

KEMI Journal of Educational Leadership and Management

KJELM

[ISSN 3079-4048]

Volume: 2 Issue: 3 | Nov-2026

Institutional pathways for the use of foundational learning assessment evidence in Kenya

Winnie Chebet Tubey[1], **Nancy Ochiba**[1], **Rosy Rodha Njiri**[2], **Daniel Wesonga**[3], **Robert Kinanga**[4]

Technical University of Kenya, Kenya[1], *Ministry of Education, Kenya*[2], *Women Educational Researchers of Kenya (WERK)*[3], *Maasai Mara University, Kenya*[4]

Corresponding Author Email: winnie.tubey@tukenya.ac.ke

Cite this article in APA

https://doi.org/10.68147/49cace1b

Copyright: ©2026 by the author(s). Open access under CC BY.

Abstract

Foundational learning assessments generate evidence for education policy and practice, yet their availability does not establish their use. This article examines how institutional arrangements shape the mobilisation and utilisation of assessment evidence in Kenya, with particular attention to institutional response. It compares three differently positioned systems: the state-led National Assessment System for Monitoring Learner Achievement (NASMLA), the citizen-led Uwezo assessment and the Foundational Literacy and Numeracy Assessment (FLANA). Guided by Weiss's research utilisation theory, the study used a documentary-led qualitative multiple-case design drawing on 61 documentary sources and six key-informant interviews documented through expanded field notes. Within-case and cross-case analysis distinguished mobilisation from conceptual/diagnostic, agenda-setting, procedural and instrumental utilisation, and from subsequent response. NASMLA showed the clearest documented procedural and selected instrumental pathways through government planning and programme structures. Uwezo's strongest documented contribution was conceptual and agenda-setting, supported by household engagement and public communication. FLANA provided wider diagnostic evidence on learning conditions and contributed principally to conceptual and agenda-setting use; its downstream instrumental pathways were less well documented. Across the cases, the evidentiary trail became less complete between utilisation and subsequent response, implementation and feedback. Response ownership emerged as an important institutional link: evidence can enter decision processes without a clear route to an actor with

responsibility, authority and implementation support to act. The findings favour connecting complementary assessment systems to identifiable decisions, institutional responsibilities and feedback without equating visibility or citation with demonstrated impact.

Key terms: foundational learning, assessment evidence, research utilization, institutional pathways, institutional response, education policy, Kenya

1.0 INTRODUCTION

1.1 Background of the Study

Foundational literacy and numeracy underpin children's subsequent learning and progression through education. As access to schooling has expanded, concern has increasingly shifted from participation alone to whether children acquire the competencies needed to continue learning. Learning assessments have consequently assumed a more prominent role in education systems, providing evidence for diagnosing learning difficulties, identifying inequalities, monitoring progress and informing decisions about policy, programmes and educational practice.

Generating assessment evidence is only part of this process. Evidence must also reach relevant actors, be interpreted within particular institutional settings and connect with decisions or practices through which a response is possible. These processes rarely follow a straightforward sequence. Research can alter how a problem is understood, contribute to its prioritisation, enter routine organisational processes or inform a specific decision without moving directly from publication to policy (Weiss, 1979; Nutley et al., 2007).

The institutional position of an assessment system matters in this regard. State-led assessments are often closely connected to formal planning, accountability and programme structures. Citizen-led assessments have developed different routes to influence by making learning outcomes accessible to households, communities, civil society and public debate. Other approaches combine measures of children's competencies with information on the household, school and system conditions in which learning takes place. These institutional differences affect the audiences an assessment can reach, the authority attached to its evidence and the routes through which that evidence may enter policy and practice.

Research on evidence use increasingly reflects this broader institutional perspective. Rickinson et al. (2022), drawing on a systematic review and narrative synthesis of 112 publications, identify two core components of quality evidence use: appropriate research evidence and thoughtful engagement and implementation. Both are supported by individual and organisational conditions and shaped by system-level influences. Research on knowledge brokerage similarly draws attention to relationships and intermediary processes through which knowledge crosses organisational and professional boundaries (Rechsteiner et al., 2024). Credible evidence may therefore remain weakly used where it is poorly connected to organisational routines, decision opportunities or actors able to respond.

Evidence from developing-country settings also cautions against defining utilisation solely in terms of direct policy adoption. Fussy's (2022) synthesis of national studies from Africa, Asia and Latin America identifies both conceptual and instrumental uses of research. Evidence can matter by changing how a problem is understood as well as by informing a particular policy decision. The distinction is important because an overly narrow definition can overlook genuine influence, whereas an overly expansive one can mistake dissemination or visibility for use.

Kenya provides a useful setting in which to examine these issues. Foundational learning evidence is

generated through state and non-state assessment systems with different institutional positions. NASMLA, implemented by the Kenya National Examinations Council (KNEC), forms part of the formal national assessment architecture (Kenya National Examinations Council, n.d.). Uwezo developed as a citizen-led, household-based assessment within Twaweza, bringing evidence on basic competencies into community and public discussion (Uwezo, 2012, 2016). Usawa Agenda emerged from a phased institutional separation of Uwezo's Kenyan work from Twaweza, retained the citizen-led assessment heritage and subsequently implemented FLANA (Usawa Agenda, n.d., 2024, 2026). Uwezo and FLANA are thus historically related, but analytically distinct assessment models with different reporting histories and diagnostic emphases.

These systems bring different institutional resources to the evidence-use process. NASMLA has formal authority and access to established planning and programme structures. Uwezo developed considerable public intelligibility and networked reach. FLANA combines learning assessment with a broader diagnostic account of household, school and system conditions. These differences do not provide a basis for ranking the assessments, nor do they make their learner results directly comparable. The populations, constructs, instruments and purposes differ. They do, however, make it possible to examine how institutional position shapes the routes through which evidence travels.

The relevance of this question is apparent within Kenya's foundational learning research ecosystem. A 2025 national workshop convened by ESSA and the REAL Centre identified weak translation of research into policy, limited coordination among researchers, policymakers and implementers, and low research uptake. Participants called for greater clarity about which actors should take ownership of findings or follow them up (Okullo et al., 2025). This concern directs attention beyond dissemination to the institutional arrangements linking findings with decisions and subsequent action.

Against this background, the study compares NASMLA, Uwezo and FLANA as institutional evidence pathways. It examines how institutional position, relationships and translation arrangements shape evidence mobilisation; the forms of utilisation that can be substantiated; and the conditions affecting movement from utilisation towards institutional response.

1.2 Statement of the Problem

Although Kenya has several sources of foundational learning assessment evidence, the institutional route from evidence mobilisation to use and then to response remains insufficiently understood. The unresolved question is comparative: how do state-led, citizen-led and broader diagnostic assessments, differently positioned within one education system, move evidence into decisions, and what makes follow-through more or less traceable? This matters for educational leadership because diagnosing a learning problem does not automatically identify a decision-maker with the mandate, resources and support to respond.

Earlier Kenyan studies establish important elements of this problem. Wamahiu et al. (2022) identified constraints affecting awareness, access, coordination, stakeholder engagement and feedback. Monk (2020) found substantial Uwezo visibility in the Kenyan press but limited engagement with its citizen-action orientation. In a randomised household intervention, Lieberman et al. (2014) found no discernible effect of learning information and participation materials on the private or collective actions examined. These studies illuminate different aspects of evidence access, communication and response; they do not establish the comparative institutional pathways of NASMLA, Uwezo and FLANA.

What remains less clearly explained is how assessments with different institutional locations mobilise evidence, which forms of utilisation can be substantiated, and where the connection to institutional response becomes uncertain. These distinctions are consequential: dissemination cannot simply be

reported as uptake, a policy citation as instrumental use, or programme association as an implemented response.

A comparison of these pathways can therefore reveal the institutional opportunities and constraints through which education evidence becomes more, or less, actionable without conflating limited documentation with evidence of non-use.

The study examines how NASMLA, Uwezo and FLANA mobilise assessment evidence, what forms of utilisation can be documented, and which institutional conditions shape movement towards educational response.

1.3 Purpose of the Study

The study examined how institutional arrangements shape the mobilisation and utilisation of foundational learning assessment evidence and its transition towards institutional response in Kenya through a comparative analysis of NASMLA, Uwezo and FLANA.

1.4 Objectives of the Study

The study sought to:

examine how institutional positioning, relationships and evidence-translation arrangements shape the mobilisation of foundational learning assessment evidence;

identify and compare the forms of evidence utilisation documented across NASMLA, Uwezo and FLANA; and

examine the institutional opportunities and constraints affecting the transition from evidence utilisation to educational response.

1.5 Research Questions

How do institutional positioning, relationships and evidence-translation arrangements shape the mobilisation of foundational learning assessment evidence?

What forms of evidence utilisation can be documented across NASMLA, Uwezo and FLANA?

What institutional opportunities and constraints affect the transition from evidence utilisation to educational response?

2.0 LITERATURE REVIEW

Research on evidence use has moved considerably beyond the assumption that high-quality research, once disseminated, will find its way into policy or practice. Work on research utilisation, evidence-informed education and knowledge mobilisation instead points to a more contingent process in which evidence interacts with organisational routines, professional judgement, relationships, institutional authority and opportunities for decision-making. For learning assessments, this raises two related questions: what counts as evidence use, and what institutional conditions enable evidence to move beyond availability and visibility?

2.1 Theoretical Framework: Research Utilisation Theory

This study draws principally on Weiss's (1979) account of research utilisation. Weiss's starting point was the conceptual ambiguity surrounding the term use. The conventional problem-solving model assumes that a decision-maker confronts a defined problem, relevant research becomes available and the findings

contribute directly to a pending decision. Weiss did not dismiss this form of utilisation, but showed that it represents only one way in which research enters policy.

Research can also influence the concepts through which a problem is understood, accumulate gradually within policy thinking, interact with professional and political knowledge, or be mobilised in support of existing positions. For assessment research, this means that a study need not produce an identifiable policy decision to have influenced the way a problem is diagnosed or prioritised. At the same time, a broad conception of influence creates an evidentiary problem: if utilisation can take several forms, the evidence required to substantiate each form must be made explicit.

Nutley et al. (2007) develop this argument by locating research use within organisational and policy settings in which research is only one among several sources of knowledge. Utilisation is often incremental and interactive, and attribution to a single research product is difficult. Evidence use therefore depends not only on the properties of the evidence but also on the organisations, relationships and professional communities through which it is interpreted.

This concern is particularly pertinent in settings where research use is sometimes assumed to be limited. Fussy's (2022) best-evidence synthesis of national studies from Africa, Asia and Latin America, explicitly drawing on Weiss's typology, identified both conceptual and instrumental uses of research. The finding cautions against treating the absence of direct policy adoption as evidence of non-use. It also reinforces the need to distinguish forms of utilisation rather than collapse them into a single measure of uptake.

Drawing on this literature, the study operationalised four forms of utilisation. These are analytical categories adapted for this comparison, rather than a four-part typology attributed directly to Weiss (1979). Conceptual or diagnostic use refers to evidence that changes or deepens understanding of a problem. Agenda-setting use concerns evidence that contributes to the visibility or prioritisation of an issue. Procedural use occurs where evidence becomes incorporated into routine institutional processes such as planning, monitoring or reporting. Instrumental use requires a more identifiable connection between evidence and a decision or response.

These categories are not stages through which evidence must pass. Conceptual use need not lead to instrumental use, and procedural incorporation does not establish implementation. They are also analytically distinct from mobilisation. A report, policy brief, stakeholder meeting, media appearance or technical dialogue may move evidence into a new institutional space without establishing that the evidence was subsequently used.

Research utilisation theory is therefore useful here for two reasons. It makes it possible to recognise forms of influence that would disappear under an exclusively instrumental definition of use, while also imposing a discipline on claims of influence. The question is not simply whether an assessment was visible or cited, but what the available evidence permits one to say about how it was used.

2.2 Conceptual Framework

The initial analytical framework distinguished five related processes:

Interacting elements of the evidence pathway: evidence production | translation and mobilisation | utilisation | institutional response | implementation and feedback.

Evidence production concerns the generation, validation and institutional positioning of assessment findings. Translation and mobilisation refer to the processes through which findings are interpreted for different audiences and moved through reports, policy products, media, networks, dialogue and formal institutional interfaces. Utilisation concerns demonstrable incorporation of evidence into understanding, agendas, institutional procedures or decisions. Institutional response refers to an organisational

commitment or course of action arising in relation to the evidence. Implementation and feedback concern the execution of that response and the return of implementation experience to subsequent assessment and decision processes.

The framework does not assume that these elements occur once or in a fixed order. Evidence can influence understanding without leading to an immediate decision, enter a programme process without being implemented at every administrative level, or circulate repeatedly between evidence producers and users. Implementation experience may also redirect subsequent interpretation, mobilisation and assessment. Institutional position, organisational capacity, relationships, brokerage and decision opportunities were treated as conditions shaping these interactions.

2.3 Empirical Studies

2.3.1 Institutional Conditions and Evidence Use

A persistent issue in evidence-use research is the relationship between the quality of evidence and the capacity of organisations to use it. Methodological rigour and credibility remain important, but they are insufficient where evidence is poorly aligned with institutional needs, decision cycles or the capabilities of intended users.

Nutley et al. (2007) locate research use within organisational settings characterised by established routines, professional knowledge, incentives and competing sources of information. Rickinson et al. (2022) make a similar argument in education. Their framework connects appropriate research evidence with thoughtful engagement and implementation, supported by individual, organisational and system-level conditions.

The distinction complicates supply-oriented approaches to evidence use. Producing more rigorous research or improving access may address important constraints without resolving whether an organisation has a decision for which the evidence is relevant, the capacity to interpret it, or the authority and resources to respond. Formal institutional location can provide advantages because evidence is closer to routine planning and programme processes. Yet proximity to decision-making does not establish that evidence will travel through administrative levels or be translated into practice.

The converse also holds. Research produced outside government may have fewer direct procedural routes into formal decision-making while exerting substantial conceptual or agenda-setting influence. Fussy's (2022) synthesis is instructive in this respect: evidence use in developing-country settings is neither absent nor confined to formal policy adoption.

The literature therefore points to a distinction between credibility and access, on the one hand, and institutional usability, on the other. The former may create an opportunity for evidence to enter decision processes; the latter affects what institutions can do with it once it arrives.

2.3.2 Translation, Brokerage and the Limits of Dissemination

Evidence frequently has to cross organisational and professional boundaries before it can inform a decision. Translation is consequently more than simplifying technical findings. It involves connecting evidence to the questions, timing and institutional circumstances of potential users.

Research on knowledge brokerage examines one part of this process. Rechsteiner et al. (2024), in a systematic review of 42 studies of brokerage in educational change, show how brokers connect actors who would otherwise be weakly linked, facilitate knowledge exchange and support capacity development. Their review also reveals substantial variation in brokerage roles, strategies and outcomes.

For the present study, the distinction between connectivity and authority is especially important. A broker may bring researchers and policymakers together, translate findings or facilitate dialogue without controlling the subsequent decision. Trusted relationships can increase the likelihood that evidence is heard without conferring the mandate or resources required to act on it. Brokerage is therefore best treated initially as a mechanism of mobilisation rather than as evidence of utilisation in itself.

The same caution applies to dissemination. Accessible information can improve awareness and understanding while leaving the conditions required for action unresolved. Lieberman et al. (2014) tested this proposition in rural Kenya. Parents received information about children's literacy and numeracy performance together with materials intended to encourage greater involvement in education. The intervention had no discernible effect on the private or collective actions examined. The authors' account of the information-to-action pathway draws attention to intervening questions of understanding, salience, responsibility, feasible action and expected institutional responsiveness.

This does not diminish the importance of communication. Rather, it locates dissemination more precisely within the evidence pathway. Communication can create the conditions under which utilisation becomes possible; it cannot establish utilisation simply by demonstrating that information reached an audience.

2.3.3 Learning-Assessment Evidence Use in Kenya

Research in Kenya already provides evidence of uneven utilisation of learning assessments. Wamahiu et al. (2022), in a formative study of past and present learning assessments, examined awareness, access and use among policymakers, education officials, assessment and curriculum institutions, teacher educators, civil-society actors, assessment specialists and teachers. Their analysis identified difficulties associated with dissemination, coordination, stakeholder engagement, accessibility and feedback.

The gap between producing assessment evidence and using it is therefore already established. The more specific question concerns how that gap is structured across assessments with different institutional locations and pathways to influence.

The Uwezo literature illustrates why this distinction is necessary. Monk's (2020) analysis of the Kenyan press found that Uwezo succeeded in placing evidence of low learning outcomes prominently in public discourse, but that coverage engaged only minimally with the initiative's social-accountability and community-action philosophy. Public visibility was therefore substantial, while the route from visibility to citizen action remained less evident.

Lieberman et al. (2014) offer a complementary but analytically different form of evidence: a randomised experiment assessing household and collective responses to information. Monk (2020), by contrast, analysed press coverage, while Wamahiu et al. (2022) examined assessment use among education stakeholders. Their methods and units of inference differ from the institutional documentary tracing undertaken here. Lieberman et al.'s results do not establish Uwezo's overall influence; they support the narrower observation that information provision does not, by itself, demonstrate subsequent action.

The institutional problem is also evident in more recent work on Kenya's foundational learning research ecosystem. Participants in the 2025 ESSA–REAL Centre workshop identified weak evidence-to-policy translation, limited coordination among research, policy and implementation actors and the need for researchers to specify who should own or follow up their findings (Okullo et al., 2025).

The literature thus identifies organisational conditions for evidence use, the relevance and limits of dissemination, and uneven learning-assessment uptake in Kenya. Yet it offers less explanation of how differently positioned assessments within the same national system generate distinct forms of utilisation and how their documented pathways approach institutional response. This is the comparative gap

addressed by the study.

3.0 METHODOLOGY

3.1 Research Design

The study used a documentary-led qualitative multiple-case design to compare NASMLA, Uwezo and FLANA. The cases were selected because they occupy different positions within Kenya's foundational learning assessment landscape while addressing a common public problem. The comparison centred on their institutional evidence pathways rather than on learner scores.

The analysis was concerned with contribution rather than exclusive attribution. Assessment evidence enters policy and practice alongside programme priorities, professional knowledge, political and administrative considerations and other research. Claims were therefore limited to forms of influence or response for which a defensible evidentiary pathway could be established.

Documents constituted the principal empirical base. They were used to establish assessment design, institutional position, translation and mobilisation processes, reported utilisation and response. KIIs served a complementary explanatory function, particularly in relation to institutional relationships, interpretation, barriers to uptake and implementation conditions.

3.2 Case Selection and Evidence Base

NASMLA, Uwezo and FLANA were purposively selected to provide institutional contrast. NASMLA represents a state-led assessment located within the formal national assessment system. Uwezo represents the citizen-led household assessment tradition, combining population-based assessment with public communication and citizen engagement. FLANA represents a more recent household-based assessment that combines measures of foundational competencies with a wider set of contextual indicators.

The underlying study registered 61 unique documentary sources, including assessment reports, government and programme documents, technical materials, policy briefs, organisational publications and academic studies. Fifty-six were classified in the research audit trail as verified and five as verified with qualification. Verification concerned whether a source could substantiate the specific claim for which it was used; qualification signalled limits concerning date, coverage or interpretation. The five qualified records therefore required narrower claims rather than being treated as unreserved corroboration. The register covered assessment cycles through 2025 and related documents available in 2026.

Six anonymised key-informant interview (KII) records supplemented the documentary evidence. Informants were purposively selected for experience in assessment, government education policy, research mobilisation, citizen-led evidence and teacher education. Interviews were recorded as expanded field notes, not verbatim transcripts; individual records are cited only where they substantiate a particular interpretation. No classroom-teacher interview or direct observation of classroom implementation was included. FLANA was analysed primarily through publicly available documentary evidence.

The cases were not treated as methodologically equivalent. Differences in populations, settings, constructs, instruments and purposes precluded direct comparison of learner scores. Uwezo also has a longer assessment history than FLANA, and the interviews did not provide equivalent direct access to all producers. The comparison was therefore bounded to documented institutional position, evidence functions and pathways to use rather than equal-length histories or equally substantiated judgements about impact.

3.3 Research Instruments

Documentary evidence was organised through structured extraction and case-analysis matrices covering assessment characteristics, institutional positioning, credibility, translation, mobilisation, utilisation and reported response.

Semi-structured Key Informant Interview (KII) guides were developed after the preliminary documentary analysis. This sequencing allowed the interviews to probe issues that were unclear, contested or insufficiently explained in the documentary record rather than reproduce information already available in published sources.

The KII records consisted of expanded field notes rather than verbatim transcripts. Their analysis generated 79 meaning units and 149 code applications. The framework's distinctions between mobilisation, utilisation and response guided the initial organisation of material; cross-case reading was also used to consider interpretations not specified in the initial framework, including response ownership. These counts refer only to the KII analysis, are not documentary coding frequencies and do not measure the importance of themes. Interview saturation was not claimed.

3.4 Data Collection Procedures

Documentary review preceded the interviews. Sources were identified through organisational repositories, government websites, academic databases, digital platforms, media archives and reference tracing. Preliminary case analysis established each assessment's characteristics and possible evidence-use pathways and informed the subsequent KIIs.

The interviews were then used to examine the institutional processes surrounding the documentary record: how evidence acquired credibility, how relationships affected its movement, how state and non-state actors interacted, and what constrained movement from evidence availability towards use and response.

This sequencing gave the two evidence streams different but complementary roles. Documentary evidence anchored traceable public, historical, institutional, technical and numerical claims. KII evidence helped explain how those processes were experienced and interpreted by actors within the wider evidence ecosystem.

3.5 Data Processing and Analysis

Analysis proceeded within cases before moving to cross-case synthesis. Within-case analysis reconstructed each assessment's institutional position and pathways of evidence production, translation, mobilisation, utilisation and response. Cross-case analysis compared institutional positioning, credibility, relationships, translation and downstream response. KII meaning units were organised in participant, case and thematic matrices to test, explain, qualify or extend documentary interpretations. The research team considered alternative readings, especially whether evidence of dissemination had been mistaken for use, whether programme association established influence, or whether documentary silence was being read as absence of action. Interview evidence is presented as paraphrase, with anonymised identifiers where individual records substantiate a finding.

Claims about evidence use were assessed against explicit thresholds. Reports, launches, policy briefs, media engagement, stakeholder meetings and technical dialogue were treated as evidence of mobilisation unless a subsequent form of use could be established. Conceptual or diagnostic use required evidence that findings informed the understanding of a problem; agenda-setting use required evidence of contribution to attention or prioritisation; procedural use required incorporation into an identifiable

institutional process; and instrumental use required a traceable connection to a decision or response.

Relationships and brokerage were analysed as mechanisms through which evidence moved between actors and institutional settings. They were not treated as utilisation in the absence of evidence of subsequent use.

3.6 Analytical Rigour and Ethical Considerations

The study used source verification, cross-source comparison and an explicit audit trail. A source was treated as verified for a specified claim where the record supported that bounded interpretation. Sources verified with qualification required caution about date, coverage or interpretation; unsupported claims remained unresolved rather than entering the findings. These verification labels describe evidentiary fitness for particular claims, not a universal quality ranking of the publications.

Triangulation considered convergence, complementarity and inconsistency across sources. Documentary silence was not treated as evidence that utilisation or response had not occurred, and frequency of mention in the KIIs was not used as an indicator of importance. Policy citation, programme association and stakeholder engagement were not taken as evidence of causal influence.

The research was conducted under authorisation from the National Commission for Science, Technology and Innovation. KII participation was voluntary and based on informed consent. Participants could decline questions or withdraw, and confidentiality was maintained in reporting. Expanded field notes and related research materials were securely maintained by the research team.

The evidence base was not equally distributed across the three cases, with FLANA relying primarily on documentary evidence. The study therefore compares documented pathways rather than claiming identical depth across cases. It did not reanalyse learner-level datasets, compare assessment scores directly or estimate causal effects of assessment evidence on policy, programmes or learning outcomes.

4.0 RESEARCH FINDINGS AND DISCUSSION

4.1 Institutional Positioning and Evidence Mobilisation

The three assessment systems occupied distinct positions within Kenya's foundational learning evidence landscape. These positions shaped how evidence reached potential users and the kinds of institutional resources available to support its movement.

NASMLA's location within KNEC provided formal authority and established links to government planning and education programmes. Documentary evidence traced assessment indicators and findings into programme baselines, monitoring, reporting and teacher-support arrangements (Ministry of Education, 2020; World Bank, 2022, 2025). Its evidence therefore moved largely through established state and programme structures.

Uwezo developed a different route. Household assessment brought evidence directly into communities, while simplified communication, county-level products, media engagement and civil-society networks extended its reach into public, research and policy spaces. Its institutional influence was less dependent on formal location within government and more on the intelligibility of its evidence and the networks through which it travelled (Uwezo, 2012, 2016). The subsequent recognition of qualifying Uwezo/Usawa Agenda citizen-generated datasets within Kenya's statistical quality-assurance arrangements further illustrates how non-state evidence can acquire formal legitimacy without becoming state-produced evidence (Kenya National Bureau of Statistics [KNBS], 2022). This recognition applies to the specified datasets, not to FLANA as an assessment system.

FLANA retained the household-based orientation of citizen-led assessment but broadened the range of evidence generated. Alongside foundational competencies, it documented household, school and system conditions associated with learning. National, regional and thematic products created several points of entry for stakeholder and policy engagement (Usawa Agenda, 2024, 2025, 2026). Its route to influence therefore combined diagnostic breadth with developing relationships across the education evidence ecosystem.

The comparison is more informative when these differences are treated as distinct institutional resources rather than as indicators of relative strength. NASMLA brings formal authority and procedural access; Uwezo brings public intelligibility and networked reach; FLANA brings diagnostic breadth and multiple points of engagement. Across the reviewed pathways, institutional authority, public reach and diagnostic breadth were complementary strengths; no single system combined the strongest documented position in all three.

Institutional position nevertheless did not determine subsequent use. Paraphrased accounts from an evidence intermediary and a government informant described fragmented evidence holdings and insufficient synthesis, while highlighting the need to connect evidence products with identifiable users and decisions (KII-001; KII-002). Relationships could help evidence move between assessment producers, government, civil society and practitioners, but connectivity did not confer authority to act.

This pattern accords with research that locates evidence use within organisational and relational settings (Nutley et al., 2007; Rickinson et al., 2022). It also places a limit on what can be inferred from brokerage. Rechsteiner et al. (2024) show how brokers can bridge otherwise disconnected actors in educational change; the Kenyan cases indicate that such connections do not remove the distinction between gaining access to a decision space and possessing the mandate or resources to respond.

4.2 Forms of Evidence Utilisation across the Assessment Systems

Once mobilisation was separated from utilisation, different patterns of evidence use became apparent across the cases.

NASMLA provided the clearest documentary evidence of procedural use. Grade 3 NASMLA indicators entered the Kenya Primary Education Equity in Learning (KPEEL) programme's results framework and subsequent progress reporting (World Bank, 2022, 2025). The programme appraisal document used assessment performance to identify low-performing counties for targeting, while the School-Based Teacher Support manual incorporated NASMLA diagnosis into the rationale for professional support (Ministry of Education, 2020; World Bank, 2022). The programme records provide a traceable observation of procedural incorporation, while targeting and support design provide selected instrumental connections. None establishes that NASMLA independently caused programme implementation or changes in learner achievement.

When interpreting the primary documents, government and programme records establish that assessment evidence entered particular institutional processes. Describing these pathways as procedural or instrumental use is an analytical classification informed by research utilisation theory, not terminology used by the documents themselves.

Uwezo's strongest documented contribution lay in conceptual and agenda-setting use. Its repeated demonstration that school participation did not necessarily correspond with mastery of basic competencies helped shift attention from access alone towards learning. The Uwezo-linked informant described this reframing of education quality and the role of household feedback, accessible messages and media engagement (KII-003). The assessment reports document the underlying learning diagnosis (Uwezo, 2012, 2016), while formal recognition of qualifying datasets provides a separate, bounded

example of procedural incorporation (KNBS, 2022).

This form of influence is readily accommodated within Weiss's (1979) account of conceptual utilisation: evidence can change the terms in which a policy problem is understood without producing a single attributable decision. Monk's (2020) analysis of Uwezo in the Kenyan press provides a useful qualification. Uwezo achieved considerable media visibility, yet press discourse engaged only weakly with its social-accountability and citizen-action orientation. Public visibility and agenda influence were therefore more readily demonstrated than downstream citizen response.

FLANA's strongest documented contribution was conceptual/diagnostic and agenda-setting. Its assessment architecture combined children's foundational competencies with information on household, school and system conditions (Usawa Agenda, 2024, 2026). Usawa Agenda's (2025) staffing policy brief reused the 2023 FLANA database to diagnose staffing inequities and frame proposals for policy attention. This is a traceable example of diagnostic reuse by the evidence-producing organisation; it is not independent evidence that a government or school recipient acted on the proposals.

The first FLANA report, based on the 2023 assessment, was published in 2024, and the second report was released in 2026 (Usawa Agenda, 2024, 2026). The shorter public reporting history limits comparison with Uwezo's longer pathway. The available record documents growing translation and stakeholder engagement but does not establish an equivalent history of downstream instrumental use. Such pathways are less well documented, not necessarily absent.

Across the three cases, four forms of utilisation proved analytically useful: conceptual/diagnostic, agenda-setting, procedural and instrumental. Their value lies less in classification for its own sake than in the different evidentiary thresholds they impose. A dissemination event can establish that evidence was mobilised; it cannot establish that a decision changed. A policy citation may indicate recognition or conceptual influence; it does not by itself establish implementation. Procedural incorporation can demonstrate institutional use without establishing that the resulting process altered classroom practice.

Table 1. Illustrative documentary pathways and limits of inference

Case and form of use	Documented observation and supporting evidence	Boundary of inference
NASMLA: procedural and selected instrumental	Grade 3 indicators entered KPEEL results and reporting; assessment evidence informed programme targeting and teacher-support rationale (Ministry of Education, 2020; World Bank, 2022, 2025).	Documents establish incorporation into specific processes or decisions, not independent causal effects on implementation or learning.
Uwezo: conceptual and agenda-setting	Learning-assessment reports and KII-003 describe the focus on learning rather than enrolment; Monk (2020) independently documents visibility in the Kenyan press (Uwezo, 2012, 2016).	Media visibility and interview accounts do not independently establish adoption of a specific policy or citizen response.
Uwezo/Usawa Agenda datasets: bounded procedural recognition	KNBS (2022) describes validation of specified citizen-generated datasets under statistical quality-assurance arrangements.	Recognition is limited to the datasets concerned and must not be attributed to FLANA as an assessment system.

FLANA: diagnostic reuse and agenda-setting	Usawa Agenda (2025) reused the 2023 FLANA database in its staffing inequity policy brief (Usawa Agenda, 2024).	Evidence is producer-authored reuse, not independent confirmation that a recipient institution changed staffing decisions.

This interpretation develops rather than displaces earlier Kenyan research. Wamahiu et al. (2022) had already established uneven utilisation of learning-assessment evidence and identified difficulties involving dissemination, coordination, stakeholder engagement and feedback. The comparison indicates that unevenness concerns the form as well as the extent of use. Differently positioned assessment systems can produce different, defensible forms of utilisation. Recognising those forms avoids both understating conceptual influence and overstating mobilisation as instrumental action.

4.3 From Evidence Utilisation to Institutional Response

The clearest cross-case difficulty appeared after evidence had been mobilised or used. Documentary pathways were generally stronger for evidence production, communication and selected forms of utilisation than for the subsequent sequence of institutional response, implementation and feedback.

NASMLA provides a traceable illustration of the distinction between utilisation and response ownership. Documentary sources establish that indicators entered KPEEL programme planning and reporting, and that assessment diagnosis informed the rationale for teacher support (Ministry of Education, 2020; World Bank, 2022, 2025). The public evidentiary trail is less complete when following those uses through county and subcounty structures to school-level interpretation and sustained classroom practice. Paraphrased accounts from an assessment-system informant and a teacher-education informant raised concerns about follow-through, professional support and feedback (KII-004; KII-005). The case therefore shows a documented institutional handover at which responsibilities and implementation evidence become harder to trace. It does not establish that responsible actors failed to act or that implementation did not occur.

Uwezo encountered a similar problem through a different route. Household feedback, accessible communication and public engagement made learning problems visible to citizens and policymakers, but visibility did not necessarily produce sustained action. Lieberman et al. (2014) provide external corroboration for this narrower proposition. Their experimental study found no discernible effect of providing learning information and participation materials on the private or collective actions examined. The study is not evidence of Uwezo's overall influence; its relevance lies in showing that information provision and action are separated by questions of salience, perceived responsibility, feasible action and expected institutional responsiveness.

For FLANA, the issue is partly a consequence of diagnostic breadth. Evidence spanning learning, staffing, disability, leadership, WASH, safety and technology implicates different actors and administrative levels (Usawa Agenda, 2024, 2026). A richer diagnosis can therefore make the response problem more, rather than less, institutionally dispersed.

Paraphrased interview accounts offered complementary perspectives on this difficulty: the evidence intermediary emphasised synthesis and routing, the government informant the usability of evidence in policy processes, the teacher-education informant professional support, and the assessment-system informant follow-through and feedback (KII-001; KII-002; KII-004; KII-005). Read alongside the documents rather than as independent proof of institutional action, these accounts suggest that responsibility for follow-through may remain unclear even where evidence circulates successfully.

Response ownership emerged as a critical institutional link. It was not specified in the initial conceptual framework; it emerged from the cross-case analysis.

Assessment organisations can identify and communicate a problem without holding the mandate, resources or operational responsibility required to address it. The practical question is therefore not exhausted by asking whether evidence reached a policymaker or was incorporated into an institutional process. It also requires asking who is expected to respond once the evidence identifies a problem, and whether that actor has the authority and means to do so.

The concern with ownership is not itself novel. Participants in the 2025 ESSA–REAL Centre workshop had already argued that researchers should specify who should own or follow up their findings (Okullo et al., 2025). What the cross-case analysis adds is an explanation of where ownership sits in the evidence pathway. The problem can arise after successful mobilisation and even after demonstrable utilisation, because the institution producing or using evidence may not be the institution responsible for implementing a response.

This distinction extends the organisational implications of research utilisation theory. Weiss (1979) explains why research influence cannot be reduced to direct application; subsequent scholarship shows how organisational and system conditions shape evidence use. The present analysis indicates that these conditions remain consequential after utilisation has occurred. Evidence may inform understanding, agendas or institutional procedures while responsibility, authority, resources and implementation support remain located elsewhere.

The initial conceptual pathway can therefore be refined as follows:

Analytical interfaces in the refined heuristic: credibility and legitimacy | prioritisation and translation | mobilisation and brokerage | decision opportunity | response ownership | implementation support | feedback (with possible feedback across interfaces).

The heuristic identifies interfaces rather than a mandatory sequence. It is derived from the bounded comparison, not a tested causal model. Its purpose is to direct attention to the institutional questions that disappear when evidence use is inferred from dissemination, policy citation or programme association alone.

Evidence-informed educational leadership depends not only on access to credible evidence but also on institutional arrangements that connect evidence to decisions, assign responsibility for response, provide the means to implement that response and allow implementation experience to inform subsequent decisions.

4.4 Implications for Educational Leadership and Evidence Governance

The comparison suggests that a plural assessment system need not be consolidated into a single evidence model to become more useful. The more immediate governance challenge is to connect complementary forms of evidence without erasing the methodological and institutional differences that give them value. Formal state assessment, citizen-generated evidence and broader diagnostic assessment can illuminate different aspects of the same learning problem.

Translation also needs to be organised around decisions rather than dissemination products alone. For a priority finding, the relevant questions include who needs the evidence, what decision it can inform, when that decision is made and whether the intended user has the authority and resources to respond. Reports, policy briefs and convenings remain important, but their value lies partly in the decisions and institutional relationships to which they are connected. The analysis therefore supports linking major findings to an identifiable user, decision, timing and practical response.

Response ownership becomes particularly important where findings cross institutional mandates. Clarifying ownership does not mean transferring implementation responsibility to assessment

organisations. It means identifying the institution or administrative level able to lead a response and making coordination, sequencing and implementation support visible.

A final requirement is feedback. Evidence-use arrangements need to document what follows a decision: what response was adopted, whether implementation occurred, what constraints arose and what subsequent evidence suggests. Without this feedback, assessment systems can become increasingly sophisticated at diagnosing recurring problems while remaining less able to learn from the responses those diagnoses prompted.

The policy and leadership challenge is therefore not to choose between state and non-state assessment evidence, but to develop a better-connected evidence system in which different forms of evidence can be related to identifiable decisions, responsibilities and implementation processes.

5.0 SUMMARY, CONCLUSION AND RECOMMENDATIONS

5.1 Summary

The three cases illustrate different routes through which assessment evidence reached potential users: formal government and programme structures, household and public networks, and broader diagnostic engagement. Across these routes, documentary evidence of selected forms of utilisation was more readily traceable than evidence of subsequent implementation and feedback.

5.2 Conclusion

The cases are complementary rather than interchangeable. NASMLA's formal position makes planning and programme incorporation more traceable; Uwezo offers documented conceptual and agenda-setting contributions; and FLANA widens the diagnostic questions available to the education evidence ecosystem. These differences are institutional, not a ranking of learner outcomes or methodological quality.

Evidence mobilisation, utilisation and response must remain distinct. Evidence can reshape a problem or enter a procedure without proving implementation. Conversely, limiting 'use' to implemented decisions would overlook conceptual and agenda-setting influence. In each case, the inference must be proportionate to its source.

The comparison identifies response ownership as a useful institutional question after utilisation: which actor is responsible and authorised to respond, what support is available and how will action be reviewed? The evidence does not demonstrate a universal failure of implementation; it reveals uneven visibility of the steps connecting documented use to response and feedback.

Improving evidence use therefore requires not only credible assessment and effective translation but clearer connections to decision opportunities, responsible institutions, implementation support and documented learning from responses.

5.3 Recommendations

1. Coordinate complementary evidence synthesis. The Ministry of Education, working with KNEC and non-state assessment producers such as Usawa Agenda, could convene periodic synthesis of priority NASMLA, Uwezo and FLANA findings. Each synthesis should state the population, constructs and limitations of the underlying assessments and identify decisions that the combined diagnosis may inform, without comparing unlike scores.
2. Link findings to planning and budget decisions. Ministry of Education planning and programme units, together with relevant county and subcounty education actors, could maintain a concise finding-to-

decision record showing the finding, intended user, decision window, proposed response and responsible office. This is a proposed tool for strengthening traceability, not an arrangement evaluated in this study.

3. Assign responsibility across institutional mandates. Where a finding requires action by more than one agency or administrative level, the relevant education-sector authority could identify a lead institution, collaborating actors and the support required for implementation. Assessment producers may assist with interpretation and communication, but should not be presumed to possess implementation authority.

4. Establish follow-up and feedback. Programme managers, county and subcounty education officers, and school leadership could review selected decisions at agreed planning or implementation intervals, recording actions taken, constraints and subsequent evidence. Such records would support accountability and organisational learning; their feasibility and effects require prospective evaluation.

5.4 Areas for Further Research

Prospective research could follow specific assessment findings from production through mobilisation, utilisation, institutional response and implementation. Such designs would make it possible to examine decisions and downstream processes that are difficult to reconstruct retrospectively.

Further work is also needed on the movement of national assessment evidence across county, subcounty, school and classroom levels, particularly the conditions under which school leaders and teachers can interpret and respond to assessment findings.

Comparative research in other education systems could examine the applicability of the response-ownership heuristic and the circumstances under which responsibility, authority and implementation support strengthen or interrupt the pathway between evidence use and institutional response.

Funding: This research was supported by the ESSA–REAL Centre Foundational Learning Research Grant (FLAFRICA001), with funding from the Gates Foundation.

References

1. Fussy, D. S. (2022). Use of research evidence in the policy domain: A best-evidence synthesis of social science research in Africa, Asia, and Latin America. International Social Science Journal, 72(246), 1181–1194. https://doi.org/10.1111/issj.12357
2. Kenya National Bureau of Statistics. (2022, October 11). Use of alternative data sources for official reporting: Uwezo Twaweza & Usawa Agenda [Citizen-generated data validation brief]. https://www.knbs.or.ke/wp-content/uploads/2023/09/Citizen-Generated-Data-Uwezo-Twaweza-Brief-on-Uwezo-Twaweza-and-Usawa-Agenda.pdf Kenya National Examinations Council. (n.d.). National Assessment System for Monitoring Learner Achievement (NASMLA). https://www.knec.ac.ke/nasmla/
3. Lieberman, E. S., Posner, D. N., & Tsai, L. L. (2014). Does information lead to more active citizenship? Evidence from an education intervention in rural Kenya. World Development, 60, 69–83. https://doi.org/10.1016/j.worlddev.2014.03.014
4. Ministry of Education. (2020). School-Based Teacher Support Initiative manual for early years education. https://www.education.go.ke/sites/default/files/2023-02/SCHOOL%20BASED%20TEACHER%20SUPPORT%20INITIATIVE%20MANUAL%20%28SBTS%29%20for%20Early%20Years%20Education%20-%2018-09-2020%20%281%29_compressed.pdf
5. Monk, J. (2020). Placing blame or a call to action? An analysis of Uwezo in the Kenyan press.

Education Policy Analysis Archives, 28, Article 185. https://doi.org/10.14507/epaa.28.5268

6. Nutley, S. M., Walter, I., & Davies, H. T. O. (2007). Using evidence: How research can inform public services. Policy Press. https://doi.org/10.1332/policypress/9781861346650.001.0001
7. Okullo, D., Omanyo, D., & Sarpong, M. (2025). Workshop report on promoting locally led and policy-relevant foundational learning research priorities in Kenya. Education Sub Saharan Africa & REAL Centre, University of Cambridge. https://www.essa-africa.org/index.php/node/2034
8. Rechsteiner, B., Kyndt, E., Compagnoni, M., Wullschleger, A., & Maag Merki, K. (2024). Bridging gaps: A systematic literature review of brokerage in educational change. Journal of Educational Change, 25(2), 305–339. https://doi.org/10.1007/s10833-023-09493-7
9. Rickinson, M., Cirkony, C., Walsh, L., Gleeson, J., Cutler, B., & Salisbury, M. (2022). A framework for understanding the quality of evidence use in education. Educational Research, 64(2), 133–158. https://doi.org/10.1080/00131881.2022.2054452 Usawa Agenda. (n.d.). Who we are. https://usawaagenda.org/who-we-are/
10. Usawa Agenda. (2024). Are ALL our children learning? 1st Foundational Literacy and Numeracy Assessment (FLANA) report: 2023. https://usawaagenda.org/wp-content/uploads/2024/05/UA_Are-ALL-Our-Children-Learning_Report-03-2023-1.pdf
11. Usawa Agenda. (2025). Equal class sizes for equal learning: Addressing staffing inequities in Kenya's primary schools (Policy Brief UAPB-001-2025). https://usawaagenda.org/wp-content/uploads/2025/09/UA-Policy-Brief_Unequal-Classrooms-Online.pdf
12. Usawa Agenda. (2026). Are all our children learning? 2nd Foundational Literacy and Numeracy Assessment report. https://usawaagenda.org/fw-portfolio/the-flana-report-2026/
13. Uwezo. (2012). Are our children learning? Uwezo Kenya 2012. Uwezo Kenya. https://twaweza.org/download/uwezo-kenya-2012-are-our-children-learning/
14. Uwezo. (2016). Are our children learning? Uwezo Kenya sixth learning assessment report. Twaweza East Africa. https://twaweza.org/download/uwezo-kenya-2016-are-our-children-learning/
15. Wamahiu, S. P., Nyamu, I., Gatwiri, W., Aminga, N., & Owino, S. O. (2022). Formative study on the utilisation of past and present learning assessments in Kenya: Research report. Global e-Schools and Communities Initiative. https://www.gpekix.org/sites/default/files/Media%20Document/ADAPT_Formative_Study_Report_Kenya.pdf
16. Weiss, C. H. (1979). The many meanings of research utilization. Public Administration Review, 39(5), 426–431. https://doi.org/10.2307/3109916
17. World Bank. (2022). Primary Education Equity in Learning Program: Program appraisal document (Report No. PAD4913; Project P176867). https://documents1.worldbank.org/curated/en/861391649085182961/pdf/IBArchive-a3efea83-58ad-4ddf-8bec-57dcadea0526.pdf
18. World Bank. (2025, December 14). Primary Education Equity in Learning Program (P176867): Implementation status & results report (ISR06078). https://documents1.worldbank.org/curated/en/099121425123534877/pdf/P176867-9d4d6b71-d8d9-481f-91f6-307cc35abad5.pdf