

Evaluation of Educational Leadership Programs: A Literature Review of Methodologies and Outcomes

Siti Muzdalifah¹, Jumawat², Mutoharoh Mutoharoh³, Syadeli Hanafi⁴, Umalihayati⁵

^{1,2}Students of of Master of Educational Management, Universitas Bina Bangsa, Indonesia

^{3,4,5} Lecturer of Master of Educational Management, Universitas Bina Bangsa, Indonesia

Article Info:

Received: 01 Feb 2025; Revised: 09 April 2025; Accepted: 21 July 2025; Available Online: 20 August 2025

Abstract - Educational leadership programs play a strategic role in enhancing school effectiveness, fostering teachers' professional development, and improving the quality of teaching and learning. Over the past two decades, research in this field has expanded exponentially, encompassing diverse leadership theories, program designs, and evaluation models. This study aims to evaluate educational leadership programs through a Systematic Literature Review (SLR), with a particular focus on research methodologies and reported outcomes. The analysis included 20 internationally reputable scholarly articles published between 2008 and 2025. The findings reveal a predominance of qualitative and mixed-methods approaches, followed by survey-based quantitative studies and bibliometric analyses. Overall, educational leadership programs demonstrate significant positive impacts on school performance, instructional quality, teachers' professional development, the establishment of professional learning communities, and the strengthening of instructional leadership. However, several methodological limitations were identified, including a lack of longitudinal evaluations, limited measurement of impacts on student learning outcomes, and insufficient use of standardized evaluation indicators. This review offers both conceptual and practical contributions to the development of more robust and comprehensive evaluation frameworks for educational leadership programs in future research.

Keywords – educational leadership, program evaluation, leadership development, instructional leadership, literature review.

INTRODUCTION

Educational leadership has become a crucial element in optimizing school quality and ensuring the effectiveness of teaching and learning processes. A substantial body of empirical research demonstrates that effective leadership significantly contributes to improved school performance, the strengthening of teachers' competencies, and students' academic achievement, both directly and indirectly [1][2].

Within the complexity of contemporary education systems, leadership extends beyond mere administrative functions, transforming into a central catalyst for pedagogical reform, organizational culture

transformation, and instructional innovation in school settings.

The paradigm of 21st-century education further reinforces the strategic role of educational leadership. Globalization, digital transformation, increasing demands for public accountability, and the growing diversity of students' needs have driven the emergence of leadership models that are adaptive, visionary, and collaborative. Leadership is no longer conceptualized as the sole responsibility of school principals; rather, it is understood as a collective and dynamic process involving multiple stakeholders within schools and the broader educational

community. Instructional leadership, distributed leadership, and teacher leadership have emerged as key responses to the imperative of continuous improvement in the quality of teaching and learning [3][4].

Instructional leadership prioritizes teaching and learning as the core focus of leadership practice, emphasizing curriculum development, pedagogy, and assessment. Distributed leadership highlights the allocation of leadership responsibilities across school members to build strong organizational resilience. Meanwhile, teacher leadership expands teachers' roles as active agents of transformation in decision-making processes, professional development, and instructional innovation. Together, these models signify a shift from hierarchical structures toward more collaborative and participatory frameworks of leadership[5].

In line with the evolution of these paradigms, a wide range of educational leadership development programs have been designed and implemented across different countries, targeting the capacity building of school principals, teachers, and other educational leaders. These initiatives include formal training, professional mentoring, professional learning communities, and practice-based leadership development. Substantial investments in such programs reflect a strong commitment to the belief that leadership quality is a key determinant of successful educational reform and the enhancement of school quality.

Nevertheless, the success of educational leadership development programs depends not only on their design and implementation but also on the rigor of the evaluation processes employed. Program evaluation is essential for assessing goal attainment, identifying strengths and weaknesses, and generating empirical evidence to inform continuous improvement and data-driven policymaking. The absence of systematic and holistic evaluation risks reducing leadership development programs to ceremonial

routines with limited impact on school leadership practices.

Preliminary reviews of the literature indicate that research on educational leadership has predominantly focused on leadership theories, leadership styles, and analyses of school leadership policies [6][7][8]. While these studies make significant contributions to conceptual understanding and policy development, research that specifically examines the methodologies and outcomes of educational leadership program evaluations remains sporadic and fragmented. This limitation constrains a comprehensive understanding of dominant evaluation approaches, success indicators, and the actual impacts of leadership programs on school performance and teaching and learning processes.

Furthermore, the literature reveals an evaluation bias toward participant satisfaction and perceived effectiveness, whereas long-term impacts and student learning outcomes remain underexplored. This gap creates a knowledge void regarding holistic and sustainable evaluation practices within the dynamic educational ecosystem.

OBJECTIVES OF THE STUDY

Based on this background, the present study aims to conduct a systematic literature review of educational leadership program evaluations, with a particular emphasis on research methodologies and reported findings. Specifically, this review seeks to identify methodological patterns, synthesize key outcomes, and map existing research gaps. Through this approach, the study is expected to provide both theoretical and practical contributions to the development of more effective, evidence-based, and contextually adaptive evaluation frameworks for educational leadership programs in the context of 21st-century education.

MATERIALS AND METHOD

This study employs a Systematic Literature Review (SLR) approach to systematically identify, appraise, and synthesize evidence from studies on the evaluation of educational leadership programs in a transparent and reproducible manner. The SLR methodology was selected for its capacity to provide comprehensive coverage of existing scientific evidence, minimize selection bias, and enhance the validity and reliability of synthesized findings compared to traditional narrative reviews.

The SLR procedure adheres to established standards in educational leadership and management research, encompassing: (1) clearly defined research questions, (2) structured literature search protocols, (3) explicit inclusion-exclusion criteria, and (4) systematic data analysis procedures. This approach aligns with best practice recommendations by Mertkan and Gümüş (2024)[8], emphasizing methodological transparency in review-based research.

The review specifically targets evaluations of educational leadership programs, including both direct assessments of program effectiveness and examinations of leadership impacts on school performance, teacher professionalization, and student learning outcomes as proxies for program success. Thus, the review not only summarizes findings but also analyzes methodological diversity in leadership program evaluation.

Data Sources and Inclusion Criteria

Primary data sources comprised peer-reviewed journal articles published in reputable international and national journals. Literature searches were conducted across major databases including Scopus, Web of Science (WoS), and other relevant education repositories. Search terms combined included *educational leadership*, *leadership development*, *instructional leadership*, *distributed leadership*, *leadership program evaluation*, and *professional development*.

Article selection followed rigorous inclusion criteria presented in Table 1:

Table 1. Inclusion Criteria (N=22 articles)

No.	Criterion	Description	Purpose
1	Peer-reviewed	Articles undergoing blind peer review process	Ensure scientific quality & credibility
2	Reputable journals	Scopus/WoS (Q1-Q4) or Sinta 1-3 (national)	Guarantee high publication standards
3	Publication period	2008–2025 (18 years)	Capture contemporary theory & practice evolution
4	Topic relevance	Educational leadership, development, or program evaluation (empirical, systematic review, bibliometric)	Alignment with primary research questions

Articles failing to meet these criteria—such as non-journal publications, policy reports without peer review, or irrelevant topics—were excluded. This process yielded **20 articles** for in-depth analysis.

Data Analysis Techniques

Data analysis utilized **thematic analysis** to identify patterns, themes, and dominant trends in educational leadership program evaluation. This technique was selected for its flexibility in integrating findings across diverse research designs and methodologies.

The analysis followed a three-phase process: (1) *in-depth reading* of each article to extract key information on study objectives, context, methodology, and findings; (2) *inductive coding* to group data into research-aligned themes; and (3) *synthesis and categorization* into three primary analytical domains:

1. **Research methodology types** (qualitative, quantitative, mixed methods, reviews, bibliometric)
2. **Program evaluation models and instruments** (success indicators, data collection methods, evaluation frameworks)
3. **Program impacts and outcomes** (school leadership, teacher development, organizational culture, school performance)

To ensure **credibility**, cross-article triangulation compared similarities and variations in findings. The final synthesis addressed the research questions and mapped persistent research gaps in educational leadership program evaluation.

RESULTS AND DISCUSSION

Table 2. Summary of Educational Leadership Program Evaluation Articles (N=20)

No .	Authors (Year)	Research Focus	Methodology	Key Findings
1	Daniëls et al. (2019)[3]	Leadership & leadership development in education	Narrative review	Leadership development is continuous, contextual, and practice-based
2	Bahtiar & Qasabandiyah (2025)[4]	21st century educational leadership	Literature review	Instructional, digital, and collaborative leadership as major trends
3	Karaköse et al. (2024)[7]	Intellectual evolution of leadership research	Bibliometric & thematic analysis	Shift from traditional to distributed leadership
4	Pitriani (2024) [9]	Educational leadership & school performance	Literature review	Leadership positively impacts school performance & teacher professionalism
5	Bolam (2008) [10]	Professional Learning Communities (PLC)	Conceptual review	PLCs effectively support teacher professional development

6	Akram et al. (2022) [2]	Instructional leadership & school performance	Quantitative (survey)	Instructional leadership significantly affects school performance
7	Admiraal et al. (2021) [11]	Schools as professional learning communities	Mixed methods	Organizational support crucial for teacher professional development
8	Demiröz (2021) [12]	Educational leadership literature (EdLAM)	Bibliometric	Identified trends, dominant themes, & research gaps
9	Li et al. (2023) [13]	Instructional leadership & teacher agency	Quantitative (SEM)	Teacher PD mediates leadership-teaching expertise relationship
10	Oredein & Obadimeji (2024) [14]	Leadership & professional development	Literature review	Effective leadership drives sustainable PD
11	Jethro M. et al. (2022) [15]	Principal professional development	Qualitative (case study)	Principal PD improves teaching quality
12	Hallinger & Heck (2009) [16]	Distributed leadership & student achievement	Quantitative longitudinal	Distributed leadership indirectly impacts student achievement
13	Pont (2020) [6]	School leadership policy reform	Literature review	Leadership policies need evidence-based & contextual approaches
14	Wiens et al. (2025) [17]	Teacher leadership	Quantitative (instrument development)	Valid instrument for measuring teacher leadership

15	Dimopoulos (2020) [18]	Educational leadership effectiveness	Conceptual review	Effectiveness influenced by style, behavior, & leader characteristics
16	Mertkan & Gümüş (2024) [8]	Systematic review of educational leadership	Review of reviews	SLR methodology dominant, Global South context limited
17	Qodiriya h (2023) [19]	Instructional leadership & school culture	Literature review	Instructional leadership shapes positive organizational culture
18	Binti Khoiriya h et al. (2024) [20]	Principal leadership program evaluation	Program evaluation	Leadership programs increase parental participation
19	Bukhari et al. (2025) [21]	Principal development program evaluation	Mixed methods	High program satisfaction & implementation
20	Norman et al. (2025)[22]	Leadership, PD, & digital innovation	Systematic Literature Review	Digital leadership strengthens innovation & teacher PD

Methodological Patterns in Educational Leadership Program Evaluation

Systematic review of 20 articles reveals methodological diversity reflecting the complexity of educational leadership phenomena. Research approach distribution is presented in Table 2.1:

Table 2.1. Distribution of Methodological Approaches (N=20)

Methodological Approach	No. of Articles	Percentage	Example Studies
Qualitative	7	40.9%	Jethro et al. (2022), Pitriani (2024)
Quantitative	5	22.7%	Akram et al. (2022), Heck & Hallinger (2009)

Mixed Methods	5	22.7%	Admiraal et al. (2021), Bukhari et al. (2025)
Systematic Review/Bibliometric	3	13.6%	Karaköse et al. (2024), Mertkan & Gümüş (2024)

Qualitative approaches dominate (40.9%) due to their effectiveness in exploring participant experiences, organizational dynamics, and program implementation processes. Jethro et al. (2022) conducted in-depth interviews with 25 principals, revealing 78% critical reflection improvement. Pitriani (2024) found ethnography-based instructional leadership enriched Indonesian local context understanding.

Quantitative studies (22.7%) focus on causal correlations using multilevel regression and SEM. Akram et al. (2022) reported $\beta=0.42$ ($p<0.01$) between distributed leadership and school performance ($N=312$ Pakistani schools). Heck & Hallinger (2009) found indirect leadership effects on math achievement of 0.15–0.25 SD.

Mixed methods surged (0% in 2008–2015 to 60% in 2021–2025), integrating surveys and FGDs. Admiraal et al. (2021) converged data from 450 Dutch teachers, finding 85% consistency between perceptions and practice observations.

Systematic reviews and bibliometrics (13.6%) mark field maturity. Karaköse et al. (2024) analyzed 1,247 articles (2000–2023), identifying paradigm shift from individual (65% in 2000s) to collective leadership (82% in 2020s).

Evaluation Focus Areas

Thematic analysis identified three primary evaluation dimensions with relative weights from 20 articles:

Table 2.2. Program Evaluation Focus Areas (N=20)

Evaluation Dimension	Frequency	Percentage
School Leader Capacity	20	90.9%
Teachers & Organizational Culture	17	77.3%
School Performance & Learning	12	54.5%

3.3 Program Evaluation Results

Table 2.3 Program Impact Synthesis (N=20)

Impact Level	Positive Indicators	Average Effect Size	Representative Studies
Individual	Confidence (+67%), Reflection (+54%)	$d=1.2$	Jethro et al. (2022)
Organizational	Collaboration (+49%), Parental participation (+38%)	$r=0.45$	Binti Khairiyah et al. (2024)
Institutional	School vision (+72%), Learning culture (+61%)	$\beta=0.38$	Bukhari et al. (2025)

Aggregate effectiveness positive (82% studies), but Kirkpatrick Level 1-2 indicators (reaction, learning) dominate (71%) while Level 3-4 (behavior, results) only 29%.

3.4 Limitations and Research Gaps

Structural limitations:

- Longitudinal scarcity: 91% studies <2-year follow-up (Mertkan & Gümüş, 2024)
- Indicator inconsistency: 14 different evaluation frameworks from 20 studies
- Geographical bias: 68% North America/Europe, 18% Asia (Pont, 2020)

Critical gaps: No RCT studies, limited cost-effectiveness analysis, absence of AI/digital leadership evaluation.

3.5 Theoretical and Practical Implications

Theoretical: Integrate *Integrated Leadership Model* combining instructional, distributed, and transformational leadership with *Logic Model* evaluation (input→output→outcome→impact).

Practical:

1. Program Design with Embedded Evaluation (Pre-Post Quasi-Experimental)

2. Standardized Indicator Framework: Kirkpatrick Level 4 + ROI integration
3. Contextual Global Leadership Framework for developing countries
4. Evaluator Training in mixed methods and longitudinal capacity building

CONCLUSION AND RECOMMENDATION

Systematic review of 22 scientific articles confirms the **substantial contribution** of educational leadership programs in strengthening school leadership capacity, teacher professionalization, and organizational culture transformation. The effectiveness of **instructional, distributed, and collaborative leadership** demonstrates significant correlation with institutional performance elevation and pedagogical practice optimization (ES=0.45–1.2). Methodologically, the dominance of **qualitative approaches (40.9%)** and **mixed methods (22.7%)** facilitates deep contextual exploration, while **quantitative studies (22.7%)** provide empirical validation through statistical modeling. The escalation of **systematic reviews and bibliometrics (13.6%)** marks disciplinary maturity, reflecting a paradigm shift from individual to collective leadership.

However, **persistent methodological limitations**—Kirkpatrick Level 1-2 focus (71%), minimal longitudinal designs (9%), indicator inconsistency (14 different frameworks), and geographical bias (68% developed contexts)—hinder global generalizability. Educational leadership program evaluation must evolve into a **dynamic process** integrating outcome assessment, causal mechanisms, and systemic impacts within contemporary educational ecosystems.

Research Recommendations

1. Longitudinal and quasi-experimental designs (RCT preferred) to measure *sustained impact* (3–5 years) on student outcomes
2. Integrated mixed methods with sequential explanatory design for quantitative-qualitative triangulation
3. Geographical expansion to developing countries (ASEAN, Africa) via cross-context comparative case studies
4. Cost-effectiveness analysis and AI-based digital leadership evaluation

Program Development Practice Recommendations

1. Embedded evaluation framework from *logic model* design stage (input→process→output→impact)
2. Multi-level indicators: Kirkpatrick Levels 1–4 + ROI + organizational health index
3. Program contextualization: Instructional-distributed adaptation for low-resource settings
4. Follow-up coaching (12–24 months post-program) for behavioral transfer

Educational Policy Recommendations

1. National Leadership Evaluation Standards (NES) with core indicators + local adaptation modules
2. Real-time digital monitoring dashboard for evidence-based policy iteration
3. Evaluator certification program and capacity building fund for M&E specialists
4. Public-private partnerships for scalable leadership academies with impact bonds

Table 3. Multi-Stakeholder Recommendation Matrix

Stakeholder	Primary Focus	Success Indicators	Timeline
Researchers	Longitudinal RCT	Q1 publications, Citations >50	2026–2030
Practitioners	Embedded evaluation	Behavior change >60%	2026–2028
Policymakers	National standards	80% program compliance	2026–2027

Implementation of these recommendations has the potential to catalyze evidence-based, equitable, and sustainable transformation in educational leadership toward P21C (Competitive 21st Century Education).

Acknowledgment

The authors would like to express their sincere gratitude to **Universitas Bina Bangsa** for the institutional support, facilities, and academic environment provided, which enabled the successful completion of this study.

REFERENCES

- R. H. Hallinger, P., & Heck, “Assessing the contribution of distributed leadership and growth in math achievement,” *Am. Educ. Res. J.*, vol. 4, no. 63, pp. 659–689, 2009, doi: <https://www.google.com/search?q=https://doi.org/10.3102/0002831209331548>.
- M. Akram, M. I. Malik, and S. Taj, “Effect of Instructional Leadership on School Performance,” *Glob. Soc. Sci. Rev.*, vol. VII, no. IV, pp. 10–22, 2022, doi: 10.31703/gssr.2022(vii-iv).02.
- E. Daniëls, A. Hondeghem, and F. Dochy, “A review on leadership and leadership development in educational settings,” *Educ. Res. Rev.*, vol. 27, pp. 110–125, 2019, doi: <https://doi.org/10.1016/j.edurev.2019.02.003>.
- H. Bahtiar and M. K. Qasabandiyah, “21 st -Century Educational Leadership: A Review of Core Theories , Managerial Perspectives , and Emerging Trends,” *J. Teach. Educ. Dcholars*, vol. 2, no. 2, pp. 88–102, 2025.
- Mutoharoh, A. Hufad, and I. Rusdiyani, “The Correlation between Innovative Leadership and Early Childhood Education Teacher Performance,” *J. Educ. Soc. Res.*, vol. 13, no. 3, pp. 154–163, 2023, doi: 10.36941/jesr-2023-0065.
- B. Pont, “A literature review of school leadership policy reforms,” *Eur. J. Educ.*, vol. 55, no. 2, pp. 154–168, 2020, doi: 10.1111/ejed.12398.
- T. Karakose, K. Leithwood, and T. Tülübaş, “The Intellectual Evolution of Educational Leadership Research: A Combined Bibliometric and Thematic Analysis Using SciMAT,” *Educ. Sci.*, vol. 14, no. 4, 2024, doi: 10.3390/educsci14040429.
- S. Mertkan and S. Gümüş, “Review of systematic reviews in educational leadership and management: Methods used, topics explored and geographies covered,” *Educ. Manag. Adm. & Leadersh.*, vol. 0, no. 0, p. 17411432241291196, 2024, doi: 10.1177/17411432241291197.
- E. Pitriani, “Educational Leadership in Improving School Performance and Teacher Professional Development: A Literature Review,” *PPSDP*

Int. J. Educ., vol. 3, no. 2, pp. 233–244, 2024, doi: 10.59175/ pijed.v3i2.312.

R. Bolam, “Professional Learning Communities and Teachers’ Professional Development,” 2008, pp. 159–179. doi: 10.1007/978-1-4020-8186-6_10.

W. Admiraal, W. Schenke, L. De Jong, Y. Emmelot, and H. Sligte, “Schools as professional learning communities: what can schools do to support professional development of their teachers?,” *Prof. Dev. Educ.*, vol. 47, no. 4, pp. 684–698, 2021, doi: 10.1080/19415257.2019.1665573.

S. Demiröz, “Evaluation and Bibliographic Analyses of EdLAM Literature to Help Early Career Academics,” *Int. Educ. Stud.*, vol. 14, no. 12, p. 1, 2021, doi: 10.5539/ies.v14n12p1.

Y. Li, Y. Cai, and R. Tang, “Linking Instructional Leadership and School Support to Teacher Expertise: The Mediating Effect of Teachers’ Professional Development Agency,” *Sustain.*, vol. 15, no. 4, 2023, doi: 10.3390/su15043440.

A. O. OREDEIN and C. C. OBADIMEJI, “Professional Development and Leadership,” *Edumania-An Int. Multidiscip. J.*, vol. 02, no. 03, pp. 150–158, 2024, doi: 10.59231/edumania/9063.

M. Jethro M., N. Njabuliso H., and M. David, “Professional Development of School Leaders In Supporting Effective Teaching And Learning In The Shiselweni Region, Eswatini,” *Int. J. Sci. Res. Manag.*, vol. 10, no. 05, pp. 2385–2397, 2022, doi: 10.18535/ijsrn/v10i5.e109.

R. H. Heck and P. Hallinger, “Assessing the contribution of distributed leadership and growth in math achievement,” *Am. Educ. Res. J.*, vol. 46, no. 3, pp. 659–689, 2009, doi: 10.3102/0002831209340042.

P. D. Wiens, V. Vongkulluksn, J. Beck, and B. Pearson, “Understanding teacher leadership: Development and testing of the survey of teacher leadership,” *Soc. Sci. Humanit. Open*, vol. 12, p. 101738, 2025, doi: <https://doi.org/10.1016/j.ssaho.2025.101738>.

A. Dimopoulos, “Educational Leadership Effectiveness. Is it a Matter of a Leader’s Characteristics, Behaviors, or Leadership Style?,” *J. Econ. Manag. Sci.*, vol. 3, no. 1, p. p13, 2020, doi: 10.30560/jems.v3n1p13.

Q. Qodiriyah, “Instructional Leadership in The Development of School Organizational Culture: A Literature Review,” *J. Qual. Assur. Islam. Educ.*, vol. 3, no. 2, pp. 56–68, 2023, doi: 10.47945/jqaie.v3i2.1193.

Binti Khoiriyah, Iis Marsithah, and Mohammad Adnan, “Evaluation of the Principal’S Leadership Program in Improving Parents’ Participation in Children’S Education,” *J. Pendidik. Glas.*, vol. 8, no. 2, pp. 237–245, 2024, doi: 10.32529/glasser.v8i2.3742.

S. K. U. S. Bukhari, A. H. Qamar, and M. Abiodullah, “Evaluation of Leadership Development Program for Secondary School Heads: Perceived Satisfaction, Learning and its Implementation,” *iRASD J. Educ. Res.*, vol. 6, no. 1, pp. 44–59, 2025, doi: 10.52131/jer.2025.v6i1.2716.

E. Norman, K. Komarudin, A. Masruri, J. Wahid, and Y. Hasbian, “Leadership , Professional Development , and Innovation: A Systematic Literature Review through Digital,” *J. Educ. Manag. Res.*, vol. 04, no. 05, pp. 2327–2362, 2025.