Volume 7 Number 2 (2025)
May-August 2025
Page: 103-122

E-ISSN: 2656-4491
P-ISSN: 2656-4548
DOI: 10.37680/scaffolding.v7i2.7390

SCAFFOLDING
Jurnal Pendidikan Islam dan Multikulturalisme

3THE APPLICATION OF THE CIPP EVALUATION MODEL IN EDUCATIONAL PROGRAMS IN INDONESIA

M. Zainuddin Alanshori[1], Ahmad Maulidi[2], Ahmad Muhlisin[3], Kusaeri[4], Suparto[5]
[1]Universitas Islam Lamongan; Indonesia
[2]Universitas Al-Amien Sumenep; Indonesia
[345]Universitas Islam Negeri Sunan Ampel Surabaya; Indonesia
Correspondence Email; zen.120888@unisla.ac.id

Submitted: 12/01/2025 Revised: 11/06/2025 Accepted: 29/06/2025 Published: 07/07/2025

Abstract

This study aims to systematically review and analyze the trends in the use of the CIPP (Context, Input, Process, Product) evaluation model across various educational programs, as well as its strengths, weaknesses, and recommendations for implementation. The research employed a systematic literature review method, consisting of four stages: identification, selection, eligibility assessment, and inclusion. Article searches were conducted through Google Scholar, ScienceDirect, and SpringerLink using the keyword "CIPP model of education evaluation," resulting in 833 articles. Most of these articles were published in international journals and conference proceedings from various countries. After applying the inclusion criteria, 26 articles were selected that specifically discussed the application of the CIPP model in education. The findings indicate that the CIPP model is widely used as a comprehensive and flexible evaluation framework, especially in Asia, between 2019 and 2025. It is effective in assessing program context, resource input, implementation processes, and outcomes. However, challenges include the complexity of mixed-method data analysis and the need for skilled evaluators. This review recommends modularizing evaluation instruments, providing standardized training for evaluators, integrating digital technology, adapting local indicators, and developing flexible policy support. These strategies aim to enhance the relevance, efficiency, and impact of the CIPP model in education evaluation practices.

Keywords Evaluation; Education; CIPP Model.

INTRODUCTION

Evaluation is an integral component of the education system, ensuring that both teaching processes and learning outcomes align with predetermined goals effectively and efficiently (Thontawi et al., 2022). However, in practice, educational evaluation in many Indonesian institutions tends to be partial and predominantly outcome-oriented. That is, evaluations often focus solely on final learning achievements, such as test scores or graduation rates, without considering other crucial elements, including program context, resource availability (input), and implementation processes (Tamsiyati et al., 2025; Wiyono et al., 2025).

This general issue limits the comprehensive understanding of the overall effectiveness of educational programs (Aziz et al., 2019; Ratnaya et al., 2022). Evaluation practices that neglect contextual and process dimensions are unlikely to yield meaningful and relevant feedback for sustainable improvements. This condition is exacerbated by the lack of evaluative resources, limited evaluator competence, and the dominance of assessment methods that emphasize memorization and theoretical academic performance (Ayyusufi et al., 2022; Bakar et al., 2023).

More specific problems also arise in the implementation of national educational programs such as the Merdeka Curriculum, inclusive education, and literacy initiatives. Although these policies aim to improve instructional quality and ensure equitable access to education, their evaluation processes still face numerous technical and methodological challenges. For instance, in evaluating the Merdeka Curriculum in elementary schools, student learning outcomes have shown promising trends; however, supporting indicators such as teacher pedagogical capacity and institutional technical support remain suboptimal. Similarly, in the context of inclusive education, many schools still lack specialized educators and adequate facilities to support the success of such programs (Hadi et al., 2019; Idrus et al., 2025).

In this context, the CIPP evaluation model (Context, Input, Process, Product) becomes highly relevant, as it provides a holistic framework that extends beyond mere outcomes to encompass background conditions, available resources, and implementation mechanisms. Unfortunately, there has been a scarcity of systematic studies reviewing the application of the CIPP model across various educational programs in Indonesia. Yet, such a review is essential to understand the extent to which this model has been effectively implemented, as well as its strengths and limitations within the national educational landscape. Therefore, the urgency of this research lies in the need to conduct a systematic literature review on the application of the CIPP evaluation model in Indonesian

educational programs between 2019 and 2025.

Many academics have proposed the CIPP (Context, Input, Process, Product) evaluation model as a comprehensive solution for holistically assessing educational programs. The CIPP model evaluates not only outcomes but also the context, input, and processes underlying the achievement of educational goals (Duan et al., 2023; Rebia et al., 2023). This model has been widely applied across various educational fields, including vocational education, medical education, language education, and early childhood education (Santoso et al., 2024; Shi, 2024; Suryanto et al., 2024).

The CIPP evaluation model, developed by Stufflebeam, offers a comprehensive evaluation approach that considers these four aspects. This model has been used in various educational contexts to evaluate programs comprehensively (Stufflebeam & Zhang, 2017). It has been extensively used in diverse educational contexts to provide a complete picture of a program's strengths and weaknesses and to support more targeted recommendations for improvement (Duan et al., 2023). Furthermore, (Sankaran & Saad, 2022) notes that the CIPP model is considered flexible and adaptable to various levels and types of education and has proven effective in identifying factors affecting educational quality.

The CIPP model is chosen as a primary solution for evaluating educational programs due to its ability to provide comprehensive information, from initial needs to outcomes. It is considered effective in helping to evaluate and improve educational programs holistically. This approach not only assesses outcomes but also investigates the processes and influencing factors, enabling more precise recommendations for improvement.

Various studies have applied the CIPP model to evaluate educational programs across different fields, such as vocational education, medical education, and civil servant training (Najafi et al., 2020; Ratnaya et al., 2022; Yoshany et al., 2025). Furthermore, research by (Muji et al., 2021) demonstrated that the CIPP model is capable of evaluating project-based learning related to the strengthening of the Pancasila Student Profile in the implementation of the "Sekolah Penggerak" curriculum in Padang City. The study (Kumalasari & Fikroh, 2023) found that the evaluation of a character education program at SD Athalia showed that the program's context was excellent; however, the input, process, and product components still required improvement. The evaluation of a literacy program also showed an increase in students' interest in reading, although challenges in program implementation remained.

All previous studies affirm the importance of applying the CIPP model for evaluating educational programs. However, these studies have generally focused on the implementation of the CIPP model without analyzing its strengths and weaknesses (Ratnaya et al., 2022; Sulkifli et al., 2024). Additionally, research by (Sankaran & Saad, 2022) highlighted the suitability of the CIPP model for evaluating Bachelor of Education (BEd) programs in Malaysian higher education institutions; however, it did not discuss its application in the Indonesian education context. Furthermore, while studies by (Kumalasari & Fikroh, 2023; Muji et al., 2021) analyzed the effectiveness of applying the CIPP model in evaluating programs at the primary and secondary education levels in Indonesia, they have not yet discussed the model's advantages, limitations, or provided recommendations for its application. However, up to 2025, no systematic literature review has been identified that critically examines the application of the CIPP evaluation model within Indonesian education.

To date, no systematic review has synthesized the various studies on the use of the CIPP evaluation model in the Indonesian education context. Existing literature remains fragmented and does not provide a comprehensive understanding of usage trends, implementation quality, or the model's contributions to the development of educational evaluation practices. Furthermore, there is a lack of research that critically examines the strengths and weaknesses of the CIPP model while offering innovative recommendations for its use in Indonesian education. This highlights an urgent need for a literature synthesis that is not only descriptive but also reflective and analytical.

Therefore, this study offers a more comprehensive model by contributing a systematic review of the implementation of the Context, Input, Process, Product (CIPP) evaluation model in Indonesian educational programs from 2019 to 2025. The novelty of this study lies in three key aspects. First, a comprehensive synthesis of previously scattered studies to form a structured analysis of best practices, common weaknesses, and recent innovations in using the CIPP model in education. Second, a critical methodological review that identifies frequent errors and areas for improvement in evaluation practices. Third, the development of an evidence-based procedural recommendation framework for researchers and practitioners to apply the CIPP model more accurately, consistently, and in alignment with local contexts.

METHOD

This research uses the method of Systematic *Literature Review* (SLR). The purpose of using this method is to minimize bias in the process of answering research questions, and mixed with community sentiment analysis with the X application through Python *Google Collab* (Pang and Lee, 2008). The research began by identifying articles relevant to the theme of model-based educational evaluation. *Context, Input, Process, Product* (CIPP) through the databases Scopus, ScienceDirect, and Google Scholar, Minimal Sinta 2. SLR in this article refers to the preferred reporting items for systematic reviews and meta-analyses (PRISMA) flow to collect data and ensure research is conducted systematically. The stages carried out by researchers are as follows:

1. Inclusion and Exclusion Criteria for Publication Selection

At the inclusion and exclusion stage, the researcher creates criteria for including articles in this study. First, search for articles in the databases Scopus, ScienceDirect, Google Scholar, and Minimum Sinta 2. Second, articles published in journals from 2019 to 2025. Third, articles are searched based on the suitability of the research theme. After obtaining articles that match the research theme, researchers conduct a screening to be analyzed in depth and comprehensively.

Table 1. Database Sources and Keywords

Database	Keywords	Number of Documents
Scopus, ScienceDirect, and Google Scholar	CIPP Model of Education Evaluation	833

The literature search was conducted on May 10, 2025, in Scopus using the keyword "CIPP model of education evaluation" and found 258 articles. While in ScienceDirect, A total of 260, and in Google Scholar, Sinta 2 found 315 articles.

To determine the eligibility of articles that will be used as a source of analysis data, categories are used for *inclusion* and *exclusion by* setting several criteria:

Table 2. Eligibility of Articles

No.	Criteria	Inclusion	Exclusion
1.	Access Type	Open Access	Not Open Access
2.	Research Context	Research on teachers in primary and secondary schools.	Research on non-formal education, such as training, and others.
3.	Research Focus	Focus on research on the use of the CIPP evaluation model in various educational programs in Indonesia.	Research beyond the CIPP Model Teacher Evaluation

4.	Research Type	Systematic Literature Review Research, Qualitative and Quantitative Studies	Research that is only opinion or studies that only focus on general education
5.	Types of research	2019-2025	Publications before 2019
6.	Source Type	Articles published in international journals, Scopus, ScienceDirect, and Google Scholar, with a minimum accredited Sinta 2	Articles published in journals or conferences that are not or have not been indexed in recognized academic databases
7.	Language	English	Besides English

The criteria for empirical research and publications in the last five years aim to ensure that the studies included in this research are recent and relevant publications.CIPP model evaluation is currently underway. Language criteria, by focusing on one language, analysis and comparison between studies become more consistent and avoid differences in meaning that may arise due to language variation or translation.

2. Data Screening and Eligibility Assessment

Article findings in the Scopus database are filtered as needed. Relevant articles are listed, and irrelevant articles are not used. Screening articles is adjusted by title, abstract, and keywords. The same theme is discarded, then selected articles are entered into the Mendeley application and saved in the Research Information System (RIS) format. The next step is to enter the VOSviewer application to map the initial network of these relationships. Strengthening relevant research arguments requires an initial analysis of the thematic associations of the articles used in this study through the VOSviewer application. This can be seen in Figure 1.

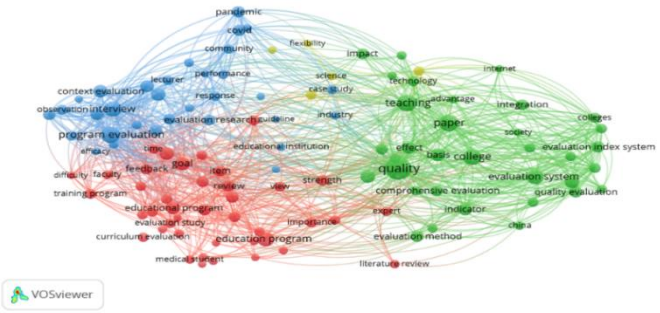

Figure 1. Intellectual Visualization Related to Evaluation Studies with the CIPP Model Approach

This visualization using VOSviewer illustrates an intellectual map related to evaluation studies using the CIPP Model approach. The map shows three main clusters that are interconnected. The red cluster focuses on the evaluation of educational programs, with dominant keywords such as educational program, goal, feedback, and curriculum evaluation, indicating great attention to aspects of program planning and implementation. The blue cluster emphasizes methodological

approaches in evaluation, such as the use of interviews, observations, and responses to the context of the pandemic (COVID-19 pandemic), which shows the importance of the contextual dimension in the application of the CIPP model.

3. Data Extraction

The process of searching and selecting literature from databases Scopus, ScienceDirect, and *Google Scholar* has been conducted since May 10, 2025. With a focus on titles, abstracts, and keywords, 833 articles published in 2019-2025 were found that are relevant to the topic that is the focus of this research, namely research on the use of the CIPP evaluation model in various educational programs, as well as the advantages and disadvantages of using CIPP in the context of education. At the identification stage, 833 articles were eliminated, 430 of which were due to duplication, corrupted files, and did not meet the inclusion criteria. Furthermore, from the remaining literature, 403 papers were screened to identify studies relevant to the research topic. A total of 125 papers were eliminated for several reasons, such as studies at universities, not in accordance with the topic, and only in the form of opinions or theories. At the final stage, an assessment of the feasibility of the results of the papers that had been screened was conducted, and as many as 26 articles that met the inclusion criteria were included in the data analysis section. In simple terms, this process can be described through the following PRISMA flow diagram.

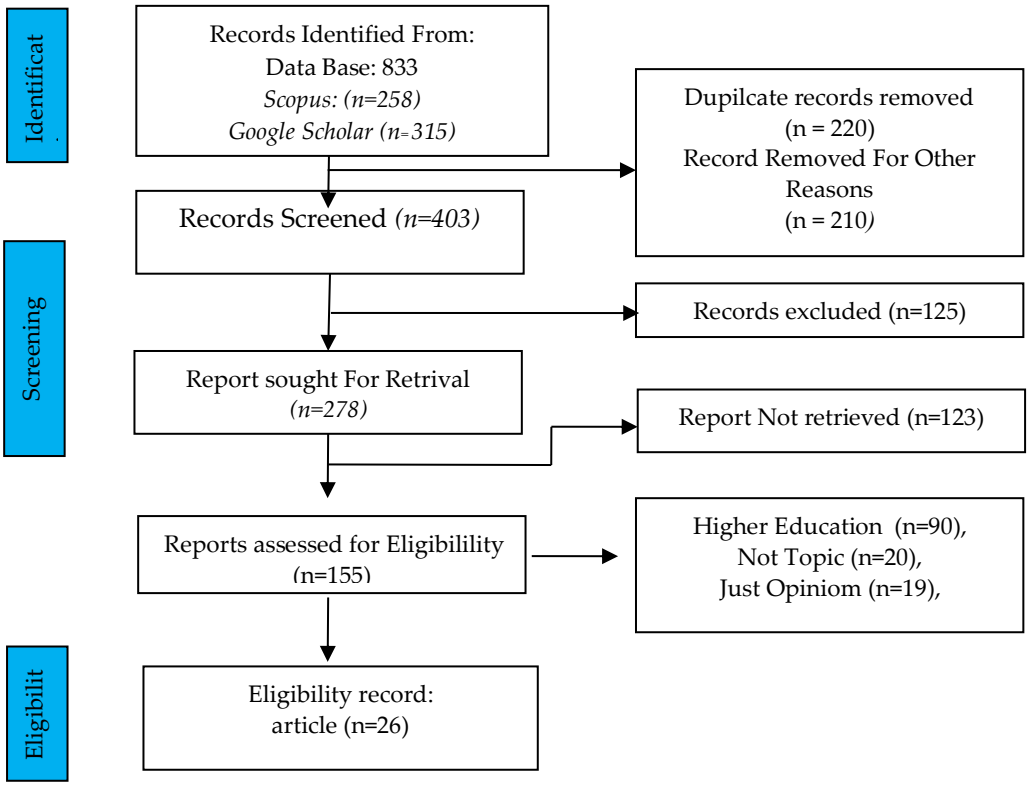

Figure 2. PRISMA Flow Diagram for Data Search

FINDINGS AND DISCUSSION

Findings

Trends in the Use of the CIPP Evaluation Model in Educational Programs

Based on a search of the Scopus, ScienceDirect, and Google Scholar (SINTA 2) databases using the keyword "CIPP Evaluation Model," a total of 833 articles were initially identified. After applying the PRISMA diagram-based identification process, 26 articles met the predetermined inclusion criteria and were selected for further analysis. The mapping of these 26 articles is detailed based on several attributes, including author names, publication year, journal information (name, volume, issue, year), type of publication, Scopus accreditation, and relevance to the research questions (RQ), which focus on how the trend of using the CIPP evaluation model in instrument validation has developed across various research studies and what common weaknesses exist in the implementation of CIPP for such validation purposes. The following chart presents the articles that met the inclusion criteria:

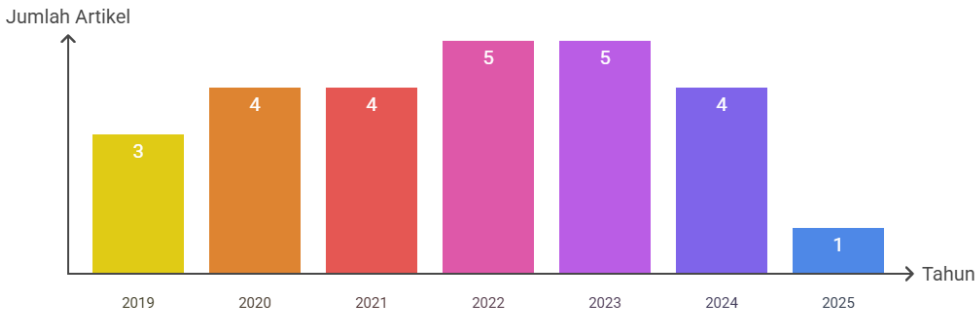

Figure 3. Illustrates The Number of Published Articles within the 2019–2025 period

The diagram above illustrates the number of published articles within the 2019–2025 period that met the inclusion criteria, totaling 26 articles. After undergoing the PRISMA-based selection process, these 26 articles were deemed eligible for deeper analysis. The publication trend reveals fluctuations in interest regarding the use of the CIPP model in educational program evaluation. The year 2022 recorded the highest number of publications (five articles), while 2025 had only one publication, likely due to time constraints in the current publication year. The data also indicate that the application of the CIPP evaluation model spans multiple academic disciplines. The following pie chart illustrates the distribution of documents by subject area related to studies on the CIPP Evaluation Model, based on Scopus data:

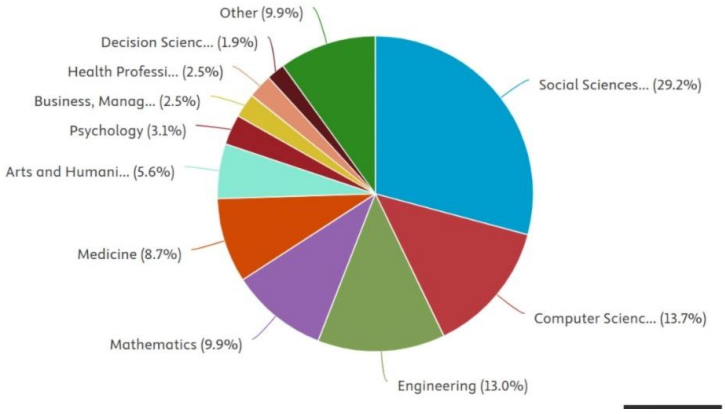

Figure 4. Evaluation of the CIPP Model in various scientific fields

This Scopus-derived data shows that 29.2% of the articles originated from the field of social sciences, followed by computer science (13.7%), engineering (13%), and mathematics (9.9%). Articles were also found in medicine, the humanities, psychology, as well as in management and health professions. These findings indicate that the CIPP model has been applied across multiple disciplines, not only in formal education but also in various scientific sectors and professional practices. Thus, the CIPP model is not only relevant for evaluating educational programs but is also effective across diverse academic and professional domains.

Strengths and Weaknesses of the CIPP Evaluation

Model in Educational Contexts The CIPP evaluation model demonstrates several notable strengths in its application across educational settings. Chief among these is its flexibility. CIPP can be adapted to various contexts, levels of education, and types of programs, ranging from formal schooling to professional training. Its ability to integrate both qualitative and quantitative data allows for a more comprehensive analysis, accommodating diverse sources of information and perspectives. Additionally, the model follows a systematic and structured framework, encompassing the evaluation of a program from the initial planning stage (Context), through resource readiness (Input), implementation (Process), and finally to the outcomes (Product). This holistic approach facilitates evidence-based decision-making and enhances accountability. Moreover, by encouraging stakeholder involvement, the CIPP model supports participatory evaluation practices, fostering ownership and relevance of the evaluation outcomes.

However, the model is not without its limitations. A frequently cited challenge is the high workload associated with conducting a full CIPP evaluation. Gathering and analyzing extensive data across all four components can be time-consuming and resource-intensive. Another critical

weakness lies in its rigid dependency on initial planning; the model tends to focus heavily on pre-established objectives, which may lead to the oversight of unanticipated outcomes or emergent issues during implementation. Furthermore, several studies have reported partial or incomplete application of the CIPP components, where evaluations disproportionately focus on only one or two dimensions, typically Process and Product, while neglecting Context and Input. This incomplete implementation undermines the comprehensive nature of the model and limits the utility of the findings for meaningful program improvement.

Discussion

Trends in the Use of the CIPP Evaluation Model in Educational Programs

The CIPP (Context, Input, Process, Product) evaluation model developed by Stufflebeam in the 1970s has evolved into a comprehensive and adaptive evaluation framework for educational settings. Unlike conventional evaluation approaches that focus solely on output, the CIPP model emphasizes a multi-layered analysis from identifying initial needs to assessing final program outcomes. Its primary strength lies in its capacity to function both as a formative and summative tool, supporting ongoing improvement throughout the program cycle. Between 2019 and 2025, the use of the CIPP model saw a significant increase, particularly in the post-pandemic period. The surge in publications during 2022–2023 reflects the academic response to Indonesia's Kurikulum Merdeka, which emphasizes contextual learning and teacher autonomy (Dizon, 2022; Thontawi et al., 2022). The CIPP model proved especially relevant as it enabled real-time monitoring of policy shifts and facilitated evidence-based feedback. Studies by (Bulhayat, 2019; Kumalasari & Fikroh, 2023) have shown that while student achievement scores were high, weaknesses in input, such as teacher training and infrastructure, were identified as core issues, thanks to the comprehensive lens of the CIPP model.

The model's flexibility is also evident in its widespread application across various educational fields. In vocational education, it supports the evaluation of industry readiness and resource availability (Ratnaya et al., 2022). In medical education, Najafi et.al. documented an eight-year use of the CIPP model to develop a practice-based medical curriculum. Here, the evaluation was not merely retrospective but formative, allowing policy adjustments to be made mid-implementation (Najafi et al., 2020). In China, Duan demonstrated how product evaluation extended beyond academic results to include industry-relevant skills, combining CIPP with competency-based and labor market-driven approaches (Duan et al., 2023).

Furthermore, digital technology has become a catalyst in accelerating the adoption of the CIPP evaluation model. The integration of tools such as VOSviewer, Google Colab, and learning analytics has enhanced the evaluation process, enabling faster and more precise decision-making (Singh, 2024). These technologies also support stakeholder engagement and accelerate feedback loops. For example, in inclusive education settings, real-time dashboards facilitate collaboration among teachers, parents, and policymakers. Geographically, the application of the CIPP model remains concentrated in Asia, particularly in Indonesia and Iran. Of the 16 selected studies between 2019–2025, most originated from these two countries. This suggests that the participatory and contextual nature of the CIPP model aligns well with the policy and cultural frameworks of Asian education systems. However, the limited distribution also signals a need to expand research into other regions, such as Africa and Europe, to test the model's cross-cultural validity and global relevance.

Despite its comprehensive framework, implementing the CIPP evaluation model presents several challenges. A significant concern is the heavy workload required to collect and analyze data across all four components: Context, Input, Process, and Product, which demands significant time, resources, and skilled personnel (Suryadin et al., 2022). Moreover, standardized indicators, particularly in the Process component, often fail to align with local educational practices, necessitating contextual adaptation (Ayyusufi et al., 2022; Muqorobin et al., 2022). To address this, many institutions have adopted a modular approach, evaluating only the most relevant components to reduce administrative burden while maintaining focus. Stakeholder involvement has become a key success factor, as the participatory nature of CIPP encourages engagement from teachers, students, parents, policymakers, and even industry players. According to (Chanthalangsy et al., 2024) such involvement enhances data validity and fosters ownership and commitment to follow-up actions. Consequently, evaluators now serve more as facilitators of collaborative reflection than as mere data collectors.

Looking ahead, the trend in CIPP-related research is expected to become increasingly integrated with advanced digital technologies. Artificial intelligence (AI) is already being used to automate qualitative data analysis, as well as predictive modeling based on historical data (Xia, 2022). Big data analytics enhances context mapping by incorporating educational policy, demographics, and socioeconomic indicators. Blockchain is also being explored as a secure and transparent data storage system for evaluation results (Bakti et al., 2023). Longitudinal

methodologies such as time-series analysis and participatory action research further reinforce the depth and reliability of continuous evaluation efforts.

In summary, the global trend of using the CIPP evaluation model from 2019 to 2025 reflects a multidimensional evolution involving deeper component analysis, broader application contexts, the integration of advanced technologies, and the strengthening of stakeholder participation. Its comprehensiveness, flexibility, and adaptability make CIPP a leading evaluation ecosystem in today's dynamic educational landscape. While implementation challenges such as workload and local indicator adaptation remain, strategies such as modular instrumentation and standardized evaluator training offer practical mitigation. With the adoption of AI, big data, and longitudinal methodologies, CIPP is poised to become a future-oriented evaluation framework that not only measures outcomes but also proactively shapes the trajectory of inclusive, innovative, and sustainable education.

Advantages, Weaknesses, and Recommendations from the Implementation of the CIPP Evaluation Model in Education

The CIPP (Context, Input, Process, Product) evaluation model offers a comprehensive framework that integrates the assessment of key dimensions within a continuous evaluation cycle, making it highly relevant for diverse educational settings. One of its notable strengths lies in its ability to systematically map the context of a program, addressing initial conditions, institutional needs, and external factors that influence implementation. For instance, (Hakan & Seval, 2019) found that applying the CIPP scale in Iranian educational institutions yielded high contextual validity, leading to more targeted policy recommendations. The model also excels in evaluating input readiness, such as human resources, facilities, and technical support, prior to implementation. Angelina demonstrated how input analysis using the CIPP framework improved the design of pre-service teacher training programs at Sanata Dharma University, resulting in more effective delivery (Angelina, 2022).

The process component of the model enables ongoing, formative assessment. Najafi et al. demonstrated how CIPP facilitated the real-time monitoring of a medical education curriculum, identifying resistance among lecturers and enabling timely interventions (Najafi et al., 2020). Meanwhile, the product dimension expands evaluation beyond academic outcomes, incorporating industry-linked achievements. Duan et al. developed product indicators for engineering programs that aligned academic success with professional application, thus bridging education with labor

market needs (Duan et al., 2023).

Stakeholder involvement further enhances the utility of CIPP. Sholeh emphasized that engaging teachers, students, parents, and policymakers increases accountability and the acceptance of results, making evaluations more relevant and actionable (Sholeh et al., 2023). This participatory aspect also facilitates communication between evaluators and implementers. Moreover, the CIPP model's flexibility in combining qualitative and quantitative methods, as noted by (Sopha & Nanni, 2019), supports both formative and summative evaluations. Aghimien et al. added that CIPP's value-based, collaborative approach helps uncover hidden strengths and weaknesses in programs (Aghimien et al., 2022). Overall, the CIPP model not only delivers detailed, evidence-based assessments but also promotes collaboration and continuous improvement, making it an invaluable tool for advancing educational quality in various contexts.

Despite its strengths, the CIPP evaluation model in educational contexts presents several notable weaknesses that require serious attention. First, the model often places a heavy administrative burden on institutions due to the large volume of instruments involved. The use of lengthy questionnaires, in-depth interviews, and mixed-method data analysis tends to be time-consuming and costly. A study by (Muqorobin et al., 2022). Comparing the CIPP and Kirkpatrick models in a physics teacher competency development program highlighted challenges, including a lack of experienced evaluators and staff resistance to administrative demands. Second, the need to adjust evaluation indicators to local contexts frequently demands expensive and time-intensive revalidation processes. Stressed that indicators effective in one academic culture may not be relevant in another, necessitating careful and culturally sensitive adaptations (Agus et al., 2023).

Third, the complexity of managing qualitative and quantitative data demands advanced analytical skills, including the use of statistical software and coding techniques, capabilities that many institutions still lack. (Musifuddin et al., 2024) Noted that many educational institutions lack evaluation teams with adequate training in advanced data analysis, resulting in a reliance on simplistic descriptive reports. This limits the evaluators' ability to reveal deeper meanings, interconnections between evaluation components, and broader policy implications. As such, while the CIPP model offers a holistic and flexible evaluation framework, its implementation must be accompanied by strategic investments in capacity building, indicator adaptation, and technological support to ensure meaningful, accurate, and contextually relevant evaluations.

To strengthen the application of the CIPP evaluation model in education, several strategic recommendations are proposed. First, modularizing evaluation instruments allows institutions to focus on specific components, such as Context, Input, Process, or Product, based on priority needs, reducing administrative workload and optimizing resources (Banister et al., 2020). Second, standardized training for evaluators is essential and can be facilitated through partnerships with universities and professional bodies to improve competencies in mixed methods and advanced analysis (Muji et al., 2021; Wiyono et al., 2025), integrating digital technologies, such as real-time dashboards, mobile data collection apps, and learning analytics—into the Process and Product phases enhances efficiency and data interactivity (Singh, 2024).

Fourth, adjusting evaluation indicators to fit local contexts is crucial. This involves forming multidisciplinary teams to design culturally and contextually sensitive indicators, with validation through pilot testing to reduce adaptation failures (Sustiawati et al., 2023). Fifth, educational institutions must provide robust administrative support, involve stakeholders, and manage evaluation scope to avoid overload (Rebia et al., 2023). At the policy level, flexible implementation guidelines from accrediting bodies are needed to balance standardization with institutional autonomy (Basaran et al., 2021; Suryadin et al., 2022). Cross-cultural research should also be expanded to validate the model's global applicability and enrich best practices.

By implementing these recommendations, educational institutions can not only maximize the strengths of the CIPP evaluation model as a holistic, participatory, and improvement-oriented framework but also mitigate practical weaknesses such as administrative burdens, the need for analytical capabilities, and challenges in local adaptation of indicators. A study by (Santoso et al., 2024) emphasizes that the successful implementation of the CIPP model is highly dependent on resource readiness and the technical capacity of evaluators to manage data complexity. Furthermore, (Suryanto et al., 2024) found that modularization and digitalization in the application of the CIPP model can accelerate the evaluation process. Thus, CIPP becomes not only a measurement tool but also a reflective framework that drives continuous institutional transformation (Irene, 2023). Ultimately, the CIPP model may evolve into an increasingly efficient, responsive, and context-sensitive evaluation ecosystem, aligning with the vision of 21st-century education that is inclusive, innovative, and adaptive to change.

The CIPP evaluation model has great potential as a tool in planning, monitoring, and evaluating educational programs as a whole. Its main advantage lies in the flexibility of use, the

integration of qualitative and quantitative approaches, as well as its participatory orientation, which involves stakeholders in each evaluation stage. This model also supports the principle of student-centered education and encourages evidence-based decision-making (Hakan & Seval, 2019). However, some challenges remain, such as the administrative burden of collecting complex data, the limited resources of evaluators for advanced analysis, and weaknesses in capturing the unexpected impact of the program being evaluated (Ninghardjanti et al., 2023; Sustiawati et al., 2023). The complexity of the data produced also demands high technical competence, both in the use of statistical software and in qualitative data processing techniques that are not yet fully understood by many educational institutions.

Thus, this study offers strategic recommendations for implementing the CIPP evaluation model in the Indonesian education context to address the challenges of modern-era education while maximizing the benefits of the CIPP model. First, the modularization of evaluation instruments enables institutions to focus their evaluation on specific components as needed (Agus et al., 2023; Duan et al., 2023). Second, the development of standardized evaluator training is crucial for enhancing data analysis capacity and methodological understanding (Muji et al., 2021). Third, the integration of digital technologies, such as real-time dashboards and mobile applications, can accelerate the evaluation process and improve data accuracy (Chanthalangsy et al., 2024; Yoshany et al., 2025).

Fourth, evaluation indicators need to be adjusted locally by multidisciplinary teams to be more contextual (Rebia et al., 2023). Fifth, flexible guidelines are needed from educational authorities to ensure that this model is adaptable yet comprehensive (Ninghardjanti et al., 2023). Finally, cross-country and cross-cultural research should be encouraged to test the global validity and enrich the best practices of the CIPP model (Santoso et al., 2024). With this approach, the CIPP model will not only be relevant in various educational contexts but can also develop as an adaptive, innovative, and sustainable evaluation framework.

This study presents significant novelty compared to previous studies, especially in terms of its scope and analytical approach. Unlike previous studies, which are generally limited to case studies of one particular level or program, this study presents a systematic synthesis of various applications of the CIPP evaluation model in the context of education programs in Indonesia and Asia from 2019 to 2025. This approach not only maps the trend of use, strengths, and weaknesses of the CIPP model but also identifies implementation innovations such as modularization of evaluation

instruments and integration of digital technologies that have not been explicitly explored in previous literature. Thus, this study broadens the understanding of the flexibility and practical challenges of the CIPP model in the context of contemporary education policy, especially post-pandemic.

In terms of contribution, this study has relevant implications both practically and theoretically. Practically, the results of this study provide concrete recommendations for educational institutions and evaluators to implement the CIPP model more efficiently, contextually, and participatively, through multi-stakeholder involvement, standardized evaluator training, and the use of technology such as evaluation dashboards and mobile surveys. Meanwhile, theoretically, this study enriches the literature on educational evaluation by proposing an integrative framework that combines formative, participatory, and data-based approaches. This opens up opportunities for the development of evaluation theories that are more adaptive to policy dynamics and socio-cultural diversity at the local and regional levels.

However, the findings in this study have several limitations that need to be considered. The dominant data coverage comes from the Asian region, especially Indonesia and Iran, making the generalization of the results to the global context must be done with caution. In addition, because this study uses a systematic literature review approach without direct observation in the field, the depth of information is highly dependent on the transparency of reporting from the articles analyzed. Therefore, although this study provides a comprehensive picture of the practices and challenges of implementing the CIPP model, the results are diagnostic and exploratory and need to be complemented by further mixed-methods studies to produce more applicable and contextually relevant recommendations.

CONCLUSION

The Evaluation model Context, Input, Process, Product(CIPP) has proven to be an effective, comprehensive framework for assessing various education programs in a global context. Analysis of usage trends for the period 2019–2025 shows a significant increase in adoption of the CIPP evaluation model, especially in the Asian region, in response to the need for more holistic evaluation after the pandemic and the emergence of new policies such as the Independent Curriculum. The CIPP model offers an evaluation framework that not only measures outcomes but also deepens understanding of the context, inputs, and processes that influence program success. By continuing to develop instruments, evaluator capabilities, and supporting digital infrastructure, CIPP will

remain the main choice in efforts to improve the quality of education in the 21st century, providing a strong evidence base for strategic decision-making and continuous improvement. The CIPP evaluation model has significant advantages in flexibility, integration of qualitative and quantitative approaches, and participatory orientation that support evidence-based decision-making in education. However, its implementation also faces weaknesses that are challenges, such as high evaluation burden, limitations in detecting unexpected impacts, and the need for advanced technical capabilities in data analysis. Therefore, it is recommended to modularize evaluation instruments, standardized evaluator training, integrate digital technologies, adjust indicators locally, and provide flexible and contextual policy support. With these steps, the CIPP model can be optimized into an adaptive, efficient, and relevant evaluation framework across levels and contexts of education.

REFERENCES

Aghimien, D. O., Aigbavboa, C., Edwards, D. J., Mahamadu, A. M., Olomolaiye, P., Nash, H., & Onyia, M. (2022). A fuzzy Synthetic Evaluation of the Challenges of Smart City Development in Developing Countries. *Smart and Sustainable Built Environment*, *11*(3), 405–421. https://doi.org/10.1108/SASBE-06-2020-0092

Agus, Juliadharma, M., & Djamaluddin, M. (2023). Application of the CIPP Model in Evaluation of The Inclusive Education Curriculum in Madrasah Aliyah. *Nidhomul Haq : Jurnal Manajemen Pendidikan Islam*, *8*(1), 31–50. https://doi.org/10.31538/ndh.v8i1.2705

Angelina, P. (2022). Evaluating English Pre-Service Teacher Professional Education Program Implementation. *LLT Journal: A Journal on Language and Language Learning*, *25*(1), 132–148.

Ayyusufi, A. M., Anshori, A., & Muthoifin. (2022). Evaluation of The CIPP Model on The Tahfidz Program in Islamic Boarding Schools. *Nazhruna: Jurnal Pendidikan Islam*, *5*(2), 466–484. https://doi.org/10.31538/nzh.v5i2.2230

Aziz, S., Mahmood, M., & Rehman, Z. (2019). Implementation of CIPP Model for Quality Evaluation at School Level: A Case Study. *Journal of Education and Educational Development*, *5*(1), 189–206. https://doi.org/10.22555/joeed.v5i1.1553

Bakar, M. A., Umroh, K. A., & Hameed, F. (2023). Improving Quality Islamic Education for Today's Generation. *At-Tadzkir: Islamic Education Journal*, *2*(2), 118–128.

Bakti, I. K., Ayuningtias, Z. Y., Rusdi, Syaifudin, M., & Syafaq, H. (2023). The Role of Artificial Intelligence in Education: A Systematic Literature Review. *International Journal on Semantic Web and Information Systems*, *19*(1), 182–197. https://doi.org/10.4018/IJSWIS.318448

Banister, G., Portney, L. G., Vega-Barachowitz, C., Jampel, A., Schnider, M. E., Inzana, R., Zeytoonjian, T., Fitzgerald, P., Tuck, I., Jocelyn, M., Holmberg, J., & Knab, M. (2020). The interprofessional dedicated education unit: Design, implementation and evaluation of an innovative model for fostering interprofessional collaborative practice. *Journal of Interprofessional Education and Practice*, *19*(11), 100–115. https://doi.org/10.1016/j.xjep.2019.100308

Basaran, M., Dursun, B., Gur Dortok, H. D., & Yilmaz, G. (2021). Evaluation of Preschool Education Program According to CIPP Model. *Pedagogical Research*, *6*(2), 91–114. https://doi.org/10.29333/pr/9701

Bulhayat. (2019). Evaluasi Model CIPP Pada Implementasi KTSP Pendidikan Agama Islam di Sekolah Menengah Atas Jombang. *Jurnal Al-Adabiya*, *14*(1), 85–102.

Chanthalangsy, P., Yeh, B. Il, Choi, S. J., & Park, Y. C. (2024). Program Evaluation of Postgraduate Obstetrics and Gynecology Training in Lao People's Democratic Republic - Using the CIPP Model. *BMC Medical Education*, *24*(1), 1–10. https://doi.org/10.1186/s12909-023-04942-6

Dizon., A. G. (2022). Historical Development of CIPP as A Curriculum Evaluation Model. *History of Education*.

Duan, P., Xiang, J., Niu, H., & Han, C. (2023). Construction of Evaluation Index for Chinese Engineering Undergraduates Based on CIPP Model. *Sage Open*, *13*(1), 1–14. https://doi.org/10.1177/21582440221149415

Hadi, S., Andrian, D., & Kartowagiran, B. (2019). Evaluation Model for Evaluating Vocational Skills Programs on Local Content Curriculum in Indonesia: Impact of educational system in Indonesia. *Eurasian Journal of Educational Research*, *2*(82), 45–62. https://doi.org/10.14689/ejer.2019.82.3

Hakan, K., & Seval, F. (2019). CIPP Evaluation Model Scale: Development, Reliability and Validity. *Procedia: Social and Behavioral Sciences*, *15*(2), 592–599. https://doi.org/10.1016/j.sbspro.2011.03.146

Idrus, A., Setiyadi, B., Syarif, M. I., Wahyu, A., & Suharto. (2025). Evaluation of education program implementation at Pusat Kegiatan Belajar Masyarakat (PKBM), Jambi City. *Evaluation and Program Planning*, *111*(1), 1–8.

Irene, E. A. (2023). Evaluation of Teacher Education Curricula and its Relevance to Licensure Examination Using Context, Input, Process and Product (CIPP) Model. *Social Sciences and Humanities Open*, *8*(1), 100607. https://doi.org/10.1016/j.ssaho.2023.100607

Kumalasari, S. M., & Fikroh, R. A. (2023). Development of Android-based Augmented Reality Learning Media on Atomic Matter. *Jurnal Pendidikan Sains Indonesia*, *11*(3), 683–696. https://doi.org/10.24815/jpsi.v11i3.30960

Muji, A. P., Gistituati, N., Bentri, A., & Falma, F. O. (2021). Evaluation of the Implementation of The Sekolah Penggerak Curriculum Using The Context,Input, Process and Product Evaluation Model in High Schools. *JPPI (Jurnal Penelitian Pendidikan Indonesia)*, *7*(3), 377–385. https://doi.org/10.29210/020211231

Muqorobin, Komarudin, Badrujaman, A., Arthur, R., & Prayogi, S. (2022). CIPP vs Kirkpatrick in the Evaluation of Physics Teacher Competency Development Programs : literature study. *Journal of Physics: Conference Series PAPER*, 1–10. https://doi.org/10.1088/1742-6596/2165/1/012039

Musifuddin, M., Yunitasari, D., & Murcahyanto, H. (2024). CIPP Model Approach to School-Based Management Program Evaluation. *Nidhomul Haq : Jurnal Manajemen Pendidikan Islam*, *9*(1), 104–116. https://doi.org/10.31538/ndh.v9i1.4440

Najafi, N. V., Khamseh, A. A., & Mirzazadeh, A. (2020). An Integrated Sustainable and Flexible Supplier Evaluation Model under Uncertainty by Game Theory and Subjective/Objective Data: Iranian Casting Industry. *Global Journal of Flexible Systems Management*, *21*(4), 309–322. https://doi.org/10.1007/s40171-020-00250-w

Ninghardjanti, P., Murtini, W., Hindrayani, A., & Sangka, K. B. (2023). Evaluation of the Smart Indonesia Program as a Policy to Improve Equality in Education. *Sustainability*, *15*(6), 1–16. https://doi.org/10.3390/su15065114

Ratnaya, G., Indriaswuri, R., Widayanthi, D. G. C., & Atmaja, I. M. P. D. (2022). CIPP Evaluation Model for Vocational Education: A Critical Review. *Education Quarterly Reviews*, *5*(3), 1–8. https://doi.org/10.31014/aior.1993.05.03.519

Rebia, P. S., Suharno, Tamrin, A. G., & Akhyar, M. (2023). Evaluation of Product-Based Education

Training Class at Vocational High School using the CIPP Model. *Journal of Curriculum and Teaching*, *12*(3), 135–146. https://doi.org/10.5430/jct.v12n3p135

Sankaran, S., & Saad, N. (2022). Evaluating the Bachelor of Education Program Based on the Context, Input, Process, and Product Model. *Frontiers in Education*, *7*(6), 1–8. https://doi.org/10.3389/feduc.2022.924374

Santoso, P., Sari, E., & Karnati, N. (2024). Evaluation of the Indonesian Languange Training Program at the Defence Languange Training Centre. *Journal of Ecohumanism*, *3*(3), 956–980. https://doi.org/10.62754/joe.v3i3.3398

Shi, X. (2024). Evaluation of Practice Courses in Preschool Education Based on CIPP Evaluation Model. *SHS Web of Conferences*, *10*(8), 190–195. https://doi.org/10.1051/shsconf/202419001018

Sholeh, M. I., Efendi, N., & Junaris, I. (2023). Evaluasi dan Monitoring Manajemen Pembelajaran Pendidikan Islam dalam Upaya Peningkatan Kualitas Pendidikan. *Refresh: Manajemen Pendidikan Islam*, *1*(3), 48–73.

Singh, M. D. (2024). Evaluation Framework for Nursing Education Programs: Application of the CIPP Model. *International Journal of Nursing Education Scholarship*, *1*(1), 101–110.

Sopha, S., & Nanni, A. (2019). The CIPP Model: Applications in Language Program Evaluation. *Journal of Asia TEFL*, *16*(4), 1360–1367. https://doi.org/10.18823/asiatefl.2019.16.4.19.1360

Stufflebeam, D., & Zhang, G. (2017). *The CIPP Evaluation Model: How to Evaluate for Improvement and Accountability*. The Guilford Press.

Sulkifli, Nade, E., Khumairah, E. S., & Riska. (2024). Pendekatan CIPP dalam Evaluasi Program Pendidikan: Tinjauan Literatur pada Program Pendidikan di Indonesia. *JMPI: Jurnal Manajemen, Pendidikan Dan Pemikiran Islam*, *1*(1), 1–15.

Suryadin, A., Nurfitriani, & Sari, W. P. (2022). Lazismu and Muhammadiyah Philanthropy in the Covid-19 Pandemic(CIPP Evaluation Model). *Jurnal Ilmiah Peuradeun*, *10*(2), 385–402. https://doi.org/10.26811/peuradeun.v10i2.674

Suryanto, A., Aman, Yuliana, L., Purwanto, N. A., Firdaus, F. M., & Fadhli, R. (2024). Context Input Process Product Evaluation of Hybrid Learning Practices in Indonesian Universities. *International Journal of Evaluation and Research in Education*, *13*(1), 475–487. https://doi.org/10.11591/ijere.v13i1.26588

Sustiawati, N. L., Negara, I. G. O. S., Sumarno, R., & Nalan, A. S. (2023). Evaluation of the Management of Performing Arts Creation Learning. *Harmonia: Journal of Arts Research and Education*, *23*(1), 195–207. https://doi.org/10.15294/harmonia.v23i1.43486

Tamsiyati, E., Agissa, A. W., Yuniar, & Junaidah. (2025). The Impact of Eliminating National Examinations on Edu- cational Evaluation Quality in Indonesia : A Policy Analysis on Equity and Accountability. *International Journal of Education and Literature*, *4*(1), 294–304.

Thontawi, M., MY, M., Chaniago, F., Fiqhi, A., Hazairin, I. N., & Afifah, Y. (2022). Tahfidz Al-Qur'an: A Study of Learning Management Systems in Higher Education. *AL-TANZIM: Jurnal Manajemen Pendidikan Islam*, *6*(1), 574–585. https://doi.org/10.33650/al-tanzim.v6i2.3535

Wiyono, B. B., Komariah, A., Hidayat, H., & Kusumaningrum, D. E. (2025). The Structural Effects of Evaluation Types in The Implementation Process of The Independent Learning Program in Higher Education. *Discover Sustainability*, *6*(1), 321–338. https://doi.org/10.1007/s43621-025-01171-3

Xia, X. (2022). Diversion Inference Model of Learning Effectiveness Supported by Differential Evolution Strategy. *Computers and Education: Artificial Intelligence*, *3*(April), 10–21. https://doi.org/10.1016/j.caeai.2022.100071

Yoshany, N., Mahmoodabad, S. S. M., Moradi, L., & Sharma, M. (2025). Beyond Traditional Training: a comprehensive CIPP evaluation of medical internships: assessing program design,

implementation, and clinical competency outcomes. *BMC Medical Education*, *25*(1), 1–8. https://doi.org/10.1186/s12909-025-07404-3