



# Quality Beyond Rankings: Mission-Led Quality Assurance as a Sovereignty Practice by African Higher Education Institutions

Nomanesi Madikizela-Madiya & John Kibwage Nyangaresi

*University of South Africa*

## Article History

Received: 28.10.2025

Revised: 10.03.2026

Accepted: 13.03.2026

Published: 30.04.2026

## Keywords

Culture

Governance

Higher education

Quality assurance

## How to cite:

Madikizela-Madiya, N., & Nyangaresi, J. K. (2026). Quality Beyond Rankings: Mission-Led Quality Assurance as a Sovereignty Practice by African Higher Education Institutions. *Eastern African Journal of Humanities and Social Sciences*, 5(1), 252-262.

Copyright © 2026



## Abstract

Global university rankings have increasingly come to function as a proxy for quality, shaping institutional priorities, internal quality assurance (IQA) practices, and scholarly labour in ways that may conflict with locally defined missions and the public good. This research is a narrative review of contemporary literature on ranking effects, audit culture, and responsible research assessment, in order to develop an evidence-based model of Mission-Led Quality Assurance (MLQA) as a sovereignty practice by African higher education institutions. A transparent search-and-screen procedure of Scopus, Web of Science, ERIC, and a specific Google Scholar search identified 14 peer-reviewed sources that meet the inclusion criteria, and one underwent thematic comparison through programme-theory-inspired synthesis. The synthesis reveals that (i) rankings strengthen audit logics and attract proxy substitutes around what is readily counted; (ii) QA units often repackage ranking indicators into internal dashboards and create mission drift and strategic gaming; and (iii) reform announcements around responsible metrics often stall aimlessly without governance redesign, plural evidence standards, and participatory judgement. The study suggests a four-pillar MLQA framework, mission-based standards, pluralised evidence, deliberative review cycles, and public-benefit accountability, and describes implications of implementation to institutional leaders and quality agencies. However, the research was limited because it was based on published literature and was not representative of the region. Future studies should embrace multi-site case studies and discussion with the stakeholders to confirm the validity of MLQA.

## Introduction

University ranking has established itself as a hegemonic evaluative system (a symbolic instrument of comparison) that enables the comparison, management, and funding of universities worldwide. Scholars indicate that in several national systems, policy bundles are now aimed at moving up international rankings and increasing reputational exposure to enhance institutional legitimacy, competitiveness, and access to funding (Guo et al., 2023; Kaidesoja, 2022). Simultaneously, literature on research assessment shows that even in those cases where organisations express their commitment to responsible evaluation, the evaluative results are tightly intertwined with journal hierarchies and ranking tools (Morgan-Thomas et al., 2024; Rushforth & De Rijcke, 2024). Such coupling is relevant, as rankings and the regimes of metrics surrounding them are more likely to support countable,



globally intelligible outputs that are not necessarily aligned with local missions or other commitments to the public (Jowi, 2024; Mfengu & Raju, 2024).

Because rankings heavily shape what counts as “quality,” their effects are especially consequential for African higher education. African universities tend to operate in environments of resource scarcity and imbalanced research provision, while at the same time being subjected to increased demands in terms of employability, digital change, and social impact (Jowi, 2024; Thondhlana & Garwe, 2021). Still, traditional rankings are biased towards indicators that recognise high-volume publication in globally indexed outlets and international reputation signals that are traditionally held in the Global North (Kaidesoja, 2022; Mfengu & Raju, 2024). Consequently, “quality” is often operationalised through proxy indicators that can undercount locally relevant scholarship, applicable knowledge, and development-facing engagement – domains in which many African university missions are anchored (Khanna et al., 2022; Thondhlana & Garwe, 2021).

Quality assurance (QA) systems are often branded as a remedy for ranking distortions in many higher education systems, especially where external accountability and competition are heightened by rankings, and they offer evidence-based solutions to improve teaching, research, governance, and the student experience. However, according to modern authors, the QA regimes can turn out to be a tool of audit culture: they extend the logics of external accountability into internal key performance indicators (KPIs) and compliance routines that might shift the orientation away from substantive educational change (Asamoah et al., 2025; Funck, 2024). Additional comparative studies of ranking systems in various contexts across countries further demonstrate that institutions restructure their internal reporting and monitoring practices to meet indicators, often at the expense of strategic gaming rather than learning on missions (Galleli et al., 2022; Morgan-Thomas et al., 2024).

In this research paper, the main contention is that, despite the pressure of ranking on African institutions of higher learning, the dilemma is not whether universities ought to be accountable, but rather why certain forms of accountability ought to take precedence and how they are to be demonstrated. An increase in research assessment reform has demanded the pluralisation of evidence of quality, integrating quantitative metrics with a qualitative judgement, contextualisation, or consideration of unintended impacts (Mfengu & Raju, 2024; Rushforth & Hammarfelt, 2023). The same reasoning has been found in higher education governance research, which highlights that the paradox of rankings is that they are powerful despite their methodological flaws and the fact that their deleterious influence is common knowledge (Kaidesoja, 2022; Rushforth & De Rijcke, 2024).

The study posits the practice of Mission-Led Quality Assurance (MLQA) as a form of sovereignty. Sovereignty here refers in a non-statist sense: the ability of institutions and systems to conceive, manifest, and tend governance of so-called quality in ways that are a priori to both local missions and the responsibility of serving the common good, but without rendering illegible to external stakeholders. The African knowledge production scholarship emphasises that governance tools, assessments, reward models, and publication institutions condition the production and assessment of the type of knowledge (Jowi, 2024; Khanna et al., 2022; Thondhlana & Garwe, 2021). Yet, per MLQA, quality assurance is considered a place where epistemic priorities are either maintained or lost, as internal standards are either mirror reflections of the rankings or embedded in the context and the mission (Mfengu & Raju, 2024; Morgan-Thomas et al., 2024).

Three theoretical perspectives overlap in their informed approach to the study. Firstly, we could use Institutional Theory to comprehend the reasons why rankings prompt universities to pursue ranking-congruent strategies: the ranking is a status symbol, which triggers isomorphic conformity to internationally recognised practices of greatness, even in places where they are known to be



incompatible with local circumstances (Guo et al., 2023; Kaidesoja, 2022). Second, the Quality Culture theory postulated that sustainable improvement is not based solely on compliance but on shared values, participatory reflection, and organisational learning; recent scholarship on QA attributes successful QA to the development of internal quality cultures and ongoing improvement routines (Asamoah et al., 2025; Pham et al., 2025). Third, Responsible Metrics and Responsible Research Assessment models focus on the idea that indicator utilisation should be context-dependent, open, and accompanied by expert judgment, as observable evidence indicates that declarations alone will rarely decouple evaluation outcomes with journal ratings (Morgan-Thomas et al., 2024; Rushforth & Hammarfelt, 2023; Rushforth & De Rijcke, 2024).

Against this background, the study addresses the following problem: despite the proliferation of QA units, standards, and evaluation cycles, many universities experience a persistent misalignment between mission statements and the indicators that drive internal decision-making, particularly under ranking pressure. Where QA functions mainly as an audit interface, it may intensify mission drift by embedding ranking logics into internal KPIs and documentation practices (Funck, 2024; Galleli et al., 2022). Conversely, a mission-led QA architecture could enable institutions to (i) defend mission-relevant definitions of quality, (ii) evidence outcomes using pluralised indicators and qualitative judgement, and (iii) negotiate external accountability without substituting rankings for improvement (Mfengu & Raju, 2024; Jowi, 2024).

Accordingly, this study asks: What does literature suggest about (1) how rankings shape QA and internal performance management, and (2) what design principles can underpin a mission-led QA architecture that supports sovereignty in African higher education?

Using a programme-theory-informed narrative synthesis, the study derives an MLQA architecture, distils practical implementation steps, and discusses implications for university leaders, national QA agencies, and funders.

### **Materials and Methods**

This section outlines the methodological approach used to develop the Mission-Led Quality Assurance (MLQA) framework. The author describes the review design, search strategy, and source identification, eligibility criteria and screening, data extraction and synthesis procedures, and the steps taken to enhance transparency and assess robustness.

#### ***Research Design: Narrative Synthesis***

The paper used the narrative synthesis method to combine a heterogeneous set of current literature regarding the question of how global rankings and ranking-based performance measures influence internal QA and the way in which improvement in evaluation, such as responsible metrics and mission-led assessment) is achieved. can be institutionalised. Narrative synthesis is used when the evidence is cross-disciplinary and cross-methodological, and its purpose is to describe mechanisms and circumstances (instead of putting estimates of pooled effects). Qualitative case studies, policy analysis, survey research, and bibliometric work that are challenging to commensurate in a meta-analytic framework may be integrated through transparent, theory-informed narrative comparison.

A four-element logic (i) formulating a preliminary programme theory; (ii) systematic searching and screening; (iii) systematic extraction and synthesising relationships and themes; and (iv) robustness testing through studying study design, the problem transferred from context, and finding convergence/divergence) was used to operationalise the study's narrative synthesis. The programme theory of this work defines the causal mechanisms between the pressure of ranking and the organisational responses: the measurable proxies as results of prioritisation under the influence of ranking pressure reconfigure routines of QA (what is audited, reported, rewarded, and improved),



and could result in mission drift unless the governance structures re-establish evaluation to organisational purpose. The entire protocol and audit trail, including the evidence-to-claim matrix used to derive the results clusters, are in Appendix A.

### ***Identification of Sources***

The search strategy focused on peer-reviewed research articles published between January 2021 and January 2026 to reflect post-2020 changes in IHE governance, digitalisation, and research evaluation reform. The search was performed using Scopus, Web of Science Core Collection, and ERIC, and supplemented with focused Google Scholar searches to reduce indexing gaps typical in higher education policy and evaluation scholarship. Search strings combined three concept blocks: (a) rankings/performance indicators (e.g., rank\*, league-table\*, "world university ranking\*", bibliometric indicator\*), (b) quality assurance/evaluation (e.g., "quality assurance", audit, accreditation, "quality culture", evaluation, assessment), and (c) mission/sovereignty/responsible metrics (e.g., mission-led, "responsible metrics", "responsible research assessment", sovereignty, decolon\*). Appendix A (Table A.1) documents databases, dates and complete search strings.

### ***Eligibility Criteria and Screening***

Inclusion criteria included: (i) peer-reviewed journal articles or in press; (ii) published between January 2021 –January 2026; (iii) substantial emphasis on rankings and/or the implications of rankings regarding QA, evaluation, governance, academic work, or the institutional strategy; (iv) an explicit empirical foundation (qualitative, quantitative or mixed methods) or a conceptual/theoretical contribution based on documented evidence; (v) and accessible in full text. Exclusion criteria were: evidence-free editorial, non-scholarly commentary, all articles that just ranked methodology with no governance/QA implications, and duplicates.

### ***Data Extraction and Narrative Synthesis Procedure***

The article used a structured extraction framework to enable consistent comparison across diverse designs. For each included study it extracted: region/country; institutional type; unit of analysis; ranking or indicator regime implicated; QA mechanism examined (internal QA, accreditation, audit, performance management, teaching evaluation); study design and data sources; reported mechanisms (how ranking pressure translated into QA practice); and reported outcomes for priorities, equity, academic labour, and knowledge production. The template used to extract the data is available in the Appendix (Table A.3), along with a sample of a filled-in record.

The process of narrative synthesis was in three steps of analysis. First, the paper has generated within-study summaries to retain contextual specificity and the causal claims made by individual authors. Second, the extracted findings underwent inductive coding to derive descriptive themes; codes were then consolidated into five general results groups aligned with the programme theory. Third, cross-study comparison by mapping clusters to (i) contexts that they appeared in, (ii) facilitating conditions, and (iii) reported outcomes using a synthesis matrix. This helped the study differentiate between repeated structures and context-dependent phenomena and provided the MLQA design proposal as an evidence-based proposal rather than a normative model (Appendix A, Table A.4).

### ***Robustness, Limitations and Transparency***

Due to a lack of pooled estimates of the effects in narrative synthesis, triangulation across methods (such as similarity of findings among qualitative, survey, and bibliometric) and reidentification of the same effect across diverse contexts, as well as focus on disconfirmation or boundary evidence, was used to assess strength. Where a claim relied on a single-context study or primarily conceptual argument, it is treated as hypothesis-generating rather than definitive. As this study used only



publicly available secondary sources, no human participants were involved, and no ethical approval was required; transparency is provided through the appendix audit trail.

## **Results**

The synthesis of the narrative provided five result clusters. Collectively, they clarify how pressure ranking is converted into QA routines and control decisions, why reform declarations tend to perform poorly, and what an evidence-based, mission-led quality assurance (MLQA) architecture would entail. The mapping between evidence and claims for these clusters is recorded in Appendix A (Table A.4), which connects each cluster to the specific sources and findings used to support it.

### ***Rankings Amplify Audit Logics and Narrow Proxies of Quality***

In the corpus, the rankings were regularly described as control mechanisms that incentivise a limited set of internationally-recognised products (publications, citations, reputation surveys) and enable institutions to construe rank movement as a measure of quality improvement (Guo et al., 2023; Kaidesoja, 2022). Quantitative indicators and journal lists, even in a post-declaration setting, are retainers of expert judgement in research evaluation, suggesting that symbolic commitments to responsible assessment are rarely implemented in evaluative outcomes (Morgan-Thomas et al., 2024; Rushforth & De Rijcke, 2024). In African settings, the issues revolve around the potential shift from metric-biased assessment to alter developmental missions by bypassing local social ills and placing research goals out of context (Jowi, 2024; Mfengu & Raju, 2024).

### ***Quality Assurance Often Becomes a Translation Layer for Ranking Indicators***

Evidence indicates that, in many cases, internal QA departments operationalise external expectations through internal dashboards and documentation systems, translating ranking indicators into institutional KPIs. This dynamic can be observed in the literature on research evaluation and organisational research on accreditation and performative technologies: documentation and measurement are ends in themselves, shaping the identities of academics and workplace practices (Funk, 2024; Morgan-Thomas et al., 2024). As a result, QA cycles tend to shift towards demonstrating compliance and competitiveness rather than identifying learning and teaching issues or reinforcing mission-related outcomes (Asamoah et al., 2025; Kaidesoja, 2022).

### ***Declarations and Reforms Have Limited Traction without Governance Redesign***

Responsible metrics and responsible research assessment reforms seek to decrease excessive dependence on journal rankings and simple indicators. However, more recent articles indicate limited behavioural change, in which incentives, promotion guidelines, and assessment templates still encode the journal hierarchy (Morgan-Thomas et al., 2024; Rushforth & Hammarfelt, 2023). Empirical data on the practice of evaluation indicate that the meaning of responsibility, as perceived by professional evaluators, is differentiated and unclear in many instances, implying that principles need to be operationalised and institutionally implemented (Rushforth & De Rijcke, 2024). Elsewhere, in South Africa, researchers explain that metric-based systems are prone to biased recognition and the manipulation of behaviour, leading to a proposal to bind institutions and nations to change (Mfengu & Raju, 2024).

### ***Mission-led QA Requires Plural Evidence and Participatory Judgement***

Across QA and research evaluation literatures, there is a uniform suggestion to pluralise evidence of quality: to integrate quantitative indices with qualitative judgement, narrative descriptions of contribution, and contextual standards (Mfengu & Raju, 2024; Rushforth & Hammerfelt, 2023). Within higher education QA scholarship, quality culture is placed as a pre-requisite of a more enduring change - getting shared values, alignment in leadership, and participatory rituals instead of top-down compliance (Asamoah et al., 2025; Pham et al., 2025). As far as Africa is concerned, mission-led



evidence should also acknowledge other types of scholarship and engagement, such as locally based research, co-production of knowledge, and societal impact (Jowi, 2024; Thondhlana & Garwe, 2021).

A four-pillar Mission-Led Quality Assurance (MLQA) architecture emerges from the evidence. Synthesising the findings, the article derives an MLQA architecture with four pillars: (i) mission-anchored standards and outcomes; (ii) pluralised evidence rules (responsible metrics plus qualitative judgement); (iii) participatory review cycles that embed quality culture; and (iv) public-value accountability, making visible contributions to development priorities. Table I operationalises these pillars into implementable QA elements and evidentiary practices.

*Table I: Mission-Led Quality Assurance (MLQA) architecture derived from thematic synthesis*

Pillar	Design Principle (What Changes)	Typical Evidence	Illustrative Sources
Mission-anchored standards	Translate mission into explicit outcomes and thresholds; avoid substituting rank movement for quality.	Mission outcome statements; programme learning outcomes; equity and inclusion indicators; engagement commitments.	Jowi (2024); Kaidesoja (2022)
Pluralised evidence rules	Use responsible metrics contextually; require narrative justification; prohibit single-metric decisions.	Contextualised bibliometrics; portfolio review; peer judgement rubrics; societal impact narratives.	Mfengu & Raju (2024); Morgan-Thomas et al. (2024)
Participatory review cycles	Embed quality culture through collegial review, learning loops, and developmental feedback.	Self-evaluation reports; peer review minutes; improvement plans; follow-up audits focused on learning.	Asamoah et al. (2025); Pham et al. (2025)
Public-value accountability	Make development-facing contribution auditable: partnerships, graduate capability, knowledge use.	Stakeholder feedback; tracer studies; policy uptake evidence; community co-production logs.	Jowi (2024); Thondhlana & Garwe (2021)

**Discussion**

The findings affirm that the rankings are an effective tool for legitimacy and coordination and motivate institutions to treat a small set of proxies as binding quality indicators. This finding echoes the theorisation of the rankings paradox: rankings do influence organisational behaviour despite the extensive distribution of limitations and unintended effects (Kaidesoja, 2022). Empirical evidence on national initiatives and ranking improvements supports the idea that rank movement has become embedded in policy programmes and institutional strategies (Guo et al., 2023), thereby adding pressure on internal systems to obtain a reading of the ranking template.

The synthesis further demonstrates that it is not permissible to assume that QA can automatically decline the effect of ranking distortions. Rather, QA units tend to become translation layers that inject ranking indicators into internal dashboards and documentation processes and recreate audit logics within the institution. Accreditation as a performative governance tool is a focus of organisational studies that clarify how measurement regimes can be used to restructure academic identities and priorities, turning measurability into part of professional legitimacy (Funck, 2024). Equivalent dynamics exist in the evaluation of research: despite the many institutions stating they favour



responsible assessment, evaluators still depend extensively on journal ranking cues (Morgan-Thomas et al., 2024).

A major lesson of mission-led QA design is the limited traction of reform announcements. According to Rushforth and Hammarfelt (2023), reform movements succeed when they turn principles into working standards, incentives, and routines that minimise reliance on simplistic indicators. Qualitative data from evaluation practice show that responsibility is construed based on local professional norms and constraints, and that capacity building, as well as clear-cut decision principles, are recommended (Rushforth & De Rijcke, 2024).

The nature of global knowledge production heightens the sovereignty wager of African universities. In places where appraisal infrastructure rewards globally indexed output and reputational indicators, scholarship that is locally relevant and other types of knowledge might be systematically undervalued, reinforcing historical disparities (Khanna et al., 2022; Thondhlana & Garwe, 2021). The South African case of metric issues highlights how indicator regimes may shift researchers' behaviour and benefit specific fields of study and publication outlets, thereby inspiring responsiveness in assessment in the Global South (Mfenge & Raju, 2024).

The architecture of MLQA established in this study reacts to these limitations by integrating three streams of scholarship: institutional theory (legitimacy pressures), theory of quality culture (learning-oriented improvement), and theory of responsible metrics (pluralised, contextual evidence). There is a practical synthesis of the four pillars. Mission-linked standards help overcome mission drift by defining outcomes that are locally important (Jowi, 2024). Pluralised evidence rules address the continued existence of journal hierarchies by restricting the use of metrics and mandating that they be described in narrative form (Morgan-Thomas et al., 2024; Mfengu & Raju, 2024). Participatory review cycles operationalise a culture of quality by integrating loops of improvement in place of sporadic compliance (Asamoah et al., 2025; Pham et al., 2025). Lastly, there is the public-value accountability, which highlights contributions to development that stand-alone conventional rankings overlook (Thondhlana & Garwe, 2021).

Implications for practice follow. Table I is a governance checklist that university leaders can use to redesign IQA systems and KPI dashboards: any adopted indicator must be traceable to the mission, accompanied by interpretive guidance and qualitative judgement. To discourage incentive gaming by institutions, national QA agencies can update audit tools to identify multiple outputs and development contributions. By linking grant and performance contracts to mission outcomes rather than league-table positions, and by encouraging data systems that reflect graduate capabilities, partnerships, and societal uptake, funders can strengthen MLQA.

### **Conclusions**

This synthesis of the narrative combined the current literature on rankings, audit culture, quality assurance, and responsible research assessment to describe how ranking incentives translate into routines of internal QA and to infer an evidence-based Mission-Led Quality Assurance (MLQA) architecture. The findings indicate that accountability is not the key issue, but rather dominance of narrow proxy measures as organisational objectives. A workable governance response proposed by MLQA is to re-ground QA in the institutional mission, broaden the evidentiary foundation to include an array of non-single-number indicators, and deepen deliberative review and learning-enabling accountability.

Future research should support the validation of MLQA in multi-site, empirical case studies in Eastern Africa, with stakeholder discussions with students, employers, communities, and QA agencies. Mixed-method assessments might also be used to determine whether the adoption of MLQA reduces



metrics gaming, promotes better teaching and learning outcomes, and enhances public-value contributions without compromising external accountability.

### Data transparency statement

The research relied on publicly available secondary sources; no primary data collection was conducted. The complete reference list is provided at the beginning.

### Acknowledgements

The author thanks colleagues and peer reviewers for constructive feedback that strengthened the argument and framing of this manuscript. The author also acknowledges the scholars and practitioners in African higher education whose work on responsible research assessment and mission-led quality assurance informed this review. Any errors or omissions remain the responsibility of the author.

### References

- Asamoah, M. K., Ansong, J. D., Mackin, E., Agyekum, B., & Eshun, S. N. (2025). The influence of quality assurance on achieving development education goals. *Discover Education*, 4(1), Article 468. <https://doi.org/10.1007/s44217-025-00694-1>
- Funck, E. K., Kallio, K.-M., & Kallio, T. J. (2024). The persuasion of performative technologies: Constructing calculating selves in universities. *Journal of Accounting & Organisational Change*, 20(6), 1-26. <https://doi.org/10.1108/JAOC-05-2023-0082>
- Galleli, B., Teles, N. E. B., Santos, J. A. R. dos, Freitas-Martins, M. S. dos, & Hourneaux Junior, F. (2022). Sustainability university rankings: A comparative analysis of UI GreenMetric and the *Times Higher Education World University Rankings*. *International Journal of Sustainability in Higher Education*, 23(2), 404-425. <https://doi.org/10.1108/IJSHE-12-2020-0475>
- Guo, C., Hao, X., Wu, J., & Hu, T. (2023). The effect of national higher education initiatives on university rankings. *Humanities and Social Sciences Communications*, 10(1), Article 527. <https://doi.org/10.1057/s41599-023-02034-w>
- Jowi, J. O. (2024). Recent developments in higher education in Africa: Partnerships for knowledge transformations. *Journal of International Cooperation in Education*, 26(1), 34-48.
- Kaidesoja, T. (2022). A theoretical framework for explaining the paradox of university rankings. *Social Science Information*, 61(1), 128-153. <https://doi.org/10.1177/05390184221079470>
- Khanna, S., Ball, J., Alperin, J. P., & Willinsky, J. (2022). Recalibrating the scope of scholarly publishing: A modest step in a vast decolonisation process. *Quantitative Science Studies*, 3(4), 912-930. [https://doi.org/10.1162/qss\\_a\\_00228](https://doi.org/10.1162/qss_a_00228)
- Mfengu, A., & Raju, J. (2024). Challenges with measures used for assessing research impact in higher education institutions. *South African Journal of Science*, 120(7/8), Article 14934. <https://doi.org/10.17159/sajs.2024/14934>
- Morgan-Thomas, A., Tsoukas, S., Dudau, A., & Gaska, P. (2024). Beyond declarations: Metrics, rankings and responsible assessment. *Research Policy*, 53(10), Article 105093. <https://doi.org/10.1016/j.respol.2024.105093>
- Naven, M., & Whalen, D. (2022). The signalling value of university rankings: Evidence from top 14 law schools. *Economics of Education Review*, 89, Article 102282. <https://doi.org/10.1016/j.econedurev.2022.102282>
- Pham, H. T., Nguyen, P. V., & Nguyen, N. D. H. (2025). IQA models: An introductory literature review and suggestions for developing an IQA model for Vietnamese higher education. *Quality in Higher Education*, 31(3), 1-25. <https://doi.org/10.1080/13538322.2025.2553987>



- 
- Rushforth, A., & De Rijcke, S. (2024). Practising responsible research assessment: A qualitative study of faculty hiring, promotion, and tenure assessments in the United States. *Research Evaluation*, 33, rvae007. <https://doi.org/10.1093/reseval/rvae007>
- Rushforth, A., & Hammarfelt, B. (2023). The rise of responsible metrics as a professional reform movement: A collective action frames account. *Quantitative Science Studies*, 4(4), 879–897. [https://doi.org/10.1162/qss\\_a\\_00280](https://doi.org/10.1162/qss_a_00280)
- Thondhlana, J., & Garwe, E. C. (2021). Repositioning of Africa in knowledge production: Shaking off historical stigmas – Introduction. *Journal of the British Academy*, 9(S1), 1–17. <https://doi.org/10.5871/jba/009s1.001>



**Appendix A. Narrative synthesis protocol and audit trail**

This appendix documents the narrative synthesis procedures used to derive the results clusters and the MLQA architecture, providing an audit trail from source identification to synthesis claims. Items are presented to enhance the transparency and replicability of this desktop narrative synthesis.

**A.1 Search strategy and databases**

Table A.1 summarises the databases searched, date limits, and the core search strings used. Search strings were adapted per database syntax and supplemented with backward and forward citation searching for key review and framework papers.

*Table A.1. Databases and core search strings used for identification of sources*

Database / index	Date limits	Core search string (illustrative)
Scopus	2021–2026	(rank* OR "league table*" OR "world university ranking*") AND ("quality assurance" OR audit OR accreditation OR "quality culture" OR evaluation) AND (mission-led OR "responsible metrics" OR sovereignty OR decolon*)
Web of Science Core Collection	2021–2026	TS=(rank* OR "world university ranking*" OR "performance indicator*") AND TS=("quality assurance" OR audit OR accreditation OR evaluation) AND TS=("responsible metrics" OR "research assessment" OR mission*)
ERIC	2021–2026	(rank* OR "university ranking*") AND ("quality assurance" OR accreditation OR evaluation)
Google Scholar (targeted)	2021–2026	"responsible metrics" rankings quality assurance mission-led; "responsible research assessment" university evaluation; rankings audit culture higher education

**A.2 Screening summary and inclusion counts**

Table A.2 provides a structured summary of identification, screening and inclusion. Counts reflect the final set of sources used for synthesis in Sections 3–4; the screening log with exclusion reasons is retained by the author and can be shared on request.

*Table A.2. Screening summary (identification to inclusion)*

Stage	Count	Notes
Records identified (databases + targeted search)	Not reported	Database and targeted search results were deduplicated; see Table A.1 for sources and strings.
Title/abstract screened	Not reported	Title/abstract screening applied scope and date limits; non-HE and ranking-method-only items excluded.
Full texts assessed	Not reported	Full-text screening applied QA/evaluation relevance and evidence criteria; exclusion reasons logged.
Studies included in synthesis	14	Final corpus used for narrative synthesis; corresponds to the 14 items in the reference list.

**A.3 Extraction template**

Table A.3 lists the fields extracted from each included study to support cross-study comparison and to enable the evidence-to-claim mapping reported in Table A.4.



*Table A.3. Data extraction fields*

Field	Description
Bibliographic details	Author(s), year, title, journal, DOI
Context	Country/region, institutional type, system characteristics
Ranking/indicator regime	Ranking system or indicator set implicated; how it operates
QA mechanism	Internal QA, accreditation, audit, performance management, teaching evaluation
Design and data	Qualitative/quantitative/mixed; data sources; sample
Mechanisms	How ranking pressure translates into behaviour or organisational routines
Outcomes	Effects on priorities, equity, academic labour, knowledge production
Boundary conditions	Conditions under which the mechanism does not hold or reverses
Author recommendations	Policy/management implications offered by the study

**A.4 Evidence-to-claim mapping for results clusters**

Table A.4 illustrates how the five result clusters were derived by mapping extracted mechanisms and outcomes from the included studies to each cluster. The ‘Key supporting sources’ column lists the principal studies that contributed evidence to each cluster, enabling readers to trace the synthesis logic.

*Table A.4. Results clusters and supporting evidence (audit trail)*

Results cluster (Section 3)	Mechanism summary (synthesised)	Key supporting sources (APA in-text)
3.1 Rankings amplify audit logics and narrow proxies of quality	Rankings privilege commensurable indicators, encouraging auditability and proxy substitution.	Funck et al., 2024; Kaidesoja, 2022; Guo et al., 2023; Galleli et al., 2022; Naven & Whalen, 2022
3.2 QA becomes a translation layer for ranking indicators	Internal QA routines and dashboards align to ranking metrics, shifting improvement work toward measured domains.	Asamoah et al., 2025; Pham et al., 2025; Funck et al., 2024; Jowi, 2024
3.3 Reform declarations have limited traction without governance redesign	Principles of responsible metrics are adopted rhetorically but lack enforcement, incentives, and deliberative infrastructures.	Morgan-Thomas et al., 2024; Rushforth & Hammarfelt, 2023; Rushforth & de Rijcke, 2024; Mfengu & Raju, 2024
3.4 Mission-led QA requires plural evidence and participatory judgement	Plural metrics, peer judgement and stakeholder participation counter single-number optimisation and support contextual quality.	Morgan-Thomas et al., 2024; Rushforth & de Rijcke, 2024; Rushforth & Hammarfelt, 2023; Khanna et al., 2022
3.5 MLQA architecture (four pillars)	Governance design re-anchors QA to mission through (i) purpose statements, (ii) plural evidence, (iii) deliberative review, (iv) learning-oriented accountability.	Morgan-Thomas et al., 2024; Rushforth & Hammarfelt, 2023; Asamoah et al., 2025; Pham et al., 2025; Mfengu & Raju, 2024