global evaluation initiative

2.5

Success Factors and Barriers in M&E Systems Strengthening: A Scoping Review

Douglas Glandon
Guilherme Cunha Lima

Abstract

Monitoring and Evaluation (M&E) systems are indispensable to evidence-informed decision-making, yet the success of efforts to strengthen those systems varies markedly across settings. This synthesis follows a scoping-review protocol grounded in JBI guidance, with reporting per PRISMA-ScR. Enabling conditions identified in the literature include sustained political leadership, broad-based stakeholder engagement, institutional integration of M&E with planning and budgeting processes, and deliberate investment in professional and organizational capacity. The reviewed sources suggest that these conditions can raise demand for credible evidence and enlarge the domestic supply of high-quality studies. Effective communication of findings, participatory evaluation processes, and institutional incentive structures in turn shape whether evidence is understood, valued, and acted upon. Major barriers — including weak political will, donor-driven agendas, low technical capacity, fragmented governance, and poor knowledge management — dampen those same dimensions. The review also underscores the importance of context-specific strategies that adapt to political economy realities, cultural norms, and levels of decentralization. Persistent knowledge gaps include the long-term impact of capacity-building, evidence use in decentralized systems, and comparative insights across political regimes. Addressing these gaps through longitudinal and mixed-method research would refine guidance for future M&E-system reforms.

Introduction

This literature review addresses the question, "What are the key factors influencing the success (or not) of M&E systems-strengthening interventions in improving the demand, supply and/or use of evaluative evidence?" Drawing on 20 sources that span conceptual frameworks, empirical case studies, and systematic reviews, the analysis proceeds in three stages: first, it examines enabling factors that strengthen demand, supply, and use; second, it analyses hindering factors that impede each dimension; and third, it synthesizes cross-cutting insights for policy and research.

Conceptual Framework

This scoping review is one of a set of twenty scoping reviews commissioned to inform the Global Evaluation Initiative (GEI) 2.0 Strategy (FY2026–30).

The review set is organized around a conceptual framework that mirrors the nested structure of GEI's theory of change and theory of action which addresses four analytical levels: country, decision-maker/unit, evaluation/decision, and individual. At the country level, the focus is on the linkage between stronger M&E ecosystems and development and governance outcomes. At the decision-maker/unit level, the emphasis is on strengthening the culture and practice of evidence-informed decision-making within specific ministries or units. At the evaluation/decision level, the focus is on how evaluations or other evidence-generating activities can inform specific policy or program decisions. Finally, at the individual level, the framework addresses how training, mentorship, peer exchange, and related supports strengthen the skills, performance, and professional development of evaluators and other evaluation actors. Within each level, the review set addresses five recurring sub-topics: how the focal construct is defined and conceptualized; the primary theories of change linking the level's activities to its intended outcomes; the metrics used to characterize status and change at that level; the empirical evidence linking activities to outcomes; and the drivers that help explain success or failure. The resulting 4×5 matrix — four analytical levels by five recurring sub-topics — structures the review set into twenty distinct but complementary scoping reviews.

The present review addresses the *drivers* sub-topic at the decision-maker/unit level.

Methods

Review design

This review was conducted as a scoping review following JBI guidance (Peters et al., 2020), with reporting per PRISMA-ScR (Tricco et al., 2018). A scoping format was appropriate because the research question is drivers-oriented and required breadth across heterogeneous, multidisciplinary material — peer-reviewed scholarship, institutional reports, country-case learning products, and ECD framework documents. Eligibility followed the JBI Population–Concept–Context (PCC) framing, with the review question and inclusion criteria anchored primarily to the concept of the factors that drive or hinder the success of M&E systems strengthening interventions in improving evaluative-evidence demand, supply, and use. Per §11.1.1 of the JBI Manual, no formal critical appraisal or risk-of-bias assessment was applied.

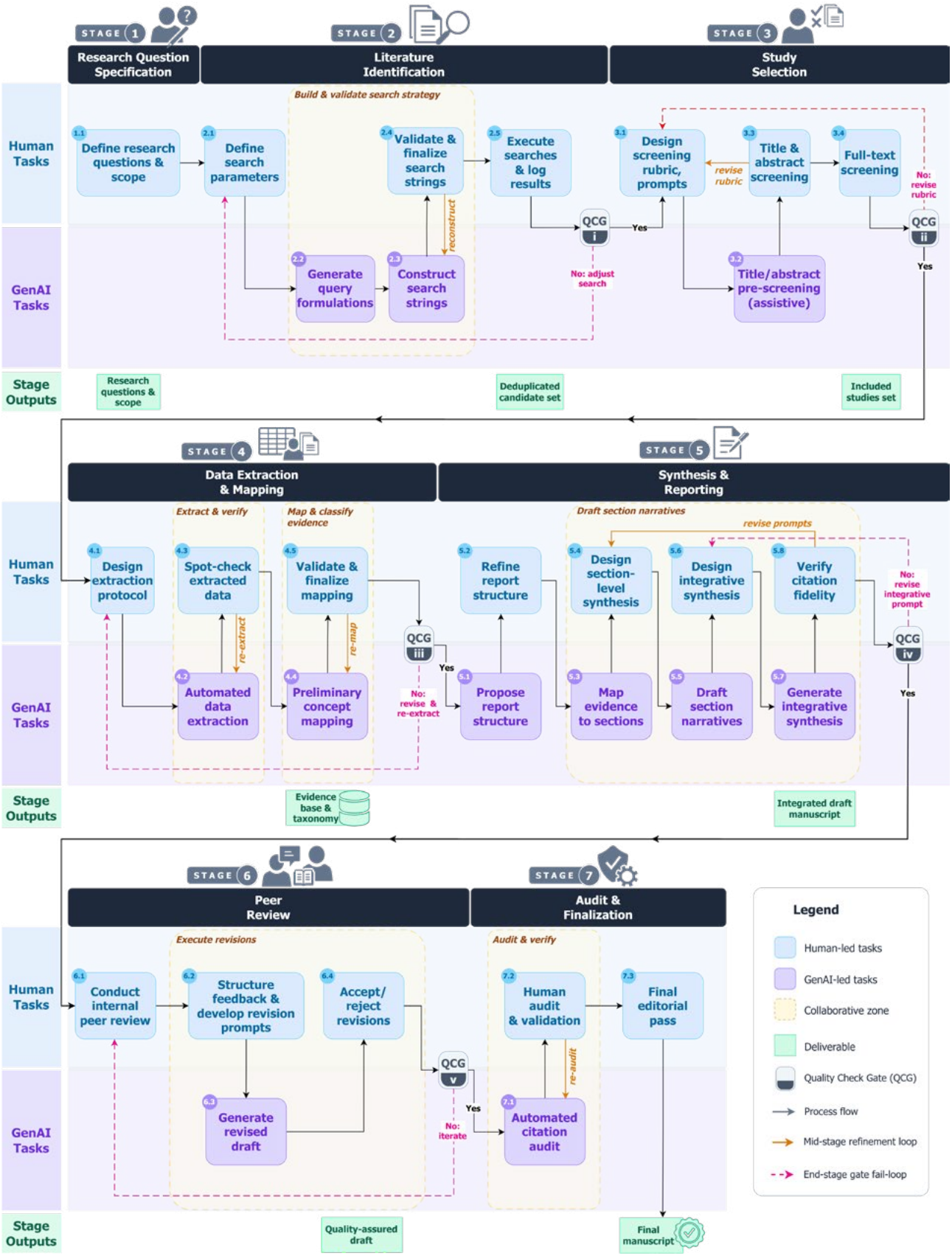
The review followed the seven-stage human-in-the-loop workflow shown in **Figure 1**, beginning with human-led specification of the research question and scope (Stage 1, Task 1.1). Across all stages, LLM tools supported defined tasks under human-designed prompts; two human reviewers retained authority over every inclusion, charting, mapping, and synthesis decision. Figure 1's five "collaborative zones" mark groupings of paired human and GenAI tasks that operated on the same evolving artifact through iterative cycles, with each actor's work shaping the other's next task.

Eligibility criteria

Sources were eligible if they identified or analyzed factors driving or hindering the success of M&E systems strengthening interventions, and connected those factors to the demand for, supply of, or use of evaluative evidence. Eligible source types included peer-reviewed journal articles, institutional reports and country-case learning products, doctoral theses, and ECD framework documents. No a priori date or language restriction was applied; included sources were entirely English-language. At full-text review, sources were excluded for not engaging with strengthening drivers, addressing drivers of unrelated phenomena, or not meeting the minimum quality threshold applied to their source type.

Two source-type-specific quality thresholds were applied. Peer-reviewed sources had to have been issued through standard editorial peer-review processes. Grey-literature sources additionally had to meet three criteria: (i) identify or analyze specific factors that drive or hinder the success of strengthening interventions, rather than describing strengthening interventions without engaging with their drivers; (ii) a documented evidence or reference base; and (iii) identifiable institutional quality control (verified pragmatically by checking whether the source was published through an official institutional channel and bore clear markers of institutional authorship, approval, or accountability).

Figure 1. GEI 2.0 Rapid Evidence Review (RER) Workflow



Information sources and search strategy

Literature identification (Stage 2 of Figure 1) followed an iterative human–GenAI cycle, beginning with the human reviewer's definition of search parameters (Task 2.1). Scopus and Web of Science Core Collection provided peer-reviewed coverage, with Google Scholar searched as a complementary academic search engine. Five institutional repositories were systematically searched: the World Bank Open Knowledge Repository, the UNEG Publications Repository, the OECD iLibrary, the 3ie Development Evidence Portal, and the institutional publication channels of CLEAR Anglophone Africa and Twende Mbele. Targeted searches of the working-paper and publication channels of these African ECD partnerships, focused on driver-bearing material, were conducted as a structured supplementary stream. Coverage was further extended through backward and forward citation chasing, expert recommendation, and key-author tracking focused on scholars whose prior work has shaped the strengthening-drivers literature.

Within the "build and validate search strategy" collaborative zone, the LLM generated semantically equivalent query reformulations and drafted database- and repository-specific search strings (Tasks 2.2–2.3), which the human reviewer audited and finalized (Task 2.4); a search log archived run dates, applied limits, and result counts. Searches were executed under a pre-specified stopping rule (Task 2.5): searching continued through successive results pages until three consecutive pages yielded no further relevant records. **Quality Check Gate (QCG) i** required the deduplicated candidate set to meet coverage and retrievability criteria; where it did not, reviewers revised search parameters or strings and re-executed searches.

Source selection

Study selection (Stage 3) operated under a human-designed screening protocol (Task 3.1). A total of 240 records were identified through databases and registers (222 from databases; 18 from registers), with a further 72 identified through backward and forward citation searching and expert recommendation and known-source inclusion (**Figure 2**). After deduplication and metadata-based screening to remove records of document types outside the eligible scope, 181 records underwent title and abstract screening: the LLM produced assistive first-pass inclusion-likelihood labels (Task 3.2), and two independent human reviewers performed the title and abstract screening (Task 3.3), informed but not bound by the LLM labels. Thirty-two records advanced to full-text screening (Task 3.4); three could not be retrieved, leaving 29 assessed, of which 15 were excluded with coded reasons. In parallel, 69 of the 72 supplementary-stream records were retrieved and assessed (three were not retrievable), and 63 were excluded with coded reasons. A total of 20 sources were included in the review.

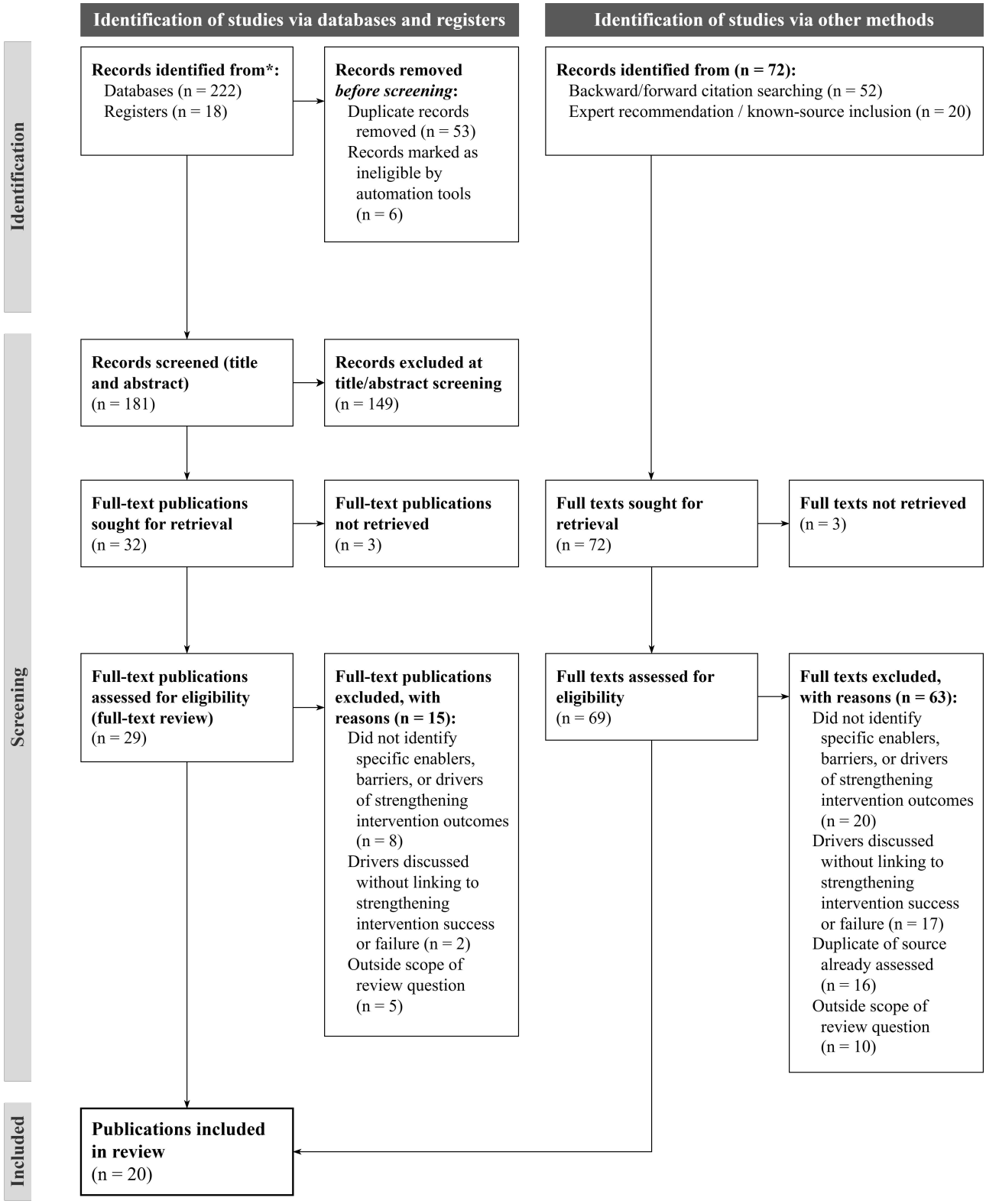
QCG ii required the included studies set to satisfy protocol-specified eligibility coverage; where it did not, the reviewers revised the screening protocol and re-screened the affected records.

Data extraction and mapping

Data extraction and mapping (Stage 4) proceeded under a human-designed extraction protocol (Task 4.1). A structured codebook captured bibliographic metadata; the driver category; the direction of influence (enabling versus hindering); the demand/supply/use pathway addressed; the country or organizational context; and the type of evidence supporting the driver claim. Within the "extract and verify" collaborative zone, LLM-assisted extraction (Task 4.2) populated the codebook with page-anchored quotations and paraphrases, with iterative spot-checks and prompt refinement (Task 4.3) used to re-extract fields where verification flagged errors — including misattributions, unsupported statements, hallucinated content, and conflation of proximate constructs. Reviewers verified all driver attributions and direction-of-influence classifications and a random 20% sample of remaining fields. Within the "map and classify evidence" zone, a preliminary concept mapping produced under structured prompts (Task 4.4) was then validated and finalized by the reviewers (Task 4.5), with mapping organized along two axes — direction of influence (enabling versus hindering) and pathway addressed (demand, supply, or use) — motivating a 2 × 3 analytical structure pairing each direction of influence with each evidence pathway.

QCG iii required the resulting evidence base and taxonomy to meet extraction-fidelity criteria — including page-anchored support for every charted statement — before synthesis. Where these criteria were not met, the process looped back to Task 4.1, where the reviewers revised the extraction protocol before the affected fields were re-extracted.

Figure 2. PRISMA Flow Diagram



Synthesis and reporting

Synthesis and reporting (Stage 5) proceeded through collaborative drafting cycles. A draft report structure was proposed under structured prompts (Task 5.1) and refined by the reviewers (Task 5.2). Within the "draft section narratives" collaborative zone, evidence was mapped to sections (Task 5.3), section-level synthesis logic was designed under human direction (Task 5.4), and section narratives were generated under human-designed prompts (Task 5.5) — producing, in this review, a 2 × 3 analytical structure pairing each direction of influence (enabling versus hindering) with each evidence pathway (demand, supply, use). An integrative prompt designed by the reviewers (Task 5.6) was then used to produce the integrated draft (Task 5.7).

A human verification of citation fidelity (Task 5.8) sampled a random 20% of cited claims and checked each against its source. Where errors, misattributions, or hallucinations appeared in the sample, the reviewer examined whether they were scattered across sections and sources or concentrated in one section or source; based on that assessment, the process looped back to Task 5.4, where the section-level synthesis prompts were revised and the affected section narratives regenerated. Where the sample was clean, the process advanced to QCG iv.

QCG iv assessed the structural and rhetorical quality of the integrated draft — whether the sections were well tied together, built on one another, and formed a coherent whole with clear flow and consistent framing. Where it did not, the reviewers revised the integrative synthesis prompt (Task 5.6) and the LLM regenerated the integrated draft.

Quality assurance

Before finalization, the integrated draft was processed through a GenAI-assisted internal peer review (Stage 6). The reviewer entered text-anchored marginal comments in the integrated draft (Task 6.1); a VBA macro converted those comments into inline, LLM-ingestible comments embedded in the body text (Task 6.2). The LLM then returned an iteratively revised draft with all changes shown in tracked changes (Task 6.3), and the reviewer accepted or rejected each proposed revision (Task 6.4). **QCG v** marked the point at which the model returned a revised draft requiring no further peer-review comments; where further comments remained, the process looped back to Task 6.1 for another cycle.

Final audit and finalization (Stage 7) used a purpose-built claim-and-citation audit pipeline (Task 7.1) within the "audit and verify" zone: the pipeline parsed every in-text citation, retrieved the corresponding extracted source passage, and used an LLM under a structured rubric to classify each cited claim as Supported, Partially Supported, or Not Supported. In the human audit and validation pass (Task 7.2), all Not Supported and Partially Supported claims were revised or removed, and a 10% random sample of the remaining Supported claims was manually re-verified against source PDFs. The audit was then re-run, with the cycle repeated until every cited claim returned a Supported verdict. A final editorial pass (Task 7.3) covered coherence and flow, formatting against the publication template, terminological consistency, and APA-compliance of the reference list.

Enabling Factors for Strengthening Demand for Evaluative Evidence

Robust demand for evaluative evidence begins with unequivocal political and institutional leadership. Kanyamuna (2021) shows that presidential and cabinet-level endorsement in Zambia catalyzed ministries to request performance data, while Chirau et al. (2020) emphasize that securing political and administrative champions is crucial for M&E systems, as these champions exert influence on ministries, departments, and agencies across Anglophone Africa (p. 2). Such leadership is reinforced when M&E outputs are woven into core policy processes; Matsiliza (2019) notes that South Africa's Government-Wide M&E system was designed to support planning, budgeting, program implementa-

tion, financial management, and reporting processes (p. 490). Legislative scrutiny and civil-society watchdogs further sustain that demand: Parliament in Ghana is increasingly demanding accountability from government ministries, departments, and agencies, with parliamentary sub-committees occasionally visiting projects of national importance to assess progress (Adams et al., 2013, p. 15). Underpinning all of these factors is growing awareness of evaluation's role in decision-making: in Ghana, there is increasing recognition at all levels of government of the important role of evaluation in sustainable development, with latent demand partly attributed to insufficient awareness that evaluation can provide relevant evidence (Adams et al., 2013, pp. 27, 30).

Enabling Factors for Strengthening Supply of Evaluative Evidence

The supply of evaluative evidence is facilitated by enabling factors — such as political commitment, institutional architecture, and resource flows — that create fertile ground for capacity-building interventions. Among these, institutionalized training pipelines, professional evaluation networks, and formal government structures each play a critical reinforcing role, though their effectiveness depends on the broader systemic conditions in which they operate.

Institutionalized training programs represent one key enabling factor. In Ghana, UNICEF collaborated with the Ghana Monitoring and Evaluation Forum (GMEF) and the University of Ghana to initiate the country's first Master of Arts program in Evaluation, a partnership designed to institutionalize professional evaluation training within the local university system (Adams et al., 2013, pp. 27, 30). Similarly, in Zambia, the well-resourced Centre of Excellence at the University of Zambia offered a foundation for strengthening supply, complemented by various think tanks and the newly merged Evaluation Association, though the university sector faces structural challenges including qualified staff leaving for better-paid positions elsewhere (Mulenga & Porter, 2013, p. ii).

Professional networks constitute another enabling factor, yet their sustainability is not guaranteed. Uganda's Evaluation Association, formed in 2001 and registered in 2002 as a professional network to facilitate knowledge-sharing among evaluators and promote professionalism in evaluation practice, illustrates both the potential and fragility of such infrastructure — while designed to connect practitioners and build capacity, the association became inactive due to limited funding (Niringiye, 2018, pp. 20, 24). Formal institutional frameworks likewise matter: Matsiliza (2019) discusses strategies to improve M&E capacity through strengthening organizational structures, training, and information systems, noting the role of a dedicated monitoring and evaluation department within the Presidency in anchoring the function within government (p. 490). Development-partner assistance, though sometimes critiqued for overshadowing domestic priorities, can still fill critical funding gaps that national treasuries cannot (CLEAR-AA, 2020, pp. v, 16).

Enabling Factors for Strengthening Use of Evaluative Evidence

Effective use of evaluative evidence is shaped by communication quality, participatory engagement, and organizational incentives. A foundational factor is how evaluation findings are communicated: evaluator communication skills and stakeholder involvement can translate technical findings into actionable insights, with reporting that is understandable and timely being critical for stakeholder use (Alkin & King, 2017, pp. 446–447). In a related vein, Kuchenmüller et al. (2022) describe how EVIPNet Europe's knowledge-translation model uses evidence briefs for policy, rapid response services, and policy dialogue exercises to facilitate the routine use of context-sensitive evidence in health decision-making, emphasizing the tailoring of information and alignment of research with policy prior-

ities (pp. 2–4). Beyond clarity, however, evidence products are often perceived as interesting but not useful when they lack clear directions for action; Morton and Seditas (2018) describe how action-focused reporting formats — including tools such as "Talking Points" and collaborative evidence use plans — help officials move beyond abstract findings toward the implementation of tangible actions (pp. 156, 161–162, 164).

Participatory designs further increase utilization by fostering process use — individual and organizational changes in thinking and culture that occur through learning during the evaluation process — as stakeholder involvement supports the evaluation's learning function and helps ensure findings are understood and applied (Donnelly & Searle, 2017, pp. 308–309). At the organizational level, institutional cultures that privilege learning and accountability also strengthen uptake; Bullock et al. (2021) argue that governance arrangements and financial incentive structures — including performance-based payments, public reporting, and enforcement strategies — are essential for translating evidence-informed policies into effectively embedded system changes (pp. 10, 12–13).

Hindering Factors for Strengthening Demand for Evaluative Evidence

Several interrelated barriers constrain demand for evaluative evidence, rooted in unfavourable political economy dynamics, limited understanding of M&E's value, donor-driven priorities, and institutional fragmentation. If leaders perceive M&E as punitive "policing" or "witch-hunting" rather than a learning tool, they may underfund or ignore evaluation units, with M&E services frequently the first to face budget cuts (Chirau et al., 2022, pp. 8, 10). Limited understanding of M&E's value compounds the problem; Kanyamuna et al. (2018) report that the Zambian parliament lacks deliberate institutional coordination with the health sector's M&E function and relies primarily on general reports, reflecting insufficient technical capacity and a lack of understanding of how M&E systems can inform parliamentary business (p. 82). Beyond domestic political dynamics, donor-driven priorities may further undermine local ownership: Chirau et al. (2020) observe that current M&E systems in African countries are driven by donors, and they recommend making M&E capacity strengthening country-led and country-owned, arguing that endogenously driven M&E findings are more likely to be used (p. 3). Cutting across all of these factors, institutional silos and coordination failures further restrict evidence demand; Liverani et al. (2013) found that the division of responsibilities within government bureaucracies compels civil servants to focus narrowly on specific policy areas, restricting their engagement with broader evidence, while institutional silos particularly limit evidence consideration for complex issues requiring multi-disciplinary thinking and cross-sectoral coordination (p. 4).

Hindering Factors for Strengthening Supply of Evaluative Evidence

Multiple reinforcing constraints hinder efforts to strengthen the supply of evaluative evidence, spanning human capacity, financing, institutional infrastructure, and coordination across actors. Supply bottlenecks often begin with scarce technical capacity. M&E and statistics capacity within ministries, departments, and agencies remains low, with high turnover rates of qualified staff in Ghana (Adams et al., 2013, p. 12) and underdeveloped analytical skills in Malawi's non-research-oriented institutions (Kumwenda & Latib, 2013, p. 14). Funding shortfalls compound the deficit: in Zambia, financial capacity remains a core challenge to implementing evaluation plans, and even when specific budgets for monitoring and evaluation are in place, the funds are rarely actually expended (Mulenga & Porter, 2013, pp. 18, 32). Where domestic capacity and funding fall short, the predominance of international evaluators temporarily fills gaps but raises concerns about long-term domestic capacity development.

In Uganda, about half of donor-led evaluations were conducted by international evaluators only, while local capacities for high-quality evaluation remain limited and the demand and supply of evaluators continue to be driven by development partners with limited country ownership (Niringiye, 2018, pp. 18, 26). Even when evaluative evidence is produced, weak knowledge-management systems restrict access to completed studies. In Malawi, the absence of a national repository for evaluations and analytical studies leaves stakeholders without a central reference point to determine the scope, quality, or number of completed assessments (Kumwenda & Latib, 2013, p. 11). Fragmentation of donor-established M&E systems further erodes coherence: state and non-state sectors operate in silos, with non-state entities often prioritizing reporting systems designed for international funders over integration with state counterparts, and the resulting lack of coordination makes the long-term effectiveness of capacity-strengthening interventions difficult to measure (Masvaure & Fish, 2022, pp. 4–5).

Hindering Factors for Strengthening Use of Evaluative Evidence

Use of evaluative evidence falters when multiple barriers — relating to accessibility, stakeholder engagement, political context, and governance structures — impede the connection between evaluation findings and decision-making. At the most immediate level, access and format pose significant obstacles: paywalls prevent access, published syntheses frequently appear in language and formats that are difficult to use, and practitioners lack time and resources to synthesize findings independently (Morton & Seditas, 2018, p. 156). Compounding these barriers, reporting must be in a form users can understand and timely enough to be useful to stakeholders (Alkin & King, 2017, p. 447). Even when findings are accessible, weak stakeholder engagement and limited organizational commitment to evaluation diminish the likelihood that they will be meaningfully applied. Johnson et al. (2009) found that while administrators may read relevant evaluations, they typically learn little from reports alone, and successful learning instead requires active involvement, interaction, and communication between evaluation clients and evaluators (Johnson et al., 2009, pp. 389, 397). Beyond organizational factors, political and ideological contestation can override even well-communicated evidence. Liverani et al. (2013) identified studies showing how normative positions — including values, moral convictions, and nationalism — can bias the selection and interpretation of evidence, with research finding that resistance to Western science grounded in nationalism helped explain why former South African President Mbeki rejected evidence on the causal link between HIV and AIDS (p. 5). Finally, fragmented governance structures often result in feedback loops that fail to inform policy. While Botswana has established central M&E structures, information utilization remains limited at 43 per cent and results are infrequently adopted in parliamentary decision-making (Makadzange, 2022, p. 8). Similarly, disconnects between specific program initiatives and main institutional systems in Ethiopia's health sector lead to non-uniform reporting and feedback mechanisms across regions (Minyihun et al., 2022, pp. 71, 79). In sum, these hindering factors illustrate that strengthening the use of evaluative evidence requires addressing not only the technical quality and accessibility of evaluation outputs but also the organizational, political, and institutional conditions that shape whether and how evidence informs decision-making.

Conclusion

This review has mapped how enabling and hindering factors shape the effectiveness of M&E system-strengthening interventions in improving the demand, supply, and use of evaluative evidence. Across the reviewed sources, political commitment emerges as a recurring precondition, with capacity and institutional frameworks providing necessary supporting structures and communication together with participatory cultures enabling evidence to gain traction in policy arenas. Converse-

ly, disengaged leadership, fragile capacity, donor dependency, and fragmentation appear to undermine reform efforts. Context matters profoundly: decentralized polities may offer more entry points for evidence but require tailored capacity support; centralized regimes may achieve rapid institutionalization but risk insulating decision-makers from critical scrutiny. A recurring theme across these sources is that no single enabling factor is sufficient in isolation; rather, effective strategies appear to require integrating political buy-in, sustainable funding, demand-driven capacity development, and explicit incentives for evidence use into a mutually reinforcing package. Significant research gaps remain, particularly on long-term capacity outcomes, sub-national systems, and cross-country comparative learning. Future studies employing mixed-method longitudinal designs could illuminate causal pathways, enabling practitioners to fine-tune interventions to the contextual enablers and barriers identified herein.

References

Adams, S., Amoatey, C., Taabazuing, J., & Feinstein, O. (2013). *Study on the demand for and supply of evaluation in Ghana*. Centre for Learning on Evaluation and Results – Anglophone Africa (CLEAR-AA), University of the Witwatersrand.

Alkin, M. C., & King, J. A. (2017). Definitions of evaluation use and misuse, evaluation influence, and factors affecting use. *American Journal of Evaluation*, *38*(3), 434–450. https://doi.org/10.1177/1098214017717015

Arksey, H., & O'Malley, L. (2005). Scoping studies: Towards a methodological framework. *International Journal of Social Research Methodology*, *8*(1), 19–32. https://doi.org/10.1080/1364557032000119616

Bullock, H. L., Lavis, J. N., Wilson, M. G., Mulvale, G., & Miatello, A. (2021). Understanding the implementation of evidence-informed policies and practices from a policy perspective: A critical interpretive synthesis. *Implementation Science*, *16*, Article 18. https://doi.org/10.1186/s13012-021-01082-7

Centre for Learning on Evaluation and Results – Anglophone Africa (CLEAR-AA). (2020). *M&E capacity-strengthening approaches and measurement in Africa*. https://unfccc.int/sites/default/files/resource/Clear%20ME%20Capacity%20report-V4b.pdf

Chirau, T., Dlakavu, A., & Masilela, B. (2022). Strengthening Anglophone Africa M&E systems: A CLEAR-AA perspective on guiding principles, challenges and emerging lessons. *African Evaluation Journal*, *10*(1), Article a601. https://doi.org/10.4102/aej.v10i1.601

Chirau, T., Masilela, B., & Dlakavu, A. (2020). *Country-owned, country-led: Strengthening African M&E systems through country-based programmes*. Centre for Learning on Evaluation and Results – Anglophone Africa (CLEAR-AA).

Donnelly, C., & Searle, M. (2017). Optimizing use in the field of program evaluation by integrating learning from the knowledge field. *Canadian Journal of Program Evaluation*, *31*(3), 305–327. https://doi.org/10.3138/cjpe.366

Johnson, K., Greenseid, L. O., Toal, S. A., King, J. A., Lawrenz, F., & Volkov, B. (2009). Research on evaluation use: A review of the empirical literature from 1986 to 2005. *American Journal of Evaluation*, *30*(3), 377–410. https://doi.org/10.1177/1098214009341660

Kanyamuna, V. (2021). Towards building a functional whole-of-government monitoring and evaluation system for Zambia: The demand side. *World Journal of Social Sciences and Humanities*, *7*(3), 83–105. https://doi.org/10.12691/wjssh-7-3-2

Kanyamuna, V., Aurick, M., Emily, N., Chinyama, M., & Ackson, M. (2018). An assessment of the demand-side of the monitoring and evaluation system of the health sector in Zambia. *World Journal of Social Sciences and Humanities*, *4*(2), 75–86.

Kuchenmüller, T., Chapman, E., Takahashi, R., Lester, L., Reinap, M., Ellen, M., & Haby, M. M. (2022). A comprehensive monitoring and evaluation framework for evidence to policy networks. *Evaluation and Program Planning*, *91*, 102053. https://doi.org/10.1016/j.evalprogplan.2022.102053

Kumwenda, H., & Latib, S. (2013). *Study on the demand for and supply of evaluation in Malawi*. Centre for Learning on Evaluation and Results – Anglophone Africa (CLEAR-AA), University of the Witwatersrand.

Liverani, M., Hawkins, B., & Parkhurst, J. O. (2013). Political and institutional influences on the use of evidence in public health policy: A systematic review. *PLoS ONE*, *8*(10), Article e77404. https://doi.org/10.1371/journal.pone.0077404

Makadzange, P. F. (2022). A descriptive narrative on the current situation against the gold standards regarding institutionalisation of national monitoring and evaluation system for Botswana and Zimbabwe. *African Evaluation Journal*, *10*(1), Article a578. https://doi.org/10.4102/aej.v10i1.578

Masvaure, S., & Fish, T. E. (2022). Strengthening and measuring monitoring and evaluation capacity in selected African programmes. *African Evaluation Journal*, *10*(1), Article a635. https://doi.org/10.4102/aej.v10i1.635

Matsiliza, N. S. (2019). Strategies to improve capacity for policy monitoring and evaluation in the public sector. *Journal of Reviews on Global Economics*, *8*, 490–499. https://doi.org/10.6000/1929-7092.2019.08.42

Minyihun, A., Getnet, M., & Tilahun, B. (2022). How to strengthen the monitoring and evaluation system to improve compassionate, respectful, and caring services in Ethiopia: An implementation study. *Ethiopian Journal of Health and Biomedical Sciences*, *12*(Special Issue), 65–76.

Morton, S., & Seditas, K. (2018). Evidence synthesis for knowledge exchange: Balancing responsiveness and quality in providing evidence for policy and practice. *Evidence & Policy*, *14*(1), 155–167. https://doi.org/10.1332/174426416X14779388510327

Mulenga, O., & Porter, S. (2013). *Study on the demand for and supply of evaluation in Zambia*. Centre for Learning on Evaluation and Results – Anglophone Africa (CLEAR-AA), University of the Witwatersrand.

Niringiye, A. (2018). *Diagnostic study on the supply and demand of evaluators in Uganda*. Twende Mbele. https://twendembele.org/wp-content/uploads/2020/01/Demand-and-Supply-of-Evaluators-Final-Report-UGANDA.pdf

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., … Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, *372*, n71. https://doi.org/10.1136/bmj.n71

Peters, M. D. J., Godfrey, C., McInerney, P., Munn, Z., Tricco, A. C., & Khalil, H. (2020). Chapter 11: Scoping reviews. In E. Aromataris & Z. Munn (Eds.), *JBI Manual for Evidence Synthesis*. JBI. https://doi.org/10.46658/JBIMES-20-12

Tricco, A. C., Lillie, E., Zarin, W., O'Brien, K. K., Colquhoun, H., Levac, D., … Straus, S. E. (2018). PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and explanation. *Annals of Internal Medicine*, *169*(7), 467–473. https://doi.org/10.7326/M18-0850