Available on: https://jmrr.ep-journals.org/ E-ISSN: 3051-3588, P-ISSN: 3051-357X

Journal of Management Research and Review

Volume: 02 Issue: 07 July-2026

Performance Measurement Frameworks to Improve Service Delivery Outcomes in Urban Public Agencies: A Systematic Review

Emmanuel Owusu Agyei
University of Memphis, USA

Matthew Oman-Amoako
Independent Researcher, Texas, USA

Published Date: July 16, 2026 **Article DOI: 10.65150/EP-jmrr/V2E7/2026-04**

ABSTRACT: Performance measurement frameworks have been widely adopted across public sector organizations as instruments of accountability and service improvement, yet the evidence on whether they reliably translate into improved service delivery outcomes in urban public agencies remains geographically uneven. This review addresses that gap by systematically examining the design characteristics, implementation conditions, and service delivery effects of performance measurement frameworks adopted by urban public agencies across diverse institutional and geographic contexts, while identifying the conditions that enable or constrain effective implementation. In line with PRISMA 2020 guidelines, searches were conducted across multiple databases, covering peer-reviewed journal articles published from 2010 to the present; 26 studies met all inclusion criteria and were retained for analysis. Thematic synthesis guided the analytical process. The evidence indicates that no single framework type predominates globally, with adoption shaped by institutional context and administrative capacity as much as by technical design considerations. The average effect of performance management on public organizational performance is modest, and improvement is conditional on the quality of framework design, purposeful managerial engagement with performance data, and the strategic alignment between measurement systems and organizational priorities. Governance design alignment, political incentive structures, and managerial culture emerged as the primary determinants of whether implementation produces substantive service improvement or ceremonial compliance. Overall, the value of performance measurement frameworks is recognized in the institutional seriousness and managerial commitment with which they are embedded and sustained.

KEYWORDS: Performance measurement frameworks, Service delivery, Urban public agencies, Public administration.

1.0 INTRODUCTION

Performance measurement frameworks have become central instruments of public sector governance over recent decades, driven by New Public Management reform waves that reoriented public organizations toward demonstrable outcomes rather than procedural compliance (Kolk, 2022). These frameworks, which consist of structured systems of indicators, metrics, and evaluation processes through which public organizations assess their efficiency, effectiveness, and accountability, translate organizational goals into measurable criteria that inform managerial decision-making (Agyei, 2026a). The post-2008 fiscal austerity environment intensified this trajectory, placing sustained pressure on public agencies to demonstrate value, particularly at the local level where service delivery obligations are most immediately experienced by citizens (Downe et al., 2010). The relationship between performance measurement and service delivery outcomes has since attracted substantial scholarly attention in public administration, though the evidence base remains uneven across governance contexts and institutional settings (Agyei, 2026b).

Urban public agencies occupy a distinctive position within this landscape. As government bodies and quasi-governmental organizations operating at the municipal or metropolitan level, they are characterized by direct citizen interaction, political accountability to local governments, and operational complexity arising from dense, diverse urban populations (Beeri et al., 2018). Service delivery in urban contexts is evaluated along dimensions of access, quality, timeliness, and equity, making the design and implementation of performance measurement frameworks consequential for urban governance outcomes in ways that extend beyond standard administrative efficiency concerns (Obodai et al., 2025). The intellectual tradition framing this inquiry draws from New Public Management and its successor frameworks, which reconceived public agencies as outcome-oriented organizations whose performance could be systematically measured and improved through structured evaluation (Lapuente & Van de Walle, 2020).

License: This is an open access article under the CC BY 4.0 license: https://creativecommons.org/licenses/by/4.0/

Despite sustained investment in performance measurement systems across public sector organizations globally, evidence on whether these frameworks reliably translate into improved service delivery outcomes remains contested (Gerrish, 2015). Meta-analytic evidence indicates that the average effect of performance management on public organizational performance is modest, and that variation in outcomes is substantially explained by differences in managerial practices, institutional design, and measurement system characteristics (Johnsen et al., 2024). Urban public agencies face measurement challenges given the layered accountability structures of municipal governance (Beeri et al., 2018). The existing literature lacks a systematic synthesis that examines, across diverse geographic and institutional contexts, what types of performance measurement frameworks urban public agencies apply, whether these frameworks measurably improve service delivery, and what conditions determine the effectiveness of their implementation. This review addresses that gap.

This review aims to systematically examine the evidence on performance measurement frameworks adopted by urban public agencies, with attention to their design characteristics, implementation conditions, and effects on service delivery outcomes.

Three central constructs are defined as they are used throughout this review. Performance measurement frameworks are structured systems of indicators, metrics, and evaluation processes through which public organizations assess the efficiency, effectiveness, and accountability of their operations and outputs; these frameworks translate organizational goals into measurable criteria and provide the basis for managerial decision-making and public accountability (Kolk, 2022). Service delivery refers to the processes and mechanisms through which public agencies provide goods, services, and regulatory functions to citizens and communities; in urban contexts, service delivery is evaluated along dimensions of access, quality, timeliness, and equity (Beeri et al., 2018). Urban public agencies are government bodies and quasi-governmental organizations 1operating at the municipal or metropolitan level, responsible for administering public services within urban jurisdictions and characterized by direct citizen interaction, political accountability to local governments, and operational complexity arising from dense, diverse urban populations (Walker et al., 2011).

2.0 METHODOLOGY

2.1 Research Design

This study adopts a systematic review design to synthesize peer-reviewed evidence on performance measurement frameworks in urban public agencies and their relationship to service delivery outcomes. This design enables the review to move beyond selective citation toward a structured account of what the evidence collectively indicates, under what conditions, and with what degree of consistency. The review was conducted in accordance with PRISMA 2020 guidelines, which establish standards for transparent reporting of systematic reviews across search strategy, study selection, data extraction, and synthesis.

2.2 Search Strategy and Data Sources

Searches were conducted across Scopus, Web of Science, and the Public Administration Abstracts database, covering the period from 2010 to 2025. The lower boundary of 2010 reflects the significant renewal of performance measurement scholarship following the post-2008 fiscal austerity period, which reoriented the field toward questions of efficiency, accountability, and service outcomes in local and urban governance. The search strategy was developed around three key conceptual clusters, with synonyms identified for each construct. These were combined using Boolean operators as follows: (“performance measurement frameworks” OR “performance management systems” OR “performance indicators” OR “accountability frameworks”) AND (“service delivery” OR “public services” OR “service outcomes” OR “citizen-facing administration”) AND (“urban public agencies” OR “local government” OR “municipal agencies” OR “metropolitan administration”).

Grey literature such as policy documents, government reports, and institutional publications were excluded from this review. While grey literature may contain relevant practice-based insight, its variable quality and limited methodological disclosure make systematic integration inconsistent with the rigor this review requires.

2.3 Study Selection Process

Study selection followed a two-stage screening process designed to minimize inclusion bias. In the first stage, titles and abstracts of all retrieved records were screened independently against the inclusion and exclusion criteria, with records clearly outside scope eliminated at this point. In the second stage, full texts of the remaining records were assessed for eligibility, with particular attention to the centrality of performance measurement frameworks and urban public agency contexts in the study's analytical focus. An initial database search yielded approximately 645 records. After removal of duplicates, 417 records proceeded to title and abstract screening, of which 104 were judged potentially eligible and advanced to full-text review. Following full-text assessment, 26 studies met all inclusion criteria and were retained for synthesis as illustrated in the figure below.

License: This is an open access article under the CC BY 4.0 license: https://creativecommons.org/licenses/by/4.0/

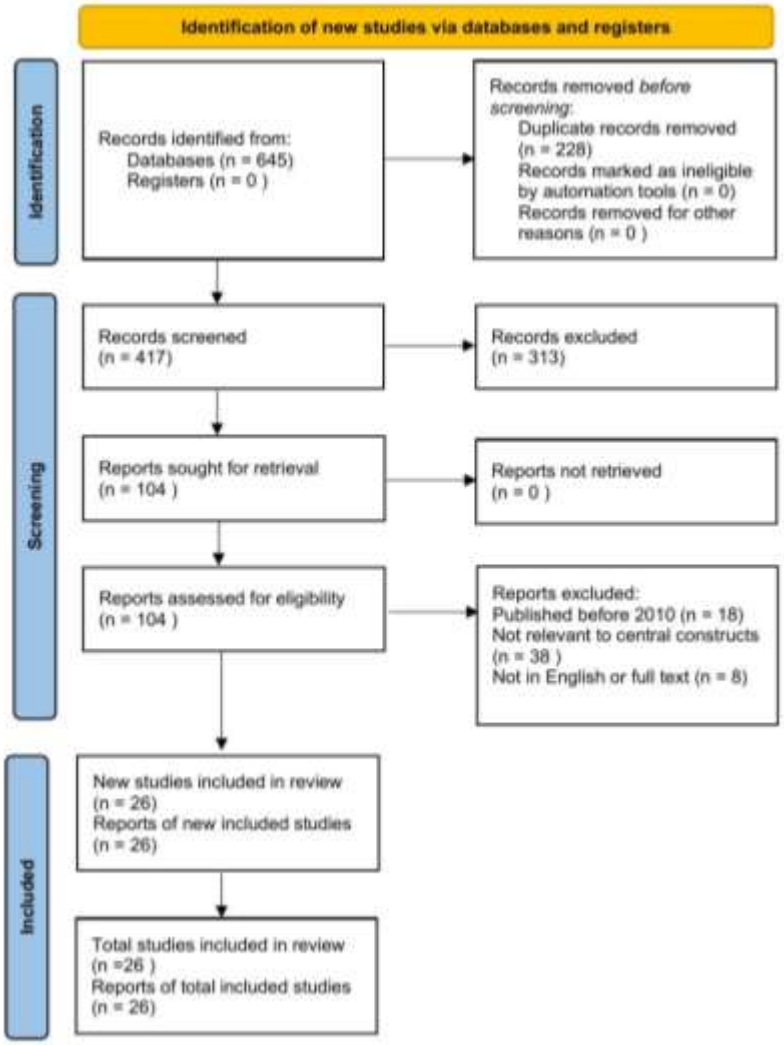


Figure 1: PRISMA 2020 Diagram
Source: Author's Construct, 2026

2.4 Inclusion and Exclusion Criteria

Included studies met the following criteria:

1. Published in peer-reviewed journals from 2010 to 2025,
2. Directly relevant to performance measurement frameworks, service delivery outcomes, or urban public agencies as defined in this review
3. Available in full text and
4. Published in English or with a verified English translation.

Excluded from consideration were

1. Sources predating 2010,
2. Non-peer-reviewed material
3. Grey literature,
4. Studies whose central focus did not engage with the key constructs of performance measurement or urban service delivery, and
5. Duplicate records identified across databases.

2.5 Data Extraction

A structured data extraction protocol was applied consistently across all 26 included studies. Extraction was guided directly by the three research objectives to ensure that information captured from each source was organized in a manner that would support systematic analysis. This approach ensured that data relevant to framework typology, service delivery effects, and institutional implementation conditions were extracted in comparable form across studies, reducing the interpretive inconsistency that can arise from unstructured reading of the literature.

License: This is an open access article under the CC BY 4.0 license: https://creativecommons.org/licenses/by/4.0/

Table 1: Study Characteristics

Author & Year	Country / Context	Study Design	Sample / Unit of Analysis	Data Collection Method	Key Focus
Andersen & Nielsen (2019)	Denmark: public sector organizations	Quantitative empirical study	Public sector units and their performance outcomes	Administrative data and performance records	Conditions under which performance information use leads to improved organizational outcomes
Andrews et al. (2011)	England / UK: local government	Quantitative empirical study	English local authorities	Survey data combined with administrative performance records	Effects of implementation style and strategic orientation on public service performance
Beeri et al. (2018)	Israel: local government authorities	Quantitative empirical study	Local government authorities over multiple years	Survey and administrative data from local authority records	Relationship between performance management frameworks, citizen satisfaction, and trust in local government
Brusca et al. (2017)	Italy and Spain: local government	Comparative policy analysis	Local governments in Italy and Spain	Document analysis and policy review	Adoption and compliance effects of performance measurement frameworks in Southern European local governments
Downe et al. (2010)	England / UK: local government	Comparative theoretical and empirical analysis	Local performance assessment frameworks in England	Document and case analysis of framework designs	Competing theories of public service improvement and local performance assessment frameworks
Jie (2015)	China: local government	Qualitative case study	Local government units in China	Document analysis and qualitative data	Political rationality versus technical rationality in China's target-based performance measurement system
Gębczyńska & Brajer-Marczak (2020)	International / multi-jurisdictional	Conceptual literature review	Performance measurement models in public administration literature	Systematic review of peer-reviewed literature	Typology and comparative review of performance measurement models used in public administration
Gerrish (2015)	International / multi-country	Meta-analysis	Studies on performance management in public organizations	Systematic literature search, coding, and statistical aggregation	Effect of performance management systems on organizational performance across public sector contexts
Gomes & de Azevedo (2024)	International / multi-jurisdictional	Bibliometric analysis and literature review	Academic publications on Balanced Scorecard in public administration	Bibliometric and content analysis	Trajectory and application of the Balanced Scorecard across public administration and local government
Han & Moynihan (2022)	USA: public sector organizations	Quantitative empirical study	Public sector managers and organizations	Survey data	Whether and how managerial use of performance information produces measurable organizational improvements
Johnsen et al. (2024)	Norway: public sector organizations	Quantitative survey study	Public sector organizations across government levels	Structured survey questionnaire	Performance measurement system design as a link between strategic planning and performance information use
Kolk (2022)	International / multi-jurisdictional	Systematic literature mapping	Survey-based research on performance measurement in public organizations	Systematic database search and literature mapping	Cumulative findings on performance measurement practices, antecedents, and effects in public organizations

License: This is an open access article under the CC BY 4.0 license: https://creativecommons.org/licenses/by/4.0/

Král (2022)	International: local public sector	Systematic review	Academic literature on performance measurement in local government	Systematic literature search across major academic databases	History, evolution, and trends in performance measurement research in the local public sector
Lapuente & Van de Walle (2020)	Multiple European countries: central government	Quantitative empirical study	Senior public executives across European countries	Cross-national survey data	Effects of New Public Management reform practices on public service quality and effectiveness
Lewis (2015)	International / conceptual	Conceptual and theoretical	Not applicable: conceptual framework	Theoretical synthesis and conceptual analysis	Politics and institutional consequences of performance measurement in public organizations
Meng & Li (2025)	International / multi-country	Meta-analysis	Studies on performance information effects in public sector	Systematic literature search, coding, and statistical aggregation	Impact of public sector performance information on stakeholder attitudes and behaviors
Moynihan & Kroll (2016)	USA: federal government agencies	Quantitative empirical study	Federal government managers and agencies	Survey of data of federal managers	Effect of performance management routines on purposeful use of performance data in public agencies
Moynihan & Pandey (2010)	USA: state and local government	Quantitative empirical study	Public sector managers across government levels	Survey data	Organisational, cultural, and individual-level factors determining whether public managers use performance information
Moullin (2017)	United Kingdom: public and voluntary sectors	Case study	Public sector organizations across organizational boundaries	Case analysis including programme documentation and evaluation data	Public Sector Scorecard as a performance improvement framework applied in public and voluntary sector contexts
Northcott & Ma'amora Taulapapa (2012)	New Zealand: local government organizations	Mixed methods	Local government organizations	Questionnaire survey followed by interviews with selected managers	Balanced Scorecard adoption rates, utility, and implementation challenges in public sector organizations
Sandfort et al. (2019)	USA: local welfare agencies	Qualitative case study	Local agencies administering social welfare programmes	Interviews and document analysis	How performance management regimes shape managerial and operational practice at the local agency level
Schoeman & Chakwizira (2023)	South Africa: municipalities	Mixed methods	South African local government and municipal planning documentation	Document analysis and theoretical framework analysis	Developing a simplified performance management tool for service delivery in local government
Tillema et al. (2010)	Less developed countries: international	Conceptual framework with empirical illustration	Public sector organizations in less developed country contexts	Conceptual framework development drawing on institutional theory	Changing role of performance measurement in public sector organizations in less developed countries
Turley et al. (2015)	Ireland: local government	Framework development and empirical application	Irish local authorities	Analysis of local government financial records and accounts	Development and testing of a financial performance measurement framework for local governments
Virani & van der Wal (2023)	International / conceptual	Conceptual and theoretical	Not applicable: causal model development	Conceptual analysis and theoretical synthesis	Causal model examining how governance design alignment enables or inhibits effective public sector performance regimes

License: This is an open access article under the CC BY 4.0 license: https://creativecommons.org/licenses/by/4.0/

Walker et al. (2011)	England / UK: local government	Quantitative empirical study	English local government councils	Survey data combined with administrative performance assessments and citizen satisfaction data	Effect of market orientation on public service performance across multiple stakeholder perspectives

2.6 Quality Appraisal

All included sources were subject to a narrative quality assessment prior to inclusion in the evidence base. Appraisal addressed four dimensions: peer-review status, as a baseline indicator of methodological scrutiny; clarity of research design and methods, to ensure findings were traceable to a defensible analytical process; relevance of the study's core focus to the central constructs of this review; and the internal rigor of the findings, including whether conclusions were proportionate to the evidence presented. This process ensured that the synthesis rests on sources whose findings can be reasonably treated as credible contributions to the evidence base.

Table 2: Quality Appraisal Summary

Author & Year	Peer-Reviewed	Methodological Clarity	Relevance to Central Constructs	Rigour of Findings	Overall Rating	Assessment Outcome
Andersen & Nielsen (2019)	Yes	High	High	High	High	Included
Andrews et al. (2011)	Yes	High	High	High	High	Included
Beeri et al. (2018)	Yes	High	High	High	High	Included
Brusca et al. (2017)	Yes	High	High	Moderate	Moderate	Included
Downe et al. (2010)	Yes	Moderate	High	Moderate	Moderate	Included
Jie (2015)	Yes	Moderate	High	Moderate	Moderate	Included
Gębczyńska & Brajer-Marczak (2020)	Yes	Moderate	High	Moderate	Moderate	Included
Gerrish (2015)	Yes	High	High	High	High	Included
Gomes & de Azevedo (2024)	Yes	High	High	High	High	Included
Han & Moynihan (2022)	Yes	High	High	High	High	Included
Johnsen et al. (2024)	Yes	High	High	High	High	Included
Kolk (2022)	Yes	High	High	Moderate	Moderate	Included
Král (2022)	Yes	High	High	Moderate	Moderate	Included
Lapuente & Van de Walle (2020)	Yes	High	High	High	High	Included
Lewis (2015)	Yes	Moderate	High	Moderate	Moderate	Included
Meng & Li (2025)	Yes	High	High	High	High	Included
Moynihan & Kroll (2016)	Yes	High	High	High	High	Included
Moynihan & Pandey (2010)	Yes	High	High	High	High	Included
Moullin (2017)	Yes	Moderate	High	Moderate	Moderate	Included
Northcott & Ma'amora Taulapapa (2012)	Yes	High	High	Moderate	Moderate	Included
Sandfort et al. (2019)	Yes	High	High	Moderate	Moderate	Included
Schoeman & Chakwizira (2023)	Yes	Moderate	High	Moderate	Moderate	Included
Tillema et al. (2010)	Yes	Moderate	High	Moderate	Moderate	Included

License: This is an open access article under the CC BY 4.0 license: https://creativecommons.org/licenses/by/4.0/

Turley et al. (2015)	Yes	High	High	Moderate	Moderate	Included
Virani & van der Wal (2023)	Yes	Moderate	High	Moderate	Moderate	Included
Walker et al. (2011)	Yes	High	High	High	High	Included

2.7 Data Analysis

Thematic synthesis was selected as the method of data analysis for this review. This approach is appropriate where the literature is qualitatively diverse, spans multiple study designs and governance contexts, and where the objective is to identify patterns, themes, and interpretive insights across sources rather than to pool statistical outcomes. The literature on performance measurement frameworks in urban public agencies is heterogeneous in design, geographically dispersed, and resistant to meta-analytic aggregation given the contextual variability of its findings. Analysis proceeded by coding findings from each included source in relation to the research objectives, grouping codes into descriptive themes, and developing analytical interpretations that extended beyond individual study conclusions to offer cross-study insight. At every stage, the analytical process remained anchored to the specific objectives of this review.

2.8 Ethical Considerations

This systematic review is based entirely on publicly available peer-reviewed literature and involves no primary data collection or human subjects. All sources are attributed accurately and cited in accordance with APA 7th edition conventions. No conflict of interest influenced the selection, appraisal, or interpretation of sources. The review is committed to representing the literature accurately, including findings that challenge dominant assumptions about the effectiveness of performance measurement frameworks.

3.0 RESULTS

Table 3: Key Frameworks, Instruments, and Tools Identified in Included Studies

Author & Year	Framework / Instrument / Tool	Type	Application Area	Notes
Gomes & de Azevedo (2024)	Balanced Scorecard	Performance Model	Public administration: local government	Reviewed across an extended period of application in the public sector; found predominantly used in local government for performance management and measurement purposes
Northcott & Ma'amora Taulapapa (2012)	Balanced Scorecard	Performance Model	Urban public agencies: local government	Adopted primarily for reporting functions rather than active performance management; implementation challenges identified in public sector contexts
Gębczyńska & Brajer-Marczak (2020)	Balanced Scorecard	Performance Model	Public administration: multiple jurisdictions	Identified as widely applied across public sector settings; frequently modified to prioritize citizen and stakeholder perspectives over financial returns
Gębczyńska & Brajer-Marczak (2020)	Common Assessment Framework (CAF)	Performance Model	Public administration: European public sector	Quality management model adapted for public sector organizations; identified alongside the Balanced Scorecard as a prominent framework in European public administration
Gębczyńska & Brajer-Marczak (2020)	Performance Prism	Analytical Tool	Public administration	Multi-stakeholder performance measurement model proposed as an alternative to the Balanced Scorecard for capturing broader value dimensions in public organizations
Jie (2015)	Target-responsibility system	Performance Model	Urban public agencies: Chinese local government	Assigns quantitative targets through political-administrative hierarchies; evidence indicates divergence between indicator compliance and substantive service delivery improvement
Turley et al. (2015)	Composite financial performance framework	Measurement Index	Local government: Irish municipalities	Developed to address the absence of standardized financial measurement tools at local government level; assesses fiscal health as a proxy for service delivery capacity

License: This is an open access article under the CC BY 4.0 license: https://creativecommons.org/licenses/by/4.0/

Moullin (2017)	Public Sector Scorecard	Performance Model	Urban public agencies: UK public and voluntary sectors	Adapts the Balanced Scorecard to align with public sector values and accountability structures; applied across organizational boundaries with a focus on citizen outcomes
Johnsen et al. (2024)	Performance measurement system design framework	Conceptual Framework	Public sector organizations: local and central government	Examine how the design of performance measurement systems mediates the relationship between strategic planning and performance information use
Virani & van der Wal (2023)	Interactional causal model for governance design alignment	Conceptual Framework	Public sector performance regimes: international	Proposes alignment between governance design choices and the institutional logic of performance regimes as a condition for effectiveness
Gerrish (2015)	Performance management effect model	Analytical Tool	Public organizations: international	Examines the conditional effect of performance management systems on organizational performance; identifies management practice quality as a significant moderating condition
Meng & Li (2025)	Information propagation framework for performance information effects	Analytical Tool	Public sector: international	Draws on social and cognitive psychology to model the conditions under which public sector performance information positively affects stakeholder attitudes and behaviors
Tillema et al. (2010)	Demand-supply framework for performance measurement in less developed countries	Conceptual Framework	Public sector: less developed countries	Distinguishes stakeholder groups by their divergent information demands on public sector organizations and maps how these shape performance measurement supply
Sandfort et al. (2019)	Performance management regime	Performance Model	Urban public agencies: local welfare agencies	Examines how a federally imposed performance regime shapes operational practice at street level; highlights tension between accountability demands and adaptive service delivery
Moynihan & Kroll (2016)	Government performance management routines	Performance Model	Public agencies: federal government	Structured performance routines associated with higher rates of purposeful performance data use in federal agencies
Schoeman & Chakwizira (2023)	Simplified performance management tool	Performance Model	Urban public agencies: South African municipalities	Developed to address measurement capacity gaps in developing country local government; incorporates service prioritization and standards-setting components

This chapter presents and interprets the findings of the systematic review. Thematic synthesis guided the analytical process, with findings grouped into coherent cross-study patterns rather than individual source summaries. Each of the three subsections addresses one objective sequentially. A discussion section follows the findings, interpreting the cumulative weight of evidence and its implications for public administration scholarship and urban governance practice.

3.1 Types of Performance Measurement Frameworks Applied in Urban Public Agencies

The literature reveals a well-populated typology of performance measurement frameworks in use across urban public agencies, though the distribution of framework types is neither uniform across governance contexts nor consistent in its design logic. The dominant frameworks documented across the included sources fall into broadly recognizable categories: balanced scorecard adaptations, target-based and indicator-driven systems, composite financial performance frameworks, and hybrid models that combine accountability reporting with strategic planning. Each carries distinct design assumptions about what performance means, how it should be measured, and who the primary audience for measurement information is.

The Balanced Scorecard has attracted the most sustained attention as a framework applied in local government and urban public agencies. Its appeal lies in its multidimensional structure, which moves beyond financial indicators to incorporate service process, stakeholder, and learning dimensions (Gomes & de Azevedo, 2024). In New Zealand local government organizations, the framework was adopted primarily for performance measurement and reporting functions, though its use for active performance management remained comparatively underdeveloped (Northcott & Ma'amora Taulapapa, 2012). The evidence suggests that in public sector contexts, the Balanced Scorecard is frequently modified to prioritize citizen and stakeholder perspectives over financial returns, reflecting the fundamentally different accountability logic of urban public agencies (Gębczyńska & Brajer-Marczak, 2020). The challenge of adapting a private sector instrument to the mission-driven, politically accountable environment of municipal government has produced a range of hybrid configurations, with no standard design consistently applied across jurisdictions (Gomes & de Azevedo, 2024).

License: This is an open access article under the CC BY 4.0 license: https://creativecommons.org/licenses/by/4.0/

Target-based systems represent a structurally distinct approach in which performance is operationalized through specific quantitative targets assigned to organizational units, with compliance monitored against predetermined thresholds. In Chinese local government, this model has been formalized through what the literature characterizes as a target-responsibility system, in which political and administrative hierarchies assign measurable objectives to subordinate agencies (Jie, 2015). This approach reflects a fundamentally political logic of performance measurement: targets are not merely technical instruments but expressions of governance priorities, and the selection of what is measured and what is excluded is itself a political act (Lewis, 2015). The Chinese case illustrates how target-based systems can produce significant divergence between formal performance compliance and substantive service delivery improvement, as agencies optimize toward measured indicators rather than underlying outcomes (Jie, 2015).

Financial performance frameworks constitute a third category, designed to assess the fiscal health and resource management efficiency of local governments as proxy indicators of their service delivery capacity. In Irish municipalities, a composite financial performance framework was developed and tested to address the absence of standardized financial measurement instruments at the local government level (Turley et al., 2015). This framework focused on solvency, liquidity, and revenue performance, reflecting the post-austerity governance context in which fiscal accountability had become a precondition for service continuity. The evidence indicates that financial performance frameworks operate at a level of abstraction from direct service delivery measurement, capturing the resource conditions that enable or constrain service provision rather than assessing its quality or outcomes directly (Turley et al., 2015).

The Public Sector Scorecard, developed as an explicit adaptation of the Balanced Scorecard for the values and accountability structures of the public and voluntary sectors, represents a more contextually attuned framework (Moullin, 2017). Case study evidence from the United Kingdom suggests that this instrument is capable of producing measurable service outcome improvements when applied across organizational boundaries, particularly when implementation maintains focus on citizen outcomes rather than internal process compliance (Moullin, 2017). The distinction between frameworks designed primarily for external accountability reporting and those designed to support internal managerial learning emerges as a structurally important typological divide in the literature (Johnsen et al., 2024). Frameworks oriented toward external reporting tend to generate compliance behavior; those oriented toward learning and internal decision-making tend to produce more adaptive managerial responses (Kolk, 2022).

Geographic variation in framework adoption is pronounced. In Southern European contexts, performance measurement in local government has been shaped by legislative mandates rather than voluntary managerial adoption, producing formal compliance with framework requirements that does not necessarily translate into substantive use of performance data in service improvement decisions (Brusca et al., 2017). In less developed country contexts, the range of frameworks applied is narrower, constrained by data infrastructure limitations and the dominance of external donor-driven reporting requirements over internally generated measurement systems (Tillema et al., 2010). In South African municipalities, the development of simplified performance management tools reflects a practical response to the gap between the complexity of established framework models and the administrative capacity available to implement them (Schoeman & Chakwizira, 2023). The literature collectively suggests that no single framework type predominates across urban public agencies globally, and that framework selection is as much a product of institutional history and political context as it is of deliberate technical design.

3.2 The Extent to Which Performance Measurement Frameworks Improve Service Delivery Outcomes

The evidence on whether performance measurement frameworks translate into improved service delivery outcomes in urban public agencies is mixed, with the overall direction of findings pointing toward conditional rather than automatic improvement. The most comprehensive synthesis of quantitative evidence indicates that the average effect of performance management on public organizational performance is small, though the magnitude of improvement increases substantially when frameworks are implemented using practices associated with high-quality design and managerial engagement (Gerrish, 2015). This finding is foundational to understanding the broader literature: the question is not whether performance measurement frameworks can improve service delivery, but under what conditions they are likely to do so.

Performance information, when actively used by managers, has a measurable association with improved organizational outcomes. The evidence indicates that integrating performance data into decision-making and problem-solving, rather than treating it as a compliance output is a key mediating variable between framework adoption and service delivery improvement (Han & Moynihan, 2022). Where managers engage with performance information as an active resource for organizational learning, the evidence suggests that framework-linked improvements in service outcomes are more consistently observed (Andersen & Nielsen, 2019). The distinction between passive data generation and active data use is therefore analytically significant: organizations that produce performance information without embedding it in managerial routines are unlikely to realize service delivery gains (Moynihan & Kroll, 2016).

The relationship between framework design characteristics and service delivery outcomes has also received sustained attention. The evidence suggests that frameworks which establish a coherent connection between strategic goals, measurable indicators, and the operational priorities of service-delivering units produce more consistent outcome improvements than those where these elements are misaligned (Johnsen et al., 2024). In local government contexts where strategy formulation and performance measurement operate as parallel rather than integrated processes, the likelihood that measurement activity will produce service improvements is reduced (Johnsen et al., 2024). The design of the measurement system emerges as a critical determinant of its effectiveness.

The impact of performance measurement on citizen-facing service delivery outcomes has been examined in urban governance contexts with mixed results. In Israeli local government, sustained engagement with performance management frameworks was associated with improvements in citizen satisfaction and trust, though the relationship was moderated by governance quality and the extent to which framework implementation was embedded in institutional accountability structures (Beeri et al., 2018). The evidence further indicates that performance information has a positive effect on

License: This is an open access article under the CC BY 4.0 license: https://creativecommons.org/licenses/by/4.0/

stakeholder attitudes and behavioral responses, particularly when information is communicated clearly and in accessible forms (Meng & Li, 2025). However, the magnitude of these effects varies by governance context and by the characteristics of the performance information itself.

The evidence from New Public Management reform contexts presents a more ambivalent picture. Market-oriented performance frameworks applied in English local government produced measurable improvements in citizen satisfaction with local services, though their effects on managerial and regulatory assessments of performance were negligible (Walker et al., 2011). This divergence between citizen-reported outcomes and internal performance assessments reflects a fundamental measurement challenge: different performance measurement frameworks privilege different stakeholder perspectives, and improvements observable from one vantage point may not be visible from another. The evidence from comparative European analysis similarly indicates that performance measurement frameworks in Italian and Spanish local governments have not consistently generated improved service delivery outcomes, suggesting that the institutional conditions of implementation bear as much influence on outcomes as the technical design of the framework (Brusca et al., 2017).

Implementation style also carries independent weight in determining service delivery outcomes. The evidence indicates that neither rational nor incremental implementation approaches reliably produce performance improvements in isolation; rather, improvements are most consistently observed when implementation style is aligned with the strategic orientation of the organization (Andrews et al., 2011). This finding suggests that the pathway from framework adoption to service delivery improvement is neither direct nor guaranteed, and that organizational fit between measurement instrument, managerial culture, and strategic direction shapes the outcome.

3.3 Institutional, Political, and Organizational Conditions Enabling or Constraining Implementation

The evidence on conditions that shape implementation of performance measurement frameworks in urban public agencies is among the most developed and analytically consistent in the literature. Three clusters of conditions emerge from the synthesis: institutional design and governance of fit, political dynamics and incentive structures, and organizational capacity and managerial behavior.

The alignment between governance design and the logic embedded in a performance measurement framework is a primary determinant of implementation effectiveness. Where frameworks are built on assumptions about accountability, incentives, and managerial discretion that conflict with the actual governance environment of the urban public agency, the result is frequently ceremonial adoption, i.e. formal compliance without substantive integration into service management (Virani & van der Wal, 2023). The evidence suggests that performance regimes are most effective when their design assumptions are congruent with the institutional conditions in which they operate, including the degree of managerial autonomy available, the structure of political oversight, and the prevailing accountability norms (Virani & van der Wal, 2023). Misalignment on any of these dimensions creates conditions under which measurement activity is decoupled from service delivery decision-making.

Political conditions shape performance measurement implementation in ways that the technical literature has historically underacknowledged. Performance measurement frameworks do not operate as neutral instruments; they are embedded in political structures that determine what is measured, whose priorities are reflected in the indicator's set, and how results are used (Lewis, 2015). In local government settings characterized by strong central oversight, the imposition of performance frameworks from above tends to produce compliance behavior oriented toward satisfying external accountability demands rather than improving internal service management (Downe et al., 2010). The evidence from China illustrates how political rationality can systematically override technical measurement logic, producing indicator systems that serve political legitimation functions rather than service delivery improvement objectives (Jie, 2015). This dynamic is not unique to authoritarian governance contexts; the evidence from English local government similarly documents how centrally mandated performance assessment frameworks generated behavior oriented toward measured indicators at the expense of unmeasured service dimensions (Downe et al., 2010).

The institutional context of less developed countries presents a distinct configuration of constraining conditions. In these settings, performance measurement frameworks face demand-side pressures from external funding bodies whose information requirements may diverge substantially from the measurement priorities of local service managers, creating systems that generate data for external audiences rather than for internal improvement (Tillema et al., 2010). The absence of reliable baseline data limited statistical capacity, and fragmented administrative systems further constrain the implementation of technically sophisticated frameworks (Schoeman & Chakwizira, 2023). The evidence suggests that in these contexts, simplified and contextually adapted measurement tools are more likely to produce implementation traction than wholesale adoption of frameworks designed for high-capacity governance environments (Schoeman & Chakwizira, 2023).

At the organizational level, managerial behavior and the cultural orientation of the agency are significant enabling conditions. The evidence indicates that managers who use performance information to interrogate service delivery processes, set priorities, and communicate results generate greater organizational performance improvements than those who treat frameworks as administrative requirements (Moynihan & Pandey, 2010). Performance management routines that create structured opportunities for managerial engagement with data, such as goal review processes, data-driven governance meetings, and cross-agency performance dialogue, are associated with higher rates of purposeful information use (Moynihan & Kroll, 2016). The capacity of the organization to learn from performance information, rather than merely report it, is a consistent enabling condition across the literature (Andersen & Nielsen, 2019).

Sandfort et al. (2019) draw attention to a further dimension: the implementation environment created by performance management regimes at the local agency level. Where performance regimes impose rigid accountability requirements on street-level agencies without building the managerial infrastructure to use performance data constructively, implementation produces compliance at the cost of adaptive service delivery. The evidence collectively suggests that enabling conditions for effective implementation include not only governance design and political will, but also the cultivation of a performance-oriented managerial culture within urban public agencies themselves.

License: This is an open access article under the CC BY 4.0 license: https://creativecommons.org/licenses/by/4.0/

3.4 Discussion

The cumulative weight of findings across the three objectives presents a picture of performance measurement in urban public agencies that is substantively more conditional, contextually dependent, and politically shaped than the normative case for these frameworks typically acknowledges. The evidence does not support a generalized conclusion that the adoption of performance measurement frameworks reliably produces service delivery improvement; rather, it supports the view that improvement is a conditional outcome, contingent on the quality of framework design, the degree of managerial engagement with performance information, and the alignment between measurement logic and the institutional environment in which urban public agencies operate (Král, 2022). This finding connects directly to the intellectual tradition of the New Public Management premise that structured performance measurement translates organizational goals into improved outputs has not been invalidated, but its validity is shown to be highly contingent on conditions that the reform literature has often treated as givens rather than variables requiring active management (Downe et al., 2010).

On the first objective, the review establishes that performance measurement frameworks applied in urban public agencies span a range of design types, including Balanced Scorecard adaptations, target-based systems, financial performance frameworks, and hybrid models, with no single approach dominating globally and significant variation driven by institutional history, political context, and administrative capacity (Lapuente & Van de Walle, 2020). On the second objective, the evidence suggests that performance measurement frameworks can improve service delivery outcomes, but that the effect is modest on average and substantially larger where frameworks are well designed, actively used by managers, and aligned with organizational strategic priorities (Gomes & de Azevedo, 2024). On the third objective, the literature is most consistent in identifying governance design alignment, political incentive structures, and organizational managerial culture as the primary determinants of whether implementation succeeds in producing performance improvement or collapses into ceremonial compliance (Tillema et al. 2010).

The gaps that remain in the literature are consequential for future scholarships. The evidence base is geographically uneven, with urban public agencies in Africa, Latin America, and much of Asia underrepresented relative to their governance challenges (Král, 2022). The mechanisms through which performance information is used by managers translates into citizen-experienced service delivery improvement remain poorly specified, and the conditions under which citizen-facing frameworks generate sustained rather than short-term behavioral change have not been systematically examined (Andrews et al., 2011). The interaction between political accountability structures and framework implementation at the municipal level warrants dedicated investigation, particularly in decentralized governance systems where local political dynamics exercise strong influence over the operational priorities of urban public agencies.

3.5 Limitations of the Study

Several limitations bear on the findings of this review. Reliance on three major academic databases, while consistent with systematic review standards, carries the risk that studies indexed elsewhere were not captured. Finally, the literature on performance measurement in public organizations tends toward positive framing, and the evidence base may reflect a publication bias against null or negative findings.

4.0 RECOMMENDATIONS

The evidence established that performance measurement frameworks produce conditional rather than automatic service delivery improvements, mediated by framework design and managerial engagement, among others. This demands that action be directed at these dimensions.

Municipal governments and the central government bodies that oversee them must move away from the assumption that any standardized performance measurement framework can be deployed uniformly across urban public agencies regardless of governance context. The evidence indicates that framework adoption is shaped by institutional history, administrative capacity, and political environment, producing outcomes that range from substantive service improvement to ceremonial compliance. Urban governance practitioners and local government reform units should therefore conduct institutional readiness assessments before selecting or mandating a framework, examining whether the measurement logic embedded in candidate frameworks is congruent with the agency's accountability structures, managerial autonomy, and data infrastructure. In less developed country contexts, particularly, the adaptation of simplified tools calibrated to available capacity is more likely to produce implementation traction than wholesale transfer of high-complexity frameworks

Agency leaders and senior managers in urban public agencies must invest in building the managerial routines and organizational culture through which performance information is actively used, rather than treating framework installation as a sufficient reform intervention. The evidence is clear that purposeful managerial engagement with performance data the mechanism through which framework adoption generates measurable service delivery improvement. Agency leaders should therefore institute structured performance review routines, aligned with the agency's strategic priorities, that create regularized opportunities for managers at all levels to interrogate performance data and act on what it reveals. Framework investment that is not accompanied by this managerial infrastructure is unlikely to translate into citizen-experienced service delivery gains. Future research should prioritize longitudinal and comparative empirical studies that examine the mechanisms through which managerial engagement with performance information produces citizen-experienced service delivery improvement in urban public agencies across diverse governance contexts. The current evidence base remains disproportionately concentrated in high-capacity, anglophone governance environments, leaving the conditions for effective framework implementation in African, Latin American, and South Asian urban contexts substantially underspecified. Addressing this geographic gap matters not merely for scholarly completeness but because the urban governance challenges of the global South are precisely the conditions in which robust, contextually validated evidence on performance measurement frameworks is most urgently needed.

License: This is an open access article under the CC BY 4.0 license: https://creativecommons.org/licenses/by/4.0/

5.0 CONCLUSION

This review sets out to clarify what the evidence actually shows performance measurement frameworks in urban public agencies. Rather than asking whether these frameworks work in general, the evidence points to a more useful conclusion: they work under certain conditions, and those conditions matter just as much as the frameworks themselves. No single framework dominates practice across cities. Adoption is shaped not only by technical design but also by institutional context and administrative capacity. For any framework to improve service delivery, its design must fit the governance environment in which it is used. The literature shows that the overall impact of performance measurement on public organizations is modest. This is not a failure of the idea itself, but a sign that outcomes depend on how frameworks are applied. Important gaps, however, remain for future research. For urban public agencies, the key message is clear: performance measurement frameworks add value when they are taken seriously and used as tools for decision-making, not just for compliance.

REFERENCES

1) Agyei, E. O. (2026). Administrative process improvement approaches for reducing service delays in city governments: A systematic review. *EPRA International Journal of Economics, Business and Management Studies (EBMS), 13*(4), 88–97. https://doi.org/10.36713/epra26808
2) Agyei, E. O. (2026). A review of digital strategies for strengthening public service delivery systems. *Sarcouncil Journal of Public Administration and Management, 5*(2), 60–67. https://doi.org/10.5281/zenodo.19545478
3) Andersen, S. C., & Nielsen, H. S. (2020). Learning from performance information. *Journal of Public Administration Research and Theory*, *30*(3), 415-431. https://doi.org/10.1093/jopart/muz036
4) Andrews, R., Boyne, G. A., Law, J., & Walker, R. M. (2011). Strategy implementation and public service performance. *Administration & society*, *43*(6), 643-671. https://doi.org/10.1177/0095399711412730
5) Beeri, I., Uster, A., & Vigoda-Gadot, E. (2019). Does performance management relate to good governance? A study of its relationship with citizens' satisfaction with and trust in Israeli local government. *Public Performance & Management Review*, *42*(2), 241-279. https://doi.org/10.1080/15309576.2018.1436074
6) Brusca, I., Manes Rossi, F., & Aversano, N. (2017). Performance measurement in Italian and Spanish local governments: Comparative policy analysis. *Journal of Comparative Policy Analysis: Research and Practice*, *19*(5), 470-486. https://doi.org/10.1080/13876988.2015.1094892
7) Downe, J., Grace, C., Martin, S., & Nutley, S. (2010). Theories of public service improvement: A comparative analysis of local performance assessment frameworks. *Public Management Review*, *12*(5), 663-678. https://doi.org/10.1080/14719031003633201
8) Gębczyńska, A., & Brajer-Marczak, R. (2020). Review of selected performance measurement models used in public administration. *Administrative Sciences*, *10*(4), 99. https://doi.org/10.3390/admsci10040099
9) Gerrish, E. (2016). The impact of performance management on performance in public organizations: A meta-analysis. *Public administration review*, *76*(1), 48-66. https://doi.org/10.1111/puar.12433
10) Gomes, R. C., & de Azevedo, C. B. (2025). Balanced Scorecard: A literature review to trace its trajectory in the public administration domain. *International Journal of Public Administration*, *48*(8), 502-518. https://doi.org/10.1080/01900692.2024.2376053
11) Han, X., & Moynihan, D. (2022). Does managerial use of performance information matter to organizational outcomes?. *The American review of public administration*, *52*(2), 109-122. https://doi.org/10.1177/02750740211048891
12) Jie, G. (2015). Political rationality vs. technical rationality in China's target-based performance measurement system: The case of social stability maintenance. *Policy and Society*, *34*(1), 37-48. https://doi.org/10.1016/j.polsoc.2015.03.005
13) Johnsen, Å., Solholm, K., & Tufte, P. A. (2024). Performance measurement system design as link between strategy formulation and performance information use in public sector organizations. *Public Performance & Management Review*, *47*(4), 813-848. https://doi.org/10.1080/15309576.2024.2360158
14) van der Kolk, B. (2022). Performance measurement in the public sector: Mapping 20 years of survey research. *Financial Accountability & Management*, *38*(4), 703-7. https://doi.org/10.1111/faam.12345
15) Král, M. (2022). 20-year history of performance measurement in the local public sector: a systematic review. *International Journal of Public Administration*, *45*(9), 726-740. https://doi.org/10.1080/01900692.2021.1891425
16) Lapuente, V., & Van de Walle, S. (2020). The effects of new public management on the quality of public services. *Governance*, *33*(3), 461-475. https://doi.org/10.1111/gove.12502
17) Lewis, J. M. (2015). The politics and consequences of performance measurement. *Policy and Society*, *34*(1), 1-12. https://doi.org/10.1016/j.polsoc.2015.03.001
18) Meng, X., & Li, C. (2025). Does public sector performance information impact stakeholders? Evidence from a meta-analysis. *Public administration review*, *85*(4), 1086-1097. https://doi.org/10.1111/puar.13924
19) Moullin, M. (2017). Improving and evaluating performance with the Public Sector Scorecard. *International journal of productivity and performance management*, *66*(4), 442-458. https://doi.org/10.1108/IJPPM-06-2015-0092
20) Moynihan, D. P., & Kroll, A. (2016). Performance management routines that work? An early assessment of the GPRA Modernization Act. *Public Administration Review*, *76*(2), 314-323. https://doi.org/10.1111/puar.12434
21) Moynihan, D. P., & Pandey, S. K. (2010). The big question for performance management: why do managers use performance information? *Journal of public administration research and theory*, *20*(4), 849-866. https://doi.org/10.1093/jopart/muq004

License: This is an open access article under the CC BY 4.0 license: https://creativecommons.org/licenses/by/4.0/

22) Northcott, D., & Ma'amora Taulapapa, T. (2012). Using the balanced scorecard to manage performance in public sector organizations: Issues and challenges. *International Journal of Public Sector Management*, *25*(3), 166-191. https://doi.org/10.1108/09513551211224234
23) Obodai, T., Tracy, T., & Singh, D., Oman-Amoako, M. (2025). The role of financial analytics in enhancing public sector accountability in the U.S. *EPRA International Journal of Economic and Business Review, 13*(1), 52–60. https://doi.org/10.36713/epra24426
24) Sandfort, J., Ong, S. J., & McKay, C. (2019). Performance management regimes in practice: Examining the local agencies implementing temporary assistance for needy families. *The American review of public administration*, *49*(2), 145-158. https://doi.org/10.1177/0275074018809202
25) Schoeman, I., & Chakwizira, J. (2023). Advancing a performance management tool for service delivery in local government. *Administrative Sciences*, *13*(2), 31. https://doi.org/10.3390/admsci13020031
26) Tillema, S., Mimba, N. P. S., & Van Helden, G. J. (2010). Understanding the changing role of public sector performance measurement in less developed countries. *Public administration and development*, *30*(3), 203-214. https://doi.org/10.1002/pad.561
27) Turley, G., Robbins, G., & McNena, S. (2015). A framework to measure the financial performance of local governments. *Local government studies*, *41*(3), 401-420. https://doi.org/10.1080/03003930.2014.991865
28) Virani, A., & Van Der Wal, Z. (2023). Enhancing the effectiveness of public sector performance regimes: A proposed causal model for aligning governance design with performance logics. *Perspectives on Public Management and Governance*, *6*(1), 54-65. https://doi.org/10.1093/ppmgov/gvac026
29) Walker, R. M., Brewer, G. A., Boyne, G. A., & Avellaneda, C. N. (2011). Market orientation and public service performance: new public management gone mad?. *Public administration review*, *71*(5), 707-717. https://doi.org/10.1111/j.1540-6210.2011.02410.x

License: This is an open access article under the CC BY 4.0 license: https://creativecommons.org/licenses/by/4.0/