

The Evolution of Curriculum Evaluation Research in Schools: A Systematic Literature Review

Niam Wahzudik*

Doctoral Student in Education Management Program, Graduate School, Universitas Negeri Semarang, Semarang, Indonesia

*Corresponding Author: niam_ktp@students.unnes.ac.id

Abstract. Curriculum evaluation plays a pivotal role in educational quality improvement and evidence-informed decision-making in schools. Nevertheless, research in this field remains fragmented across diverse evaluation models, methodological approaches, and educational contexts, limiting a comprehensive understanding of its development. This study systematically examines the evolution of school-based curriculum evaluation research by identifying publication trends, dominant evaluation models, methodological developments, paradigm shifts, major and emerging themes, research gaps, and future directions. A systematic literature review was conducted following the PRISMA 2020 framework. Studies published between 2016 and 2026 were retrieved from Google Scholar. Of the 200 initially identified records, 44 journal articles and conference papers met the inclusion criteria after duplicate removal, title and abstract screening, full-text assessment, and document authentication. The data were analysed using descriptive analysis, VOSviewer-assisted bibliometric mapping, and thematic synthesis. The findings indicate a substantial increase in publications between 2022 and 2026. The CIPP model was the most frequently identified evaluation framework, while descriptive studies, single-case designs, and document analyses remained methodologically prevalent. Major themes included curriculum implementation and effectiveness, evaluation models, inclusive education, and the conceptual development of curriculum evaluation. The field has gradually shifted from objective- and outcome-oriented evaluation towards more contextual, adaptive, participatory, and governance-oriented perspectives. However, multisite, longitudinal, comparative, and participatory studies, as well as technology-supported evaluation, remain limited. Future research should develop adaptive, data-informed, participatory evaluation models that are embedded in school governance and decision-making systems.

Keywords: curriculum evaluation; systematic literature review; CIPP model; inclusive education; school governance

INTRODUCTION

Curriculum evaluation is an essential component of educational quality improvement because it provides systematic evidence concerning the relevance, feasibility, effectiveness, and consequences of a curriculum. It is not limited to determining whether predetermined objectives have been achieved. Rather, it also examines the alignment between curriculum design, student needs, school resources, instructional processes, and educational outcomes. In this broader sense, curriculum evaluation involves the systematic collection, analysis, interpretation, and use of information to judge the quality, value, sustainability, and necessary improvement of a curriculum (Ornstein & Hunkins, 2018; Stufflebeam, 2003).

Curriculum evaluation should be distinguished from student learning assessment. Learning assessment primarily measures students' knowledge, skills, and competencies, whereas curriculum evaluation treats the curriculum itself as the object of inquiry. This object may include curriculum ideas and objectives, content structure, formal documents, implementation processes, learning experiences, outcomes, and broader

effects on students and schools. Accordingly, curriculum evaluation should generate not only evidence of student achievement but also recommendations concerning the continuation, modification, expansion, or termination of curriculum policies and programmes (Fitzpatrick et al., 2011; Ornstein & Hunkins, 2018).

Historically, early approaches to curriculum evaluation were strongly influenced by objective-oriented evaluation. Within this tradition, curriculum success was judged according to the extent to which previously formulated educational objectives had been attained (Tyler, 1949). Although this approach provided a systematic foundation for evaluation, it tended to privilege formal objectives and intended outcomes. Consequently, it could overlook unintended consequences, implementation dynamics, differences across school contexts, and the experiences of stakeholders. Critiques of these limitations stimulated the development of more responsive, goal-free, context-sensitive, and utilization-oriented approaches (Patton, 2008; Scriven, 1967; Stake, 1975).

One of the most influential models in educational evaluation is the Context, Input, Process, and Product (CIPP) model. The model

was developed to provide systematic information for decision-making and programme improvement. Context evaluation identifies needs and problems; input evaluation examines strategies, resources, and programme readiness; process evaluation monitors implementation; and product evaluation assesses results and impacts. Through this structure, evaluation is intended not merely to prove success or failure but to facilitate continuous programme improvement and accountability (Stufflebeam, 2003; Stufflebeam & Shinkfield, 2007).

The comprehensiveness of the CIPP model has made it widely applicable to curriculum and educational programme evaluation. However, using a comprehensive model does not automatically ensure substantive change. In practice, evaluations may end with a description of strengths, weaknesses, indicator attainment, and recommendations. The mechanisms through which findings influence decisions, instructional practices, teacher capacity, and school governance are not always clearly explained. Thus, evaluation quality depends not only on the completeness of the components assessed but also on whether the findings are meaningfully used by relevant stakeholders (Patton, 2008).

Utilization-focused evaluation emphasises that evaluation should be designed with specific primary users and intended uses in mind. Stakeholders should be involved in defining evaluation priorities, questions, evidence, interpretations, and follow-up actions. Evaluation is therefore not a purely technical activity separated from the organisation; it is a process through which stakeholders understand problems, consider alternatives, and make evidence-informed decisions (Patton, 2008). This perspective is highly relevant to schools, where curriculum evaluation findings should inform planning, instruction, resource allocation, teacher support, and student services.

The need for improvement-oriented evaluation is particularly important because curriculum implementation is not determined solely by the quality of curriculum documents. It is shaped by the interaction between curriculum characteristics, school capacity, teacher practices, resources, leadership, organisational culture, and external support. Rogan and Grayson (2003) proposed three constructs for understanding curriculum implementation: the profile of implementation, the capacity to support innovation, and support from outside agencies. Their framework indicates that implementation should be analysed together with

a school's capacity to sustain innovation and the external assistance available to it.

This contextual perspective implies that schools implementing the same curriculum may experience markedly different outcomes. Variations in resources, student characteristics, leadership, community support, and organisational culture can produce different implementation patterns. An evaluation that merely measures compliance with standards may therefore fail to explain why a curriculum succeeds in one school but encounters difficulties in another. Curriculum evaluation should consequently integrate the analysis of documents, implementation processes, organisational capacity, stakeholder experiences, and the broader social context of schooling.

Attention to context is especially important in inclusive education. Inclusion does not simply mean placing students with diverse backgrounds or needs in the same school. It requires education systems to identify and reduce barriers to access, participation, learning, and achievement. Curriculum, pedagogy, assessment, learning environments, and support mechanisms must be responsive to student diversity (UNESCO, 2020). In inclusive schools, curriculum evaluation should therefore assess flexibility, curriculum adaptation, teacher support, access to learning resources, and the involvement of students and parents. Success should be judged not only by average academic attainment but also by participation, access, equity, relevance, and responsiveness to diverse needs.

Digital transformation has further expanded the scope of curriculum evaluation. Learning management systems, digital platforms, online resources, learning analytics, and blended learning generate new sources of evidence, including participation records, engagement patterns, resource use, digital assessment results, and user experiences. However, the availability of data does not necessarily lead to better decisions. Evidence must be interpreted in relation to curriculum aims, student experiences, school conditions, and principles of educational equity. Curriculum evaluation in the digital era therefore requires the integration of quantitative and qualitative evidence as well as clear mechanisms for using data in decision-making.

These developments suggest that curriculum evaluation has gradually moved beyond an exclusive focus on outcome measurement and compliance towards more contextual, participatory, adaptive, and use-oriented

approaches. Responsive evaluation gives priority to stakeholder concerns and context (Stake, 1975); goal-free evaluation seeks to identify actual effects beyond stated objectives (Scriven, 1967); and utilization-focused evaluation treats intended use as a central criterion of quality (Patton, 2008). These developments indicate that evaluation involves not only measurement techniques but also participation, values, context, decisions, and organisational change.

Despite its continuing development, curriculum evaluation research remains fragmented. Some studies apply specific models, particularly CIPP, while others investigate curriculum implementation, programme effectiveness, textbook evaluation, teacher readiness, inclusive education, or policy reform. This diversity enriches the field but also blurs the boundaries among curriculum evaluation, learning assessment, programme evaluation, and institutional evaluation. As a result, a coherent picture of the development of curriculum evaluation research in school settings is still lacking.

Previous reviews have frequently been limited to a particular model, subject area, school level, or policy context. Reviews centred on CIPP, for instance, commonly discuss its components and applications but do not necessarily explain how the orientation of curriculum evaluation research has changed over time. There remains a need for a review that integrates publication trends, dominant models, methodological developments, evaluation objects, stakeholder participation, emerging themes, and the use of findings.

Another gap concerns the relationship between curriculum evaluation and school governance. Evaluation is often positioned as a final stage after a curriculum or programme has been implemented. Such a view separates evaluation from planning, organisation, implementation, monitoring, and decision-making. In contrast, an improvement-oriented perspective treats evaluation as part of a continuous cycle through which schools reassess needs, adapt strategies, improve instruction, develop teacher competence, allocate resources, and justify decisions (Patton, 2008; Stufflebeam, 2003).

The novelty of this review lies in reconstructing the evolution of curriculum evaluation research specifically within school contexts. It not only maps the models used but also examines methodological development, objects of evaluation, stakeholder involvement, the orientation of result utilization, paradigm shifts,

thematic patterns, research gaps, and future directions. The study positions curriculum evaluation as a component of school governance and evidence-informed decision-making rather than merely as an exercise in measuring programme attainment.

Accordingly, this study aims to systematically review the evolution of curriculum evaluation research in schools and to identify publication trends, dominant evaluation models and approaches, methodological developments, paradigm shifts, major and emerging themes, research gaps, and future directions. It addresses six questions: (1) How has school-based curriculum evaluation research developed over time? (2) Which evaluation models and approaches have been used most frequently? (3) How have methodological approaches developed? (4) How has the curriculum evaluation paradigm shifted over time? (5) What major and emerging themes characterise the field? and (6) What research gaps and future directions can be identified?

METHOD

Research Design

This study employed a systematic literature review guided by the PRISMA 2020 framework. The review integrated descriptive analysis, bibliometric mapping, and thematic synthesis to identify the development, structure, and conceptual evolution of curriculum evaluation research in school contexts (Page et al., 2021).

Data Source and Search Strategy

Google Scholar was used as the sole data source because of its broad multidisciplinary coverage and its capacity to retrieve journal articles, conference papers, and regionally published studies that are not always indexed in commercial databases. The search covered publications from 2016 to 2026 in English and Indonesian. Search terms included “curriculum evaluation,” “school curriculum evaluation,” “curriculum implementation evaluation,” “curriculum evaluation model,” “evaluasi kurikulum,” and “evaluasi implementasi kurikulum.”

Eligibility Criteria

Studies were included when curriculum evaluation was a principal or substantial focus; the context involved schools or K–12 education; and the document was a verifiable journal article or

conference paper. Studies were excluded when they focused only on student assessment, classroom learning evaluation, higher education, or when they were books, book chapters, theses, dissertations, reports, or duplicate records.

Study Selection

The initial search identified 200 documents. Two duplicates were removed, leaving 198 records for title and abstract screening. Records that did not address curriculum evaluation, were outside school contexts, or failed to meet the publication-type criteria were excluded. Full-text and source verification was subsequently undertaken through journal websites, DOI records, and official publisher pages. This process identified one book chapter and one dissertation that had initially been classified as articles. Both were excluded, resulting in a final sample of 44 journal articles and conference papers.

Quality Appraisal and Data Extraction

The methodological quality of the included studies was appraised using criteria concerning the clarity of the research aims, appropriateness of the methods, adequacy of the data sources, transparency of analysis, coherence between findings and conclusions, and relevance to curriculum evaluation. Data were extracted into a structured matrix containing author, year, country, school level, research objective, evaluation model, method, participants or data sources, object of evaluation, principal findings, evaluation paradigm, and recommendations.

Data Analysis

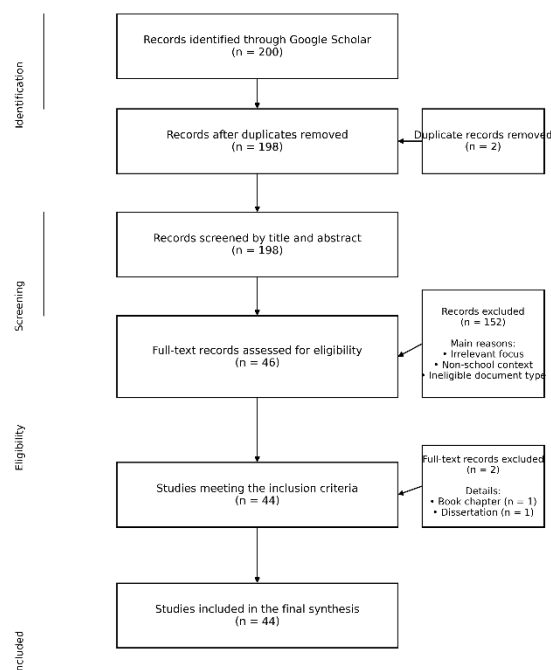
Descriptive analysis was used to summarise publication years, evaluation models, school contexts, research objects, and methodological approaches. Bibliometric analysis was conducted in VOSviewer using term co-occurrence in titles, abstracts, and keywords. A thesaurus file was used to merge synonyms, standardise variant terms, and remove generic or irrelevant expressions. Network, density, and overlay visualisations were generated. Thematic synthesis involved reading each study, assigning codes to its principal findings, grouping codes into themes, and comparing themes across publication periods. The combined analyses were used to identify dominant models, methodological development, paradigm shifts, thematic structures, research gaps, and future directions.

RESULTS AND DISCUSSION

Study Selection

The Google Scholar search yielded 200 documents related to curriculum evaluation. After two duplicate records were removed, 198 documents were screened by title and abstract. Records that did not focus on curriculum evaluation, were outside school contexts, or failed to meet the publication-type criteria were excluded. Full-text assessment identified one book chapter and one dissertation that had previously been classified as articles. These documents were removed because only journal articles and scientific conference papers were eligible. Consequently, 44 studies met all inclusion criteria and were included in the final synthesis, representing 22% of the records initially identified.

The selection process confirms that Google Scholar offers broad coverage but requires rigorous cleaning and authentication. Many records contained terms such as “curriculum,” “evaluation,” “assessment,” or “inclusive education” without substantively examining school curriculum evaluation. Article selection therefore required attention to the actual research focus, educational setting, evaluation object, and publication type rather than keyword occurrence alone.



Note. The flow diagram was developed based on Google Scholar records published between 2016 and 2026. After document authentication, the final dataset comprised 44 journal articles and conference papers.

Figure 1. PRISMA flow diagram of the study selection process

Publication Trends

The distribution of publications indicates a marked increase in school-based curriculum evaluation research in recent years. Two studies were published in each year from 2017 to 2021. The number rose to five in 2022, nine in 2023, six in 2024, nine in 2025, and five in 2026.

Table 1. Publication Trends

Year	Number of studies	Percentage
2017	2	4.55
2018	2	4.55
2019	2	4.55
2020	2	4.55
2021	2	4.55
2022	5	11.36
2023	9	20.45
2024	6	13.64
2025	9	20.45
2026	5	11.36
Total	44	100

Thirty-four of the 44 studies (77.27%) were published between 2022 and 2026, whereas only 10 studies (22.73%) appeared between 2017 and 2021. This pattern suggests a substantial rise in scholarly attention to curriculum evaluation during the most recent five-year period. The increase may be associated with curriculum reform, inclusive education, digital learning, and stronger demands for school accountability. However, the high number of studies in 2025 and 2026 should be interpreted cautiously because recent publications have had less time to accumulate citations. The trend therefore indicates intensified research activity rather than necessarily greater scholarly influence.

Dominant Curriculum Evaluation Models

Coding of titles and abstracts showed that CIPP was the most frequently named evaluation model. Seven studies explicitly used CIPP, one applied the expanded CIPPO model by adding an outcome component, and one used a discrepancy model. In 35 studies, no specific evaluation model was explicitly named in the available metadata.

Table 2. Dominant Curriculum Evaluation Models

Evaluation model	Number	Percentage
CIPP	7	15.91
CIPPO	1	2.27
Discrepancy model	1	2.27
Not explicitly stated	35	79.55
Total	44	100

The prominence of CIPP among explicitly named models indicates a continuing preference for frameworks that connect needs, resources, implementation, and outcomes. Its structure allows researchers to examine curriculum programmes, textbooks, inclusive education, and policy implementation in a systematic manner (Stufflebeam, 2003; Stufflebeam & Shinkfield, 2007). Nevertheless, this dominance also suggests limited diversification of evaluation theory. Responsive evaluation, goal-free evaluation, illuminative evaluation, participatory evaluation, and utilization-focused evaluation were not strongly represented in study metadata.

The large “not explicitly stated” category suggests that many studies were oriented towards practical evaluation rather than the testing or development of evaluation theory. These studies typically assessed implementation, effectiveness, content alignment, teacher readiness, or school responses to curriculum policy. A clearer articulation of the evaluation model is important because it shapes the evaluation questions, evidence sources, criteria of merit, and intended use of findings. Moreover, the use of CIPP did not necessarily mean that evaluation results were integrated into decision-making. In several studies, CIPP components structured the presentation of findings, but the processes through which recommendations influenced subsequent curriculum decisions were not described in detail.

Methodological Development

Curriculum evaluation research in schools employed a range of methodological approaches, including qualitative, quantitative, and mixed methods, as well as evaluation research, case studies, document analysis, literature reviews, and developmental research. This diversity demonstrates that curriculum evaluation is understood not only as outcome measurement but also as a process for examining implementation, context, and the use of findings for educational improvement (Fitzpatrick et al., 2011; Stufflebeam & Shinkfield, 2007).

Table 3. Methodological Development

Approach	Strength	Limitation
Qualitative/case study	Provides an in-depth understanding of implementation and school context	Usually limited to one setting
Quantitative	Produces measurable patterns and enables comparison	Often depends heavily on respondent perceptions
Mixed methods	Combines breadth and	Integration of qualitative

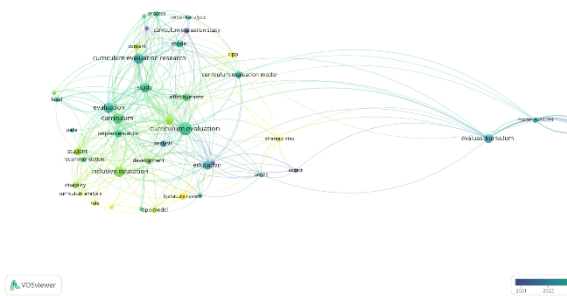


Figure 4. Overlay visualisation of emerging curriculum evaluation research themes

Taken together, the main themes were curriculum implementation and effectiveness, the CIPP model, inclusive education, and the development of curriculum evaluation scholarship. Emerging themes concerned student roles, recommendations, and strategic actions. In contrast, school governance, stakeholder participation, digitalisation, adaptive evaluation, and learning analytics were not yet prominent, indicating important opportunities for further research.

Paradigm Shifts in Curriculum Evaluation

The temporal synthesis indicated that curriculum evaluation research developed through three broad periods. These periods should not be interpreted as mutually exclusive stages; rather, they represent changes in the dominant focus, actors, and uses of evaluation evidence.

During 2017–2021, studies primarily focused on objective attainment, effectiveness, and curriculum implementation. Evaluation was commonly used to examine correspondence among objectives, processes, and outcomes, with CIPP serving as one of the main frameworks (Stufflebeam, 2003; Tyler, 1949). During 2022–2023, attention expanded to school context, student needs, inclusive education, curriculum adaptation, teacher capacity, and variations in school conditions. During 2024–2026, themes increasingly concerned stakeholder roles, recommendations, strategic actions, accountability, and the use of evaluation findings. This development suggests movement towards participatory and governance-oriented evaluation, although empirical application remains limited (Cousins & Whitmore, 1998; Patton, 2008).

Table 4. Paradigm Shifts in Curriculum Evaluation

Period	Dominant paradigm	Primary focus	Main actors	Use of results
2017–2021	Objective-based and implementation-oriented	Objective attainment, effectiveness, context, input, process, and outcomes	Evaluators, teachers, and school managers	Assessing success and identifying weaknesses
2022–2023	Contextual and adaptive	Student needs, inclusive education, adaptation, and school conditions	Teachers, schools, and students	Adjusting curriculum implementation
2024–2026	Participatory and governance-oriented	Actor roles, recommendations, strategic actions, decisions, and accountability	Principals, teachers, students, and other stakeholders	Continuous improvement and decision-making

The literature has moved from an emphasis on measuring programme success towards evaluation that is more contextual, inclusive, participatory, and linked to decision-making. Nevertheless, implementation-oriented evaluation remains dominant, while governance-oriented evaluation is still emerging.

Research Gaps and Future Directions

The bibliometric findings showed that curriculum evaluation research continues to centre on implementation, effectiveness, evaluation models, CIPP, and inclusive education. By contrast, “school governance,” “accountability,” “stakeholder participation,” “adaptive evaluation,” “digital learning,” and “learning analytics” did not form strong nodes in the thematic network. This pattern indicates that curriculum evaluation is still more commonly framed as programme assessment than as an integrated component of school governance and decision-making.

The methodological synthesis likewise showed that descriptive research, single-school or single-programme evaluation, and case-study designs predominated. Multisite, longitudinal, comparative, and strongly integrated mixed-methods research remained limited. As a result, much of the literature provides a cross-sectional picture but does not explain curriculum change, the sustainability of follow-up actions, or the long-term consequences of evaluation.

Stakeholder involvement was also uneven. The terms “student” and “role” emerged in recent publications, but parent, community, and broader stakeholder participation did not become dominant themes. Evaluation therefore appears to

be conducted more often on stakeholders than with them. More participatory designs are needed in which students, teachers, school leaders, parents, and community members contribute to defining criteria, interpreting evidence, and determining follow-up actions.

Table 5. Research Gaps and Future Directions

Type of gap	Evidence base	Current condition	Future research agenda
Theoretical	Network visualisation and thematic synthesis	CIPP and implementation evaluation dominate; adaptive and governance-oriented models remain weak	Develop adaptive, responsive, participatory, and governance-oriented evaluation models
Methodological	Method extraction from 44 studies	Descriptive and single-case designs predominate	Use multisite, longitudinal, comparative, and integrated mixed-methods designs
Participatory	Bibliometric mapping and thematic synthesis	Teachers and managers dominate; students, parents, and communities are less involved	Engage stakeholders from criterion setting through interpretation and follow-up
Contextual	Thematic synthesis	Inclusive education is growing, but marginal, remote, and resource-constrained schools remain underrepresented	Extend research across diverse and resource-limited school contexts
Technological	Network and density visualisations	Digital learning, learning analytics, and AI are not strong themes	Develop ethical and context-sensitive data-informed curriculum evaluation
Governance	Network visualisation and thematic synthesis	Evaluation is not strongly integrated with school decisions and accountability	Connect evaluation with planning, resource allocation, accountability, and continuous improvement

Future research should therefore move from episodic and descriptive evaluation towards continuous, participatory, data-informed evaluation embedded in school governance. Such an agenda requires cross-context and longitudinal testing so that evaluation can demonstrate not only what recommendations are produced but also whether they lead to meaningful changes in curriculum practice. Digital tools, learning analytics, and artificial intelligence may support this development, but their use must be guided by ethical principles, contextual relevance, accessibility, and educational equity.

CONCLUSION

This systematic review shows that school curriculum evaluation research has developed from objective- and outcome-oriented approaches towards more implementation-focused, contextual, inclusive, participatory, and governance-oriented perspectives. CIPP remains the most prominent explicitly named framework because it enables systematic analysis of context, input, process, and product. However, many studies have not clearly linked evaluation findings to decision-making or sustained curriculum improvement.

Methodologically, the literature continues to rely heavily on descriptive research, single-case studies, document analysis, and programme evaluation, while multisite, longitudinal, comparative, participatory, and strongly integrated mixed-methods designs remain limited. Bibliometric mapping shows that curriculum implementation, effectiveness, evaluation models, and inclusive education are central themes. School governance, accountability, stakeholder participation, adaptive evaluation, and technology-supported evaluation have not yet developed into strong thematic areas.

The review therefore highlights the need for adaptive, participatory, data-informed curriculum evaluation models that are integrated with school decision-making systems. Theoretically, the study broadens the understanding of curriculum evaluation as a component of school governance and organisational learning. Practically, it provides a foundation for schools and policymakers to design evaluation processes that are more responsive, continuous, equitable, and relevant to diverse student needs.

ACKNOWLEDGEMENT

The author gratefully acknowledges Universitas Negeri Semarang for the academic and institutional support provided during the preparation of this study. The author also appreciates the Postgraduate School of Universitas Negeri Semarang for providing an intellectually supportive environment, academic guidance, and opportunities for scholarly development during the author's doctoral studies. This support contributed significantly to the completion of the systematic review.

REFERENCES

- Bowen, G. A. (2009). Document analysis as a qualitative research method. *Qualitative Research Journal*, 9(2), 27–40. <https://doi.org/10.3316/QRJ0902027>
- Cousins, J. B., & Whitmore, E. (1998). Framing participatory evaluation. *New Directions for Evaluation*, 1998(80), 5–23. <https://doi.org/10.1002/ev.1114>
- Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). SAGE Publications.
- Creswell, J. W., & Plano Clark, V. L. (2018). *Designing and conducting mixed methods research* (3rd ed.). SAGE Publications.
- Creswell, J. W., & Poth, C. N. (2018). *Qualitative inquiry and research design: Choosing among five approaches* (4th ed.). SAGE Publications.
- Fitzpatrick, J. L., Sanders, J. R., & Worthen, B. R. (2011). *Program evaluation: Alternative approaches and practical guidelines* (4th ed.). Pearson.
- Ornstein, A. C., & Hunkins, F. P. (2018). *Curriculum: Foundations, principles, and issues* (7th ed.). Pearson.
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., et al. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, 372, n71. <https://doi.org/10.1136/bmj.n71>
- Patton, M. Q. (2008). *Utilization-focused evaluation* (4th ed.). SAGE Publications.
- Rogan, J. M., & Grayson, D. J. (2003). Towards a theory of curriculum implementation with particular reference to science education in developing countries. *International Journal of Science Education*, 25(10), 1171–1204. <https://doi.org/10.1080/09500690210145819>
- Scriven, M. (1967). The methodology of evaluation. In R. W. Tyler, R. M. Gagné, & M. Scriven (Eds.), *Perspectives of curriculum evaluation* (pp. 39–83). Rand McNally.
- Stake, R. E. (1975). *Evaluating the arts in education: A responsive approach*. Merrill.
- Stufflebeam, D. L. (2003). The CIPP model for evaluation. In T. Kellaghan & D. L. Stufflebeam (Eds.), *International handbook of educational evaluation* (pp. 31–62). Kluwer Academic Publishers. https://doi.org/10.1007/978-94-010-0309-4_4
- Stufflebeam, D. L., & Shinkfield, A. J. (2007). *Evaluation theory, models, and applications*. Jossey-Bass.
- Tyler, R. W. (1949). *Basic principles of curriculum and instruction*. University of Chicago Press.
- UNESCO. (2020). *Global education monitoring report 2020: Inclusion and education—All means all*. UNESCO.
- Van Eck, N. J., & Waltman, L. (2010). Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics*, 84(2), 523–538. <https://doi.org/10.1007/s11192-009-0146-3>
- Yin, R. K. (2018). *Case study research and applications: Design and methods* (6th ed.). SAGE Publications.
- Adnan, M. (2017). Evaluasi kurikulum sebagai kerangka acuan pengembangan pendidikan Islam. *Al-Idaroh: Jurnal Studi Manajemen Pendidikan Islam*.
- Agus, A., Juliadharma, M., et al. (2023). Application of the CIPP model in evaluation of the inclusive education curriculum in Madrasah Aliyah. *Nidhomul Haq: Jurnal Manajemen Pendidikan Islam*.
- Aini, K. A. N., & Inayah, I. (2026). CIPP's model of curriculum evaluation focused on the Arabic reading skills textbook KSKK 2019. *Aphorisme: Journal of Arabic Language, Literature, and Education*.
- Almalky, H. A., & Alqahtani, S. M. (2025). Teachers' evaluation of vocational curriculum for secondary students with intellectual disabilities: An exploratory analysis from Saudi Arabia. *Research in Developmental Disabilities*.
- Ananda, R., Yusadi, A., & Zein, W. L. (2024). Tujuan, prosedur, dan model evaluasi kurikulum: Perspektif kajian sistematis. *PRODU: Prokurasi Edukasi Jurnal Manajemen Pendidikan Islam*.
- Anis, M. (2023). Assessing the effectiveness of a curriculum in education: A review of evaluation models. *Galaxy: International Multidisciplinary Research Journal*.
- Aristya, S. (2026). The CIPP model for curriculum evaluation in Islamic education: A comprehensive literature review. *Al-Hayat: Journal of Islamic Education*.
- Aybek, B., & Oğuz, O. (2025). Sustainable and inclusive education reform in Türkiye: A CIPP evaluation of the primary Turkish language curriculum. *Sustainability*, 17(19), 8659.

- Chen, M. J., Guo, C. Y., & Tseng, L. T. (2023). Issues and trends in international curriculum evaluation research from 2000 to 2020. *Journal of Educational Research and Development*.
- Fauziah, N., Munsyifah, A., et al. (2021). Evaluasi kurikulum pendidikan inklusi di Madrasah Aliyah Negeri 2 Sleman Yogyakarta. *At-Thullab: Jurnal Mahasiswa Studi Islam*.
- Gifari, M. K., & Arifin, Z. (2024). Literature review: Evaluation of the Kurikulum Merdeka using the CIPP model. *Inovasi Kurikulum*.
- Gündüz, N. (2023). Content analysis of curriculum evaluation model studies during 2000–2020. *Near East University Online Journal of Education*.
- Hafizh, M., Hanif, S., & Nurhaliza, S. (2024). Diskursus evaluasi kurikulum dalam pendidikan formal. *Geneologi PAI: Jurnal Pendidikan Agama Islam*.
- Ibrahim, I., Zainuri, A., Yunita, Y., et al. (2023). Evaluasi kurikulum pendidikan: Memastikan proses pembelajaran sesuai dengan tujuan pendidikan. *El-Idare: Journal of Islamic Education Management*.
- Indrianto, N., & Nurdin, A. (2024). Evaluation of primary school inclusive education curriculum based on context, input, process, and product (CIPP) model. *Al-Hayat: Journal of Islamic Education*.
- Ipek, D. S., Gezer, E. T., Tunçer, F. K., Baltacı, H. S., & Ok, A. (2023). A systematic meta-evaluation of curriculum evaluation research studies conducted between 2004–2022 on K–12 school curricula in Türkiye. *International Journal of Curriculum and Instruction*.
- Khalik, Z. (2025). Integration of the paradigm of inclusivity in the 21st-century competency-based PAI curriculum evaluation model. *Al-Insan: Islamic and Humanities Perspectives Journal*.
- Lestari, B. D., Samta, S. R., Nisak, H., & Rahayu, S. S. (2022). Kurikulum pendidikan inklusi di masa pandemi ditinjau dari evaluasi program pembelajaran. *Sentra Cendekia*.
- Lestari, B. D., Sutarto, J., & Raharjo, T. J. (2019). The evaluation of inclusive education curriculum program in PAUD Talenta Semarang. *Journal of Primary Education*.
- Mihajlovic, C. (2019). Towards inclusive education? An analysis of the current physical education curriculum in Finland from the perspective of ableism. *International Journal of Physical Education*. <https://doi.org/10.5771/2747-6073-2019-4-29>
- Mustika, D., Assyifa, A., et al. (2025). Evaluasi kurikulum sekolah dasar dalam mendukung implementasi pendidikan inklusi. *Menulis: Jurnal Penelitian Nusantara*.
- Nada, A. Q., & Prastowo, A. (2023). Implementasi evaluasi kurikulum sekolah: Studi kasus sekolah dasar di Kota Semarang. *MODELING: Jurnal Program Studi PGMI*.
- Nadrah, N. (2023). Model evaluasi pendidikan inklusif. *Innovative: Journal of Social Science Research*.
- Nasarudin, N., & Syafii, A. H. (2022). Evaluasi kurikulum madrasah inklusi pada era kenormalan baru. *INKLUSI: Journal of Disability Studies*.
- Nazir, S., & Fatima, G. (2026). Transforming primary education for hearing-impaired learners: A curriculum review grounded in social justice and sustainable development. *Annals of Human and Social Sciences*.
- Nurhamidin, B., Saleh, H., et al. (2025). Inclusive language and social representation of people with disabilities in the educational curriculum: A literature review. *Sosioedukasi: Jurnal Ilmiah Ilmu Pendidikan dan Sosial*.
- Nurhayati, F., Khotimah, I., Eviyani, A. H., et al. (2026). Integrasi analitik pendidikan dan Q-methodology dalam pengembangan model evaluasi kurikulum multidimensi. *Dirasah: Jurnal Studi Ilmu dan Manajemen Pendidikan Islam*.
- Özüdoğru, F. (2018). Analysis of curriculum evaluation studies conducted in foreign language education: 2005–2016. *Journal of Language and Linguistic Studies*.
- Pratama, M. I. L. (2024). English curriculum analysis for mental retardation students. *Journal of Development and Innovation in Language and Literature Education*.
- Rasid, S. A., Ma'ruf, J. A., Shoha, S. I., et al. (2026). The effectiveness of Independent Curriculum on learning interest and character development of students with special needs: CIPPO analysis. *Cendikia Journal of Multidisciplinary Studies*.
- Sari, A. V., Pariyasto, S., Simamora, W., et al. (2025). Evaluasi kurikulum. *Qomaruna: Journal of Multidisciplinary Studies*.
- Sastrawati, E. (2025). Efektivitas kurikulum pada siswa berkebutuhan khusus. *Pendas: Jurnal Ilmiah Pendidikan Dasar*.
- Sholeh, M. I. (2024). The effectiveness of inclusion curriculum in meeting the needs of diverse students. *Journal of Asian Primary Education*.

- Sianturi, E. R., Simangunsong, F. A., et al. (2022). Pengawasan dan evaluasi kurikulum. *Pediaqu: Jurnal Pendidikan Sosial dan Humaniora*.
- Subando, J., Kartawagiran, B., et al. (2021). Development of curriculum evaluation model as a foundation in strengthening the ideology of Al-Irsyad education. *Journal of Research and Educational Research Evaluation*.
- Süer, S. (2022). A content analysis of English curriculum evaluation studies in Turkey between 2005–2021. *E-Kafkas Journal of Educational Research*.
- Sumarni, S. (2017). Evaluasi implementasi Kurikulum 2013 di madrasah. *Edukasi: Jurnal Penelitian Pendidikan Agama dan Keagamaan*.
- Ummah, A. (2025). Evaluasi dan tindak lanjut perbaikan kurikulum di sekolah berbasis inklusi. *EDUSAINTEK: Jurnal Pendidikan, Sains dan Teknologi*.
- Wahzudik, N., & Istyarini, W. (2020). Implementasi evaluasi kurikulum sekolah: Studi kasus sekolah dasar di Kabupaten Kudus. *Teknodika*.
- Yazçayır, N., & Selvi, K. (2020). Curriculum evaluation model: KÖNDEM. *Elementary Education Online*.
- Yüksel, D. K. (2023). Evaluation of life studies curriculum by Stufflebeam's context, input, process, and product evaluation model. *International Journal of Education, Technology and Science*.
- Yusuf, S., & Andriani, T. (2025). Meta-analisis model evaluasi kurikulum CIPP, Stake, dan Scriven untuk peningkatan kualitas pembelajaran: Tinjauan sistematis. *Jurnal Tawadhu*.
- Yuwono, I. (2018). Evaluation of curriculum implementation and curriculum modifications in inclusive schools. In *Proceedings of the 1st International Conference on Creativity, Innovation and Technology in Education*. Atlantis Press.