# What are the key measures or approaches used to assess whether and to what extent a specific source of evidence (e.g., an evaluation) influenced policy and/or programmatic decision-making?

## Executive Summary

This literature review examines how to assess whether – and to what extent – specific sources of evidence, such as evaluations, influence policy and programmatic decision-making. It synthesizes key frameworks, methodologies, and case studies to support policymakers and researchers in selecting appropriate tools for evaluating evidence use.

A central finding is that evidence influences decision-making through instrumental (direct), conceptual (indirect), and symbolic (political) pathways. Recognizing these distinctions is critical for selecting appropriate measurement strategies. The review categorizes assessment approaches into three main types: quantitative, qualitative, and mixed methods. Quantitative tools – such as citation tracking, indicators, and surveys – are scalable and suited for contexts with strong documentation. Notable examples include Tudisca et al.’s (2018) 25 indicators and Abadie et al.’s (2023) Value of Evidence metric. However, these often miss more subtle or delayed forms of influence.

To address these limitations, qualitative and mixed-methods approaches – like case studies, process tracing, stakeholder interviews, and social network analysis – are essential for capturing deeper shifts in thinking, institutional learning, and relationship dynamics. Case examples from Africa, Europe, and Asia demonstrate how these methods can be adapted to varied decision-making environments.

The review also highlights key challenges, including attribution difficulties, time lags, and political or cognitive biases that complicate measurement. It concludes that there is no one-size-fits-all approach. Instead, organizations must tailor their assessment strategies based on available resources, capacity, and the type of evidence in question. Ultimately, the review provides a roadmap for more deliberate, context-sensitive evaluations that go beyond tracing influence to fostering it.

## Introduction

Understanding how evidence influences policy and programmatic decision-making is a critical endeavor in public administration and policy research, as it bridges the gap between knowledge production and practical application. This literature review addresses the research question: “What are the key measures or approaches used to assess whether and to what extent a specific source of evidence (e.g., an evaluation) influenced policy and/or programmatic decision-making?” Drawing on a diverse set of studies, the review synthesizes methodologies, practical examples, and challenges identified in the literature to provide a comprehensive analysis. It begins with an overview of evidence influence, exploring its conceptual foundations and the varied ways it manifests in decision-making processes. The discussion then shifts to an in-depth examination of methodologies and measures, ranging from quantitative indicators to qualitative frameworks, used to evaluate this influence. Following this, case studies and examples illustrate how these approaches are applied in real-world contexts, offering tangible insights into their utility. Finally, the review analyzes the challenges and limitations that complicate these assessments, setting the stage for a concluding synthesis of findings and implications. This structure aims to clarify the tools and strategies available to researchers and policymakers while highlighting areas for further refinement.

## Overview of Evidence Influence

The influence of evidence on policy and programmatic decision-making is a multifaceted phenomenon, encompassing both direct applications and subtler, long-term shifts in understanding. Across the literature, a recurring distinction emerges between instrumental, conceptual, and symbolic uses of evidence, providing a foundational framework for assessing influence. Instrumental use, as described by Herbert (2014), involves the direct application of research findings to shape specific policies or programs, such as when evaluation results lead to immediate programmatic adjustments. In contrast, conceptual use, emphasized by Johnson et al. (2009) and Martinaitis et al. (2018), reflects a more indirect influence, where evidence reshapes policymakers’ perspectives or informs broader strategic thinking without immediate action. Symbolic use, noted in studies like Nduku et al. (2024), occurs when evidence is cited to legitimize pre-existing decisions rather than drive new ones, highlighting its political utility. These distinctions underscore the complexity of influence, as evidence may not always lead to observable outcomes but can still play a critical role in framing debates or justifying actions.

The significance of understanding these pathways lies in their implications for measurement. Direct influence, such as a policy change tied to a specific evaluation, is often easier to trace than the gradual diffusion of ideas through conceptual use. Studies like Tudisca et al. (2018) and Edwards et al. (2019) stress that evidence uptake varies across contexts, influenced by factors such as institutional capacity, political priorities, and stakeholder engagement. For instance, Edwards et al. (2019) illustrate how evidence in African health systems can shift mental health onto the policy agenda conceptually, even if immediate action lags. Similarly, Martinaitis et al. (2018) argue that evaluation systems themselves shape how evidence is framed and used, with some prioritizing accountability (symbolic) over learning (conceptual). This diversity suggests that assessing influence requires approaches that capture both tangible outputs and intangible shifts in discourse or capacity. By recognizing these varied roles, researchers can better design tools to evaluate not just whether evidence is used, but how deeply it permeates decision-making processes, setting the stage for a detailed exploration of methodologies.

## Methodologies and Measures for Assessment

Evaluating the influence of evidence on policy and programmatic decision-making demands a robust set of methodologies and measures capable of capturing both the tangible and intangible dimensions of this complex process. The literature presents a spectrum of approaches, broadly categorized into quantitative measures, which emphasize numerical precision and replicability, and qualitative or mixed-methods strategies, which prioritize contextual depth and narrative insight. These methodologies are not mutually exclusive; rather, they often complement one another, offering a multifaceted lens through which to assess whether and to what extent a specific source of evidence – such as an evaluation – shapes decision-making. This section delves into these approaches, exploring their theoretical underpinnings, practical applications, and comparative strengths, while weaving together insights from across the source material to address the research question comprehensively.

Quantitative measures form a cornerstone of evidence influence assessment, focusing on observable, countable indicators that signal uptake in policy or programmatic contexts. Tudisca et al. (2018) exemplify this approach with their development of 25 indicators for evidence-informed policy-making (EIPM) in health systems, validated through a rigorous Delphi process involving 82 experts across six countries. Indicators such as “Citation of peer-reviewed research articles in policy documents” (Indicator 8) and “Available evidence briefs for policy” (Indicator 10) are measured through content analysis of official documents, providing a direct metric of instrumental use. For instance, counting citations involves systematically reviewing policy papers to identify explicit references to research, offering a clear, replicable way to quantify evidence integration. Similarly, Collado et al. (2017) employ citation counts within their Pyramid Approach, tracking mentions in legislative reports, media, and gray literature to assess awareness, influence, and impact. Their methodology extends beyond simple counts to include web analytics – such as page views and downloads – quantifying the reach of research outputs as a precursor to policy influence. Siar (2023) builds on this with the Results Chain Approach, proposing indicators like the number of policymaker inquiries or legislative references, which can be tallied over time to map a trajectory from research dissemination to policy adoption. These indicators are particularly valuable in structured settings where documentation is robust, enabling longitudinal tracking of evidence uptake.

Another quantitative approach comes from Abadie et al. (2023), who focus on figuring out how much value evidence adds to decision-making. They use a method called empirical Bayes estimation to compare decisions made with extra evidence – like from experiments – to decisions made without it. Their work, based on data from over 8,000 randomized trials in the Cochrane database, looks at whether gathering more evidence is worth the effort and cost. For example, they found that having more precise evidence – like reducing uncertainty about a policy’s effects – can noticeably improve decision outcomes, such as by about 11.6% when uncertainty drops by half. This approach helps policymakers weigh the practical benefits of evidence against the resources needed to collect it, offering a straightforward way to measure its usefulness in real-world choices. Nduku et al. (2024) complement this with formalized EIDM indicators, such as those in the Global EIDM Index, which measure how often policymakers turn to research before deciding, using surveys or administrative data. These quantitative tools excel in providing objective, comparable data, especially for spotting direct influence – like when a policy clearly relies on an evaluation. However, they can miss quieter, undocumented effects, calling for a wider range of methods.

Transitioning to qualitative and mixed-methods approaches, the literature reveals a suite of strategies designed to capture the nuanced, often indirect pathways of evidence influence. Herbert (2014) champions case studies as a primary method, arguing that retrospective analysis of specific evaluations can uncover both instrumental and process use. For instance, a case study might trace how an evaluation of a health intervention led to program continuation (instrumental) or shifted organizational priorities (process), using interviews and document reviews to reconstruct influence narratives. This approach shines in its ability to contextualize findings, revealing the “why” behind uptake that citation counts alone cannot address. MacKillop et al. (2020) and Edwards et al. (2019) advance this with social network analysis (SNA), a method that maps relationships among policymakers, researchers, and knowledge brokers to visualize evidence flows. Edwards et al. (2019) apply SNA in African health systems, identifying key actors – like intermediaries in South Africa – who bridge research and policy, facilitating both instrumental and conceptual use. By quantifying network ties (e.g., frequency of interactions) and qualifying their nature (e.g., trust levels), SNA offers a hybrid lens, blending numerical rigor with qualitative richness to assess systemic influence.

Stakeholder interviews, a staple in qualitative research, are widely utilized to probe decision-makers’ perceptions of evidence use. Johnson et al. (2009) review 41 studies employing interviews to assess evaluation influence, finding that policymakers often report conceptual shifts – e.g., reframing a problem – more readily than instrumental changes. Nelson et al. (2023) echo this, using interviews alongside Knott and Wildavsky’s six-stage typology (reception to impact) to trace how academic research progresses from awareness to policy adoption. However, both caution against self-reporting bias, where respondents may overstate evidence use to appear rational, a limitation mitigated by triangulating with other data sources like policy documents. Process tracing, highlighted by Williams & Lewis (2021) and Nduku et al. (2024), takes this further by dissecting causal pathways. For example, Nduku et al. (2024) propose tracing how a systematic review on climate adaptation influences national regulations, using archival records and interviews to link evidence to policy design. This method excels in attributing influence amidst multiple factors, though it demands significant time and expertise to execute effectively.

Mixed-methods approaches integrate these qualitative insights with quantitative precision, offering a balanced assessment of evidence influence. Siar’s (2023) Influencing Approach exemplifies this, using the RAPID Outcome Assessment – combining document reviews, stakeholder workshops, and follow-up interviews – to evaluate five types of change: attitudinal, behavioral, procedural, policy content, and discursive. An episode study might backtrack a procedural shift, like a new governance rule, to research inputs, blending qualitative narratives with citation data. Similarly, Tudisca et al.’s (2018) Delphi method bridges methodologies by quantitatively scoring expert consensus (e.g., relevance ratings) while qualitatively refining indicators through iterative feedback, ensuring they reflect real-world applicability. Collado et al.’s (2017) Pyramid Approach operationalizes this synthesis, tracking awareness via downloads (quantitative), influence via collaborations (mixed), and impact via policymaker briefings (qualitative), validated through tools like Researchfish and Altmetric. Nelson et al. (2023) further illustrate this with hybrid measures, combining bibliometric analysis (citations in policy papers) with case studies and surveys, revealing how research moves through stages of utilization. This integrative strategy captures both the breadth of evidence reach and the depth of its impact, addressing the limitations of singular approaches.

The diversity of these methodologies reflects the varied nature of evidence influence, from immediate policy shifts to long-term learning. Quantitative measures, like Tudisca et al.’s indicators or Abadie et al.’s value assessment, provide a snapshot of explicit uptake, ideal for settings with strong documentation, such as health or legislative domains. Their precision and scalability make them valuable for comparative analyses, as seen in Siar’s (2023) Results Chain, which tracks inputs to impacts across projects. Yet, they risk oversimplification, missing the subtle conceptual shifts that qualitative methods uncover. Case studies and SNA, for instance, reveal how evidence permeates networks or reshapes priorities, as Edwards et al. (2019) demonstrate in healthcare, where mental health gained salience through researcher-policymaker ties. Mixed-methods approaches, like Collado et al.’s (2017), strike a balance, offering a comprehensive view that aligns with Martinaitis et al.’s (2018) typology of evaluation systems – symbolic, instrumental, and conceptual. For example, a symbolic system might prioritize citation counts for accountability, while a conceptual one leverages process tracing for learning, illustrating how methodology aligns with purpose.

Practical application further highlights these distinctions. Tudisca et al.’s (2018) indicators, measured via surveys and content analysis, assess human resources (e.g., research-trained staff) and documentation (e.g., literature reviews), providing a checklist for EIPM in health policy. In contrast, MacKillop et al.’s (2020) proposed ethnographic studies of knowledge brokering delve into the lived experiences of brokers, uncovering barriers like political context that quantitative metrics overlook. Nduku et al.’s (2024) use of experimental designs, such as RCTs testing evidence briefings, adds a causal dimension, quantifying behavioral changes while qualitative interviews contextualize findings. Recognizing the methodological diversity of these assessment enables researchers to tailor the tool to the task, the specific evidence types (e.g., evaluations vs. systematic reviews) and the decision-making contexts (e.g., rapid response vs. strategic planning). However, their effectiveness hinges on data quality, resource availability, and the ability to mitigate biases – challenges explored later – underscoring the potential value in triangulation to validate findings across methods where such added insight justifies the additional cost. Together, these approaches provide a robust framework to assess not just the presence of influence, but its scope, depth, and sustainability, aligning with the research question’s dual focus on “whether” and “to what extent.”

## Case Studies or Examples

Practical applications of these methodologies illuminate their real-world utility and adaptability across diverse policy contexts. Tudisca et al. (2018) provide a compelling example through their development of 25 indicators for evidence-informed policy-making in European health systems. Using the Delphi method, they engaged 82 experts from six countries in two iterative rounds to refine indicators like “Procedures for ensuring a review of scientific literature relevant to the policy” (Indicator 6). This participatory approach ensured the indicators were both relevant and feasible, subsequently applied through content analysis of policy documents and surveys of policymakers to measure evidence integration. The process demonstrated how a structured, consensus-driven methodology can produce actionable tools to assess instrumental use, such as the presence of research-trained staff in policy teams, offering a model replicable in other health policy settings.

Edwards et al. (2019) offer a contrasting example from South Africa, where a case study approach revealed both instrumental and conceptual evidence use in healthcare policy. Their analysis of knowledge translation strategies showed how research directly informed the development of a primary healthcare service model (instrumental) while also elevating mental health onto the policy agenda (conceptual). By employing stakeholder interviews and document analysis, they traced these influences within a resource-constrained African context, highlighting the role of social networks in facilitating uptake. This case underscores the value of qualitative methods in capturing nuanced impacts that quantitative metrics, like citation counts, might overlook, particularly in settings where formal documentation is limited.

Siar (2023) further enriches this narrative with three frameworks – Pyramid, Influencing, and Results Chain – applied across various policy domains. The Pyramid Approach, for instance, was used by the Philippine Institute for Development Studies to track research impact through downloads (awareness), media mentions (influence), and legislative citations (impact), offering a tiered assessment of evidence uptake. The Influencing Approach, with its associated theory of change, was operationalized through episode studies in development policy, backtracking key policy shifts to research inputs, such as attitudinal changes among policymakers. Meanwhile, the Results Chain, adopted by the Asian Development Bank, mapped inputs like research funding to outputs like policy briefs and outcomes like program adjustments, providing a linear yet comprehensive evaluation. These examples showcase how diverse methodologies can be tailored to specific contexts, from health to development, revealing both direct policy changes and broader ecosystem shifts, and reinforcing the need for flexible, multi-method strategies to capture evidence influence effectively.

## Challenges and Limitations

Despite the robustness of these methodologies, measuring evidence influence faces significant obstacles, as the literature consistently identifies. A primary challenge is the lack of standardized metrics, which hampers comparability across studies and contexts. MacKillop et al. (2020) argue that the absence of agreed-upon frameworks leads to fragmented findings, a sentiment echoed by Edwards et al. (2019) in their review of African health systems, where KT outcomes are inconsistently reported. Herbert (2014) similarly notes methodological fragmentation in evaluation research, with studies employing varied definitions of influence, from citations to attitudinal shifts, complicating synthesis. Tudisca et al. (2018) succinctly capture this issue, stating, “The absence of universally agreed-upon measurement frameworks” (p. 9) limits the ability to benchmark evidence use systematically, a gap that undermines efforts to establish best practices.

Attribution poses another formidable barrier, as policy decisions rarely stem from a single evidence source. Cantarelli et al. (2023) highlight how multiple factors – political priorities, stakeholder interests, and competing information – obscure the causal link between evidence and outcomes. Williams & Lewis (2021) reinforce this, noting that impact often results from cumulative knowledge rather than isolated studies, making it difficult to isolate a specific evaluation’s role. Siar (2023) concurs, emphasizing that policymaking’s non-linear nature, driven by diverse actors and interests, confounds efforts to attribute change directly to research, especially when evidence is one of many influences shaping a decision.

Time lags further complicate measurement, as the influence of evidence may not manifest immediately. For example, Nelson et al. (2023) observe that short-term studies miss long-term conceptual shifts, such as gradual changes in policy discourse, requiring longitudinal approaches that are resource-intensive and rare. This temporal disconnect challenges researchers to design assessments that capture delayed or incremental effects, rather than focusing solely on immediate outcomes.

Political and cognitive biases also distort evidence uptake and measurement. Cantarelli et al. (2023) detail how framing effects and confirmation bias lead policymakers to interpret evidence selectively, aligning it with pre-existing beliefs rather than objective analysis. Langer et al. (2016) add that political feasibility often trumps research findings, with evidence used symbolically to justify decisions rather than inform them. Oliver & Cairney (2019) extend this critique, noting that power dynamics and institutional constraints can sideline evidence, particularly when it conflicts with dominant agendas. These biases not only affect how evidence is used but also how its influence is perceived and reported, introducing subjectivity into even the most rigorous methodologies. Collectively, these challenges underscore the need for adaptive, context-sensitive approaches that account for both the measurable and the elusive aspects of evidence influence.

## Conclusion

The literature demonstrates that a broad range of approaches – including quantitative, qualitative, and mixed-methods – can be used to assess whether and how evidence informs policy and programmatic decision-making. These range from citation-based metrics and formal indicators (e.g., Tudisca et al. 2018; Collado et al. 2017) to more contextualized strategies such as stakeholder interviews, process tracing, and social network analysis (e.g., Johnson et al. 2009; Edwards et al. 2019). Each approach offers distinct strengths: standardized indicators can generate comparable data points across programs, while qualitative or mixed methods capture subtler, longer-term conceptual shifts. Importantly, Abadie et al.’s (2023) Value of Evidence metric, though promising for exploring how precision in evidence might affect decision outcomes, remains largely conceptual and exploratory.

From a practical standpoint, deciding which approaches to adopt hinges on each program’s resources, time horizons, and primary goals. Some methods (e.g., citation tracking, surveys) are relatively streamlined and can be readily integrated into existing monitoring routines. Others, such as ethnographic studies or social network analysis, may require specialized expertise and more substantial investments in data collection, but yield deeper insights into how and why evidence is adopted – or overlooked. Resource constraints, organizational capacity, and specific decision-making contexts should guide the choice of methods rather than a “do everything” approach. While the methodological pluralism highlighted in this review underscores the flexibility to select or combine multiple approaches, these selections must be deliberately prioritized. For instance, teams with limited funding or tight timelines might opt for basic quantitative indicators (e.g., citation counts, references in policy documents), supplemented by a few targeted interviews to capture conceptual shifts. By contrast, teams with more robust capacity may integrate mixed methods designs, triangulating across quantitative metrics, social network data, and process tracing to generate comprehensive evidence of impact.

Ultimately, this review aims to help programs make actionable decisions about which tools best fit their operational needs. Organizations can use simpler quantitative metrics to establish a baseline for evidence use, then add qualitative or network-based methods if they seek richer insights into why evidence does or does not gain traction. Building on these findings, programs can more confidently tailor their evaluation strategies, ensuring that assessments of evidence use align with their strategic goals, resource realities, and organizational cultures. In doing so, they lay the groundwork for evidence-informed practices that not only measure influence but also help nurture it.

## **References**

Abadie, A., Agarwal, A., Imbens, G., Jia, S., McQueen, J., & Stepaniants, S. (2023). Estimating the value of evidence-based decision making. arXiv preprint. <https://arxiv.org/abs/2306.02447>

Belita, E., Squires, J. E., Yost, J., Ganann, R., Burnett, T., & Dobbins, M. (2020). Measures of evidence-informed decision-making competence attributes: A psychometric systematic review. BMC Nursing, 19(44). <https://doi.org/10.1186/s12912-020-00436-8>

Cantarelli, P., Belle, N., & Hall, J. L. (2023). Information use in public administration and policy decision-making: A research synthesis. Public Administration Review, 83, 1667–1686. <https://doi.org/10.1111/puar.13735>

Collado, M., Gerlach, L., Ticse, C., & Hempstead, K. (2017). Considerations for measuring the impact of policy-relevant research. The Foundation Review, 9(4). <https://doi.org/10.9707/1944-5660.1386>

Edwards, A., Zweigenthal, V., & Olivier, J. (2019). Evidence map of knowledge translation strategies, outcomes, facilitators and barriers in African health systems. Health Research Policy and Systems, 17(16). <https://doi.org/10.1186/s12961-019-0419-0>

Herbert, J. L. (2014). Researching evaluation influence: A review of the literature. Evaluation Review, 38(5), 388–419. <https://doi.org/10.1177/0193841X14547230>

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>

Langer, L., Tripney, J., & Gough, D. (2016). The science of using science: Researching the use of research evidence in decision-making. EPPI-Centre, Social Science Research Unit, UCL Institute of Education, University College London.

MacKillop, E., Quarmby, S., & Downe, J. (2020). Does knowledge brokering facilitate evidence-based policy? A review of existing knowledge and an agenda for future research. Policy & Politics, 48(2), 335–353. <https://doi.org/10.1332/030557319X15740848311069>

Martinaitis, Ž., Christenko, A., & Kraučiūnienė, L. (2018). Evaluation systems: How do they frame, generate, and use evidence? Evaluation, 25(1), 46–61. <https://doi.org/10.1177/1356389018802135>

Nduku, P., Ategeka, J., Madonsela, A., Mdlalose, T., Stevenson, J., Shisler, S., Pande, S., & Mahlanza-Langer, L. (2024). What works to increase the use of evidence for policy decision-making: A systematic review protocol. Campbell Systematic Reviews. <https://doi.org/10.1002/cl2.1435>

Nelson, J. P., Lindsay, S., & Bozeman, B. (2023). The last 20 years of empirical research on government utilization of academic social science research: A state-of-the-art literature review. Administration & Society. <https://doi.org/10.1177/00953997231172923>

Oliver, K., & Cairney, P. (2019). The dos and don’ts of influencing policy: A systematic review of advice to academics. Palgrave Communications, 5, Article 21. <https://doi.org/10.1057/s41599-019-0232-y>

Siar, S. (2023). The challenges and approaches of measuring research impact and influence on public policy making. Public Administration and Policy, 26(2), 169–183. <https://doi.org/10.1108/PAP-05-2022-0046>

Tudisca, V., Valente, A., Castellani, T., Stahl, T., Sandu, P., Dulf, D., Spitters, H., Van de Goor, I., Radl-Karimi, C., Syed, M. A., Loncarevic, N., Lau, C. J., Roelofs, S., Bertram, M., Edwards, N., & Aro, A. R. (2018). Development of measurable indicators to enhance public health evidence-informed policy-making. Health Research Policy and Systems, 16(47). <https://doi.org/10.1186/s12961-018-0323-z>

Williams, K., & Lewis, J. M. (2021). Understanding, measuring, and encouraging public policy research impact. Australian Journal of Public Administration, 80(3), 554–564. <https://doi.org/10.1111/1467-8500.12506>