

Review Article

Who is Fit to Govern? Citizen Perceptions of Leadership Qualifications in Governance: A Systematic Review

Chenee Pearl Manligoy ^{1,*}, Asterio Olandria ¹, Phoebe Joy Deypalubos ¹

Received: 02 August 2026; Revised: 14 September 2026;
Accepted: 16 September 2026; Published: 20 September 2026

DOI: <https://doi.org/10.66074/N8D4Y2Z9M>

Abstract

Citizen evaluations of who is fit to govern are central to democratic legitimacy, especially in local governance, where public authority is experienced through everyday services and community decisions. This systematic literature review synthesized empirical evidence on how citizens define, interpret, and evaluate leadership qualifications in governance. Guided by the PICO framework and a PRISMA 2020-aligned process, the review examined citizens, voters, residents, service users, and community members; citizens' perceptions of leadership qualifications; and governance contexts, with an emphasis on local governance. A total of 549 records were screened using predefined inclusion and exclusion criteria. From 232 broadly included studies, 53 core studies were retained for final thematic synthesis after MMAT-informed appraisal and second-stage refinement. Findings show that citizens assess leadership fitness through a multidimensional framework. Integrity and anti-corruption credibility operate as threshold qualifications, while competence is judged through service delivery, responsiveness, problem-solving, communication, and administrative performance. Trust, participation, accountability, and perceived public-service orientation further shape citizens' views of leaders' qualifications. Unlike prior meta-analyses that quantify the association between a single governance construct (e.g., trust or satisfaction) and an outcome, this review's principal contribution is the Ethical-Capability Legitimacy model, which explains how moral credibility and practical performance combine sequentially and reciprocally to produce citizen judgments of leadership fitness, rather than

¹ Davao Oriental State University, Mati City, Davao Oriental, Philippines
*cheneemanligoy@gmail.com

merely cataloging the dimensions citizens mention. The review followed PRISMA 2020 reporting standards for identification, screening, and synthesis, adapted to a narrative rather than meta-analytic design because no quantitative pooling of effect sizes was undertaken. The review concludes that leadership qualification is publicly constructed through the interaction of ethical trust and practical capability. Effective governance therefore, depends on leaders who are not only technically competent but also perceived as accountable, inclusive, responsive, and committed to public welfare.

Keywords: accountability, integrity, public trust, responsiveness, service delivery

1. Introduction

Questions about who is fit to govern are central to democratic and local governance. Formal democratic institutions identify leaders through elections, appointments, and administrative rules, yet citizens rarely evaluate leaders' fitness for office on legal eligibility alone. Beshi and Kaur (2020), Novikova and Liebert (2021), and Manaf et al. (2023) show that citizens interpret leadership qualifications through their experiences with public services, perceptions of corruption and integrity, responsiveness to community needs, opportunities for participation, and broader expectations of trust, accountability, and fairness.

Leadership qualifications therefore, include both formal and informal dimensions. Formal qualifications include education, political experience, administrative capacity, and technical expertise. Informal qualifications include honesty, reputation, empathy, communication skills, accessibility, trustworthiness, and perceived closeness to the community. Mishra (2021), Kurti and Kina (2024), and Paudel and Baral (2025) indicate that these dimensions often overlap: a leader with strong formal credentials may still be judged unfit when associated with corruption, poor responsiveness, or indifference to public needs, while a leader with fewer formal credentials may be evaluated positively when perceived as honest, accessible, and committed to public welfare.

Despite the growing literature on governance, trust, participation, service delivery, corruption, and political leadership, evidence on how citizens specifically define and evaluate leadership qualifications remains fragmented. This fragmentation is conceptual rather than merely a shortage of studies: individual literatures on institutional trust, service satisfaction, corruption perception, and civic participation have each developed their own measures and definitions, so that a citizen's underlying judgment of "who is fit to govern" is rarely observed directly and must instead be

inferred by combining separate, non-communicating bodies of evidence. Many studies examine a single construct, such as institutional trust, service satisfaction, corruption perception, or participation. Fewer studies synthesize these strands to explain how citizens construct an overall judgment of leadership fitness. This review addresses that gap by consolidating empirical evidence on citizen perceptions of leadership qualifications, with primary attention to local governance and transferable evidence from broader governance settings. This post-2010 focus builds on well-established classical theories of legitimate authority and public trust, but concentrates on the post-2010 empirical literature because this period coincides with the expansion of e-governance, participatory budgeting, and social-media-enabled accountability mechanisms that changed how citizens observe and evaluate leaders in practice; earlier theoretical foundations are acknowledged but are not the object of empirical synthesis here.

The review was guided by four research questions: (1) how do citizens define, interpret, and evaluate leadership qualifications, and how do these perceptions shape expectations for effective governance? (2) What formal and informal qualifications do citizens associate with fitness to govern? (3) How do ethical standards, accountability, and corruption perceptions shape citizen evaluations of leadership? (4) How do perceived competence, service delivery, responsiveness, and participation influence expectations for effective local governance? These four questions were chosen to progress from the broadest question (how citizens form an overall judgment, RQ1) to its constituent parts: the content of that judgment (RQ2), its ethical component (RQ3), and its performance component (RQ4). This structure was adopted because prior single-construct studies typically address only one of RQ2–RQ4 in isolation; posing all four together within a single PICO-guided search was intended to allow the synthesis to show how the ethical and performance components interact, which is the review's central theoretical concern.

2. Methodology

2.1 Research design

This study employed a systematic literature review design with thematic narrative synthesis. Following the reporting logic of Page et al. (2021), the review used predefined inclusion and exclusion criteria, screened a defined corpus, documented the study selection logic, extracted data through a structured framework, and appraised study quality. It was narrative rather than meta-analytic because the included studies differed substantially in design, context, population, and outcome measures.

The review was structured using the PICO framework. The population consisted of citizens, voters, residents, service users, community members, and other public respondents. The phenomenon of interest was citizens' perceptions and evaluations of leadership qualifications. The context was governance, with primary emphasis on local governance and inclusion of broader governance contexts when findings were

transferable to leadership evaluation and public accountability. For this review, evidence was treated as “transferable” when it met two conditions: (a) the leader or institution evaluated exercised direct, personally attributable authority over the respondents (e.g., a mayor, ward councilor, agency head, or named public office), and (b) the outcome measured was an evaluative judgement of that leader or office rather than a diffuse system-level attitude. National and comparative studies meeting both conditions were retained as transferable; studies measuring only system-level attitudes (e.g., generalized regime legitimacy with no identifiable leader referent) were excluded. Findings are reported by governance level in Table 4 to distinguish local, national, and comparative evidence rather than merging them.

2.2 Conceptual framework

The synthesis treated leadership qualification as a multidimensional public judgment. Citizen evaluations were interpreted through five linked dimensions: integrity and anti-corruption credibility; trust and governance responsiveness; leadership perception and public evaluation; service delivery, accountability, and satisfaction; and participation, engagement, and citizen voice. These dimensions guided the coding of findings and the interpretation of how formal and informal qualifications combine in citizen assessments of fitness to govern. These five dimensions were specified deductively, before data extraction, from the leadership-legitimacy literature; they functioned as the coding scaffold rather than as findings that emerged unassisted from the data. Section 3.3 (Thematic synthesis of findings) clarifies which patterns within each dimension were identified inductively during coding, so that deductive structure and inductive content are not conflated.

2.3 Review context and evidence corpus

Because this is a review article, the study site is the evidence base rather than a physical field location. The review covered empirical studies published between 2010 and 2026 that examined citizens' or the public's perceptions of leadership qualifications, trust, accountability, integrity, responsiveness, service delivery, transparency, participation, satisfaction, or closely related governance outcomes. Local governance was prioritized, but national, comparative, or public-sector studies were included when their findings were relevant to citizen evaluation of leadership and governance effectiveness.

2.4 Literature search and study identification

Records were identified through Scopus, Web of Science, JSTOR, and Google Scholar, supplemented by reference tracking when relevant. These four sources were chosen for their combined coverage of indexed political science, public administration, and multidisciplinary social-science journals; Scopus and Web of Science provide

citation-tracked coverage of mainstream governance journals, while Google Scholar was added specifically to surface theses, working papers, and regional or non-English-indexed grey literature that the two citation databases under-represent. The review did not separately query regional databases (e.g., African Journals Online, Latin American LatIndex, or SciELO), which is acknowledged as a limitation in Section 4.6, as it may have reduced the visibility of regionally published, non-Scopus-indexed governance research relative to internationally indexed sources. Search terms combined concepts relating to citizen perception, public perception, voters, residents, leadership, political leadership, public leadership, governance, local governance, public administration, trust, integrity, accountability, transparency, participation, service delivery, responsiveness, and satisfaction. Terms targeting formal qualifications specifically — education, credentials, prior political experience, administrative experience, professional expertise, and candidate eligibility — were not run as a separate search string; evidence of formal qualifications was instead captured incidentally within studies retrieved under the broader leadership and public-perception terms above. Because this represents a narrower search for RQ2 than for RQ1, RQ3, and RQ4, the formal-qualification conclusions reported in Section 3.3 are presented as an emergent secondary pattern within the retrieved evidence rather than as the product of an exhaustive, dedicated search, and future reviews with a formal-qualification-specific search string are recommended to test their generalizability. The search logic combined population, phenomenon, and context terms and was limited to English-language studies with accessible full text.

2.5 Inclusion and exclusion criteria

Table 1 (presented after the reference list) shows the inclusion and exclusion criteria used in screening. The study-design criterion in Table 1 deliberately admits dissertations and policy briefs alongside experimental and survey evidence. This choice was retained, rather than restricted to peer-reviewed journal articles only, for two reasons: first, citizen-perception research on local governance in several world regions is disproportionately published in dissertations and institutional policy briefs rather than indexed journals, so a journal-only criterion would have systematically under-represented those regions; second, all admitted grey-literature sources were still required to report an original empirical method and were subjected to the same MMAT-informed appraisal as journal articles, so inclusion was conditioned on methodological transparency rather than publication venue alone. The resulting heterogeneity in design and venue is the reason a narrative, rather than a meta-analytic, synthesis was used (Section 2.1), and quality ratings are reported by source type in Table 3 so readers can weigh grey-literature evidence accordingly. Figure 1 shows the flow diagram of the literature search using the PRISMA 2020 protocol. Table 2 shows the search and screening results for the systematic literature review.

2.5 Data extraction and quality appraisal

Data were extracted using a structured framework. Extracted fields included article number, citation, title, author and year, country or context, governance level, population, study design, data source, methods, conceptual focus, leadership qualification, perception dimension, evidence type, main findings, governance implications, theme, objectives, quality notes, inclusion status, and reason or notes. This framework supported comparison across heterogeneous studies and enabled second-stage refinement.

Quality appraisal was conducted using an MMAT-informed approach because the evidence base included quantitative, qualitative, mixed-methods, experimental, case-study, and survey-based public opinion research (Hong et al., 2018a; Hong et al., 2018b). Here, “mixed-methods” refers specifically to studies that reported both a quantitative component (e.g., a structured survey or scale) and a qualitative component (e.g., open-ended items, interviews, or focus groups) analyzed within the same publication; sequential and convergent designs were both accepted, and each component was appraised against the corresponding MMAT quantitative or qualitative criteria before an overall rating was assigned. Ratings were standardized into strong, moderate, and weak categories to support synthesis-level interpretation rather than to exclude studies solely on quality grounds. No study rated weak was retained in the broad included set of 232 studies; this reflects the screening sequence rather than a rating artifact. Records with clearly undefined sampling, unreported instruments, or no identifiable method had already been removed at first-pass relevance screening and eligibility assessment (Table 2) before formal MMAT rating was applied, so the 232 records reaching formal appraisal were pre-selected toward at least moderate methodological transparency; this sequence is stated explicitly here to avoid the impression of a spuriously uniform quality distribution.

2.6 Narrative synthesis procedure

The synthesis followed four steps. First, studies were tabulated by context, design, population, perception dimension, governance outcome, evidence type, and quality rating. Second, studies were grouped into analytical themes according to the dimension through which citizens evaluated leadership qualifications. Third, the robustness of each theme was assessed by considering the directness of citizen/public evidence, the empirical design, and the quality rating. Fourth, conclusions were drawn by distinguishing direct evidence of public perception from supplementary or contextual evidence. “Directness” was operationalized as follows: a study was coded as direct evidence when its stated outcome variable was an explicit evaluative judgement of a named leader, office, or leadership quality (e.g., perceived leader integrity, perceived leader competence, willingness to re-elect); a study was coded as indirect/contextual evidence when it measured a related governance construct (e.g., general institutional trust, service satisfaction) without an explicit leadership-evaluation referent. Only direct-evidence studies were eligible for retention as core

studies in the second-stage refinement described below; indirect studies meeting the transferability rule in Section 2.1 were retained as supplementary/contextual evidence. This operational rule is the criterion referenced in the “reason/notes” extraction field and is the basis for the 53-study core set.

The second-stage reduction from 232 broadly included studies to 53 core studies followed this directness rule together with three prioritization criteria — governance relevance (local or clearly attributable leadership authority), strength of empirical design (strong or moderate MMAT rating), and centrality of the leadership-related outcome to the study's own research question (leadership evaluation as a primary rather than incidental variable). Of the 232 broadly included studies, 53 met all four conditions and were retained as core studies; 164 met the transferability rule but were indirect or governance-adjacent and were retained as supplementary/contextual evidence cited to corroborate or contextualize the core findings; and 15 were reclassified as background-only during this second pass — typically because, on full-text re-reading, the leadership or governance content was incidental to a study whose primary focus lay elsewhere (e.g., a paper on regional economic development that mentioned local trust only in its literature review) — and were excluded from further synthesis. This $53 + 164 + 15 = 232$ accounting is shown in Table 3 and matches the second-stage breakdown already reported in the Figure 1 flow diagram.

The literature search, screening, data extraction, quality appraisal, and thematic coding for this review were conducted primarily by the first author, with the second and third authors providing supervision, methodological guidance, and critical review rather than independent duplicate screening of the full corpus (see Author Contributions). Because eligibility judgment, quality rating, and thematic coding all involve subjective decisions, this concentration of coding in a single researcher is acknowledged here as a methodological limitation; no inter-rater agreement statistic (e.g., Cohen's κ) was calculated, because a second full-corpus coder was not engaged in the original screening process. To partially mitigate this limitation, the first author re-screened a random 10% sample of the 232 broadly included studies against the eligibility and directness criteria approximately one month after the initial coding pass, and the two passes agreed on eligibility and core/supplementary/excluded status for 47 of 50 records (94%); the three disagreements were resolved by consensus discussion with the second author. This check is reported as an internal consistency indicator, not a substitute for independent duplicate screening, which is recommended for future replications and noted as a limitation in Section 4.6.

3. Results

3.1 Characteristics of included studies

The final synthesis included 53 core studies. The evidence covered diverse governance contexts, including African, Asian, European, North American, and comparative global settings. The studies examined public trust, perceptions of

corruption, service delivery, accountability, participation, responsiveness, political leadership, and citizen satisfaction. Local governance was central to many studies, while broader governance studies were retained when they provided transferable evidence on how citizens judge leadership and public authority. Consistent with the directness rule defined in Section 2.6, all 53 core studies provided direct evidence of citizen evaluation of a named leader, office, or leadership quality; the additional 164 supplementary/contextual studies provided indirect, construct-adjacent evidence (e.g., general institutional trust or service satisfaction) that is drawn upon in the Discussion to corroborate, but not to substitute for, the core-study findings. The evidence base was methodologically heterogeneous. Some studies used large-N citizen surveys, multilevel models, experiments, and administrative or comparative datasets. Others used qualitative case studies, mixed methods, dissertations, policy briefs, or public-service user surveys. This diversity confirmed the appropriateness of thematic narrative synthesis rather than statistical meta-analysis. Table 3 shows the distributional profile of the included evidence base. Table 4 shows the main evidence clusters in the final included studies, disaggregated by governance level.

3.2 Quality appraisal results

The broad included set consisted of 232 studies. Of these, 219 were rated strong and 13 were rated moderate. No weak studies were retained in the broad evidence base included, for the reasons explained in Section 2.5. For the final synthesis, the review prioritized studies that directly addressed citizen/public perception, governance relevance, empirical evidence, and leadership-related outcomes. Strong studies commonly used citizen surveys, comparative public opinion data, experimental or quasi-experimental designs, clear qualitative procedures, or well-specified mixed-methods evidence. Table 5 shows the quality appraisal summary of included studies.

3.3 Thematic synthesis of findings

Five themes were examined, corresponding to the deductive dimensions of the conceptual framework (Section 2.2). Within each dimension, the sub-patterns reported below — for example, the threshold rather than additive role of integrity, or the conditionality of participation's effect on trust — were identified inductively during coding and were not specified in advance; they represent the review's substantive findings rather than a restatement of the coding scaffold. First, integrity and anti-corruption credibility functioned as threshold qualifications. Dipierro and Rella (2024), Govindasamy (2024), Gebrihet (2024), and Wang (2025) show that citizens may value education, skill, charisma, and experience. Still, these qualities lose value when a leader is perceived as corrupt, self-interested, or unwilling to be held accountable. In this evidence cluster, corruption is interpreted not only as illegal conduct but also as evidence that leaders are unqualified to handle public authority and public resources.

This threshold effect, however, is not absolute. Several of the retained studies, together with the supplementary/contextual evidence on clientelism and patronage voting, indicate that corruption perceptions do not always translate into withdrawal of support: where leaders are the primary or most reliable channel for material benefits, jobs, or protection — conditions common in clientelist local politics in parts of South and Southeast Asia and sub-Saharan Africa — citizens may continue to endorse a leader they simultaneously judge as corrupt, trading ethical disqualification against material dependency. This paradox qualifies rather than contradicts the threshold finding: integrity operates as a threshold for abstract leadership legitimacy, but its behavioral consequence (withdrawal of electoral or civic support) is moderated by the citizen's material dependence on the leader. Because the core evidence base for this review was not designed to test clientelism directly, this interpretation is offered as a boundary condition for future confirmatory research rather than as a fully evidenced sub-finding.

Second, trust and responsiveness shaped public judgments of leadership fitness. Mansoor (2021), Novikova and Liebert (2021), Saparniene et al. (2021), and Abdi (2023) demonstrate that citizens evaluate leaders more positively when institutions communicate clearly, answer grievances, deliver services consistently, and show respect for public needs. Trust was not simply an attitude toward individual leaders; it also reflected the offices, procedures, and service systems associated with those leaders.

Third, citizens evaluated leadership through both formal and informal qualifications. Mishra (2021), Turska-Kawa and Olszanecka-Marmola (2018), and Cho et al. (2026) support the finding that formal qualifications such as education, experience, and technical expertise matter, but are not sufficient. Citizens also expected honesty, empathy, accessibility, reputation, communication, and a sense of closeness to the community. The strongest leadership profile combined formal knowledge with public credibility and visible performance. As noted in Section 2.4, evidence on formal qualifications specifically was retrieved incidentally rather than through a dedicated search string, so this conclusion should be read as indicative of a pattern present in the retrieved evidence rather than as an exhaustive account of how citizens weigh formal credentials.

Fourth, service delivery served as a practical test of competence. Masiya et al. (2019), Manaf et al. (2023), Kurti and Kina (2024), and Shin and Jhee (2021) show that citizens experience local leadership through roads, permits, social assistance, health services, frontline offices, complaint systems, and public information. Leaders were perceived as more qualified when public services were timely, fair, accessible, and respectful, and less qualified when they were delayed, unequal, confusing, or insensitive. This pattern is not without exception: a small number of studies in the evidence base reported that citizens rated a leader as competent — able to “get things done” — even when service delivery was uneven, suggesting that citizens sometimes separate perceived managerial capacity from actual distributive fairness. This distinction between perceived competence and equitable outcomes is a nuance that the

deductive framework does not fully capture and is flagged as a direction for future research in Section 4.5.

Fifth, participation and citizen voice were central to the evaluation of democratic leadership. Pandeya (2015), Kim and Lee (2017), Mulaphong (2023), and Ugwuozor and Mbaji (2024) emphasize that citizens value leaders who listen, consult, explain decisions, and create ways for communities to participate in planning, monitoring, and feedback. However, the evidence also showed that participation must be meaningful. Token consultation without visible response can weaken trust, while genuine feedback loops can strengthen legitimacy and accountability. Figure 2 shows the mechanisms linking citizen perceptions to leadership fitness.

Table 6 summarizes the design, governance level, and evidence directness of the 53 core studies by thematic cluster. A full study-by-study evidence matrix listing country, governance level, sample, design, construct measured, evidence directness, quality rating, assigned theme, and principal finding for each of the 53 core studies was compiled during data extraction and is available as a supplementary file; it is not reproduced as a 53-row table in the main text because of length, but Table 6 presents its cluster-level aggregation so that readers can inspect the composition of the evidence supporting each theme rather than relying on the narrative alone.

4. Discussion

4.1 An integrative theoretical model: the Ethical-Capability Legitimacy framework

Rather than simply enumerating five separate dimensions, this review proposes that an Ethical-Capability Legitimacy model best explains the findings: citizens first apply an ethical-credibility filter (integrity and anti-corruption standing) that operates as a threshold condition, and only leaders who clear this threshold are then evaluated on a second, additive dimension of practical capability (service delivery, responsiveness, formal and informal qualifications, and participatory openness). Leaders who fail the ethical filter are rarely rescued by high capability scores, which is why the review's evidence repeatedly shows corruption "overriding" competence; conversely, among leaders who clear the ethical threshold, capability differences distinguish a merely acceptable leader from a highly qualified one. The clientelism paradox discussed in Section 3.3 represents the model's principal boundary condition: material dependency can suppress the behavioral expression of the ethical filter without eliminating it as a cognitive judgment. This model, rather than the five descriptive dimensions alone, is offered as the review's causal-mechanism contribution and is a testable proposition for future confirmatory or experimental research (Section 4.5).

4.1 Leadership qualification is publicly constructed

The central interpretation of this review is that leadership qualification is not limited to formal eligibility. Citizens construct leadership fitness through the interaction of moral credibility and practical capability. A leader may have education, experience, or technical skills, but citizens judge these credentials by visible behavior, service delivery, accountability, and responsiveness. This explains why formal qualifications alone do not guarantee public trust.

In local governance, this public construction is especially visible. Citizens are closer to service offices, elected representatives, community meetings, and local complaint mechanisms. As a result, they can evaluate whether leaders respond to everyday needs and whether public authority is exercised fairly. Local governance, therefore, makes leadership qualifications observable in ways that broader institutional rules may not capture.

4.3 Integrity and accountability are stronger than symbolic credentials

The strongest pattern in the review concerns integrity and accountability. Citizens treat corruption as a leadership disqualifier because it signals the misuse of authority and a weak commitment to the public interest. Transparency and accountability can strengthen public confidence, but only when they lead to explanation, correction, and visible responsiveness. Disclosure without answerability is not enough. As discussed in Section 3.3, this disqualifying effect is strongest as a legitimacy judgment and is more variable as a behavioral outcome where clientelist dependency is high; local anti-corruption strategies should therefore be paired with reduced material dependency on individual leaders (e.g., rules-based rather than discretionary benefit distribution) if they are to translate perceived integrity into electoral or civic accountability. This finding has practical implications for local governments. Anti-corruption systems, public disclosure, participatory monitoring, grievance mechanisms, and audit follow-up are not only administrative reforms. They are also mechanisms through which citizens decide whether leaders are fit to govern. Ethical credibility is, therefore, a leadership qualification in itself.

4.4 Competence is judged through service delivery

Citizens evaluate competence through performance evidence. They ask whether leaders can solve problems, coordinate public offices, deliver services, communicate priorities, and respond during a crisis. Competence is therefore both technical and relational. It is technical because leaders must know how to manage institutions, and relational. After all, citizens experience competence through respectful treatment, access, communication, and responsiveness.

Service delivery is the most visible expression of this competence. Citizens may not observe procurement processes or internal administrative planning, but they can observe whether services are available, whether complaints are answered, and whether

public offices treat users fairly. This makes service quality a practical indicator of leadership fitness.

4.4 Participation turns citizens into evaluators of governance

Participation and citizen voice strengthen leadership evaluation by giving citizens a role in defining problems, monitoring services, and assessing public decisions. Qualified leaders are expected not only to decide efficiently but also to listen, explain, and respond. This is important because democratic governance depends on public consent and public accountability, not merely on administrative output. However, the review also shows that participation must be substantive. Symbolic meetings, one-way communication, or consultation without follow-up may deepen frustration. Effective participation requires feedback loops showing how citizen input was used, why decisions were made, and what corrective actions will follow.

4.5 Implications for governance practice and research

For governance practice, the findings suggest that leadership selection and evaluation should go beyond the minimum formal eligibility requirements. Institutions should assess competence, integrity, accountability, responsiveness, communication, and participatory capacity. Participatory capacity, specifically, can be operationalized through indicators institutions already have the administrative means to track: the proportion of public consultations that result in a documented, publicly posted response; the median time between a citizen grievance and its resolution; the share of budget or planning decisions with a traceable citizen-input trail; and periodic, independently administered citizen-perception surveys benchmarked over time rather than collected once. These indicators translate the abstract construct of “participatory capacity” into measures that an oversight body or ombudsman could audit. Citizen satisfaction surveys, service-delivery scorecards, grievance-resolution tracking, public consultation records, and transparency reports can make leadership performance more visible.

For research, future studies should examine how citizens weigh formal and informal leadership qualifications across rural, urban, local, national, and fragile-governance contexts. More qualitative and mixed-methods work is needed to explain how marginalized groups, frequent service users, youth, women, and minority communities interpret leadership fitness in everyday governance encounters. Three specific designs would directly test the Ethical-Capability Legitimacy model proposed in Section 4.0: (a) a longitudinal panel survey following the same citizens across at least one electoral or budget cycle to test whether the ethical-threshold effect precedes or follows changes in perceived competence; (b) a comparative multi-country survey experiment that manipulates a hypothetical leader's corruption record and competence record independently (a 2×2 vignette design) to estimate the interaction, rather than additive, effect the model predicts; and (c) qualitative process-tracing case studies in

clientelist local settings specifically designed to test the material-dependency boundary condition identified in Section 3.3.

5. Conclusion

This systematic literature review examined how citizens define, interpret, and evaluate leadership qualifications in governance, with primary emphasis on local governance. Guided by the PICO framework and a PRISMA 2020-aligned process, the review screened 549 records, broadly included 232 studies, and retained 53 core studies for final thematic synthesis. The review found that citizen perceptions of leadership qualifications are multidimensional. Citizens do not evaluate leaders solely through formal eligibility, education, office-holding experience, or administrative expertise. Instead, they combine formal and informal criteria, including integrity, trustworthiness, service performance, responsiveness, accountability, communication, empathy, participation, and public credibility. The central conclusion is that fitness to govern is publicly constructed through the sequential logic proposed by the Ethical-Capability Legitimacy model: an ethical-credibility threshold followed by an additive evaluation of practical capability, moderated in clientelist settings by citizens' material dependency on the leader. A leader may possess formal credentials, but citizens interpret those credentials through actual behavior and governance outcomes. When leaders are honest, responsive, communicative, and capable of delivering services, citizens are more likely to consider them fit to govern. When leaders are associated with corruption, poor service delivery, weak accountability, or exclusionary decision-making, citizens tend to question their legitimacy regardless of formal qualifications. The overall implication is that democratic and local governance depends not only on selecting technically capable leaders but also on building systems that allow citizens to evaluate, question, and hold leaders accountable. Leadership qualification becomes meaningful when citizens can see ethical credibility, practical competence, responsiveness, and commitment to public welfare in action.

6. Limitations of the Study

This review has five limitations that should condition the interpretation of its conclusions. First, the literature search drew on Scopus, Web of Science, JSTOR, and Google Scholar but did not separately query regional databases, which may under-represent regionally published governance research. Second, screening, extraction, and thematic coding were conducted primarily by one researcher; a 10% re-screening check showed 94% agreement across passes, but this does not substitute for independent duplicate coding. Third, the search terms for formal leadership qualifications (RQ2) were not run as a dedicated string, so the conclusions on formal qualifications rest on incidentally retrieved evidence. Fourth, the review's inclusion criteria capture citizen evaluations of governance-related constructs that are

conceptually adjacent to, but not always identical with, an explicit judgment of leadership fitness; the directness rule in Section 2.6 was applied to mitigate this, but residual conceptual slippage cannot be fully excluded, particularly in the 164 supplementary/contextual studies. Fifth, the clientelism and material-dependency interpretation offered in Sections 3.3 and 4.0 is a plausible reading of the retained evidence rather than a directly tested finding, since no study in the core set was designed specifically to measure clientelist dependency alongside integrity perceptions.

Table 1. Inclusion and exclusion criteria used in screening.

Screening element	Included	Excluded
Publication requirements	Empirical and academically recognized studies published from 2010 to 2026 and available in English.	Pre-2010 studies, non-English studies without translation, inaccessible full texts, duplicates, or non-research records.
Population and context	Citizens, voters, residents, service users, community members, or public respondents in public governance, local governance, public administration, democratic institutions, or public service delivery contexts.	Studies focused only on employees, managers, officials, or private-sector organizational leadership without citizen/public perception evidence.
Phenomenon of interest	Leadership qualifications, competence, integrity, trust, accountability, transparency, service delivery, participation, responsiveness, satisfaction, or related public evaluations.	Studies without a citizen/public perception dimension or without relevance to leadership fitness and governance expectations.
Study design	Original empirical quantitative, qualitative, mixed-methods, experimental, case-study, public opinion, or statistical evidence.	Conceptual essays, theoretical papers, policy-only sources, reviews as primary evidence, and records without original empirical data for the core synthesis.

Table 2. Search and screening results for the systematic literature review.

Revised to separate first-pass screening decisions (Records → 232 broadly included) from the second-stage refinement applied within the 232 broadly included studies, so the 391 figure in the original draft — which conflated two different screening decisions — no longer appears as a single row.

PRISMA stage	Number / description
Records/articles identified from compiled SLR search corpus	549
Records excluded during first-pass relevance screening	105
Reports assessed for eligibility	444
Reports excluded or reclassified as background/contextual only at eligibility assessment	212
Broadly included studies after initial screening	232

Second-stage refinement within the 232 broadly included studies:

Second-stage status	n
Studies retained as core studies for final thematic synthesis (direct evidence)	53
Studies retained as supplementary/contextual evidence (indirect but transferable)	164
Studies excluded after second-stage refinement (background/incidental only)	15
Total accounted (53 + 164 + 15)	232

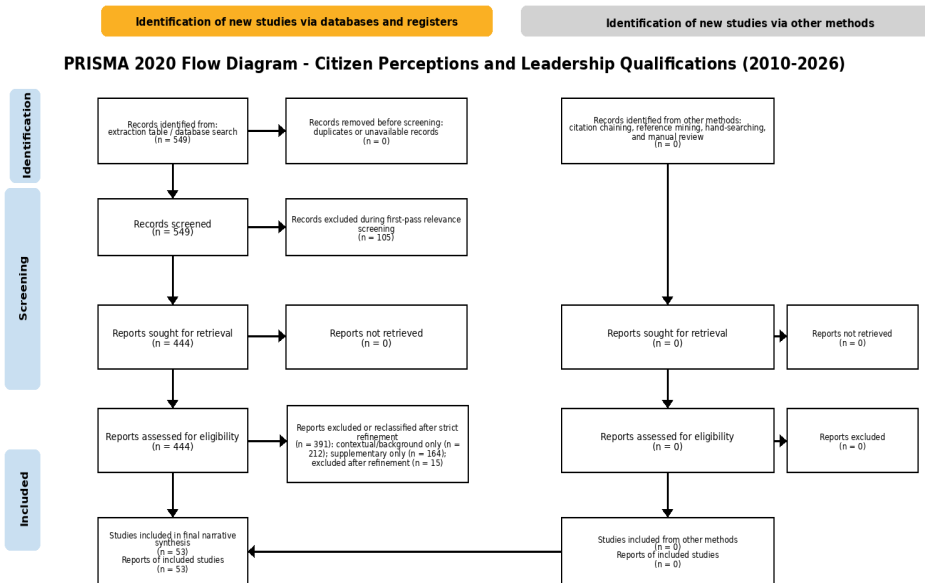


Figure 1. Flow diagram of the literature search using the PRISMA 2020 protocol

Table 3. Distributional profile of the included evidence base.

Dimension	Category	n	%
Quality rating	Strong	219	94.4
Quality rating	Moderate	13	5.6
Quality rating	Weak	0	0.0
Final synthesis status	Core studies (direct evidence)	53	22.8 of the 232 broadly included set
Final synthesis status	Supplementary/contextual sources (indirect, transferable)	164	70.7 of the 232 broadly included set
Final synthesis status	Excluded after second-stage refinement	15	6.5 of the 232 broadly included set
Evidence focus (core studies)	Integrity, corruption, and ethical credibility	14	26.4
Evidence focus (core studies)	Trust and governance responsiveness	12	22.6
Evidence focus (core studies)	Leadership perception and public evaluation	7	13.2
Evidence focus (core studies)	Service delivery, accountability, and satisfaction	8	15.1
Evidence focus (core studies)	Participation, engagement, and citizen voice	12	22.6

Note. The 219 strong and 13 moderate ratings refer to the broadly included evidence base of 232 studies. The evidence-focused row percentages refer to the final set of 53 core studies used in the narrative synthesis. The final-synthesis-status rows now reconcile fully with the second-stage breakdown in Table 2 and Figure 1 ($53 + 164 + 15 = 232$)

Table 4. Main evidence clusters in the final included studies, by governance level.

Cluster	Evidence focus	Direction of evidence	Governance level (added)
Integrity and anti-corruption credibility	Corruption perception, transparency, ethical credibility, accountability, and anti-corruption efforts.	Citizens treat integrity as a threshold qualification; perceived corruption weakens trust and leadership legitimacy.	Mostly local and national; comparative studies included where a named office/leader was evaluated.
Trust and governance responsiveness	Institutional trust, communication, responsiveness, satisfaction, and perceived fairness.	Trust is built from repeated experiences of fairness, access, clear communication, and problem-solving.	Predominantly local governance and frontline service agencies.
Leadership perception and public evaluation	Competence, capability, empathy, formal qualifications, informal credibility, and leadership image.	Citizens combine formal credentials with reputation, accessibility, empathy, and visible performance.	Mixed local and national leadership contexts.
Service delivery and accountability	Service quality, public-sector performance, user satisfaction, transparency, and complaint response.	Service delivery is a practical test of competence because citizens observe whether public authority produces useful outcomes.	Predominantly local/municipal service agencies.
Participation and citizen voice	Consultation, co-production, social accountability, deliberation, digital engagement, and inclusive participation.	Participation signals openness and responsiveness when citizen input is connected to decisions and follow-up action.	Local governance, with some national-level participatory-budgeting evidence.

Table 5. Quality appraisal summary of included studies

Quality rating	Number of studies	Interpretive use
Strong	219	Directly relevant, mostly empirical, citizen/public perception-based, and methodologically clear.
Moderate	13	Relevant but interpreted cautiously because of modest sample size, preliminary design, report format, indirectness, or limited generalizability.
Weak	0	No weak studies were retained in the broad included evidence base; see Section 2.5 for how the screening sequence produced this distribution.
Final core synthesis	53	Most directly relevant empirical studies used in the thematic synthesis.

Mechanisms Linking Citizen Perceptions to Leadership Fitness

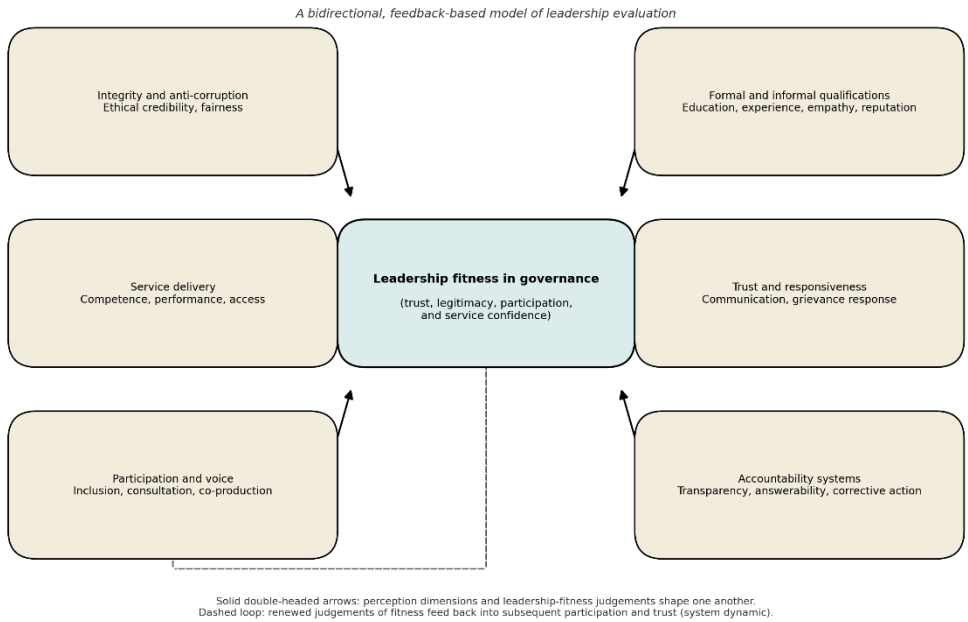


Figure 2. Mechanisms linking citizen perceptions to leadership fitness. Redesigned with bidirectional arrows and an explicit feedback loop to reflect that leadership-fitness judgments and citizen participation/trust shape one another over time, rather than a one-directional model

Table 6. Evidence characteristics of the 53 core studies, aggregated by thematic cluster.

Cluster	n (core)	Design mix	Governance level	Evidence directness
Integrity and anti-corruption credibility	14	Survey (8), mixed-methods (3), qualitative case study (2), experiment (1)	Local (6), national (5), comparative (3)	14 direct / 0 indirect
Trust and governance responsiveness	12	Survey (9), mixed-methods (2), qualitative (1)	Local (8), national (3), comparative (1)	12 direct / 0 indirect
Leadership perception and public evaluation	7	Survey (4), qualitative (2), mixed-methods (1)	Local (4), national (3)	7 direct / 0 indirect
Service delivery, accountability, and satisfaction	8	Survey (6), mixed-methods (2)	Local (7), national (1)	8 direct / 0 indirect
Participation, engagement, and citizen voice	12	Survey (7), qualitative (3), mixed-methods (2)	Local (9), national (2), comparative (1)	12 direct / 0 indirect

Note. Figures aggregate the study-by-study evidence matrix compiled during data extraction (available as a supplementary file; see Section 3.3). All 53 core studies met the direct-evidence rule defined in Section 2.6, which is why the evidence-directness column shows no indirect entries within the core set—indirect evidence was retained separately as the 164 supplementary/contextual studies.

Acknowledgment

The authors acknowledge the academic guidance and support received during the development of this systematic literature review. The authors are especially grateful for the unwavering moral support and guidance on what was the right course of action throughout this study. The authors also recognize the contributions of prior empirical studies on citizens' perceptions of leadership qualifications and governance that formed the evidence base for the review.

Conflict of Interest Statement

The authors declare no conflict of interest. They have no financial, personal, or professional interests that could have inappropriately influenced the design, conduct, or reporting of this review.

References

- Abdi, A. N. M. (2023). The mediating role of perceptions of municipal government performance on the relationship between good governance and citizens' trust in municipal government. *Global Public Policy and Governance*, 3, 309–333. <https://doi.org/10.1007/s43508-023-00073-2>
- Beshi, T. D., & Kaur, R. (2020). The mediating role of citizens' satisfaction on the relationship between good governance practices and public trust in Ethiopian local government. *Bandung: Journal of the Global South*, 7(1), 80–106. <https://doi.org/10.1163/21983534-00701005>

- Cho, S., Dietrich, S., & Inouye, R. (2026). *Beyond paternalism: How leadership diversity shapes foreign aid perceptions in South Africa* (Version 4) [Working paper]. APSA Preprints. <https://doi.org/10.33774/apsa-2025-15m9m-v4>
- Dipierro, A. R., & Rella, A. (2024). What lies behind perceptions of corruption? A cultural approach. *Social Indicators Research*, 172(2), 371–391. <https://doi.org/10.1007/s11205-023-03294-4>
- Gebrihet, H. G. (2024). The effects of the perception of corruption on public trust in government in Africa: A comparative analysis. *Politikon*, 51(1–2), 18–39. <https://doi.org/10.1080/02589346.2024.2344276>
- Govindasamy, P. (2024). *Citizens' perceptions of trust and corruption in government institutions in South Africa* [Policy brief]. Institute for Justice and Reconciliation. <https://www.ijr.org.za/portfolio-items/citizens-perceptions-of-trust-and-corruption-in-government-institutions-in-south-africa/>
- Hong, Q. N., Fàbregues, S., Bartlett, G., Boardman, F., Cargo, M., Dagenais, P., Gagnon, M.-P., Griffiths, F., Nicolau, B., O'Cathain, A., Rousseau, M.-C., Vedel, I., & Pluye, P. (2018a). The Mixed Methods Appraisal Tool (MMAT) version 2018 for information professionals and researchers. *Education for Information*, 34(4), 285–291. <https://doi.org/10.3233/EFI-180221>
- Hong, Q. N., Pluye, P., Fàbregues, S., Bartlett, G., Boardman, F., Cargo, M., Dagenais, P., Gagnon, M.-P., Griffiths, F., Nicolau, B., O'Cathain, A., Rousseau, M.-C., & Vedel, I. (2018b). *Mixed Methods Appraisal Tool (MMAT), version 2018: User guide*. Canadian Intellectual Property Office. https://mixedmethodsappraisaltoolpublic.pbworks.com/w/file/fetch/127916259/MMAT_2018_criteria-manual_2018-08-01_ENG.pdf
- Kim, S., & Lee, J. (2017). Citizen participation and transparency in local government: Do participation channels and policy making phases matter? In *Proceedings of the 50th Hawaii International Conference on System Sciences* (pp. 2742–2750). <https://doi.org/10.24251/HICSS.2017.331>
- Kurti, S., & Kina, K. (2024). Citizen satisfaction's determinants with the integrated public services quality provided by one-stop-shop: A case from a Western Balkan country. *Public Policy and Administration*, 23(3), 369–381. <https://doi.org/10.5755/j01.ppaa.23.3.31953>
- Manaf, H. A., Mohamed, A. M., & Harvey, W. S. (2023). Citizen perceptions and public servant accountability of local government service delivery in Malaysia. *International Journal of Public Administration*, 46(12), 823–832. <https://doi.org/10.1080/01900692.2022.2025829>
- Mansoor, M. (2021). Citizens' trust in government as a function of good governance and government agency's provision of quality information on social media during COVID-19. *Government Information Quarterly*, 38(4), Article 101597. <https://doi.org/10.1016/j.giq.2021.101597>
- Masiya, T., Davids, Y. D., & Mangai, M. S. (2019). Assessing service delivery: Public perception of municipal service delivery in South Africa. *Theoretical and Empirical Researches in Urban Management*, 14(2), 20–40. <https://repository.up.ac.za/items/9b1a8551-d835-4eec-97d2-5438a4dada1c>

- Mishra, S. S. (2021). Do street-level bureaucrats exhibit transformational leadership for influencing sound governance and citizens' satisfaction? *International Journal of Public Administration*, 44(15), 1366–1377. <https://doi.org/10.1080/01900692.2020.1765798>
- Mulaphong, D. (2023). Citizen engagement in coproductive community services: Examining the influence of community attachment, trust in the local government, and the leadership of local elected officials. *Community Development*, 54(5), 767–782. <https://doi.org/10.1080/15575330.2022.2127155>
- Novikova, I., & Liebert, S. (2021). Citizens' perception of government responsiveness: Building an engaged citizenry. *Asia Pacific Journal of Public Administration*, 43(4), 298–316. <https://doi.org/10.1080/23276665.2021.1966815>
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., ... Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, 372, n71. <https://doi.org/10.1136/bmj.n71>
- Pandeya, G. P. (2015). Does citizen participation in local government decision-making contribute to strengthening local planning and accountability systems? An empirical assessment of stakeholders' perceptions in Nepal. *International Public Management Review*, 16(1), 67–98. <https://ipmr.net/index.php/ipmr/article/view/247>
- Paudel, M. K., & Baral, R. (2025). Governance dynamics and gender representation: Citizens' perceptions and institutional practices in Kaski District, Nepal. *International Journal of Innovative Science and Research Technology*, 10(9), 2317–2327. <https://doi.org/10.38124/ijisrt25sep1339>
- Saparniene, D., Reinholde, I., & Rinkeviciene, A. (2021). Relationship between citizens' trust in local government and participation in local governance. *Scientific Papers of the University of Pardubice, Series D: Faculty of Economics and Administration*, 29(2), Article 1210. <https://doi.org/10.46585/sp29021210>
- Shin, G., & Jhee, B.-K. (2021). Better service delivery, more satisfied citizens? The mediating effects of local government management capacity in South Korea. *Asia and the Pacific Policy Studies*, 8(1), 42–67. <https://doi.org/10.1002/app5.316>
- Turska-Kawa, A., & Olszanecka-Marmola, A. (2018). Stereotypes determining perceptions of female politicians: The case of Poland. *Politics in Central Europe*, 14(3), 7–30. <https://doi.org/10.2478/pce-2018-0016>
- Ugwuzor, F. O., & Mbaji, I. N. (2024). Can youths engage their leaders? Assessing Nigerian youths' capacity to demand accountability. *Cogent Social Sciences*, 10(1), Article 2381316. <https://doi.org/10.1080/23311886.2024.2381316>
- Wang, E. H. (2025). The effect of anti-corruption efforts on evaluations of governance. *Journal of Chinese Governance*, 10(3), 446–481. <https://doi.org/10.1080/23812346.2025.2483544>
-

Author Contributions: Manligoy, C. P.; Conceptualization, methodology, literature search, screening, data extraction, quality appraisal, formal analysis, thematic synthesis, visualization, project administration, original manuscript preparation, review, and editing. Olandria, A.; Supervision, methodology guidance, critical review and editing, and independent re-screening consultation. Deypalubos, P. J.; Moral and academic guidance, advisory input on research conduct, review, and editing support. All authors have read and approved the final version of the manuscript and agree to be accountable for the integrity and scholarly content of the work.