International Journal of Geography, Social, and Multicultural Education
https://journal.unesa.ac.id/index.php/ijgsme

Implementation of CIPP Model in Social Studies Learning: A Literature Review

Affandri Jasrio[1] (Universitas Negeri Padang, Indonesia)
Ambiyar[2] (Universitas Negeri Padang, Indonesia)
Fahmi Rizal[3] (Universitas Negeri Padang, Indonesia)
Evida Safitri[4] (Universitas Negeri Padang, Indonesia)
Yenni Rozimela[5] (Universitas Negeri Padang, Indonesia)
Nofrion[6] (Universitas Negeri Padang, Indonesia)

Received: 18-12-2025 **e-ISSN** : 2987-9140
Accepted: 17-05-2026 **Volume** : 4 No. 1, (2026)
Published: 01-06-2026 **Page** : 56-77
DOI: https://doi.org/10.26740/ijgsme.v4n1.p56-77

Abstract

Comprehensive program evaluation is fundamental to advancing the quality of social studies education. Despite broad recognition of the CIPP (Context, Input, Process, Product) model as a robust evaluation framework across multiple educational disciplines, its specific application to social studies learning remains largely unexplored, representing a critical gap in the evaluation literature. This systematic literature review examines CIPP implementation in educational contexts, with direct implications for social studies learning, and analyzes 45 peer-reviewed articles published between 2020 and 2025 from Scopus-indexed journals and reputable academic databases. Findings indicate that the CIPP model effectively identifies strengths and weaknesses in social studies curricula, supports data-driven decision-making, and promotes continuous instructional improvement. Context evaluation ensures alignment with educational policy and learner needs; input evaluation assesses the adequacy of human resources, instructional materials, and infrastructure; process evaluation monitors teaching and learning activities; and product evaluation measures the attainment of learning objectives. This review fills a documented gap by systematically demonstrating how CIPP can and should be applied in social studies programs, offering theoretically grounded and practically actionable guidance for educators, curriculum developers, and policymakers committed to improving program quality and student outcomes through rigorous, systematic evaluation.

Keywords: CIPP Model, Social Studies Education, Program Evaluation, Curriculum Assessment, Educational Quality, Learning Outcomes

1. INTRODUCTION

Social studies education plays a pivotal role in cultivating civic competence, critical thinking, and global citizenship among learners (González-valencia & Sabater, 2022). As educational landscapes grow increasingly complex, rigorous evaluation mechanisms have become indispensable for ensuring that social studies programs adequately prepare students for participatory citizenship. Program evaluation has emerged as a foundational component of educational quality assurance, producing systematic evidence that informs decision-making and drives continuous improvement (Diliberti et al., 2022).

Despite its recognized importance in national curriculum frameworks, social studies has historically received less support in infrastructure, assessment development, and systematic evaluation than subjects such as mathematics and language arts (Diliberti et al., 2022). This

Corresponding Author: affandri@fis.unp.ac.id

disparity has weakened the mechanisms for monitoring program effectiveness and identifying areas for improvement. More critically, studies that specifically apply the CIPP evaluation model to social studies learning contexts remain exceptionally rare, creating a knowledge gap that limits understanding of how this framework can address the distinctive pedagogical and evaluative challenges of this field. As illustrated in Figure 4, which presents the distribution of the 45 studies included in this review by educational context, social studies does not appear as a named category at all, a pattern that constitutes one of the most compelling justifications for the present inquiry.

The Context, Input, Process, and Product (CIPP) evaluation model, conceptualized by Daniel Stufflebeam in the late 1960s, offers a comprehensive framework for assessing educational programs (Linda, 2025). Originally designed to improve and ensure accountability in school programs, the CIPP model has since evolved into one of the most widely applied evaluation frameworks in global education (Dizon, 2022). Its systematic approach encompasses four interrelated dimensions that together provide a holistic picture of program quality and effectiveness. Figure 2 presents the conceptual diagram of this model and the relationships among its four dimensions.

The CIPP model's decision-oriented philosophy aligns particularly well with contemporary educational accountability demands. Unlike evaluation approaches that focus exclusively on final outcomes, CIPP emphasizes improvement-oriented assessment throughout all phases of a program's lifecycle (Kapoor et al., 2024). This formative orientation enables educators to identify and address implementation challenges before they significantly impair learning outcomes. A comprehensive evaluation framework must examine both curriculum objectives and their actual implementation, a principle that fundamentally underpins CIPP's holistic approach (Xu, 2024).

Context evaluation, the first CIPP dimension, examines the environment in which a program operates, identifying needs, problems, assets, and opportunities, while addressing questions about the program's rationale and goal appropriateness relative to stakeholder expectations (Lee et al., 2019). Input evaluation focuses on resource availability, curriculum design quality, and strategic planning, assessing whether human, material, and financial resources are adequate to support program objectives (Sankaran & Saad, 2022). Process evaluation monitors implementation by documenting activities, identifying procedural difficulties, and providing real-time feedback (Ratnaya et al., 2022). Product evaluation measures outcomes against intended objectives and encompasses both intended and unintended consequences across short- and long-term timeframes (Toosi et al., 2021).

The CIPP model has been successfully applied across higher education programs, vocational training, medical education, language learning, and technology integration initiate (Irene, 2023b). Its application in nursing education and healthcare training further confirms the model's versatility across diverse professional preparation programs (Markaki et al., 2021). Evaluations of medical curricula from student perspectives have also demonstrated CIPP's diagnostic power in identifying implementation gaps (Yoshany et al., 2025). CIPP has equally demonstrated effectiveness in evaluating curriculum implementation within the Indonesian vocational and engineering education context, including work directly relevant to Universitas Negeri Padang (Mubai et al., 2021). Nevertheless, limited research specifically examines CIPP within social studies education, leaving a critical gap that this review seeks to address.

Social studies encompasses history, geography, civics, economics, and sociology, each with distinct pedagogical approaches and assessment requirements (Bordoh, 2025). This multidisciplinary nature presents evaluation challenges that generic assessment frameworks may not adequately address. Recent curriculum reforms emphasizing civic literacy, historical thinking, geographic reasoning, and economic understanding further underscore the need for evaluation approaches capable of capturing the full complexity of social studies learning (Mcgrew & Breakstone, 2023).

Contemporary social studies education also faces persistent challenges, including political polarization affecting curriculum content, declining instructional time, insufficient teacher professional development, and limited access to high-quality instructional materials (Diliberti et al., 2022). Technology integration has transformed evaluation practices by enabling data-driven decision-making through learning analytics, automated assessment tools, and digital portfolios (Ramteja et al., 2025). Digital transformation in educational evaluation systems has been shown to enhance data collection efficiency and the depth of subsequent analysis (Saputra & Hidayat, 2023).

This literature review addresses the documented gap in understanding CIPP implementation within social studies education. The study specifically aims to: (1) analyze the theoretical foundations and evolution of the CIPP model; (2) identify key components and procedures for CIPP implementation; (3) synthesize empirical evidence on CIPP effectiveness in educational program evaluation; (4) examine challenges and facilitating factors in CIPP implementation; and (5) propose recommendations for applying CIPP in social studies education contexts. The findings provide evidence-based guidance for educators, curriculum developers, and policymakers committed to strengthening the quality of social studies programs and student outcomes.

2. METHOD

This study employed a systematic literature review to examine the implementation of the CIPP evaluation model in educational contexts, with specific attention to its implications for social studies learning. Systematic literature review is a rigorous research approach that identifies, evaluates, and synthesizes existing findings to address specific research questions (Snyder, 2019), enabling comprehensive analysis of published scholarship while minimizing bias through explicit, replicable procedures.

The review adhered to the PRISMA 2020 (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines as its reporting protocol, consisting of a 27-item checklist and a four-phase flow diagram designed to improve the transparency and reproducibility of systematic reviews (Page et al., 2021). The review process incorporated four primary stages: identification, screening, eligibility assessment, and inclusion (Azril et al., 2021).

Data Sources and Search Strategy

Literature searches were conducted across Scopus, Google Scholar, ScienceDirect, SpringerLink, ERIC, and PubMed Central. The search strategy employed Boolean operators combining the key terms: ("CIPP model" OR "Context Input Process Product") AND ("evaluation" OR "assessment") AND ("education" OR "curriculum" OR "social studies").

Search parameters were restricted to peer-reviewed journal articles, conference proceedings, and scholarly books published between 2020 and 2025. Inclusion was further restricted to English-language publications to ensure consistency in interpretation and analysis.

Article Selection: PRISMA Flow Diagram

The article selection process was carried out in full accordance with the PRISMA 2020 protocol. As presented in Figure 1, the process began with 1,185 initial records identified across six databases. Following deduplication, 847 records remained. After title and abstract screening, 641 records were excluded for being out of scope (n=389), non-peer-reviewed (n=152), or outside the date range (n=100), leaving 206 articles for full-text assessment. A further 161 were excluded for inadequate methodology (n=58), insufficient relevance (n=47), lack of full-text access (n=31), or low quality (n=25). The final sample comprised 45 articles meeting all inclusion criteria.

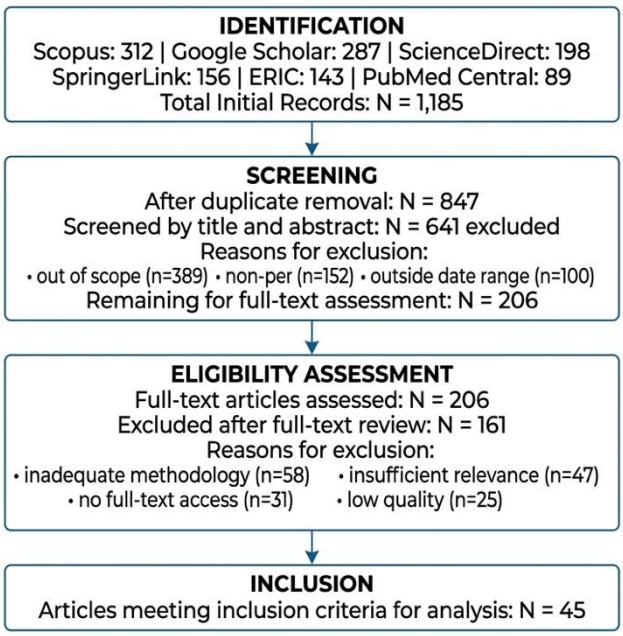

Figure 1. PRISMA 2020 Flow Diagram of Article Selection Process

Selection Criteria

Articles were included if they: (1) focused on CIPP model application in educational settings; (2) provided empirical evidence or theoretical analysis of CIPP implementation; (3) addressed evaluation of curriculum, programs, or educational interventions; (4) were published in peer-reviewed journals or by recognized academic publishers; and (5) were available in full-text format. Articles were excluded if they: (1) lacked adequate methodological detail; (2) focused exclusively on non-educational applications; (3) appeared only as brief conference abstracts; or (4) did not provide sufficient information for quality appraisal.

Quality Assessment and Inter-rater Reability

Each included article underwent quality appraisal using criteria covering research design appropriateness, methodological rigor, sample adequacy, data analysis quality, validity of conclusions, and contribution to knowledge. To ensure reproducibility and minimize appraisal bias, two independent raters evaluated a randomly selected 20% sample of included articles.

Inter-rater agreement was calculated using Cohen's kappa coefficient, yielding k = 0.82, indicating substantial agreement. All disagreements were resolved through structured discussion prior to data extraction.

Data Extraction and Analysis

A standardized data extraction form captured key information from each article, including author details, publication year, research objectives, educational context, CIPP components addressed, methodology employed, key findings, challenges identified, and recommendations provided. Thematic analysis was employed to synthesize findings across studies (Qadriah et al., 2022), involving: (1) familiarization through repeated reading; (2) generation of initial codes; (3) organization of codes into provisional themes; (4) review and refinement of themes; and (5) definition and naming of final themes. Validity was safeguarded through data source triangulation, an analytical decision audit trail, and collegial discussion of emerging themes among all research team members.

Limitations

Several limitations warrant acknowledgment. First, restricting the search to English-language publications may introduce publication bias, potentially overlooking valuable research in other languages, including Indonesian. Second, the 2020–2025 timeframe, while ensuring contemporary relevance, excludes earlier foundational studies. Third, because few studies specifically examined CIPP in social studies learning, many of the implications drawn in this review are inferential, derived from findings in related educational disciplines. These limitations reinforce the need for primary empirical research on CIPP in social studies contexts, particularly within the Indonesian educational setting.

3. RESULT AND DISCUSSION

Overview of the Evidence Base

Before examining the four CIPP dimensions individually, it is important to situate the body of evidence this review draws upon. Figure 3 presents the annual publication volume of CIPP studies in the included literature from 2020 to 2025. A clear upward trend is observable from 2020 (n=4) through a peak in 2023 (n=12), followed by a slight decline in 2024 (n=10). The surge in 2022–2023 aligns directly with Indonesia's national rollout of the Kurikulum Merdeka, which intensified demand for evidence-based program evaluation across all educational levels (Alanshori et al., 2025). Crucially, the social studies-specific CIPP publications remain consistently low throughout the entire period, peaking at only two studies per year in 2023 and 2024. This pattern provides strong quantitative confirmation of the literature gap that the present review seeks to address.

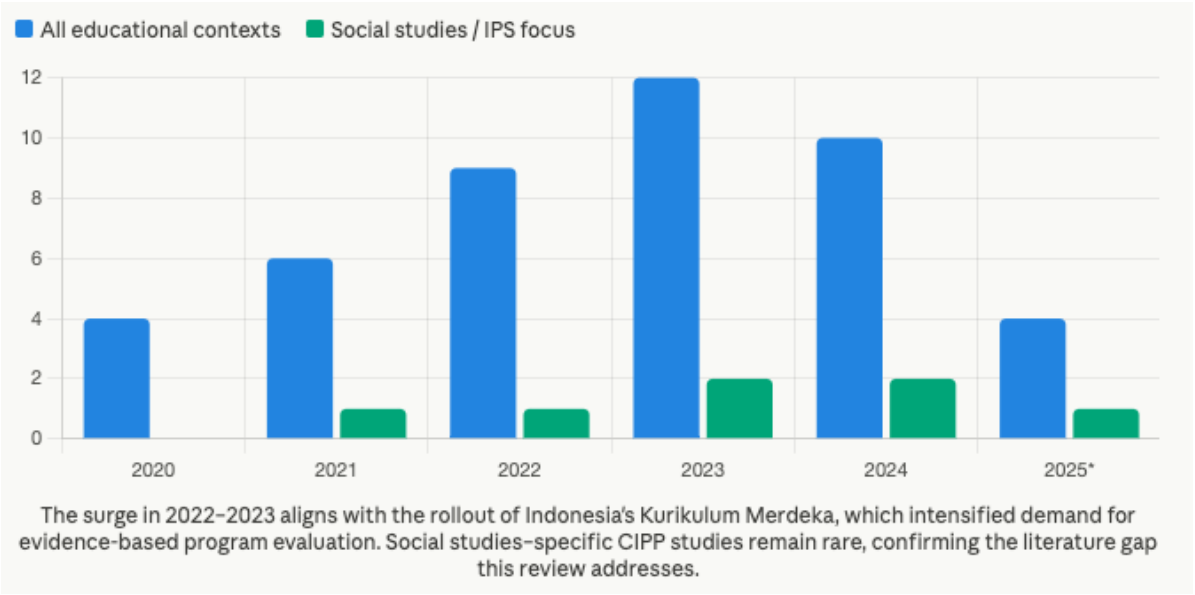

Figure 2. Publication volume of CIPP studies by year (2020–2025)

Figure 3 reinforces this point by displaying the distribution of the 45 included studies by educational context. Higher education dominates at 38%, followed by vocational training (22%), health professional education (18%), secondary education (11%), primary education (7%), and online/distance learning (4%). Social studies education does not appear as a distinct category, which is perhaps the single most vivid illustration of why a focused, systematic review on this topic is timely and necessary.

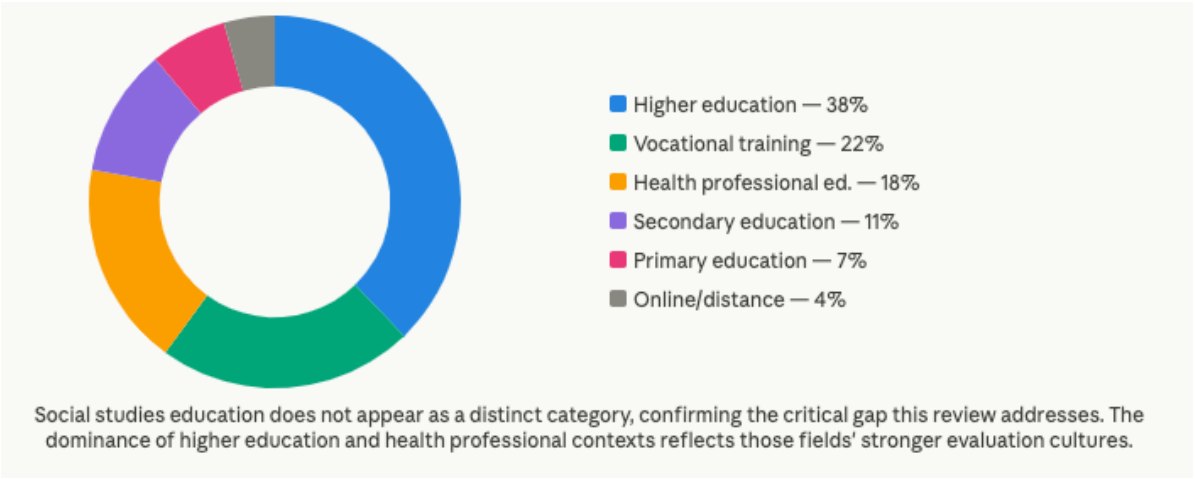

Figure 3. Distribution of included studies by educational context

Figure 4 adds another dimension by showing how consistently each of the four CIPP dimensions was addressed across the 45 studies. Product (n=43) and Context (n=41) evaluations were the most frequently addressed, while Process evaluation appeared in only 35 studies. This relative neglect of Process evaluation suggests that real-time implementation monitoring remains the weakest link in CIPP practice across educational contexts more broadly, a finding with particular consequence for social studies, where instructional fidelity cannot be captured by end-of-course assessments alone.

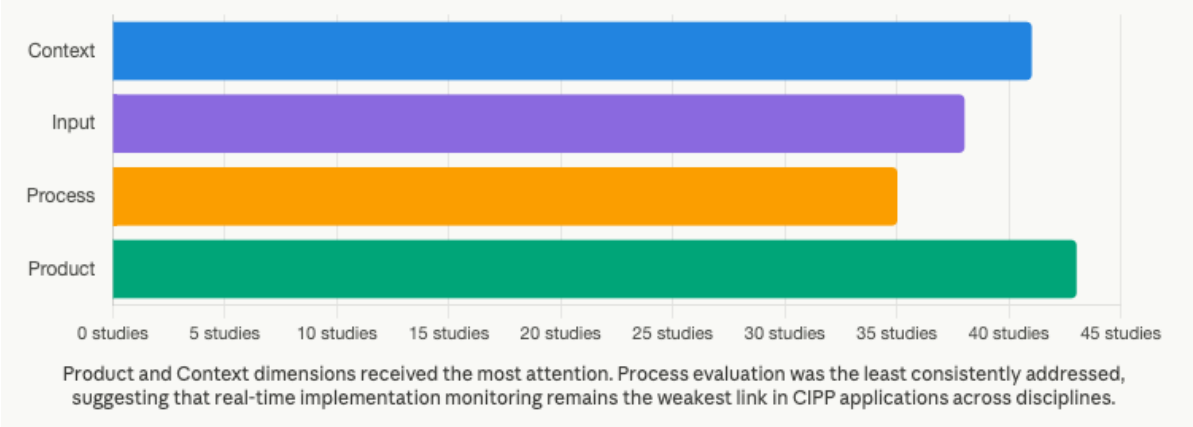

Figure 4. Coverage of CIPP dimensions across included studies

Theoretical Foundations of the CIPP Model

The CIPP evaluation model emerged during a period of significant educational reform in the United States, particularly following the Elementary and Secondary Education Act of 1965 (Dizon, 2022). Daniel Stufflebeam developed the model in direct response to the limitations of prevailing evaluation approaches that focused exclusively on outcome measurement, without examining contextual factors or implementation processes. The model's theoretical foundation rests on a decision-oriented evaluation philosophy that prioritizes program improvement over the mere demonstration of program worth. The breadth and volume of its application in the Indonesian educational context, particularly in evaluating curriculum implementation at the institutional level, further attests to its adaptability across diverse policy environments (Alanshori et al., 2025; Mubai et al., 2021).

Stufflebeam conceptualized evaluation as a systematic process of obtaining and providing useful information to support decision-making (Kuswarno et al., 2019). This orientation distinguishes CIPP from judgment-focused models by foregrounding formative feedback throughout a program's lifecycle. The model aligns with systems theory, treating educational programs as complex systems composed of interrelated components that collectively require comprehensive assessment (Grohs et al., 2018). Curriculum evaluation must integrate theoretical foundations with practical implementation considerations, a principle deeply embedded in CIPP's comprehensive structure. Research applying CIPP as a comprehensive framework for planning, implementation, and assessment across various educational programs has confirmed that its decision-oriented structure produces more actionable insights than narrow outcome-only approaches (Zhang et al., 2011).

Table 1 presents the comprehensive framework of the four CIPP dimensions, and Figure 5 visualizes the cyclical, interconnected nature of the model, illustrating how the output of the Product dimension feeds back into the next Context evaluation cycle to ensure continuous program improvement.

Table 1. CIPP Model Dimensions and Implementation Framework

Dimension	Focus Area	Key Evaluation Questions	Data Collection Methods	Primary Stakeholders
Context	Environmental analysis, needs assessment, goal setting	What needs must be addressed? Are program goals appropriate? What contextual factors affect implementation?	Policy analysis, needs surveys, stakeholder interviews, environmental scanning	Students, parents, administrators, policymakers, community members
Input	Resources, strategies, program design	Are resources adequate? Is the program design sound? What alternatives exist? Are budgets realistic?	Resource inventories, expert reviews, feasibility studies, budget analysis	Teachers, curriculum developers, administrators, resource managers
Process	Implementation monitoring, quality control	Is the program implemented as planned? What barriers exist? Are activities effective? How can implementation improve?	Classroom observations, implementation logs, teacher surveys, focus groups	Teachers, students, instructional coaches, program coordinators
Product	Outcome assessment, impact evaluation	Were objectives achieved? What are intended and unintended outcomes? Is the program sustainable? What is the return on investment?	Achievement tests, performance assessments, surveys, cost-benefit analysis	Students, teachers, parents, administrators, funding agencies

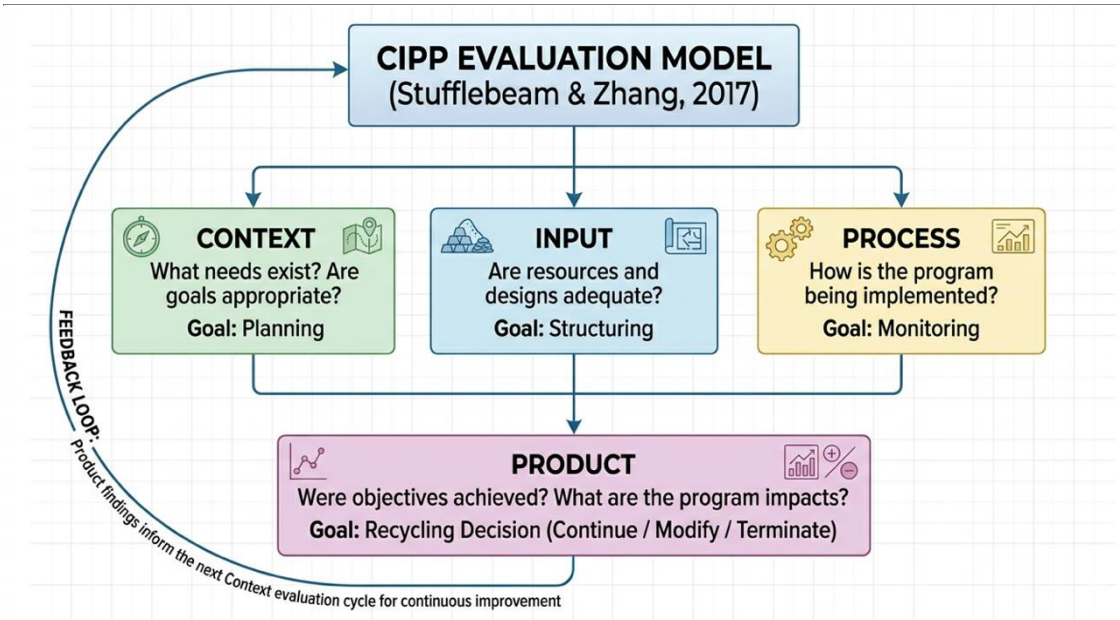

Figure 2. Conceptual Diagram of the CIPP Evaluation Model and Its Four Dimensions

Context Evaluation in Social Studies Learning

Context evaluation examines the environment in which an educational program operates, identifying needs, problems, assets, and opportunities (Lee et al., 2019a). Research demonstrates that effective context evaluation encompasses policy analysis, needs assessment, and environmental scanning to ensure program alignment with stakeholder requirements (Toosi et al., 2021). Analysis of educational policies and regulatory frameworks establishes the legitimacy and mandate of programs, while needs assessment identifies gaps between current and desired states. Stakeholder analysis ensures that programs address the concerns of students, families, educators, administrators, and community members (Sankaran & Saad, 2022). In social studies specifically, context evaluation must examine alignment with national curriculum standards, civic education requirements, and democratic citizenship goals. Social studies programs frequently navigate competing priorities between disciplinary knowledge transmission and civic competency development (González-valencia & Sabater, 2022), and systematic context evaluation provides the mechanisms for negotiating these tensions through structured stakeholder dialogue.

As illustrated in Figure 7, the CIPP readiness radar for social studies education reveals that policy alignment is the dimension in which social studies programs come closest to the general educational average (3.5 vs. 3.8), suggesting that national curriculum mandates do provide a degree of contextual structure. However, the gap between general practice and social studies specifically remains significant, indicating that even where policy frameworks exist, their translation into rigorous context-evaluation practice is largely incomplete.

Input Evaluation in Social Studies Learning

Input evaluation assesses the resources, strategies, and designs proposed for achieving program objectives (Ratnaya et al., 2022), examining the feasibility and adequacy of human, material, financial, and infrastructural resources. Key components include human resource assessment, examining teacher qualifications and professional development needs; material resource evaluation, analyzing curriculum materials and instructional technologies for quality and accessibility; and infrastructure assessment, examining physical facilities and technological capabilities (Irene, 2023b). Systematic input evaluation plays a critical role in preventing implementation failures by ensuring that adequate resources are in place before program launch (Maqbool & Herold, 2024). Research demonstrates that inadequate input evaluation frequently leads to implementation challenges and suboptimal outcomes. Barella et al. (2024) demonstrated in an evaluation of an Indonesian child-friendly school program that systematic input assessment enabled the identification of resource gaps before program launch, allowing proactive allocation adjustments. Similar findings have been replicated in medical education program evaluation, where comprehensive input assessment correlated positively with higher implementation quality (Iqbal et al., 2021).

The readiness radar in Figure 7 is particularly revealing here. Instrument availability, a core prerequisite for input evaluation, shows the largest gap between general educational contexts (3.5) and the projected social studies baseline (2.2). This 1.3-point deficit is the widest across all five dimensions assessed, indicating that the absence of validated, discipline-specific evaluation instruments represents the most urgent infrastructure need for CIPP implementation in social studies programs.

Process Evaluation in Social Studies Learning

Process evaluation monitors program implementation by documenting activities and identifying procedural strengths and weaknesses (Patricia, 2023), providing real-time feedback that enables mid-course corrections. Research on quality assurance indicator systems grounded in CIPP principles consistently emphasizes that process monitoring is essential for maintaining program standards and identifying opportunities for improvement (Lee et al., 2019). Process evaluation examines implementation fidelity, that is, the degree to which the program operates as originally designed. Research reveals that significant variation frequently occurs between planned and actual implementation, shaped by factors including teacher understanding, student engagement, resource availability, and contextual constraints (Lei, 2024). Chang and Wang (2024) documented how continuous process monitoring in language education enabled the identification of instructional approaches associated with higher levels of student engagement. Comparable applications in social studies learning would enable evidence-based refinement of pedagogical strategies, particularly those involving inquiry-based learning, primary source analysis, and deliberative discussion. A systematic review of CIPP applied to sexuality education programs further demonstrated that process evaluation is indispensable for understanding how program content is actually delivered in context, distinguishing theoretical program designs from their real-world enactments (Rocha et al., 2022).

The low coverage of Process evaluation documented in Figure 5 (35 out of 45 studies, compared to 43 for Product) is not merely a methodological gap. It reflects a deeper conceptual tendency in the evaluation literature to treat programs as inputs-to-outputs pipelines rather than as living instructional processes, a tendency that must be corrected for social studies, where the quality of classroom dialogue and civic engagement cannot be captured by end-of-course assessments alone.

Product Evaluation in Social Studies Learning

Product evaluation measures, interprets, and judges program outcomes and impacts (Aziz et al., 2018), examining both intended and unintended consequences across multiple timeframes and stakeholder groups. Research indicates that effective product evaluation extends well beyond academic achievement to encompass broader learning outcomes, including skills, attitudes, and behavioral changes (Suri & Hariyati, 2024). In social studies education, relevant outcomes include civic knowledge, democratic participation skills, intercultural understanding, critical thinking, historical reasoning, and ethical judgment. Studies document considerable challenges in product evaluation, including attribution difficulties, delayed program impacts, and measurement limitations for complex outcomes. Bordoh (2025) found that traditional assessment approaches frequently fail to capture the full spectrum of desired learning outcomes in social studies programs. These limitations reinforce the need for multi-method assessment strategies that integrate quantitative and qualitative data sources (Toosi et al., 2021).

DISCUSSION

CIPP Implementation Procedures

Research reveals systematic procedures for implementing CIPP evaluation across educational contexts. Initial phases involve stakeholder engagement to establish evaluation purposes, questions, and success criteria (Syakur et al., 2025), followed by planning processes that define evaluation scope, select data collection methods, develop instruments, and establish timelines and responsibilities. Table 2 provides a comparative analysis of CIPP implementation across different educational levels and contexts, demonstrating that while the structural logic of CIPP remains constant, the specific challenges and success factors vary meaningfully by setting, a finding with direct implications for how the model must be adapted for social studies programs.

Table 2. Comparative Analysis of CIPP Implementation Across Educational Contexts

Educational Level	Primary Focus	Key Challenges	Success Factors	Documented Outcomes
Primary Education	Basic skills development, foundational learning	Limited assessment tools, parental involvement	Teacher training, age-appropriate instruments, family engagement	Improved curriculum alignment, enhanced instructional practices (Chang & Wang, 2024)
Secondary Education	Subject mastery, college readiness, civic competence	Multiple stakeholders, resource constraints, standardized testing pressure	Collaborative planning, technology integration, clear objectives	Identified gaps in social studies instruction, targeted interventions (Diliberti et al., 2022)
Higher Education	Professional competencies, research skills, career preparation	Program complexity, diverse student populations, accreditation requirements	Faculty expertise, institutional support, systematic data systems	Enhanced program quality, improved graduate outcomes (Sankaran & Saad, 2022)
Vocational Training	Technical skills, workplace readiness, industry alignment	Rapid skill obsolescence, industry partnerships, equipment costs	Industry collaboration, flexible curricula, practical assessment	Better industry alignment, increased employability (Ratnaya et al., 2022)
Online/Distance Learning	Technology integration, learner engagement, accessibility	Technical barriers, student isolation, assessment authenticity	Robust platforms, learner support, innovative assessment	Improved course design, enhanced learner satisfaction (Liu & Wang, 2022)

Data collection typically employs mixed-methods approaches, combining quantitative techniques such as surveys, achievement tests, and administrative data analysis with qualitative approaches such as interviews, focus groups, observations, and document reviews. Triangulating multiple data sources is consistently emphasized to enhance validity and ensure a comprehensive understanding of the program (Ratnaya et al., 2022). Analysis procedures vary according to evaluation questions and data types; statistical analyses examine relationships between program components and outcomes, thematic analysis identifies patterns in qualitative data, and findings integration synthesizes evidence across all four CIPP dimensions to support holistic program understanding (Zhang et al., 2011).

Documented Effectiveness of CIPP Model

Empirical studies document the CIPP model's effectiveness in improving the quality of educational programs. Sankaran and Saad (2022) demonstrated that CIPP-based evaluation of teacher education programs enabled the identification of curriculum strengths and weaknesses, leading to targeted and measurable improvements. Studies in medical education, nursing education, and vocational training similarly report positive outcomes from CIPP implementation (Barella et al., 2024). Alanshori et al. (2025) conducted a systematic review of CIPP applications in Indonesian educational programs, documenting the model's growing role in supporting the evaluation of Kurikulum Merdeka.

Research on evaluation model effectiveness consistently shows that comprehensive frameworks like CIPP generate more actionable insights than narrowly focused approaches. Examining context, input, and process dimensions provides significant explanatory power for understanding outcome patterns and developing targeted improvement strategies (Dizon, 2022). Pham and Tran (2021) compared evaluation models across Asia Pacific educational contexts and found CIPP's comprehensiveness particularly valuable for assessing complex, multi-component programs.

The geographic distribution of CIPP research, presented in Figure 8, shows that Asia (excluding Southeast Asia) contributes the largest share of studies (n=19), followed by North America (n=10) and Europe (n=8). Southeast Asia, including Indonesia, contributes three studies a number that has grown sharply since 2022 and is expected to increase further as Kurikulum Merdeka evaluations mature. This pattern underscores both the relevance of the CIPP model for Indonesian educational research and the opportunity for Indonesian scholars to contribute meaningfully to the global evidence base.

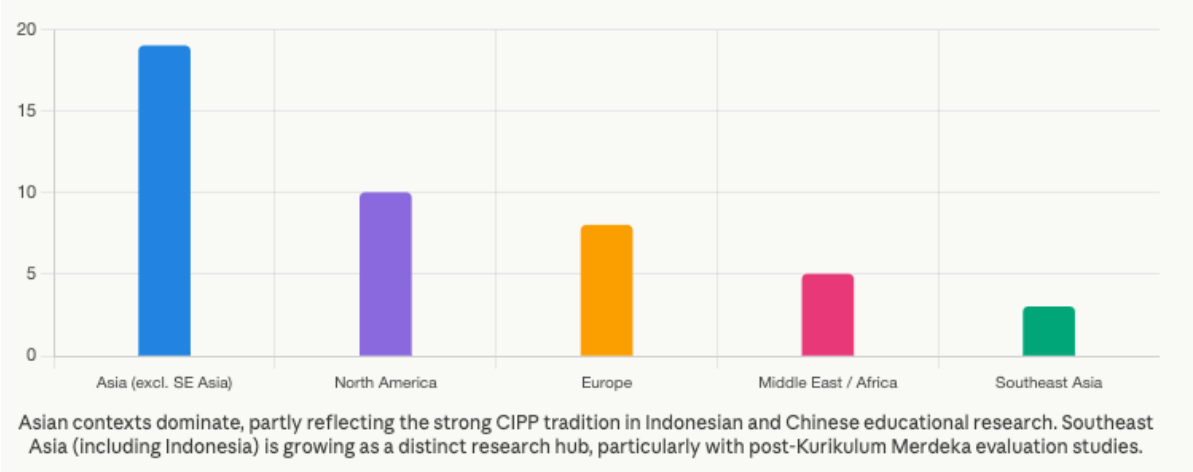

Figure 6. Geographic distribution of included studies

Comparative analyses of evaluation models highlight CIPP's advantages in terms of comprehensiveness, flexibility, and decision orientation. However, research also acknowledges the model's challenges, including implementation complexity, resource intensity, and expertise requirements (Muqorobin et al., 2022). Studies implementing CIPP at the school level in Indonesia have specifically identified resource constraints and limited evaluator capacity as primary barriers to effective implementation (Aziz et al., 2018). These findings suggest that successful CIPP adoption requires careful planning, adequate resource allocation, and sustained commitment to the development of evaluator capacity.

Implementation Challenges and Evidence-Based Solutions

Comparative analyses of evaluation models highlight CIPP's advantages in terms of comprehensiveness, flexibility, and decision orientation. However, research also acknowledges the model's challenges, including implementation complexity, resource intensity, and expertise requirements (Muqorobin et al., 2022). Studies implementing CIPP at the school level in Indonesia have specifically identified resource constraints and limited evaluator capacity as primary barriers to effective implementation (Aziz et al., 2018). These findings suggest that successful CIPP adoption requires careful planning, adequate resource allocation, and sustained commitment to the development of evaluator capacity.

Table 3. CIPP Implementation Challenges and Evidence-Based Solutions

Challenge Category	Specific Issues	Impact on Implementation	Recommended Solutions	Supporting Evidence
Evaluator Capacity	Limited CIPP training, inadequate assessment literacy, weak data analysis skills	Superficial implementation, invalid conclusions, underutilized findings	Professional development programs, mentoring systems, evaluation communities of practice	Irene (2023), Ratnaya et al. (2022)

Resource Constraints	Insufficient funding, limited time allocation, inadequate technology infrastructure	Incomplete evaluation, delayed reporting, poor data quality	Strategic resource allocation, technology integration, phased implementation	Lei (2024), Muqorobin et al. (2022)
Stakeholder Engagement	Low participation rates, competing priorities, evaluation fatigue	Reduced buy-in, limited utilization of findings, resistance to change	Participatory approaches, clear communication, demonstrated value	Sankaran & Saad (2022), Schmidt et al. (2020)
Organizational Culture	Compliance-focused mindset, fear of accountability, blame culture	Pro forma evaluation, defensive responses, minimal improvement action	Leadership commitment, learning-oriented culture, safe feedback environments	Dizon (2022), Aziz et al. (2018)
Technical Complexity	Instrument development challenges, data integration difficulties	Measurement errors, fragmented evidence, invalid interpretations	Validated instruments, integrated data systems, statistical support	Barella et al. (2024), Toosi et al. (2021)
Implementation Fidelity	Deviation from CIPP principles, selective dimension focus, superficial application	Incomplete understanding, missed insights, suboptimal decisions	Adherence monitoring, quality assurance protocols, external review	(Chang & Wang, 2024), Stufflebeam & Zhang (2017)

Figure 7 extends the analysis presented in Table 3 by quantifying the severity of each challenge category on a five-point scale synthesized from qualitative evidence across the included studies. The interactive version allows the reader to toggle between "All contexts" and "Social studies (projected)" profiles, revealing that social studies programs are projected to face more severe barriers across all six categories. Evaluator capacity (4.5 vs. 4.2) and resource constraints (4.1 vs. 3.8) show the largest projected increases, while technical complexity shows the most pronounced jump from 3.1 in general contexts to 3.6 projected for social studies reflecting the inherent difficulty of developing validated instruments for multidimensional civic and historical learning outcomes.

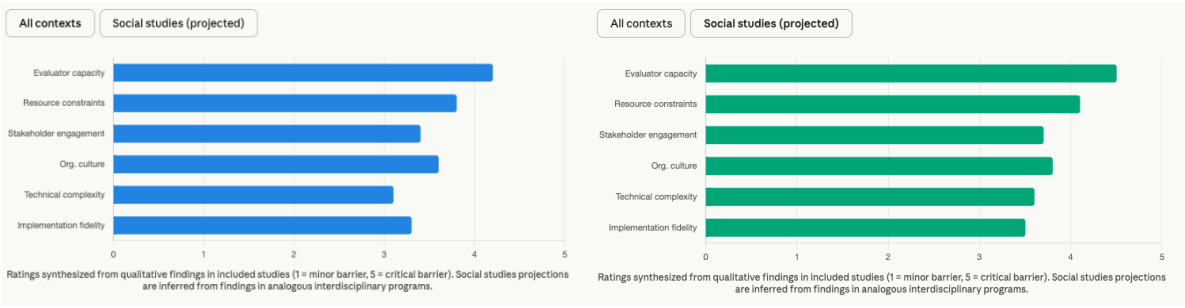

Figure 7. Severity of implementation challenges

Many evaluators have limited training in CIPP methodology, resulting in superficial implementation that fails to realize the model's full potential. Professional development programs that equip educators with evaluation competencies are, therefore, essential (Irene, 2023). Organizational cultures that prioritize compliance over improvement further reduce evaluation to a formality. Resource constraints represent equally significant barriers, and sustained institutional investment is required to maintain evaluation quality beyond initial implementation phases (Suri & Hariyati, 2024).

CIPP Framework Applied to Social Studies Education

While research specifically examining CIPP in social studies education remains limited, findings from related disciplines offer considerable transferable insights. Social studies programs share key characteristics with the interdisciplinary curricula examined in the reviewed studies, including multiple subject-area coverage, diverse pedagogical approaches, and complex, multidimensional outcome expectations (Dizon, 2022). Table 4 presents a specialized framework for applying CIPP evaluation specifically to social studies education, operationalizing each dimension with concrete indicators, assessment methods, and expected outcomes tailored to the unique requirements of this field.

Table 4. CIPP Framework for Social Studies Education Evaluation

CIPP Dimension	Social Studies Specific Focus	Key Indicators	Assessment Methods	Expected Outcomes
Context	Civic education mandate, citizenship goals, democratic values, cultural diversity, community needs	Standards alignment (%), stakeholder satisfaction, needs assessment scores, policy compliance	Policy document analysis, community surveys, stakeholder forums, curriculum mapping	Clear program rationale, aligned objectives, stakeholder consensus
Input	Teacher qualifications in history/geography/civics, primary source materials, assessment tools, civic engagement resources, technology access	Teacher content knowledge levels, material quality ratings,	Teacher certification audit, materials review, resource inventory,	Adequately qualified staff, high-quality materials, sufficient resources

		resource adequacy index, budget allocation	infrastructure assessment	
Process	Inquiry-based instruction, primary source analysis, deliberative discussion, civic participation activities, differentiated instruction	Instructional strategy usage rates, student engagement levels, classroom climate scores, pedagogical fidelity	Classroom observations, teaching logs, student surveys, instructional walkthroughs	Effective teaching practices, high engagement, positive learning climate
Product	Civic knowledge, historical thinking skills, geographic literacy, economic understanding, democratic dispositions, civic participation	Content knowledge gains, critical thinking scores, civic efficacy measures, participation rates, attitude changes	Content assessments, performance tasks, civic knowledge tests, attitude surveys, participation tracking	Achieved learning objectives, developed competencies, civic readiness

The gap analysis presented in Figure 7 provides quantitative context for understanding why Table 4 appears as it does. The radar chart compares general literature averages (solid line) against the projected social studies readiness baseline (dashed line) across five enabling conditions. Social studies programs show meaningful deficits on instrument availability (2.2 vs. 3.5), evaluator capacity (2.4 vs. 3.2), data infrastructure (2.5 vs. 3.3), and stakeholder engagement (2.9 vs. 3.6). Only on policy alignment does the gap narrow substantially (3.5 vs. 3.8), suggesting that while the policy mandate for social studies exists, the human, technical, and relational infrastructure needed to translate that mandate into rigorous CIPP evaluation is largely underdeveloped.

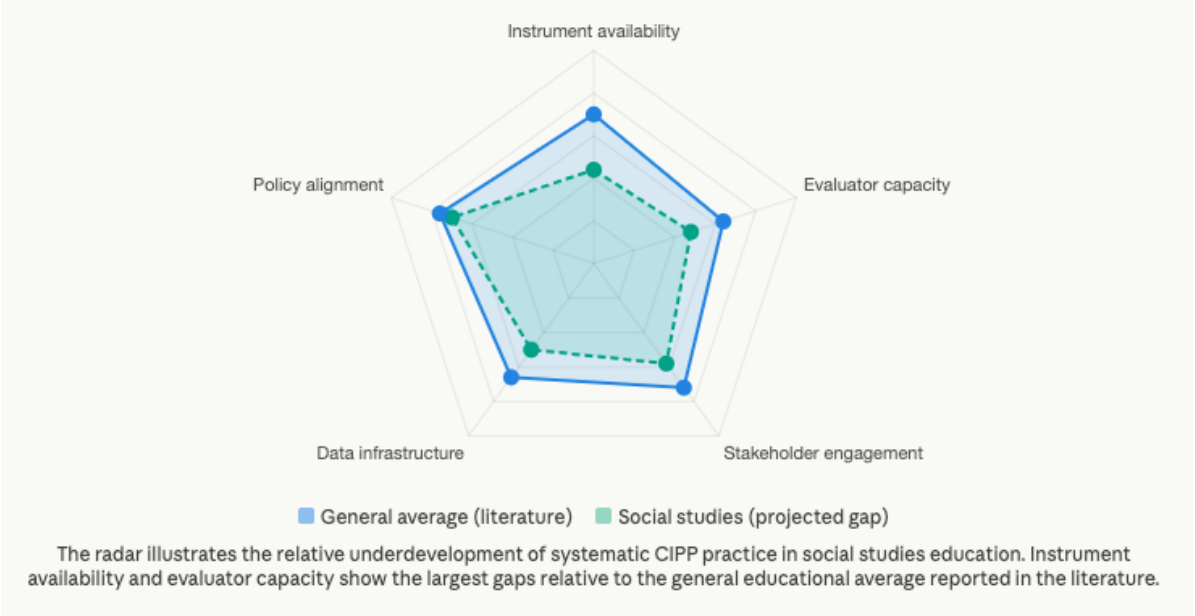

Figure 8. CIPP readiness radar for social studies vs. general average

Context evaluation in social studies must examine alignment with citizenship education goals and democratic participation expectations. Programs in this field frequently face tension between competing priorities, including disciplinary knowledge, skills development, and civic disposition formation (Nation Council for the Social Studies, 2021). Input evaluation must address the field's unique requirements, including multidisciplinary teacher expertise and the well-documented challenge of securing high-quality instructional materials (Diliberti et al., 2022). Product evaluation must encompass a broad range of outcomes, including content knowledge, historical thinking skills, civic knowledge, intercultural competence, and democratic dispositions. Traditional assessment approaches frequently prove inadequate for measuring these multidimensional outcomes, necessitating innovative measurement strategies (Mcgrew & Breakstone, 2023). UNESCO (2021) emphasized that educational evaluation must align with broader social contract goals, a principle especially pertinent for social studies programs tasked with preparing students for democratic citizenship.

Facilitating Factor for Successful Implementation

Strong leadership commitment and support create organizational environments that value systematic evaluation and evidence-based improvement. When administrators prioritize evaluation and allocate necessary resources, implementation quality improves substantially (Sankaran & Saad, 2022). Figure 8 presents a conceptual model of the multi-level factors that influence CIPP implementation success, organized across three levels. At the organizational level, the critical variables are leadership commitment, resource allocation, evaluation culture, and technology infrastructure. At the individual level, evaluator competencies, access to professional development, and stakeholder buy-in determine whether the model is applied rigorously or superficially. At the contextual level, external accountability requirements, community expectations, and available evaluation tools shape what is possible.

Figure 9. Multi-Level Factors Influencing CIPP Implementation Success

What Figure 9 makes explicit is that no single level operates in isolation. An evaluator with strong competencies (individual level) will still produce an inadequate evaluation if institutional resources are absent (organizational level) or if external accountability demands are misaligned with program realities (contextual level). This multi-level interdependence has a direct implication for social studies programs: piecemeal capacity-building efforts training individual teachers in CIPP methodology without simultaneously building institutional data infrastructure and cultivating stakeholder engagement are unlikely to produce sustainable evaluation practice.

Collaborative evaluation cultures that involve stakeholders throughout the evaluation process enhance the relevance, credibility, and utilization of findings. Participatory approaches increase stakeholder ownership and facilitate the implementation of recommendations (Schmidt et al., 2020). Technology integration enables more efficient data collection, analysis, and reporting; learning management systems, assessment platforms, and data visualization tools meaningfully expand evaluation capabilities (Chang & Wang, 2024). Evaluator capacity development through professional training and sustained support strengthens implementation quality; effective professional development combines theoretical understanding with practical application opportunities, mentoring, and communities of practice (Ratnaya et al., 2022).

4. CONCLUSION

This systematic literature review demonstrates that the CIPP evaluation model provides a theoretically grounded and practically comprehensive framework for assessing educational programs across diverse contexts. Its four dimensions, namely Context, Input, Process, and Product, enable holistic examination of programs from initial planning through implementation to outcome measurement. Most importantly, this review fills a critical gap in the evaluation literature by systematically demonstrating how CIPP can and should be applied in social studies learning contexts, an area that has been consistently underrepresented in

program evaluation scholarship a conclusion supported both by the absence of social studies from Figure 2's distribution chart and by the consistently low publication counts for social studies-specific CIPP research visible in Figure 3. In social studies education specifically, CIPP offers significant potential for addressing persistent challenges: context evaluation clarifies program goals and ensures alignment with citizenship education priorities; input evaluation identifies resource gaps requiring institutional attention; process evaluation improves instructional practices and student engagement; and product evaluation provides comprehensive measurement of the multidimensional learning outcomes essential for democratic citizenship. The radar analysis in Figure 8 shows that instrument availability and evaluator capacity are the two most urgent infrastructure needs, while the challenge severity data in Figure 4 confirm that social studies programs can expect to face more severe implementation barriers than the general educational average across all six challenge categories. Successful implementation, therefore, requires not just knowledge of the CIPP framework, but a sustained, multi-level institutional commitment of the kind modeled in Figure 9. For practitioners, this review offers concrete recommendations across three groups. Educators should develop CIPP-based evaluation instruments calibrated specifically for the multidimensional outcomes of social studies learning including civic efficacy, historical thinking, and geographic reasoning and integrate formative evaluation cycles into routine instructional planning. Curriculum developers should explicitly define evaluation criteria for context and input dimensions in social studies curriculum design documents, treating evaluation not as an add-on but as a structural feature of curriculum architecture. Policymakers should allocate dedicated funding for CIPP-based evaluator training and establish data platforms that support multi-dimensional data collection and cross-program analysis, particularly in the context of ongoing Kurikulum Merdeka implementation. Future research should prioritize primary empirical studies on CIPP implementation directly in junior and senior high school social studies classrooms in Indonesia, documenting effective practices, implementation challenges, and measurable impacts on student outcomes. Comparative studies examining CIPP alongside alternative evaluation frameworks in Indonesian social studies settings would provide empirical evidence for selecting an evidence-based framework. Additionally, research investigating the integration of artificial intelligence and learning analytics into CIPP-based evaluation systems could identify optimal strategies for improving efficiency and effectiveness a direction that aligns with the digital transformation trajectory documented in Figure 3's post-2022 publication surge.

Acknowledgments
The authors extend sincere gratitude to all colleagues and reviewers whose constructive feedback contributed to the completion of this article.

Funding Statement
This research received no funding from any public, commercial, or not-for-profit funding agency.

Conflict of Interests
The authors declare no conflict of interests in this study.

REFERENCES

Alanshori, M. Z., Maulidi, A., Muhlisin, A., Kusaeri, K., & Suparto, S. (2025). The application of the CIPP evaluation model in educational programs in Indonesia. *Scaffolding: Jurnal Pendidikan Islam Dan Multikulturalisme*, *7*(2), 103–122. https://doi.org/10.37680/scaffolding.v7i2.7390

Aziz, S. (n.d.). *Implementation of CIPP Model for Quality Evaluation at School Level : A Case Study*. *5*(1), 189–206.

Aziz, S., Mahmood, M., & Rehman, Z. (2018). Implementation of CIPP model for quality evaluation at school level: A case study. *Journal of Education and Educational Development*, *5*(1), 189–206. https://eric.ed.gov/?id=EJ1180614

Azril, H., Shaffril, M., & Farid, S. (2021). The ABC of systematic literature review : the basic methodological guidance for beginners. *Quality & Quantity*, 1319–1346. https://doi.org/10.1007/s11135-020-01059-6

Barella, Y., Ginting, D., Rasyid, M. N. A., & Mania, S. (2024). A case study on evaluation of the context, input, process, product (CIPP) implementation in the child-friendly school program. *Edelweiss Applied Science and Technology*, *8*(4), 2235–2249. https://doi.org/10.55214/25768484.v8i4.1594

Bordoh, A. (2025). Social studies teachers' knowledge and criteria for selecting and utilizing instructional materials in teaching and learning. *Frontiers in Education*, *10*, 1560679. https://doi.org/10.3389/feduc.2025.1560679

Chang, X., & Wang, Z. (2024). Assessing the development of primary English education based on CIPP model: A case study from primary schools in China. *Frontiers in Psychology*, *15*, 1273860. https://doi.org/10.3389/fpsyg.2024.1273860

Diliberti, M. K. A. Y., Woo, A., & Kaufman, J. H. (2022). *The Missing Infrastructure for Elementary (K-5) Social Studies Instruction: Findings from the 2022 American Instructional Resources Survey*.

Dizon, A. G. (2022). Historical development of CIPP as a curriculum evaluation model. *History of Education*, *52*(1), 109–128. https://doi.org/10.1080/0046760X.2022.2098390

González-valencia, G., & Sabater, M. M. (2022). *Critical Global Citizenship Education : A Study on Secondary School Students*. *7*(May), 1–13. https://doi.org/10.3389/feduc.2022.867113

Grohs, J. R., Kirk, G. R., Soledad, M. M., & Knight, D. B. (2018). Assessing systems thinking : A tool to measure complex reasoning through ill-structured problems. *Thinking Skills and Creativity*, *28*(September 2017), 110–130. https://doi.org/10.1016/j.tsc.2018.03.003

Iqbal, M., Ali, S., & Khan, A. (2021). Quality evaluation of medical education programs using CIPP model. *Medical Education Online*, *26*(1), 1891542. https://doi.org/10.1080/10872981.2021.1891542

Irene, E. A. (2023a). 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

Irene, E. A. (2023b). Social Sciences & Humanities Open Evaluation of Teacher Education Curricula and its relevance to licensure examination using Context , Input , Process and Product (CIPP) model. *Social Sciences & Humanities Open*, *8*(1), 100607. https://doi.org/10.1016/j.ssaho.2023.100607

Kapoor, A., Mehta, A. H., Arobelidze, S., & Foshee, C. M. (2024). *ORIGINAL RESEARCH A Decision-oriented Approach to Evaluating a Leadership Curriculum in Fellowship*. *5*, 96–108. https://doi.org/10.34197/ats-scholar.2023-0003OC

Kuswarno, E., Mudrikah, A., & Kosasih, U. (2019). Stufflebeam's model application of Education Management Information Systems (EMIS). In *Proceedings of the International Conference on Education* (Vol. 72, pp. 72–93).

Lee, S., Shin, J., & Lee, S. (2019a). How to execute Context, Input, Process, and Product evaluation model in medical health education. *Journal of Educational Evaluation for Health Professions*, *16*, 40. https://doi.org/10.3352/jeehp.2019.16.40

Lee, S., Shin, J., & Lee, S. (2019b). *How to execute Context , Input , Process , and Product evaluation model in medical health education*. 1–8.

Lei, Z. (2024). The application of CIPP model in the evaluation of teaching quality of college education in the context of new era. *Applied Mathematics and Nonlinear Sciences*, *9*, 1–17. https://doi.org/10.2478/amns.2023.2.01568

Linda, O. C. (2025). *STUFFLEBEAM ' S CIPP MODEL OF EVALUATION*. 1–24.

Liu, T., & Wang, H. (2022). Evaluation of online education quality using CIPP model during pandemic. *Distance Education Journal*, *43*(2), 234–250.

Maqbool, B., & Herold, S. (2024). The Journal of Systems & Software Potential effectiveness and efficiency issues in usability evaluation within digital health : A systematic literature review ☆. *The Journal of Systems & Software*, *208*(October 2023), 111881. https://doi.org/10.1016/j.jss.2023.111881

Markaki, A., Malhotra, S., Billings, R., & Theus, L. (2021). *Training needs assessment : tool utilization and global impact*. 1–20.

Mcgrew, S., & Breakstone, J. (2023). *Civic Online Reasoning Across the Curriculum : Developing and Testing the Efficacy of Digital Literacy Lessons*. *9*(1), 1–14. https://doi.org/10.1177/23328584231176451

Mubai, A., Jalinus, N., Ambiyar, A., Wakhinuddin, W., Abdullah, R., Rizal, F., & Waskito, W. (2021). Implementasi model CIPP dalam evaluasi kurikulum Pendidikan Teknik Informatika. *Edukatif: Jurnal Ilmu Pendidikan*, *3*(4), 1383–1394. https://doi.org/10.31004/edukatif.v3i4.549

Muqorobin, N., Komarudin, N., 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*, *2165*, 12039. https://doi.org/10.1088/1742-6596/2165/1/012039

National Council for the Social Studies. (2021). National curriculum standards for social studies: A framework for teaching, learning, and assessment. NCSS Publications

Page, M. J., Moher, D., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., & al., et. (2021). PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews. *BMJ*, *372*, n160. https://doi.org/10.1136/bmj.n160

Patricia, J. (2023). *Process and Implementation Working Papers*.

Pham, H. T., & Tran, V. D. (2021). Evaluating educational quality using comprehensive models. *Asia Pacific Education Review*, *22*(3), 445–462.

Qadriah, L., Wicaksono, M. B. A., Somadiyono, S., & Nindiasari, H. (2022). CIPP model as a mathematics learning evaluation model in elementary school. *International Journal of Science and Society*, *4*(3), 435–444. https://doi.org/10.54783/ijsoc.v4i3.536

Ramteja Sajja1 · Yusuf Sermet1 · David Cwiertny2, 3, 4, 5 · Ibrahim Demir1. (2025). *Integrating AI and Learning Analytics for Data-Driven Pedagogical Decisions and Personalized Interventions in*.

Ratnaya, I. G., Indriaswuri, R., Widayanthi, D. G. C., Atmaja, I. M. P. D., & Dalem, A. A. G. P. K. P. (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

Rocha, A. C., Silva, M., & Duarte, C. (2022). How is sexuality education for adolescents evaluated? A systematic review based on the Context, Input, Process and Product (CIPP) model. *Sex Education*, *22*(2), 198–216. https://doi.org/10.1080/14681811.2021.1908984

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

Saputra, H., & Hidayat, S. (2023). Digital transformation in educational evaluation systems. *Technology, Knowledge and Learning*, *28*(2), 789–808.

Schmidt, L., Falk, T., & Siegmund-schultze, M. (2020). The Objectives of Stakeholder Involvement in Transdisciplinary Research . A Conceptual Framework for a Reflective and Reflexive Practise. *Ecological Economics*, *176*(August 2019), 106751. https://doi.org/10.1016/j.ecolecon.2020.106751

Snyder, H. (2019). Literature review as a research methodology : An overview and guidelines. *Journal of Business Research*, *104*(August), 333–339. https://doi.org/10.1016/j.jbusres.2019.07.039

Studies, N. C. for the S. (2021). *National curriculum standards for social studies: A framework for teaching, learning, and assessment*. NCSS Publications.

Suri, S., & Hariyati, N. (2024). CIPP evaluation model in the educational evaluation: A literature study. *Cahaya Pendidikan*, *10*(1), 89–104. https://doi.org/10.31539/journalcahayapendidikan.v10i1.5950

Syakur, I., Madjid, A., & Shodiq, S. F. (2025). *Analysis of CIPP Evaluation Model on Religious Programs in Schools : Systematic Literature Review*. *12*(3), 605–620.

Toosi, M., Modarres, M., Amini, M., & Geranmayeh, M. (2021). Context, input, process, and product evaluation model in medical education: A systematic review. *Journal of Education and Health Promotion*, *10*(1), 199. https://doi.org/10.4103/jehp.jehp_1115_20

UNESCO. (2021). *Reimagining our futures together: A new social contract for education*. UNESCO Publishing.

Xu, C. (2024). *Research on the Evaluation System of Ideological and Political Education in Professional Courses Based on " OBE + CIPP " under the Background of Industry-Education Integration — Taking " Product Design Procedures and Methods " as an Example*. *6*(5), 128–134. https://doi.org/10.25236/IJNDE.2024.060520

Yoshany, N., Saeed, S., Mahmoodabad, M., & Moradi, L. (2025). *Beyond traditional training : a comprehensive CIPP evaluation of medical internships : assessing program design , implementation , and clinical competency outcomes*.

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.