*(19), 1-31. https://doi.org/10.3390/su17198659

Bachman, L. F., & Damböck, A. (2018). *Language assessment for classroom teachers*. Oxford University Press.

Brown, J. D. (2016). *Introducing needs analysis and English for specific purposes*. Routledge.

Dang, T. T., Nguyen, T. V. T., & Huynh, L. T. (2026). Contestation of ideologies in English language education in Vietnam: Internationalisation vs. decolonisation. In Chew, S. Y., Shangeetha, R. K., & Pillai, S. (Eds), *Language use in Southeast Asia: Present realities and future trends*. Routledge. https://doi.org/10.4324/9781003656623-14

Dang, T., & Le, Q. (2022). Learning place control: Vietnamese EFL students' appreciation and trust. *International Journal of Language Education and Applied Linguistics, 11*(2), 56-63. https://doi.org/10.15282/ijleal.v11i2.6597

Ergin, A., & Kambur, S. (2025). Evaluation of English language teaching via objective-based curriculum evaluation model. *International Journal of Research in Education and Science*, *11*(4), 810-830. https://doi.org/10.46328/ijres.1323

House, E. R. (2003). Introduction. In T. Kellaghan & D. L. Stufflebeam (Eds.), *International handbook of educational evaluation* (pp. 1-6). Kluwer Academic Publishers. https://doi.org/10.1007/978-94-010-0309-4_1

Madaus, G. F., & Stufflebeam, D. L. (2000). Program evaluation: A historical overview. In D. L. Stufflebeam, G. F. Madaus, & T. Kellaghan (Eds.), *Evaluation models: Viewpoints on educational and human services evaluation* (2nd ed.; pp. 3-18). Kluwer Academic Publishers. https://doi.org/10.1007/0-306-47559-6_1

Nation, I. S. P., & Macalister, J. (2020). *Language curriculum design* (2nd ed.). Routledge. https://doi.org/10.4324/9780429203763

Nurjanah, L., Wicaksono, B. H., Andini, T. M., & Effendi, M. I. (2023). Evaluation of teacher professional development program: A review of pre-service English teacher professionalism. *Journal of English Language Teaching and Learning (JETLE)*, *5*(1), 20-27. https://doi.org/10.18860/jetle.v5i2.23820

Owston, R. (2008). Models and methods for evaluation. In J. M. Spector, M. D. Merrill, J. van Merriënboer, & M. P. Driscoll (Eds.), *Handbook of research on educational communications and technology* (3rd ed.; pp. 605-617). Routledge.

Popham, W. J. (1995). *Educational evaluation* (3rd ed.). Allyn and Bacon.

Scriven, M. (1991). *Evaluation thesaurus* (4th ed.). Sage.

Solis, M. (2026). Deployment of Context-Input-Process-Product (CIPP) on evaluating Project LIGHT: An embedded exploratory-explanatory analysis on the development of reading comprehension skills among junior high school learners. *International Journal of Education, Research, and Innovation Perspectives*, *2*(2), 1020-1035. https://doi.org/10.5281/zenodo.18765370

Stake, R. E. (1975). *Evaluating the arts in education: A responsive approach*. Merrill.

Stufflebeam, D. L. (2000). The CIPP model for evaluation. In D. L. Stufflebeam, G. F. Madaus, & T. Kellaghan (Eds.), *Evaluation models: Viewpoints on educational and human services evaluation* (2nd ed.; pp. 279-317). Kluwer Academic Publishers. https://doi.org/10.1007/0-306-47559-6_16

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

Stufflebeam, D. L. (2014). Daniel Stufflebeam's CIPP model for evaluation: An improvement- and accountability-oriented approach. In D. L. Stufflebeam & C. L. S. Coryn (Eds.), *Evaluation theory, models, and applications* (2nd ed.; pp. 310-339). Jossey-Bass.

Tomlinson, B. (2013). Materials evaluation. In B. Tomlinson (Ed.), *Developing materials for language teaching* (2nd ed.; pp. 21-48). Bloomsbury Academic. https://doi.org/10.5040/9781474211826.ch-001

Tomlinson, B. (Ed.). (2016). *SLA research and materials development for language learning*. Routledge. https://doi.org/10.4324/9781315749082

Vo, T. K. A. (2018). Evaluation models in educational program: Strengths and weaknesses. *VNU Journal of Foreign Studies*, *34*(2), 140-150. https://doi.org/10.25073/2525-2445/vnufs.4252

Zhang, G., Zeller, N., Griffith, R., Metcalf, D., Williams, J., Shea, C., & Misulis, K. (2011). Using the Context, Input, Process, and Product Evaluation Model (CIPP) as a comprehensive framework to guide the planning, implementation, and assessment of service-learning programs. *Journal of Higher Education Outreach and Engagement*, *15*(4), 57-84.