



Bringing Learning to Light: The Role of Citizen-led Assessments in Shifting the Education Agenda

Cover photos: Children being assessed as part of citizen-led assessments. Clockwise from top left: Beekunko in Mali, Jàngandoo in Senegal, ASER in India, Uwezo in Tanzania.

Results for Development Institute (R4D) is a nonprofit organization whose mission is to unlock solutions to tough development challenges that prevent people in low- and middle-income countries from realizing their full potential. Using multiple approaches in multiple sectors, including global education, global health, governance, and market dynamics, R4D supports the discovery and implementation of new ideas for reducing poverty and improving lives around the world.

This report was authored by a team at R4D led by Molly Jamieson Eberhardt under the general direction of Nicholas Burnett and including Michelle Engmann, Tara Hill, and Daniel Plaut. For more information, please contact Molly Jamieson Eberhardt at meberhardt@r4d.org.

Copyright © 2015
Results for Development Institute
1100 15th Street, N.W., Suite #400, Washington, DC 20005

Table of Contents

Acknowledgements	ii
Executive Summary	1
Introduction	8
Chapter 1. What is a Citizen-led Assessment?	11
Theory of change	12
Who is involved?	13
What activities take place?	16
How much do citizen-led assessments cost?	22
Chapter 2. How Well do Citizen-led Assessments Work?	23
Evaluation methodology	23
How well do citizen-led assessments measure learning?	26
How well do citizen-led assessment processes work?	33
How well do citizen-led assessments stimulate awareness and action?	39
Chapter 3. Recommendations	52
Linked Annexes: Portfolio of Related Research	58
Annexes	59
References	65
Figures and Tables	
Figure 1. Citizen-led Assessments Currently Underway	8
Figure 2. Roadmap of this Citizen-led Assessment Report	9
Figure 3. Citizen-led Assessment Theory of Change	12
Figure 4. Basic Structure of Citizen-led Assessments	13
Figure 5. Examples of Reading Tasks from Testing Tools	17
Figure 6. Examples of Math Tasks from Testing Tools	18
Figure 7. Excerpt from Uwezo Pamphlet Left with Parents after the Survey	21
Figure 8. Story Task (English) — From Uwezo Reading Test	31
Figure 9. Action Spectrum for Citizen-led Assessments	42
Table 1. 2013 Citizen-led Assessment Coverage	11
Table 2. Reading Tasks Included in Citizen-led Assessments	17
Table 3. Math Tasks Included In Citizen-Led Assessments	18
Table 4. 2013 Citizen-Led Assessment Costs	22
Table 5. Resource Allocation by Activity as a Percentage of Total Budget	22
Table 6. Test Characteristics Related to Comparability	30
Table 7. Uwezo Participation in National Dialogue on Learning	45
Table 8. Examples of Actions Taken by States in India	47

Acknowledgments

The authors would like to thank the organizations featured in this report for the time and extensive inputs they contributed to the evaluation: ASER Centre in India, *Œuvre Malienne d'Aide à l'Enfance du Sahel* (OMAES) in Mali, *Laboratoire de Recherche sur les Transformations Économiques et Sociales* (LARTES) in Senegal, and Twaweza East Africa's offices in Kenya, Tanzania, and Uganda. The authors are grateful to these teams for the resources and information shared throughout the course of the evaluation and for welcoming the evaluation team to conduct site visits, which enriched the analysis in this report. Each organization's extended team provided invaluable support, with particular thanks due to: Rukmini Banerji, Suman Bhattacharjea, and Wilima Wadwha of ASER Centre; Massaman Sinaba and Bréhima Traoré of OMAES; Rokhaya Cissé and Abdou Salam Fall of LARTES; and Zaida Mgalla, John Mugo, Mary Goretti Nakabugo, Zipporah Ongwenyi, and Sara Ruto of Twaweza.

The evaluation team included in-country and technical partners whose analysis and findings were critical to this report. These partners include: Catalyst Management Services (CMS) in India, OWN & Associates Ltd. in East Africa, Mr. Abdoulaye Bagayogo in Mali, Mr. Souleymane Barry in Senegal, and the Australian Council for Educational Research (ACER). Particular thanks are due to Seyed Fayez and Vijay Kulkarni at CMS, Dr. Okwach Abagi at OWN & Associates Ltd., and Charlotte Waters at ACER, for their leadership in the data collection and analytical processes.

Qdata Enterprises served as the local implementation partner for the quasi-experimental studies in Kenya, which were also supported by the Uwezo regional team in Nairobi, Uwezo District Coordinators, and RTI International team members who generously contributed time and resources to the studies.

The evaluation team would like to thank The William and Flora Hewlett Foundation for supporting this evaluation, in particular to Ruth Levine, Patricia Scheid, and Dana Schmidt for their ongoing engagement. The report also benefited from peer review provided by an Advisory Committee assembled by the Foundation. Thanks are due to committee members Amy Arbretton, Penelope Bender, Peter Colenso, Christopher Elbers, Lucy Heady, Robin Horn, Elizabeth King, Ruth Levine, Kate Anderson Simons, and Waly Wane for their valuable input.

Results for Development Institute (R4D) conducted the evaluation and authored this report. The evaluation team was led by Molly Jamieson Eberhardt under the general direction of Nicholas Burnett and included Michelle Engmann, Tara Hill, and Daniel Plaut. Nathaniel Heller and Courtney Tolmie, also of R4D, provided additional guidance and review.

Great thanks are also due to Anne Girardeau, who translated this and other evaluation documents into French.

Executive Summary

“Are children learning?” is a question that should inform all education policy-making. Yet in many countries, the answer to this question has remained largely unknown. The pursuit of an answer lies at the heart of the citizen-led assessment movement.

As governments and donors focused on increasing access in the wake of the Millennium Development Goals, the issue of learning received comparatively little concerted attention. Some organizations working in countries where access was rapidly increasing took notice of the fact that, while rising enrollment rates were being celebrated, there was little evidence of whether or not learning was taking place. One of the results of this realization was the emergence of the citizen-led assessment movement, initiated by Pratham in India in 2005. The movement is an attempt by civil-society organizations to gather evidence on learning and use it for two main purposes: first, to increase awareness of low learning outcomes and second, to stimulate actions that are intended to address the learning gap.

Today, ten years after the first citizen-led assessment was conducted, it is widely anticipated that the Sustainable Development Goals that will be announced in September

2015 will include a goal that addresses learning. The inclusion of such a goal raises the challenge of measuring learning outcomes in a manner that is both country-relevant and globally-applicable. Debates over how to do so have brought attention to various models of national learning assessments, including citizen-led assessments, which are now being undertaken in nine countries.

This innovative approach to assessment has attracted interest and raised questions about the potential for non-traditional assessments to play a role in not only monitoring learning but also advocating for more focus on educational outcomes within countries and at the international level. In an effort to more deeply understand the nuts and bolts of the citizen-led assessment model and to evaluate its ability to measure learning, disseminate findings, and stimulate awareness and action, Results for Development Institute (R4D) evaluated four citizen-led assessments between May 2013 and November 2014: The Annual Status of Education Report (ASER) in India, Beekunko in Mali, Jàngandoo in Senegal, and Uwezo in Kenya, Tanzania, and Uganda.

Methodology

The evaluation aimed to answer three key questions, each of which was addressed by a separate but complementary component of the evaluation methodology:

How well do citizen-led assessments measure learning?	Technical review of the testing tools, sampling design, and analytical processes used
How well do citizen-led assessment processes work?	Process evaluation
How well do citizen-led assessments stimulate awareness and action about learning outcomes?	Non-experimental evaluation of impact ¹



Uwezo: Child in Uganda is assessed in the presence of his parent.

¹ The evaluation of impact was only conducted for ASER and Uwezo, as Beekunko and Jàngandoo have only been conducting assessments for three and two years respectively, and such an exercise would be premature given that they are still refining their models. R4D developed an evaluation framework for each that can be used to inform the design of a future evaluation of impact.

Technical review: How well do citizen-led assessments measure learning?

The technical review was led by the Australian Council for Educational Research (ACER). It included a desk review for each of the four citizen-led assessments that examined the sample design, the test development process, the testing tools themselves, test administration methods, and data analysis, including the way in which assessment results are reported. It also included two quasi-experimental studies of the Uwezo Kenya testing tools: a concurrent validity study and an inter-rater reliability study (ACER 2015). Findings from earlier studies of the validity and inter-rater reliability of ASER testing tools (Vagh 2013) given the importance of the country context to the potential for impact were also incorporated.

Process evaluation: How well do citizen-led assessment processes work?

The process evaluation was carried out by R4D in partnership with in-country partners: OWN & Associates in East Africa, Catalyst Management Services in India, Mr. Abdoulaye Bagayogo in Mali, and Mr. Souleymane Barry in Senegal. It assessed the design and implementation of activities conducted by ASER, Beekunko, Jàngandoo, and Uwezo. The process evaluation addressed six key areas: the suitability and effectiveness of key players in the survey process, training of trainers and volunteer surveyors, execution of household sampling, survey administration, quality assurance, and dissemination and engagement. The process evaluation entailed direct observation in a small sample of sites in each country, as well as interviews and desk reviews.²

Evaluation of impact: How well do citizen-led assessments stimulate awareness and action?

The evaluation of impact was only conducted for ASER and Uwezo, as Beekunko and Jàngandoo have only been conducting assessments for three and two years respectively, and such an exercise would be premature. The evaluation of impact used a non-experimental approach, as identifying a valid control group for a national-scale initiative is not possible. It assessed the extent to which the ASER and Uwezo initiatives have increased awareness of learning outcomes and influenced actions to address poor learning achievement. The evaluation methodology was designed to gather data at the national, state, district, and community levels. Primary data was collected through key-informant interviews and focus group discussions with key stakeholders at these levels. Secondary research included a review of ASER and Uwezo data, media coverage, national and state education policies to track references to citizen-led assessments, use of testing tools by other organizations, and findings of earlier studies of Uwezo's impact.³

The evaluation of impact included a sample of four states in India and three districts in each of Kenya, Tanzania, and Uganda. The selection of evaluation sites was carried out through purposive sampling, seeking to capture diversity in perspective from a broad spectrum of stakeholders that may have been impacted by the ASER and Uwezo initiatives.

The full evaluation report describes the citizen-led assessment model in detail. This summary briefly states the findings of the evaluation, framed around the key questions it aimed to answer.

² A detailed list of observation sites is provided in Annex 2.

³ Lieberman et al. (2012, 2014, 2014b) and Uwezo's own 2013 monitoring reports

HOW WELL DO CITIZEN-LED ASSESSMENTS MEASURE LEARNING?

What they measure, they measure well. Citizen-led assessments test a very limited set of competencies in reading and mathematics. The testing tools they use yield valid results, which cast a spotlight on limited achievement in these basic competencies. In this way they are important and useful, but broadening the testing tools to assess a broader range of skills would allow citizen-led assessments to do more to inform policy and practice.

Volunteers can reliably assess children's basic competencies. Inter-rater reliability studies of ASER and Uwezo indicate a high level of agreement in volunteers' scoring of children's responses.

Sub-national results should include information about context. All citizen-led assessments report results below the national level (e.g., district, region, *commune*); this invites comparisons that may be misleading when there are differences in context across the different districts, regions, or *communes*, such as socioeconomic status, which is known to have a relationship with children's performance on learning assessments.

Comparability of findings across years needs to be carefully investigated. Formal equivalence exercises need to be undertaken to ensure the comparability of tests, especially when findings are compared over multiple years and test forms.

HOW WELL DO CITIZEN-LED ASSESSMENT PROCESSES WORK?

The sheer scale of the surveys that are orchestrated each year is commendable, especially considering that such a feat requires mobilizing hundreds of organizations and thousands of volunteers for the cause of measuring learning.

Volunteers are key to the citizen led assessment model. The evaluation found that volunteers carried out the tasks assigned to them effectively, with the possible exception of household sampling which needs further testing.

Partner organizations are valuable assets, but good ones are hard to find and retain. Systematic partnership models like partnering with District Institutes of Education and Training (DIETs) in India, Centres d'Animation Pédagogique in Mali, or a single strong NGO across many districts are promising in that they can be scaled up through existing structures.

Training related to test administration was very strong, but in most cases should include more field practice that mimics volunteers' responsibilities during the actual survey. Trainers and volunteers were well-equipped to conduct the community survey, household survey, and administer the tests to children. With the exception of Jàngandoo, whose training lasts five days, training time was insufficient according to trainers, trainees, and the evaluation team. There are of course significant resource implications to increasing the length of training; in all cases citizen-led assessments indicated that training time is one of the key areas affected by resources constraints.

It is important for the quality and credibility of the survey that each initiative develop a systematic approach to quality assurance. Good examples of this are ASER's and Uwezo's quality assurance mechanisms which involve extensive monitoring and re-check activities, including field re-checks for a subset of villages and districts.

Citizen-led assessments have been very successful in using the media to disseminate survey findings, particularly at the national level.

Including key stakeholders at the national level as advisors in the survey design process increases institutional buy-in. Beekunko and Jàngandoo are currently experimenting with pre-survey engagement at lower levels as well (regional, *commune*, and community). This may be a powerful strategy for building support among key constituents (including potential critics); further testing is needed to fully understand the impact of such engagement.

HOW WELL DO CITIZEN-LED ASSESSMENTS STIMULATE AWARENESS AND ACTION?

These findings only apply to the evaluations of ASER and Uwezo. Beekunko and Jàngandoo were not included in the evaluation of impact.

Increased awareness of learning outcomes is one of the key pillars of citizen-led assessments' theory of change: by providing key education stakeholders with quality information regarding the state of learning within their country, district, and community, these initiatives hope to incite broad-ranging action targeted at improving literacy and numeracy skills. **Awareness of learning outcomes is a necessary prerequisite to action. Action, then, is required for the concern generated by increased awareness to translate through various mechanisms into improvements in learning outcomes.** The evaluation team has deliberately adopted a broad definition of action as any shift in behavior, policy, or practice by an education stakeholder related to the improvement of learning outcomes.

This summary describes the impact that ASER and Uwezo have had on awareness and action at the international, national, state, district, and community levels.

International

Both ASER and, more recently, Uwezo, have contributed to an increased focus on learning outcomes in global discourse and agenda-setting. Their contribution has included both providing evidence of the seriousness of the learning crisis (i.e., revealing major deficiencies in even the most basic competencies) and demonstrating how a low-resource model can be used to assess learning on a national scale. ASER and Uwezo did not single-handedly cause this shift toward learning outcomes, but they did contribute to it by raising the visibility of the crisis by quantifying it in very simple, stark terms.

National assessments have for many years revealed low learning levels in many countries, but very few assess children in the early grades or use a population-based sample to ensure that out-of-school and rarely-in-school children are included. The lack of even the most basic reading and math skills that citizen-led assessments reveal is in some ways more powerful in stimulating debate than similarly dismal results of more formal assessments of older children because it reflects such widespread failure of the system to deliver even the most basic education. **The evaluation finds that ASER and Uwezo have been quite successful at generating awareness of low learning levels at the international level.**

National

Increasing awareness of the learning crisis at the national level is one of the main successes of both ASER and Uwezo. Success in generating awareness at the national level has been achieved largely by a media-centric approach to dissemination of results. Each year's survey results are prominently featured in multiple languages in print media and, to a lesser extent, in radio and television broadcasts as well.

ASER and Uwezo have also stimulated debate at the national level. While national-level stakeholders reported to have already been generally aware of low learning levels in their countries, they indicated that ASER and Uwezo have made low learning levels highly visible in the public sphere.

There is some evidence that ASER and Uwezo have played a contributory role, among other factors, in the prioritization of learning in national education policy documents. For example, India's 2012-2017 national planning document and 2014 resource allocations indicate an increased focus on measuring and improving learning. Similarly, Uwezo results have been cited in government reports and strategy documents, noted as supporting evidence for renewed government focus on learning outcomes.

Generating concrete action to improve learning outcomes has proven challenging for both initiatives. Sporadic but powerful examples of direct action do exist (e.g., programs focused on measuring learning at the state level in India), and they can provide insight into the types of action that are possible and what factors support such action.

State

The level of impact that ASER has had at the state level varies by state, but **in the four states included in this evaluation, there was evidence that ASER contributed to stimulating debate, engagement, and in some cases direct action for the improvement of learning outcomes.** The potential for ASER's impact at the state level was found to be greater than at the national level, largely because education implementation and practice are state-level responsibilities. Examples of state-level action to which ASER may have contributed include the development of state-wide learning assessments and the use of ASER survey results by officials to identify weak areas (both geographic and issue-based) around which state programs can be designed.

District

Both ASER and Uwezo aim to influence two key groups at the district level: their district partner organizations and district-level government officials. **For both initiatives, only sporadic evidence of impact at the district level was found. This can largely be attributed to the lack of resources available for systematic involvement of the network of district partner organizations in dissemination activities, and, relatedly, to the limited capacity of these organizations.**

Outside of the government, ASER and Uwezo have triggered the uptake of the testing tools in education programs run by NGOs and CSOs. The evaluation identified dozens of groups that use the testing tools in various ways including initial assessment of children's learning levels, tracking children's progress over time, and monitoring the impact of their work as an organization. As these tools are easy to use and adapt, they can be used by a wide range of organizations with varying capacities.

Community

The evaluation uncovered only limited anecdotal evidence that participation in the survey stimulates awareness or action at the community level. Much more testing is needed of ways to engage at the community level before, during, and after the survey if citizen-led assessments aim to close the feedback loop between collecting data and inciting action.

Key lessons from sporadic examples of success

From the set of success stories where action was generated, two key lessons emerge that shed light on what is needed, in addition to the provision of information, to generate action:

- Those few who are motivated to act by information alone, or who have already demonstrated commitment to the issue of learning outcomes, are critical partners. While limited in number, these champions can be very powerful.
- In order for action to take place, there needed to be an entity available to provide suggestions for what those actions might look like. In India, Pratham often played this role by collaborating with government officials to design and implement interventions. Other non-governmental organizations (NGOs), CSOs, and agencies also contributed to augment the capacity of a champion that wanted to act, but did not know what the first step might look like. In Kenya, Tanzania, and Uganda, the lack of an obvious partner or set of partners to play this role limited Uwezo's ability to generate action.

These lessons align with the recently emerging school of thought in the transparency and accountability space that strategic engagement with government, and particularly with champions within the government, as opposed to (or in combination with) a whistle-blowing approach, may be most effective at achieving common goals of governments and civil society (Fox 2014). But these lessons do not just apply to influence of government officials; they also apply to district partners, CSOs, NGOs, community leaders, and community members. Champions exist within these groups, and if linked with each other and provided with suggestions and support for actions they can lead, they can be powerful forces in bringing about broader action. If more systematically encouraged, the impact of currently small-scale actions (such as the chain of libraries started by a district partner in Bihar, or the incorporation of ASER's and Uwezo's testing tools into NGOs' education programs) could be exponentially increased.

Contextual considerations for the success of citizen-led assessments

The evaluation sheds light on the contexts in which citizen-led assessments are likely to have the most traction. Broadly defined, these contexts are ones in which target audiences have autonomy to make decisions about policy and practice. This manifests in different ways at different levels. For example, in India, where state officials have significant control over policy and district officials over planning, citizen-led assessments are likely to have most traction at those levels. In countries whose political

systems are highly centralized, the national level is a critical target audience. Targeting district officials in such a context is less likely to lead to systematic reform.

With regard to the community level, citizen-led assessments should, in theory, be able to gain traction in various contexts, as long as the type of information provided and the engagement surrounding that information is designed to educate, instruct, and empower citizens. In a decentralized context, engagement might be designed to generate collective action to put pressure on local policymakers and service providers. Even in a centralized context, individual or collective action at the community level could impact the performance of service providers. In practice, the amount of traction is dependent not just on contextual factors, but on how citizen-led assessments design their engagement activities to respond to those contextual factors appropriately.

Can assessments be expected to change learning outcomes?

One of the goals of this evaluation was to calibrate expectations for the type of impacts that citizen-led assessments can achieve. One of the ultimate intended outcomes of these initiatives is improved learning outcomes. ASER and Uwezo survey data clearly shows that this outcome has not been achieved: an analysis of the ASER survey data between the years 2006 and 2012 shows that the number of states showing a declining trend in learning outcomes is increasing while the number of states showing improvement is declining. Uwezo is also unlikely to achieve its goal of increased learning outcomes; they aim for a ten percent increase in the literacy and numeracy of 6-16 year olds in Kenya, Tanzania, and Uganda after four years of the Uwezo assessment — by 2015. Attainment of this goal seems unlikely: Uwezo's own findings indicate that learning outcomes have remained largely consistent across the three countries over the last three years of surveying.

It may not be reasonable to expect that learning outcomes would go up, or stop declining, as a result of actions motivated by citizen-led assessment results. Even flawless implementation of a sample-based survey and related engagement activities, even at the large scale that the ASER and Uwezo initiatives have achieved and should be commended for, leaves hundreds of thousands of communities (the places where education actually occurs) untouched. This is not a criticism of the citizen-led assessment model, but a check on the theory that such a model could cause a significant shift in national learning outcomes.

Priority Recommendations

Recommendations for existing and new citizen-led assessments

This summary includes a subset of recommendations that draw on the key evaluation findings. The complete set of recommendations can be found in the full report below.

- A) Consider broadening the range of skills that is tested by citizen-led assessments in order to better inform policy and practice.** While it is important to maintain the simplicity of the testing tools, foundational concepts that are not currently tested could be added without greatly increasing the complexity or length of the test. For example, the reading test could assess listening comprehension. This would allow for distinguishing between children who have inadequate decoding skills and children who have deficits in the ability to make meaning in the test language. It would also give the assessment more power to inform policy, resourcing, and classroom practice. All citizen-led assessments should also test foundational mathematical competencies such as place value, which Beekunko and Uwezo already do, and measurement and shape, which Jångandoo already does.
- B) Take steps to investigate equivalence of testing tools, in order to facilitate the comparison of results.** In order to enable the exploration of factors that lead to high performance, it is important for citizen-led assessments to support the making of comparisons. The first step in supporting the making of comparisons is to demonstrate that it is valid to make them in the first place, and this is achieved in part by ensuring that the assessment tools are at an acceptable level of linguistic and psychometric equivalence. Linguistic and psychometric equivalence can be investigated by drawing on expert judgment, by conducting equating studies, and by analysis of assessment data. Investigating and confirming equivalence across years is particularly important if citizen-led assessments aim to play a role in efforts to track learning outcomes over time.
- C) Consider reducing the frequency of citizen-led assessments, moving from an annual to a biannual or triennial assessment.** Given that significant changes in learning outcomes over the course of one year are not expected, the massive annual investment in data collection may not generate enough marginal impact to justify its cost. According to the majority of stakeholders interviewed for the evaluation of impact for both ASER and Uwezo (including education experts, policymakers, and CSOs), the impact of the annual survey has diminished over time, as anticipation of results has waned. Even for Beekunko and Jångandoo, which have not released data as many times, the frequency

could still be reduced. There are undoubtedly tradeoffs to implementing this recommendation. Surely, the level of awareness and debate about learning that ASER and Uwezo have generated at the national level has been partially due to the momentum its yearly assessments have generated. But there is no evidence that new data (especially if it is not notably different data) is required to launch a new awareness-building and action-generating campaign. **The important rhythm that the current annual survey provides could be maintained with annual campaigns or events, without the need for annual data collection.** Distinct changes in learning outcomes could be anticipated over the course of two or three years, giving stakeholders a reasonable timeline for pursuing goals aligned to the survey findings. Most importantly, a biannual or triennial survey would free-up resources to use on dissemination and engagement activities during the off-year(s)—activities that have so far been limited due in part to resource constraints.

- D) Continue to use volunteers as surveyors, but explore ways to extend their role beyond that of data collectors** in order to increase their potential to serve as champions within their communities. The survey “infrastructure,” which identifies partner organizations in every district/region and deploys volunteers to communities, is not currently leveraged. Citizen-led assessments could consider training partners and volunteers in survey follow-up engagement activities, and then use community/volunteer pairings to implement community engagement activities.
- E) Capitalize on the large networks of partner organizations that have been built. Engage partners beyond data collection,** and identify networks of organizations that could become partners in multiple locations. Systematic partnership models like partnering with DIETs in India, *Centres d'Animation Pédagogique* in Mali, or a single strong NGO across many districts are promising in that they can be scaled up through existing structures. Such engagement must be tailored to align with the capacity levels of partner organizations, or to build their capacity through strategic engagement.
- F) The network of citizen-led assessments could consider experimenting with systematic ways to develop awareness and action about low learning outcomes at the lower levels (district, community, etc.).** At the community level, impact evaluations of transparency and accountability interventions indicate that a common feature of successful ones is **facilitation of dialogue and action development between citizens and providers** (Bjorkman Nyqvist et al. 2014 and others). This is an example of the type of evidence-based engagement activity that could be tested.

- G) **In addition to aiming for general awareness, systematically work to identify champions** among each of the audiences that the citizen-led assessments wish to influence. Committed individuals have been a key part of many of the successes achieved to date in generating action.
- H) **Beekunko, Jàngandoo, and Uwezo could consider developing partnerships with providers of solutions.** This is intentionally distinct from providing solutions themselves, as it is important for citizen-led assessments to retain their neutrality. The ASER and Pratham combination has proven to be effective at enabling action.
- I) **Build on the observed uptake of citizen-led assessment testing tools by NGOs to share tools more systematically with NGOs, private schools, and teachers.**

Recommendations for supporters and researchers of citizen-led assessments

- J) **Support citizen-led assessments' dissemination and engagement activities, so they have enough resources left over after data collection to conduct meaningful engagement.** This could manifest as a continued level of support but for a biannual or triennial survey, with the off years' support going toward engagement activities as described above, especially at the district and community levels.
- K) **Support citizen-led assessments to experiment with techniques that transcend the boundaries of "dissemination" and move toward strategic engagement.**
- L) **Disseminate lessons from the above experimentation widely.** These will not only inform citizen-led assessments, but the transparency and accountability field more broadly, as well. **Coordinate such experimentation across the growing network of citizen-led assessments,** so a variety of approaches can be tested and lessons shared across the group. Include experimentation with similar approaches in different contexts to enable identification of contextual factors that may enable or hinder effectiveness.

Introduction: The Challenge of Measuring Learning

“Are children learning?” is a question that should inform all education policy-making. Yet in many countries, the answer to this question has remained largely unknown. The pursuit of an answer lies at the heart of the citizen-led assessment movement.

As governments and donors focused on increasing access to primary education in the wake of the Millennium Development Goals, the issue of learning received comparatively little concerted attention. Some organizations working in countries where access was rapidly increasing took notice of the fact that, while rising enrollment rates were being celebrated, there was little evidence of whether or not learning was taking place. One of the results of this realization was the emergence of the citizen-led assessment movement, initiated by Pratham in India in 2005. The movement is an attempt by civil-society organizations to

gather evidence on learning and use it to provoke greater attention to this critical gap.

Today, ten years after the first citizen-led assessment was conducted, it is widely anticipated that the Sustainable Development Goals that will be announced in September 2015 will include a goal that addresses learning. In the proposal of the Open Working Group of the General Assembly (United Nations 2014), one of the proposed goals is to “Ensure inclusive and equitable quality education and promote lifelong learning opportunities for all,” with a sub-goal of, “By 2030, ensure that all girls and boys complete free, equitable and quality primary and secondary education leading to relevant and effective learning outcomes.”

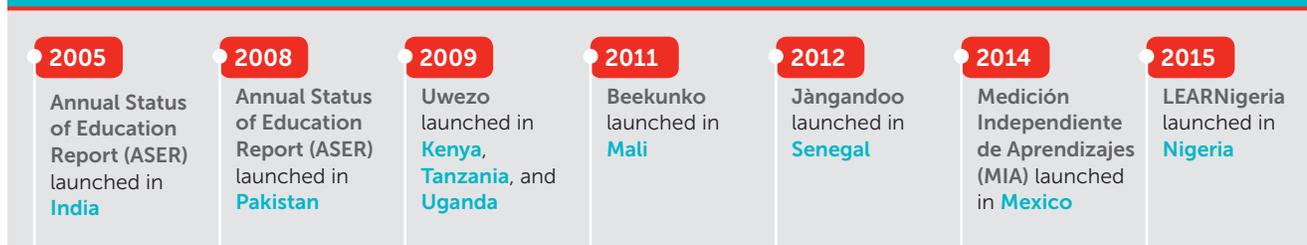
The challenge of measuring learning outcomes in a manner that is both country-relevant and globally-applicable has brought attention to various models of national and regional learning assessments, including citizen-led assessments, which are now being undertaken in nine countries.

The key characteristic that distinguishes citizen-led assessments from others is that they combine learning measurement approaches (from the education field) with citizen-monitoring approaches (from the transparency and accountability field) to engage ordinary citizens in the assessment of children’s learning. Local volunteers are trained to conduct a household survey during which a short reading and math test is administered orally and one-on-one to all school-aged children. This participatory method is designed to broaden the audience that usually consumes assessment data (policymakers, pedagogues, education authorities) to include a wider range of people—all of whom have a stake in the educational outcomes of the country’s children. The following citizen-led assessments are currently underway (see Figure 1).



ASER: Assessing a child at her home

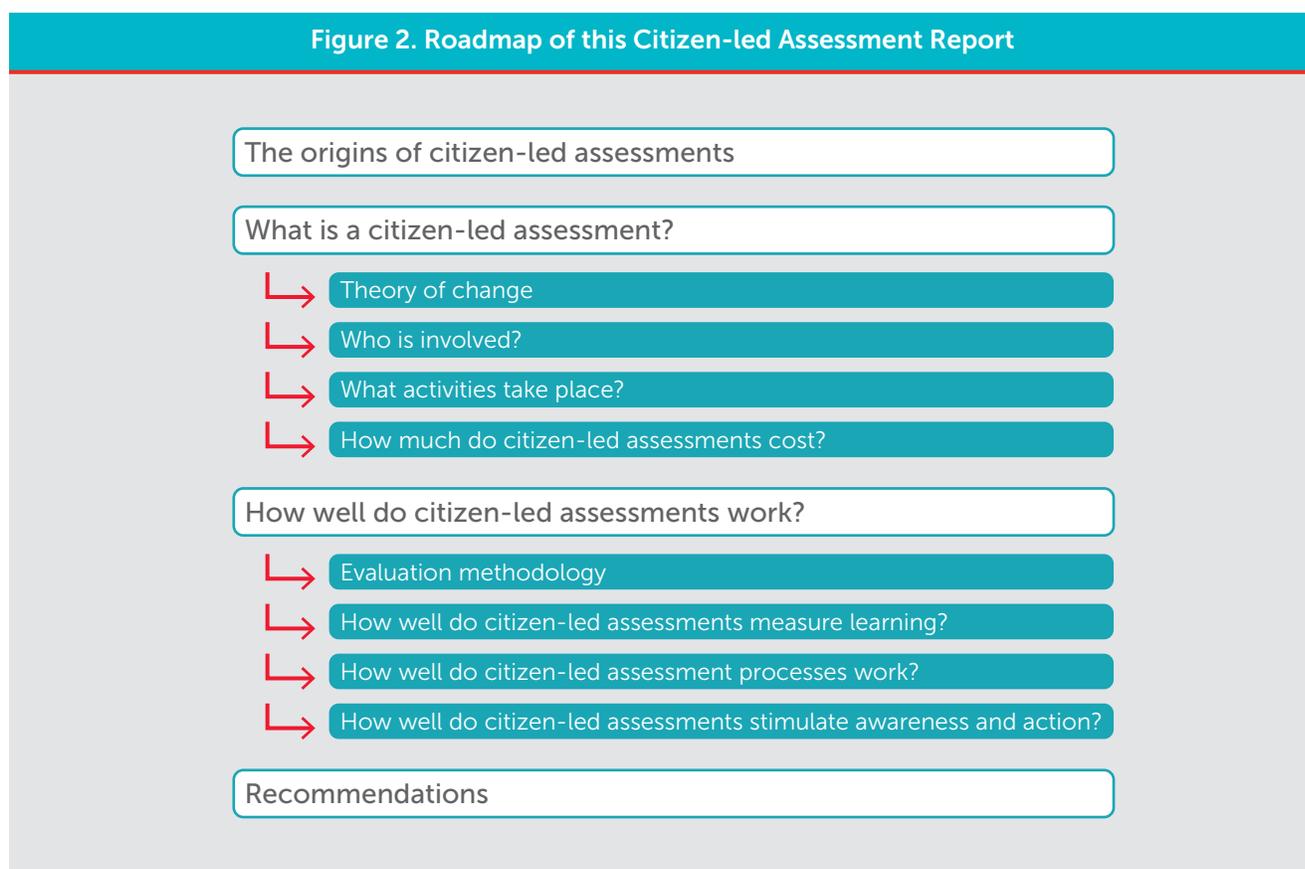
Figure 1. Citizen-led Assessments Currently Underway



This innovative approach to assessment has attracted interest and raised questions about the potential for non-traditional assessments to play a role in not only monitoring learning but also advocating for more focus on educational outcomes within countries and at the international level. In an effort to more deeply understand the nuts and bolts of the citizen-led assessment model and to evaluate its ability to measure learning, disseminate findings, and stimulate awareness and action, Results for Development Institute (R4D) evaluated four citizen-led assessments between May 2013 and November 2014: ASER India, Beekunko, Jàngandoo, and Uwezo.⁴

The evaluation assessed the more mature initiatives, ASER and Uwezo, with a focus not only on technical validity and execution of key processes, but also on the impact that these initiatives have had in their respective countries since inception. The evaluation of the younger citizen-led assessments, Beekunko and Jàngandoo, was a largely formative assessment focusing on the emerging capacity of these initiatives to execute key processes in their early years.

This report describes the citizen-led assessment model in detail, and presents the findings of the evaluation. A roadmap to the report is provided below (Figure 2).



⁴ These four citizen-led assessments were evaluated because they are funded in part by the William and Flora Hewlett Foundation, which commissioned the evaluation. Hewlett also funds MIA and LEARNigeria, but they were not included in the evaluation as it was only recently launched in a pilot phase. R4D is currently conducting an evaluation of ASER Pakistan in partnership with the Open Society Foundations.

THE ORIGINS OF CITIZEN-LED ASSESSMENTS

ASER

EST. 2005

Pratham is a non-governmental organization (NGO) that has worked in communities and schools throughout India since 1995 to improve enrollment, attendance, and learning. ASER was inspired by the Pratham practice of developing village report cards before beginning work in a given village. The report card was developed by visiting every household in the village and recording whether or not children went to school. Then, a simple, short reading and math test was administered to each child. The purpose of the assessment activity was to enable ordinary citizens to participate so they could understand the reality of their children's learning and take appropriate action. The interest that was generated in communities during the development of village report cards, together with the levying of a 2% tax for elementary education in 2004, brought to light the importance of including citizens in the monitoring of outcomes delivered by the education system. In 2005, the Pratham team initiated ASER, a sample-based national household survey that assesses children's basic reading and math skills.

ASER inspired the development of several citizen-led assessments in other countries. Leaders of NGOs and research organizations in these countries were similarly motivated by an observed need to include citizens in the monitoring of educational outcomes. ASER provided a concrete example of how this could be done on a large scale. Each of the citizen-led assessments below visited India to witness the ASER survey early in their planning phases.

Uwezo

EST. 2009

In the early 2000s, Kenya, Tanzania, and Uganda were aiming for universal primary education. Governments focused on enrollments and other schooling inputs. On the other hand, civil society organizations were engaged in agitating for better quality of education, but there was no consistent and compelling narrative about raising learning outcomes in particular. Parents and citizens expressed indifference, celebrating more schooling but feeling powerless to advocate for better quality education for their children. A team of NGO leaders from all three countries, convened by Rakesh Rajani, came together and decided to adopt simple tools to generate large-scale, rigorous evidence on literacy and numeracy, and this initiative was named Uwezo, Kiswahili for "capability."

Beekunko

EST. 2011

While the extensive decentralization that has taken place since 1992 reflects the Malian government's willingness to empower citizens to manage their own development, civil society has not been heavily involved in the development of education policies and programs. Further, parents and communities are not aware of the actual level of learning among their children. Therefore, there has not been the opportunity for informed citizens to demand better services—a fundamental characteristic of a strong democracy. Given this situation, Œuvre Malienne d'Aide à l'Enfance du Sahel (OMAES), an NGO based in Bamako, aimed to build an initiative that engaged more actors in addressing the education challenges in Mali. OMAES determined that there was a need for civil society to build an independent mechanism for the assessment of learning. It named this initiative Beekunko, which means "the concern of everyone" in Bamanakan.

Jàngandoo

EST. 2012

In 2011, Laboratoire de Recherche sur les Transformations Économiques et Sociales (LARTES), a research organization located within l'Université Cheikh-Anta Diop in Dakar that specializes in the study of economic and social transformations and the governance of social policy, conducted a study on the dynamics of education and poverty. Their research uncovered both entrenched structural inequalities and cyclical vulnerability among specific groups. LARTES used these findings to assess the extent to which current policies aiming to address education and poverty were relevant given the patterns they had identified. This work led them to see that a paradigm shift was needed that would encourage a focus on outcomes—so that inequality could be identified and addressed. LARTES therefore proposed to develop a tool to measure the quality of education in Senegal. It named this initiative Jàngandoo, which means "let us learn together" in Wolof.

Chapter 1. What is a Citizen-led Assessment?

Citizen-led assessments are tailored to each country context, but they share a set of defining characteristics. Citizen-led assessments are:

- **An assessment of basic reading and math competencies:** The survey is an assessment of children’s mastery of basic reading and math, using tools that are simple to administer and easy to understand.
- **Conducted in households:** The assessment is structured as a survey and is conducted in households, not schools, so as to include all children—not just those enrolled in and present in government and other recognized schools on testing day. This means that out-of-school children are included, as well as children in private and unrecognized schools.
- **Conducted orally and one-on-one:** Unlike a written test, this method allows children to demonstrate skills even if they cannot read the questions.
- **Statistically representative:** Survey findings are statistically representative at various levels (e.g., national, state, district) with the goal of enabling discussions, planning, and action at each level.
- **Independent, as in, organized by civil society:** The assessment is organized by an NGO, not the government, and carried out through a network of partners. Assessors are volunteers from the districts in which they are conducting the survey.

Citizen-led assessments differ from many other large-scale assessments in several key ways:

- They test basic competencies, and include only a small number of test items for each competency
- All children are tested at the same level, regardless of age or grade-level
- Citizens are involved in data collection, and the data collection process itself is an important component of the initiative, not just sharing data after the assessment
- They have a broad audience, not just authorities and policymakers
- They are conducted in households and not in schools

Citizen-led assessments aim to achieve national scale while still remaining statistically representative at lower levels such as state, region, and district. Most citizen-led assessments start at a sub-national scale and plan to grow to national coverage over time. They are typically conducted annually.

Table 1. 2013 Citizen-led Assessment Coverage

	Districts/Regions	Communities	Households	Children
ASER	All 550 rural districts	15,941	327,397	569,664
Beekunko	5 of Mali’s 8 regions plus the district of Bamako	1,080	21,251	79,079
Jängandoo	All of Senegal’s 14 regions	103	5,000	15,277 ⁵
Uwezo ⁶				454,367
Kenya	156 of approx. 158 districts	4,507	89,553	189,552
Tanzania	131 of 133 districts	3,833	76,463	148,000
Uganda	80 of 80 districts	2,391	47,345	116,815

⁵ In 2014, Jängandoo increased its sample to 10,000 households and over 26,000 children.

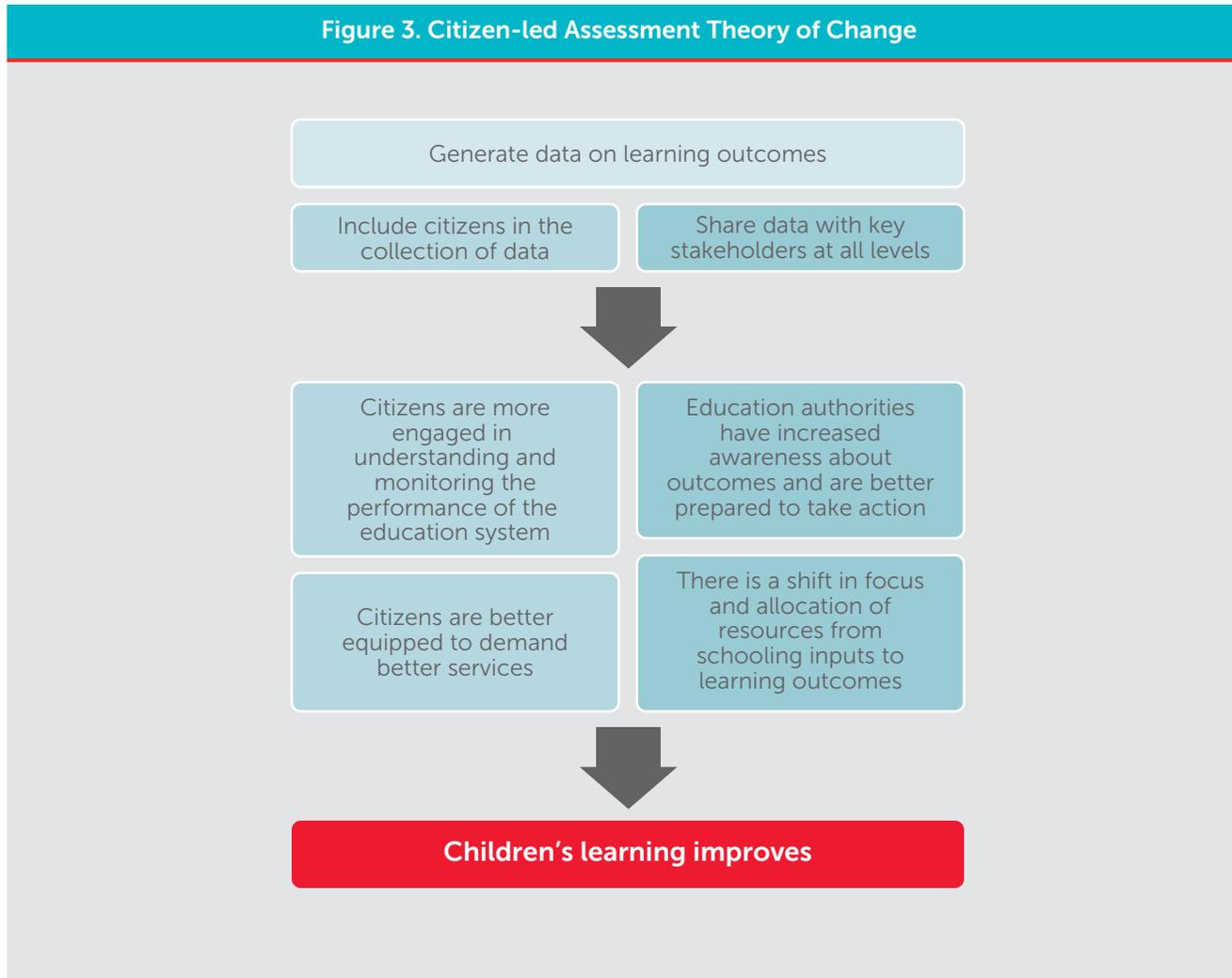
⁶ Uwezo plans to increase its sample size in 2015 to 158 districts in Kenya, 159 districts in Tanzania, and 112 districts in Uganda. This corresponds to 100% of the districts in each country according to the 2009 census in Kenya, the 2012 census in Tanzania, and the 2014 census in Uganda.

Theory of change

The goals of citizen-led assessments go beyond the generation of data on learning outcomes. They intend to stimulate a movement to improve learning outcomes.

Citizen-led assessments are meant to not only influence those who participate in them directly (partner organizations, volunteers, and parents whose children are tested), but also others who hold responsibility for children’s schooling and learning. The relative focus that each citizen-led assessment assigns to a given audience varies, but, in general, the target audiences are:

- Citizens (parents, teachers, students, other community members)
- Citizen groups (including but not limited to parents’ associations), NGOs
- Universities, colleges, teaching institutes
- Elected representatives (at national, state/regional, district, and community levels, including traditional leaders)
- Government (national, state, and district officials, public sector agencies related to education)
- International education stakeholders



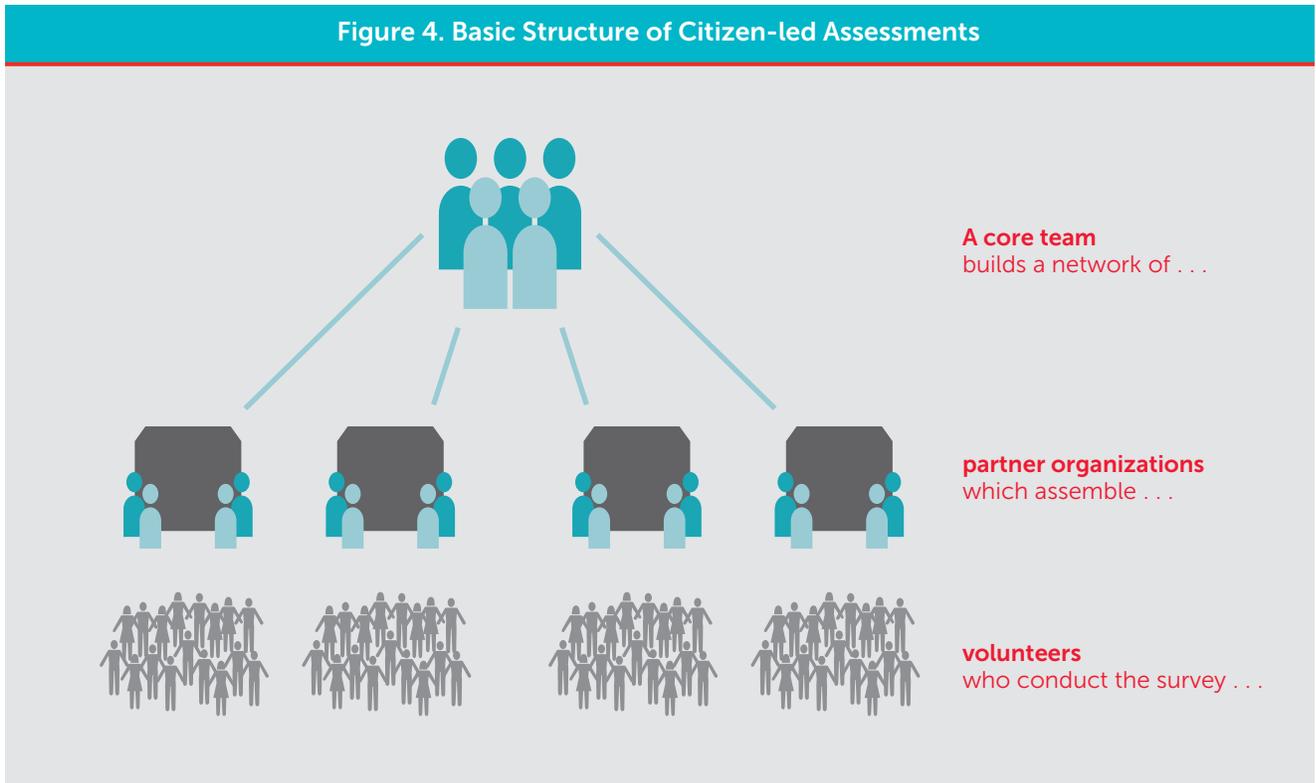


ASER: Conducting the household survey in India.

Who is involved?

Each citizen-led assessment is led by a core national team that oversees survey design and implementation, conducts data analysis, develops reports, and coordinates national and international dissemination. The core teams build networks of district-level or regional institutions to carry out the survey. In every district or region, a partner organization is identified to coordinate the survey. They are responsible for assembling volunteer⁷ surveyors who will be trained to conduct the survey.

Figure 4 illustrates the basic structure common to all four citizen-led assessments. The details of the key players involved in each are provided below.



⁷ Volunteers are not compensated for their time but they do receive a stipend to cover their expenses including travel to and from training sessions and survey sites.

ASER

The **ASER Centre national team** has led ASER since 2008. Pratham conducted ASER annually from 2005 until 2008, when the ASER Centre was established as an autonomous organization within the Pratham family with a mission of not only conducting ASER each year, but also generating data on learning through separate research projects, building the capacity of organizations at the state and district levels to conduct their own assessments, and helping to create an enabling environment for evidence-based planning and policymaking. The ASER Centre and Pratham remain key partners in each other's work, with Pratham focused on implementing education programs, often in partnership with the government and almost always in collaboration with communities themselves. The ASER Centre oversees survey design and implementation, runs the national training, conducts data analysis, develops reports, and coordinates national and international dissemination.

ASER Centre state teams are located in all 29 states. Because India is so large, ASER Centre state teams in some ways perform the same function as the core team in smaller countries. Their role includes:

- ✓ Identify district partners and coordinate survey implementation in partnership with them
- ✓ Train master trainers
- ✓ Lead outreach to state policymakers and other stakeholders; disseminate state reports (in collaboration with Pratham state teams)
- ✓ Coordinate dissemination at the state level through print and broadcast media

District partners are NGOs, CSOs, National Service Scheme volunteer groups, private and government colleges and universities, teacher training institutes, etc. ASER has 500-600 district partners each year. Their role includes:

- ✓ Recruit and assemble 30-60 volunteers per district
- ✓ Coordinate training of volunteers by master trainers

Master trainers are often NGO staff members. ASER has roughly 1,000 master trainers each year. Their role includes:

- ✓ Train volunteers to sample households and conduct the survey
- ✓ Conduct the first round of monitoring and recheck tasks

Volunteers come from the districts where they will conduct the survey, but ideally not from within the village to which they will be assigned. Volunteers are typically young adults including CSO members, local CSO/NGO staff, college or university students, occasionally high school students, etc. 20,000-30,000 volunteers participate in ASER each year.

Beekunko

OMAES oversees survey design and implementation, supports training activities, conducts data analysis, develops reports, and coordinates national and international dissemination.

Partner organizations work closely with OMAES to implement Beekunko. OMAES has three partner organizations: two regional NGOs and one regional network of NGOs. Their role includes:

- ✓ Oversee data collection in their region
- ✓ Facilitate partnerships between Beekunko and regional authorities and decentralized technical government departments
- ✓ Coordinate dissemination of Beekunko data through workshops and regional broadcasting
- ✓ Participate in the development and monitoring of action plans for improving learning

Local government technical agencies and community organizations such as *Centres d'Animation Pédagogique*, School Management Committees, Parents' Associations, and Students' Mothers' Associations support the implementation of Beekunko. Their role includes:

- ✓ Help recruit volunteers
- ✓ Provide logistical support during trainings and the survey itself

Focal points are the equivalent of master trainers. They are ideally young university graduates and at least 25 years old. They are often recruited from CSOs, government divisions, or apply on an individual basis. Their role includes:

- ✓ Train and supervise volunteers as they sample households and conduct the survey
- ✓ Conduct awareness-raising activities about Beekunko within communities

Relais, as Beekunko refers to volunteer surveyors, are generally recruited from volunteer associations, universities, or training institutes. Relais are required to have at least a high school diploma and be between the ages of 23 and 28. They come from the communities to which they are assigned to conduct the survey.

Jàngandoo

LARTES oversees survey design and implementation, supports all training activities, conducts data analysis, develops reports, and coordinates national and international dissemination.

The **Pedagogical Team** is responsible for developing the survey assessment tools. It is comprised of respected and experienced individuals from the education sector (e.g., school inspectors and teachers). The team has both French- and Arabic-speaking members.

The **Pedagogical Task Force** is made up of similar individuals as the Pedagogical Team. The role of the Task Force is to review and provide feedback on the testing tool on a frequent basis.

The **Quantitative Team** consists of a team of statisticians and economists who conduct the survey data analysis.

Partner organizations have an existing on-the-ground presence in their regions. LARTES has ten partner organizations that collectively cover the 14 regions of Senegal. Their role includes:

- ✓ Recruit supervisors and volunteers according to criteria
- ✓ Sensitize local authorities and community organizations about Jàngandoo
- ✓ Organize logistics of household survey
- ✓ Train supervisors and volunteers to sample households and conduct the survey
- ✓ Perform quality control during survey implementation
- ✓ Organize dissemination events in communities

Supervisors are individuals that work through the partner NGO to oversee and coordinate the implementation of the survey.

They come from the communities in which they will supervise the survey. Their role includes:

- ✓ Supervise volunteers
- ✓ Monitor data collection during the survey
- ✓ Report data to core team after survey

Animateurs, as Jàngandoo refers to volunteers, also come from the communities in which they will conduct the survey.

Uwezo

Uwezo is a program of **Twaweza East Africa**, a Tanzania-based organization focused on facilitating large-scale change in East Africa through increased accountability and information sharing. Within Twaweza's new strategy (2015-2018), Uwezo falls under the Directorate of Data and Voice, utilizing large-scale infrastructure to generate convincing evidence through data, ideas, and stories. Twaweza's Learning, Monitoring, and Evaluation Unit provides support to Uwezo activities.

The **Uwezo East Africa Regional Office** is based in Nairobi. The responsibilities of the Regional Office include:

- ✓ Provide technical support to Uwezo's teams in three country offices
- ✓ Lead data management and analysis
- ✓ Conduct quality assurance processes
- ✓ Promote Uwezo assessment findings both regionally and internationally.

Uwezo team members in Twaweza **country offices** in Dar es Salaam, Kampala, and Nairobi orchestrate the survey in their respective countries, operating collaboratively to oversee survey design and implementation, supports training activities, conduct data analysis, develop reports, and coordinate national and international dissemination.

Each country has an **Advisory Committee** which includes education stakeholders such as government officials and teachers' union representatives; the committees convene quarterly and provides feedback on Uwezo processes and tools.

Master Trainers generally include long-standing participants in Uwezo, who, along with regional coordinators, facilitate the training-of-trainers in each country. Trainers conduct the volunteer training.

Regional Coordinators play a key role in managing assessment implementation and training in each region (a cluster of several districts). They serve as a key link between country offices and District Coordinators.

District Coordinators are individuals selected from among the staff of Uwezo's partner CSO in each surveyed district. The role of the District Coordinator includes:

- ✓ Recruit, assemble, and train volunteers
- ✓ Pre-survey engagement with District Education Officers, village elders, and school leaders
- ✓ Conduct household sampling and village survey

Village Coordinators were an addition to the Uwezo model in 2014. They are generally volunteers who have served with Uwezo for multiple years. Their role includes:

- ✓ Assist District Coordinators in engaging with community leaders, teachers, and parents
- ✓ Assist with volunteer training sessions, sampling, and monitoring survey implementation

Volunteers are typically young adults who have completed secondary school. They come from the communities in which the volunteers will conduct the survey. Roughly 20,000 volunteers participate in Uwezo each year.

Note that in the case of ASER, volunteers come from the district where they will conduct the survey, but not their own village. In the other citizen-led assessments, volunteers are, as often as possible, assigned to conduct the survey in their own village. This is a result of each group weighing the benefits of having the volunteer be a known member of the community against the potential bias of having volunteers survey households and children whom they know.

What activities take place?

This section describes the components of a citizen-led assessment, including:

- Sample design
- Test development
- Training
- The survey itself
- Quality assurance
- Data analysis and reporting
- Dissemination and engagement

Sample design of citizen-led assessments

ASER

Rural

Urban

All rural districts in all states are included in the survey (550 districts total in 2013). ASER uses a 2-stage sampling design:

Stage 1: 30 villages from each district are sampled with probability proportional to size from the village directory in the national census.

Stage 2: 20 households are randomly sampled in each village. The sampling of households is undertaken by volunteers on the day the survey is administered. Volunteers walk around the village and construct a map with the help of local people. They then use the map to divide the village into quadrants, and, starting in the middle of each quadrant, they select every fifth household on the volunteer's left-hand side until five households per quadrant have been selected.

Results are statistically representative at the

- National level State level District level

Jàngandoo

Rural

Urban

All 14 regions of Senegal are included in the survey. Jàngandoo uses a 2-stage sampling design:

Stage 1: The number of census districts¹¹ to be sampled from each region is proportional to the size of the region. After being categorized as either rural or urban, census districts from each category are sampled with probability proportional to size. A total of 103 census districts were included in the sample in 2013.¹²

Stage 2: 20 households in each census district are randomly sampled by volunteers on the day of the survey. They choose a landmark in the census district from which to begin, and then sample every fifth household (in urban areas) or every fourth household (in rural areas).

Results are statistically representative at the

- National level Regional level (2015 and beyond)

Sample design

Across all four assessments, the sampling design and methodology has been developed in consultation with experts from the national statistical institutes in the respective countries.⁸

Each citizen-led assessment has a unique sampling design tailored to the country's structure and the scale of the survey, but they share many common features.

Beekunko

Rural

Urban

Each year a subset of the 8 regions in Mali are included in the survey.⁹ Beekunko uses a 3-stage sampling design:

Stage 1: *Communes*¹⁰ are sampled randomly from each region (216 total in 2013)

Stage 2: 5 villages are sampled with probability proportional to size from each *commune*.

Stage 3: 20 households are randomly sampled from each village by volunteers on the day of the survey. Volunteers construct a map with the help of the village head. Beginning at the house of the village head, volunteers sample households at intervals until 20 households are selected.

Results are statistically representative at the

- Commune* level

Uwezo

Rural

Urban

Nearly every district in all three countries is included in the survey.¹³ Uwezo uses a 2-stage sampling design:

Stage 1: 30 enumeration areas¹⁴ are randomly sampled from each district using the national census.

Stage 2: 20 households are randomly sampled in each enumeration area by District Coordinators in advance of the survey. This process includes meeting with a village elder to create a household listing, referring to an enumeration area map provided by the Bureau of Statistics. The District Coordinator selects every *n*th household depending on the size of the enumeration area to arrive at a total of 20 sampled households. Volunteers are provided with the map including names of heads of households for those selected, which they use along with the household listing to identify the randomly selected households on the day of the survey.

Results are statistically representative at the

- National level District level

⁸ For ASER, this includes the Indian Statistical Institute, the Planning Commission of India, and the National Sample Survey Organisation. For Beekunko, L'Institut National de la Statistique. For Jàngandoo, L'Agence National de la Statistique et de la Démographie. For Uwezo, the National Bureau of Statistics in each country: Kenya, Tanzania, Uganda

⁹ 2013 regions included Kayes, Koulikouro, Sikasso, Ségou, Mopti, and the District of Bamako. 2014 regions included Kayes, Koulikouro, Sikasso, and the District of Bamako.

¹⁰ *Communes* are roughly equivalent to districts; there are 703 *communes* in Mali.

¹¹ Census districts are "geographical entities with clearly identifiable boundaries;" on average, each census district has 100 households (LARTES-IFAN, 2013c).

¹² In 2014, Jàngandoo increased its sample to 239 districts.

¹³ In 2014, this design continued in Kenya, but the number of districts covered was reduced to 50 districts in Tanzania (down from 131 in 2013) and 30 districts in Uganda (down from 80 in 2013). The reduction in coverage in Tanzania and Uganda was motivated by a desire to strengthen implementation at a smaller scale, including finding strong partners and experimenting with new training structures. In 2015, all districts will again be included using the original sampling design described here.

¹⁴ Enumeration areas are administrative units set by the National Bureau of Statistics. They sometimes but not always have the same boundaries as villages. EAs are used as the community level sampling unit because they are the unit used in the census and for which maps exist.

Each citizen-led assessment has detailed sampling instructions for unique circumstances in the household sampling stage, such as accounting for multi-unit apartment buildings in urban areas or villages in which houses are arranged linearly and should therefore be divided into four adjacent sections as opposed to quadrants.

Test development and testing tools

Test development

Test development for each citizen-led assessment is the responsibility of a panel comprised of people from various parts of the country's education sector including teachers, school inspectors, and ministry representatives specializing in curriculum and examinations.¹⁵ In some cases a consultative committee of experts reviews the tests once they have been developed. The tests are designed to be

simple and easy to administer, and to assess only a basic set of reading and math competencies. The tests assess skills included through each country's Grade 2 curriculum, with the exception of Jàngandoo which is aligned to Senegal's Grade 3 curriculum. In contrast to most assessments, all children are given the same test, regardless of their age. This approach enables the collection of data that shows how many children are performing below grade level and how far behind they are; it also fits with citizen-led assessments' goal of collecting basic data that is easy to understand: for example, they do not aim to report on children's ability to perform on grade level, but simply to report the percentage of children who can complete a basic task, like reading a short paragraph of simple sentences.

Table 2. Reading Tasks Included in Citizen-led Assessments

	ASER	Beekunko	Jàngandoo	Uwezo
Letter recognition	✓	✓	✓	✓ ¹⁶
Syllable recognition (read aloud 2- or 3- letter syllables)			✓	
Word recognition	✓	✓	✓	✓
Fluency (read aloud a 4-sentence text)	✓	✓	✓	✓
Fluency (read aloud a 6-15 sentence text)	✓	✓		✓
Comprehension (answer questions based on text read)	2006, 2007	✓	✓	✓

Figure 5. Examples of Reading Tasks from Testing Tools

From Jàngandoo reading tasks 2014

From ASER reading tasks (Hindi and English)

¹⁵ Note that in the case of ASER, ministry representatives are not involved in test development.

¹⁶ Uwezo combines a letter (English) and syllable (Kiswahili and local languages) task: from a set of 10 letters or syllables, children select 5 to read aloud.

Table 3. Math Tasks Included In Citizen-Led Assessments

	ASER	Beekunko	Jàngandoo	Uwezo
Counting and numeration (match a number of objects to the number symbol, draw the number of dots that correspond to a number symbol)			✓	✓
1-2 digit number recognition	✓			✓
Comparing numbers (order numbers 0-9 and 10-99, select the larger of pairs of 2-digit numbers)		✓		✓
Addition of 1-2 digit numbers without carrying		✓	✓	✓
Addition of currency (add two 3-digit dollar amounts without carrying)	2008			✓ TZ only
Addition of 3- 4 digit numbers with carrying			✓	
Subtraction of 1-2 digit numbers without borrowing			✓	✓
Subtraction of 2-4 digit numbers with borrowing	✓	✓	✓	✓ TZ only
Multiplication of two 1-digit numbers or a 1-digit number and a 2-digit number				✓
Division of 1-3 digit number by 1-digit number	✓			✓ KE and UG
Problem solving (word problem involving 1-3 digit addition or subtraction)	2007, 2010	✓	✓	
Geometry (identify shapes by name, complete a partially drawn shape, draw a shape) and measurement (identify whether a given object should be measured in units of length, volume, or mass)				

Figure 6. Examples of Math Tasks from Testing Tools

From Jàngandoo math tasks 2014	From Uwezo Kenya math tasks 2012

Testing tools

ASER, Uwezo Kenya, and Uwezo Tanzania follow a test administration method in which a child starts at a middle difficulty task and progresses either up or down depending on whether or not he or she is successful at this first task. This means that all tasks are not administered to all children. For example, a child who successfully completes the “which is greater” task on the Uwezo test shown above will not be administered the counting and number recognition tasks but will instead proceed to the addition task

Beekunko and Uwezo Uganda follow a test administration method in which a child starts at the lowest-level task and progresses up through the tasks in increasing order of difficulty. In these two assessments, the administration stops when the child is not able to successfully complete a task or once he or she has attempted the final task, whichever comes first. Jàngandoo follows a similar method but, if a child can successfully read words, he or she will attempt both of the more difficult tasks (fluency and comprehension) including attempting the comprehension questions even if the fluency task is not successful (see the Jàngandoo example tool in Figure 5).

School and village survey

In each household, volunteers conduct a household survey in addition to the reading and math tests. This survey is developed by the core teams and includes information on the enrollment status of each child in the household, what type of school he or she attends, parents’ age and education levels, and questions on household assets. Each citizen-led assessment also has a village survey and a school survey. The village survey collects information on the schools in the village¹⁷ and other



Jàngandoo: A village elder in Senegal helping to map the village before the survey.

characteristics such as electricity and water sources, health services, etc. The school survey collects information on the number of students and teachers present, facilities, etc.

Training

Beekunko, Uwezo, and ASER follow a training-of-trainers approach. For Beekunko and Uwezo, master trainers are trained at a national training workshop lasting 3 days. Uwezo has an additional set of regional trainings in which the same master trainers are trained in smaller groups for 2.5 days to go into greater detail and depth on the assessment skills covered in the national training. For ASER, the national workshop lasts 5 days and is a training of ASER state teams, which then return to their states and hold their own trainings for master trainers. State trainings typically last 4-5 days.

For all three initiatives, the master trainers then (shortly after they have been trained, to facilitate retention) conduct the training of volunteers who will conduct the survey in each district (ASER and Uwezo) or region (Beekunko). This final round of trainings lasts for 2-3 days.

Jàngandoo’s structure is somewhat different. It does not use a tiered approach, but holds seven trainings around the country which together cover the 14 regions of Senegal. Supervisors are trained at the same time as volunteers by the Jàngandoo national team and staff members from the partner NGOs that are responsible for coordinating the survey in the regions covered by the training. These trainings last for 5 days.

In all cases, training sessions include both classroom sessions and field practice sessions where trainers and volunteers are instructed in the following areas:

- How to introduce the citizen-led assessment to parents and other village members
- How to conduct the village and school surveys
- How to conduct the sampling of households (for Uwezo this is only covered in the training-of-trainers, as District Coordinators, not volunteers, conduct the household sampling)
- How to conduct the household survey
- In the cases of Beekunko, Jàngandoo, and Uwezo, how to give “instant feedback” to parents.
- How to administer the assessment to children
- For master trainers and supervisors, how to provide support and conduct monitoring activities during the survey

¹⁷ The term “village” is used for consistency’s sake, but Jàngandoo refers to this unit as a “census district” and Uwezo refers it as an “enumeration area.”

Conducting the survey

Immediately after they are trained, two volunteers are dispatched to each sampled village to conduct the survey. Uwezo aims to have one female and one male volunteer make up each pair, and Jàngandoo aims to have one Arabic speaker and one French speaker. The survey is designed to be conducted in as short a time period as possible, typically with all data collected within 1-2 months. One key reason for this is to ensure that children who are tested are at nearly the same point in the school year. Data is collected by the volunteers, who turn it in to partner organizations, which send it to core teams for processing and analysis.

Quality assurance

For all four citizen-led assessments, master trainers, partner organizations, and core teams support the survey administration by visiting sampled villages and having support available for volunteers via mobile phone. Beyond that, the quality assurance mechanisms in place for monitoring data collection vary for each initiative.

For ASER, a monitoring and recheck process is carried out by trainers and other supervisors and includes a desk check of data collection forms, a phone recheck of a subset of schools and households, and a field recheck in a subset of villages and households in each district. For Beekunko, focal points are meant to monitor and support volunteers while they are conducting the survey. There is not to date a set monitoring system in place that indicates exactly what such monitoring entails or if a fixed number of communities and households are checked by focal points to ensure that they were in fact visited.

For Jàngandoo, master trainers are given the responsibility of monitoring the activities of the volunteers and are required to fill in various monitoring tools such as a field book and supervision matrix.



Master Trainers in Rajasthan, India conducting the re-check process during a training session

For Uwezo, district and village coordinators travel to communities within their district during the survey to check in on the volunteers, provide them with support if needed, and check to make sure they are following the prescribed methodology. In 2013 Uwezo conducted follow-up visits to a subset of districts to assess the implementation of the survey, and even conducted a resurvey in one district. This is an excellent way to self-evaluate the strengths and weaknesses in implementation that can be addressed in the next year's training and survey design.

Data analysis and reporting

Across all four citizen-led assessments, frequency analyses are conducted which determine the percentages of children successfully completing different *tasks* in the assessment (e.g., letter level, word level, story level, etc.). Jàngandoo and Uwezo conduct frequency analyses for children's performance on the assessments *overall* in order to give a "pass" rate (i.e., the percentage of children successfully completing all tasks in the assessment). Beekunko also calculates average scores in the reading and mathematics assessments for the entire dataset and then disaggregates the data to provide average scores by region, *commune*, and community. Beekunko then categorizes these averages as representing a "critical situation" (which corresponds to an average score less than 2.5/5), "warning situation" (between 2.5 and 2.9/5), and "situation in which children are performing" (between 3 and 5/5).

Dissemination – How is the data shared?

Dissemination is a vital link in the chain from data collection to action, and the way in which assessment results are disseminated and communicated is critical to the likelihood of such assessments achieving the impact they strive for. While the four initiatives all prioritize slightly different audiences, broadly speaking each of them seeks to increase awareness about low learning outcomes at the community, sub-national, national, and international levels, and to encourage key stakeholders (parents, schools, government and education officials, national policymakers, etc.) to take action. Each citizen-led assessment has a portfolio of dissemination activities targeted to these various audiences.

At the **international level**, citizen-led assessments disseminate the results of their own surveys, and work to raise the profile of the citizen-led assessment movement. The core teams frequently present survey findings at international conferences and meetings.

Citizen-led assessments are active in disseminating their results at the **national level**. This generally includes three main activities:

- Developing a national report and other briefs, hosting national launch events
- Widespread media campaigns (print, TV, radio)
- Presenting assessment findings to key stakeholders (including Ministries of Education, technical government agencies, NGOs, etc.)

Notably, Jàngandoo holds pre-survey launch events to kick off the survey cycle at the national, regional, and *commune* levels. These events introduce the Jàngandoo assessment and its goals, not only raising the visibility of the initiative among key stakeholders but also obtaining their buy-in to the assessment

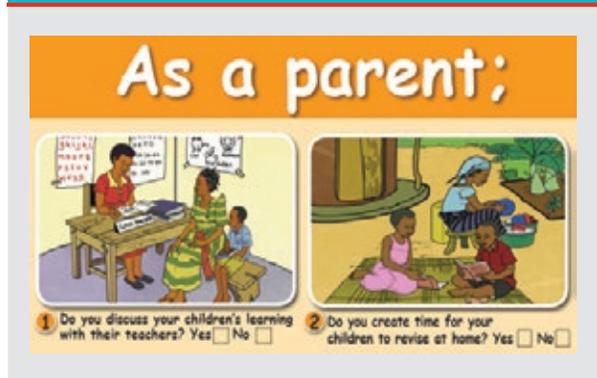
At the **sub-national level**, dissemination strategies across the four citizen-led assessments vary widely. ASER and Uwezo focus their sub-national efforts on media: core and state teams work with media houses to get the ASER findings in print and broadcast media in various languages.

Other communications materials are produced each year. For example, in some years ASER and Uwezo have made district report cards showing the results of a particular district. In Tanzania, a district ranking poster is made each year showing each district's results beside the photograph of the district representative. Some state teams and district partners hold launch events or meet with education authorities; but these activities are not part of a systematic dissemination plan.

In the case of Jàngandoo, local NGO partners organize dissemination events at the regional level and facilitate working groups which convene for the purpose of developing action plans to address certain problems identified from the survey results. Beekunko follows a similar model, whereby Beekunko's technical team in collaboration with master trainers organize workshops at the regional level for local policymakers (including mayors and local councils), and other key local education actors. The purpose of these sessions is to present the results from the Beekunko survey and to generate debate around the key underlying problems in the education system, following which the convened stakeholders work together to develop action plans.

In order to include ordinary citizens in the movement they aim to build, dissemination at the **community and household level** is needed. Beekunko and Uwezo conduct outreach to communities before the survey. For Beekunko, focal points (similar to master trainers) either visit or call village leaders to tell them that the survey will be happening in their community and to introduce the survey process and objectives. For Uwezo, District Coordinators

Figure 7. Excerpt from Uwezo Pamphlet Left with Parents after the Survey



visit the village and work with the village leader to conduct the household sampling, as described above.

In most citizen-led assessments, the first form of dissemination at the community level occurs during the survey itself; community members observing the assessment witness children's learning levels first hand. Additionally, Beekunko, Jàngandoo, and Uwezo provide "instant feedback" to parents on their children's performance on the assessment. Schools and village leaders are also involved when volunteers conduct the school survey and engage with the village leader before beginning to collect data.

Beekunko is unique from the other three assessments in that it takes the process of instant feedback one step further. Before leaving the communities in which they have conducted the survey, Beekunko volunteers are responsible for convening a community-level meeting where they gather local authorities (village heads, for example) and community members to share information about the survey.

Uwezo volunteers also leave behind calendars and fliers (Figure 7) with pictorial and text suggestions for actions parents can take to support their child's learning ("read with your child," "talk to your child's teacher," etc.). Jàngandoo volunteers provide suggestions orally to parents regarding how they can better support their children's learning.

None of the four citizen-led assessments systematically return to the village level to disseminate findings after the survey, but examples of very small scale dissemination activities do exist. In these cases, volunteers, district partners, or even staff of core teams will go to communities and conduct local engagement. These activities are sporadic and not formally organized or structured. Significantly more resources would be required for such ongoing engagement at the community level, due to the large scale of the surveys.

How much do citizen-led assessments cost?

Citizen-led assessments are meant to be low-cost alternatives to more formal assessments or household surveys, making them feasible initiatives for civil-society organizations to implement. One of the principle sources of cost-savings is the use of volunteers as data collectors. Of course, the motivation behind using volunteers is also that they represent the “citizen-led” approach; the engagement of ordinary citizens is meant to contribute to a growing awareness of learning levels among the citizenry, and to promote action.

Table 4 describes the costs associated with each citizen-led assessment for the 2013 assessment cycle. The costs are presented side-by-side, but it is important to bear in mind that there are various factors that affect the large differences between the initiatives’ per-child costs, including but not limited to the following:

- Large discrepancies in the funding allocated to activities that are not related to the number of children assessed, such as dissemination.
- Unit costs for earlier-stage initiatives are often higher than for more mature ones.

- In-kind contributions are not represented in these costs. For example, some of ASER’s costs are subsidized by Pratham (though to a lesser extent each year) which provides training venues and staff support in select states.
- The cost of goods and services varies between countries.

The value of the “citizen-led” component of these assessments is revealed by their costs per child assessed. When compared to other assessment models, these are mostly low unit costs: a study by Wagner et al. (2011) which compares four major types of assessments (national, regional, international, and hybrid) over 13 different contexts finds an average cost per child surveyed of \$42.

When investigated more deeply, these numbers are even more impressive: ASER’s very low cost per child of \$2 is roughly one fifth of the cost of the least expensive assessment included in Wagner’s study (Uruguay’s national assessment which cost about \$8 per child to implement).

It is illustrative to consider citizen-led assessments’ allocation of resources to various activities. Because data collection is the foundation for all other activities, when citizen-led assessments face budget constraints their priority must necessarily be to fund data collection, often leaving limited funding for dissemination and engagement. The breakdown is shown below.

Table 4. 2013 Citizen-Led Assessment Costs

	ASER	Beekunko	Jàngandoo	Uwezo (overall) ¹⁸
Number of children	569,664	79,079	15,277	454,367
Other coverage details	550 districts 327,397 households	5 regions + Bamako 21,251 households	14 regions 5,000 households	367 Districts 213,361 Households
Total annual cost (USD)	\$881,123	\$753,022	\$770,152	\$7,620,701
Cost per child assessed (USD)	\$1.55	\$9.52	\$50.41	\$16.77

Table 5. Resource Allocation by Activity as a Percentage of Total Budget

Activity Type	ASER	Beekunko	Jàngandoo	Uwezo
Core: salaries for core team, administrative costs	21%	31%	37%	22%
Data collection: partner recruitment, training, survey materials, master trainer and supervisor pay, volunteer stipends	74%	47%	57%	52%
Dissemination: report production, launch events, other communications materials, dissemination activities	5%	22%	6%	26%
Total	100%	100%	100%	100%

¹⁸ Uwezo’s overall cost for all three countries is provided here to account for the significant resources allocated to the Uwezo regional office, which contributes to its survey implementation across countries as well as dissemination. Data on children surveyed and households visited is also combined from all three countries.

Chapter 2. How Well do Citizen-led Assessments Work?

Evaluation methodology

Citizen-led assessments have attracted attention in recent years as measuring and improving learning outcomes have become priority issues in the field of education. R4D's evaluation contributes to the conversation by assessing the extent to which the citizen-led assessment model is able to measure learning, disseminate findings widely, and stimulate awareness and action to improve learning outcomes.

The evaluation was composed of three parts:

- A technical review of the testing tools, sampling design, and analytical processes used
- A process evaluation
- A non-experimental evaluation of impact¹⁹

These three parts map to the sections in this chapter:

- How well do citizen-led assessments measure learning?
- How well do citizen-led assessment processes work?
- How well do citizen-led assessments stimulate awareness and action about learning outcomes?

Technical review: How well do citizen-led assessments measure learning?

The technical review was led by the Australian Council for Educational Research (ACER). It included a desk review for each of the four citizen-led assessments that examined the sample design, the test development process, the testing tools themselves, test administration methods, and data analysis, including the way in which assessment results are reported. As part of the desk review, survey data were analyzed in an attempt to replicate reported findings. The technical review also included two quasi-experimental studies of the Uwezo Kenya testing tools: a concurrent validity study and an inter-rater reliability study (ACER 2015). Findings from earlier studies of the validity and inter-rater reliability of ASER testing tools (Vagh 2013) were also incorporated.

¹⁹ The evaluation of impact was only conducted for ASER and Uwezo, as Beekunko and Jàngandoo have only been conducting assessments for three and two years respectively, and such an exercise would be premature given that they are still refining their models. R4D developed an evaluation framework for each that can be used to inform the design of a future evaluation of impact.

²⁰ A detailed list of observation sites is provided in Annex 2.

Process evaluation: How well do citizen-led assessment processes work?

The process evaluation was carried out by R4D in partnership with in-country partners: OWN & Associates in East Africa, Catalyst Management Services in India, Mr. Abdoulaye Bagayogo in Mali, and Mr. Souleymane Barry in Senegal. It assessed the design and implementation of activities conducted by ASER, Beekunko, Jàngandoo, and Uwezo. The process evaluation addressed six key areas: the suitability and effectiveness of key players in the survey process, training of trainers and volunteer surveyors, execution of household sampling, survey administration, quality assurance, and dissemination and engagement. The process evaluation entailed direct observation in a small sample of sites in each country, as well as interviews and desk reviews:²⁰

- Direct observation of training-of-trainers and training-of-volunteers
- Direct observation of household sampling during training field sessions and the actual survey administration
- Direct observation of household, community, and school survey administration during training field sessions and the actual survey
- Direct observation of reading and math assessments being administered in households during training field sessions and the actual survey
- Direct observation of dissemination events
- Interviews with staff, partner organizations, coordinators, master trainers, and volunteers
- Interviews with external stakeholders about reach and effectiveness of products and events
- Review of citizen-led assessments' own internal monitoring data
- Mapping of planned dissemination activities, desk review of dissemination products
- Desk review of training manuals and agendas
- Review of an earlier set of studies of Uwezo, including an impact evaluation (Lieberman et al. 2012, 2014)

The process evaluation was not a comprehensive audit of all activities. Instead, observations, interviews, and reviews of existing data were triangulated in order to identify recurring strengths and challenges. In particular, the process evaluation was focused on identifying process constraints that are likely to affect the impact of the citizen-led assessments.

Evaluation of impact: How well do citizen-led assessments stimulate awareness and action?

The evaluation of impact was carried out for ASER and Uwezo in partnership with the same in-country partners as the process evaluation. It assessed the extent to which the ASER and Uwezo initiatives have increased awareness of learning outcomes and influenced actions to address poor learning achievement. Citizen-led assessments are meant to influence a wide variety of audiences, ranging from national policymakers to parents. In order to assess impact across this spectrum, the evaluation methodology was designed to gather data at the national, state, district, and community levels. The evaluation of impact used a non-experimental approach, as identifying a valid control group for a national-scale initiative is not possible given the importance of the country context to the potential for impact. A contribution analysis approach was used to explore the extent to which the ASER and Uwezo initiatives have contributed to observed results in the external environment (i.e., increased awareness about low learning levels and action to address them).

Primary data was collected through key-informant interviews and focus group discussions with key stakeholders including:²¹

- National policymakers, technical experts, NGO leaders, and academics
- State- and district-level education stakeholders including government authorities and NGOs
- ASER Centre national and state teams, Pratham state teams
- Uwezo regional and country teams, Twaweza staff

In both East Africa and India, key informants were interviewed about their experience with and opinion of Uwezo or ASER, their perception of the initiatives' impact on raising awareness and action to improve learning outcomes, as well as their suggestions about what could be done differently in the future. In India, a total of 87 key informants were interviewed for the evaluation. In East Africa, 84 key informants were interviewed.

In Kenya, Tanzania, and Uganda, a national stakeholder forum was convened in each of the three countries' capitals, bringing together a diverse group of prominent education players (Ministry of Education officials, teachers' union representatives, academics, NGO leaders) for a discussion of Uwezo's role and impact. Between the three countries, 65 national-level stakeholders participated in these forums, which sought to synthesize a variety of perspectives to generate an informed consensus on Uwezo's influence and impact. Observations and findings from these forums were triangulated with data from key-informant interviews.

Secondary data review included:

- Review of citizen-led assessment data
- Review of media coverage
- Review of national and state education policies to track references to citizen-led assessment data and activities
- Mapping of use of citizen-led assessment testing tools by other organizations
- Review of findings of earlier studies of Uwezo's impact²²

The evaluation of impact included a sample of four states in India and three districts in each of Kenya, Tanzania, and Uganda. The selection of evaluation sites was carried out through purposive sampling, seeking to capture diversity in perspective from a broad spectrum of stakeholders that may have been impacted by the ASER and Uwezo initiatives.

In India, the following criteria were taken into account when conducting the purposive sampling: change in states' learning levels since ASER started, regional representation, and the intensity of Pratham presence in the state. Aiming for a sample of states with diversity according to these criteria, the following states were selected for the evaluation:

²¹ A detailed list of informants is provided in Annex 3.

²² Lieberman et al. (2012, 2014, 2014b) and Uwezo's own 2013 monitoring reports

State	Criteria
Bihar	Learning levels have remained low; Eastern Region; Pratham works closely with the government, including in all 37 districts in collaboration with district education offices
Karnataka	Learning levels have remained stable; Southern Region; Pratham is active in 2 of 27 districts
Madhya Pradesh	Learning levels have shown a sharp declining trend since 2006; Central Region; Pratham is active in 3 of 51 districts
Uttarakhand	Learning levels have gone up (the only state in India in which learning levels have improved); Northern Region; Pratham is active in all 13 districts

While in-depth primary data was only collected in these four states, the evaluation team also interviewed ASER State teams from 14 states when they gathered for the 2013 national training.

Two districts in each state were selected as the focus areas for key informant interviews. In some cases, the evaluation team was able to interview individuals from additional districts as well. The focus districts included in the evaluation were:

- Bihar: Siwan, Vaishali
- Karnataka: Belgaum, Mandya
- Madhya Pradesh: Chhindwara, Panna
- Uttarakhand: Harwar, Almora

In East Africa, a purposive sample of three districts per country was selected. While not representative of each country as a whole, the sample was designed to be illustrative of Uwezo's impact in diverse settings. Seeking to account for regional and geographic diversity, a mix of rural, urban, and peri-urban districts were selected in each country. The districts included in the evaluation were:

- Kenya: Nairobi North, Pokot North, Rongo
- Tanzania: Kibaha, Morogoro, Mpwapwa
- Uganda: Kayunga, Sironko, Wakiso

Two to three enumeration areas in each district were chosen in collaboration with District Coordinators. In total, 28 enumeration areas (8 each in Kenya and Tanzania and 12 in Uganda) were included in the evaluation. A total of 39 focus group discussions comprised of parents, community leaders, and teachers were conducted.

Limitations

Timeline

Data collection occurred during 2013 and 2014. While the evaluation did include a review of each program's documentation since inception, it nonetheless evaluated each citizen-led assessment according to its design at one point in time. Some have long histories of implementation and evolution—for example, ASER is a 10-year initiative, of which this evaluation only captured a window at the end of its course in 2013 and 2014. Moreover, citizen-led assessments are constantly evolving, and have even made program design changes during the course of this evaluation.

Coverage

Citizen-led assessments are being implemented in large countries and in an enormous number of villages and districts (particularly in India, Kenya, Uganda, and Tanzania). The evaluation team collected data in a sample of these sites, a very small sample relative to the number of sites in which the surveys are being implemented.

Boundaries of the evaluation

The purpose and scope of this evaluation was to evaluate the citizen-led assessment model. In some cases these initiatives have extensive partnerships with other organizations, whose work and impact are closely linked to the impact of the citizen-led assessment itself. The main example of this is the case of ASER, which grew out of and continues to work in a symbiotic relationship with Pratham, building relationships with education authorities and working with communities and schools to improve learning. This evaluation did not capture Pratham's activities and impact, and since Pratham and ASER are to some extent inextricably linked, this approach may risk underestimating the impact of the citizen-led assessment (ASER) alone.

How well do citizen-led assessments measure learning?

Citizen-led assessments use unconventional testing tools and data collection methods, intentionally prioritizing simplicity and the involvement of ordinary citizens. This approach has raised questions about the extent to which their findings are valid measures of learning outcomes. The technical review aimed to answer such questions by assessing the following areas, which make up the sections of this chapter:

- Sample design
- Test development processes
- Testing tools
- Test administration
- Data analysis and reporting

KEY FINDINGS

What they measure, they measure well. Citizen-led assessments test a very limited set of competencies in reading and mathematics. The testing tools they use yield valid results, which cast a spotlight on limited achievement in these basic competencies. In this way they are important and useful, but broadening the testing tools to include a broader range of skills would allow citizen-led assessments to do more to inform policy and practice.

Volunteers can reliably assess children’s basic competencies. Inter-rater reliability studies of ASER and Uwezo indicate a high level of agreement in volunteers’ scoring of children’s responses.

Sub-national results should include information about context. All citizen-led assessments report results below the national level (e.g., district, region, *commune*); this invites comparisons that may be misleading when there are differences in context across the different districts, regions, or *communes*, such as socioeconomic status, which is known to have a relationship with children’s performance on learning assessments.

Comparability of findings across years needs to be carefully investigated. Formal equivalence exercises need to be undertaken to ensure the comparability of tests, especially when findings are compared over multiple years and test forms.

Sample design

All four citizen-led assessments work with their respective national statistical institutes. This is a strength of the sampling methodology that lends credibility to the assessment results.

Overall, the 2- or 3-stage sample designs used by the citizen-led assessments are technically sound. One challenge to the validity of the sample is in the final stage: the sampling of households. For ASER, Beekunko, and Jàngandoo, it is the volunteers who sample the households on the day of the survey. ASER volunteers first map the village and divide it into quadrants; they then sample every fifth household until they have sampled five in each quadrant. Beekunko and Jàngandoo volunteers begin at the house of the village head and then sample every fourth or fifth household until they have sampled 20.

There is the potential for this sampling technique to lead to a biased sample, as the population of households is not known. The ASER model of sampling five households per quadrant is stronger than the Beekunko and Jàngandoo model of starting from the village head’s house, because it provides some assurance that more of the population is included in the sampling frame. In situations where complete and accurate household listings are not available, the approaches used are a reasonable solution. A more technically sound method would be for volunteers to create a complete listing of all households in the village before beginning to sample, but the amount of time this would take is not reasonable given the resource constraints of citizen-led assessments and the number of days volunteers are willing and able to devote to the survey. Uwezo is the exception in that it does develop complete household listings, but in order to address these constraints, it is District Coordinators that work with village elders in advance of the survey to develop the household listing. They then randomly sample 20 households from that list for volunteers to survey.



ASER: Master Trainers practicing village mapping in Rajasthan, India.

It should be noted that both household sampling approaches presented significant challenges in practice, which are described below.

The selection of which school or schools to visit for the school survey in each village also has the potential to lead to a biased sample. The specifications that guide the selection of the school vary across the four citizen-led assessments. In the cases of ASER, Beekunko, and Uwezo, the volunteer is advised to try to identify the largest school or the school attended by the majority of children in the village. This means that the overall sample of schools participating in the survey is biased towards larger schools from larger villages (because villages are sampled with probability proportional to size), and as a result, some school-level outcomes might be under- or over-stated. In Jàngandoo the volunteer visits all schools in the sampled census district. In this case the school sample is also biased, but only towards schools from larger census districts (again, because census districts are sampled with probability proportional to size). While sampling schools probabilistically would enable the schools to be weighted and therefore address any selection bias, in the context of the citizen-led assessments, this may not be feasible or practical, and therefore the current method is quite appropriate from a fitness-for-purpose lens, especially given that children, and not schools, are the primary focus of the survey.

Sample size

In the cases of ASER and Uwezo, one question that arose during the evaluation was whether such a large sample is necessary: a much smaller sample might not be representative at the district level, but it would still be representative at the national level—and it might result in significant cost savings while potentially generating similar levels of awareness about learning outcomes at the national level that could trickle down to lower levels. However, the evaluation finds that the larger sample is critical to the potential impact of citizen-led assessments. This is an important part of not only the sample design but the design of these initiatives generally: in addition to allowing for sub-national comparisons and findings specific to subsets of the sample (e.g., girls, children in a certain grade), the very large sample offers tens or hundreds of thousands of “contact points” where participation in the survey has the potential to plant the seeds for action in both volunteers and parents. As will be discussed, this potential has not so far been achieved due to limited sustained engagement at the community level, but the theory of change of citizen-led assessments’ indicates that these initiatives aim to be far more than research efforts; they aim to use survey participation to build movements focused on children’s learning—for this reason the evaluation team finds the sample design to be appropriate.

Test development processes

There are three key findings about the processes used by citizen-led assessments to develop their reading and math tests:

- Citizen-led assessments conduct a variety of small-scale pre-testing activities to validate their testing tools:

PRE-TESTING ACTIVITIES

ASER

- Field piloting in households
- Piloting in classrooms
- Qualitative data collected informs refinements to administration instructions

Beekunko

- Pre-testing activities
- Anecdotal data informs test revision

Jàngandoo

- Pre-test trials to check that items are targeted at the right level
- Qualitative and quantitative data informs item revision

Uwezo

- Three rounds of pre-testing in households followed by revisions
- Anecdotal data collected on tasks children struggle with; used to inform revisions
- District-wide pilot; qualitative and quantitative data collected
- Validation meeting, final tools selected

It is important for all citizen-led assessments to use such activities to generate both qualitative and quantitative data that are used to inform test revisions.

- Most of the individuals who participate in the test-development process of citizen-led assessments are members of the core team who have little formal training in the development of test items. As a result standard procedures for developing tests are not always followed. The technical review of testing tools has identified key areas for improvement and many groups are actively pursuing capacity building in those areas.
- In order to build their credibility, it is important for citizen-led assessments to document and make public their test development processes and assessment frameworks. These are standard documents for assessments, which support public understanding of and confidence in the results.

Testing tools

Citizen-led assessments test a limited range of basic skills. While it is important to maintain the simplicity of the testing tools, foundational concepts that are not currently tested could be added without greatly increasing the complexity or length of the test.

In the current reading tools, only letter recognition, word recognition, and reading fluency are tested across all four citizen-led assessments. In the case of ASER, reading comprehension was tested in the 2006 and 2007 testing tools but not in other years. For Beekunko, reading comprehension is tested at a literal level only. Both Uwezo and Jängandoo test reading comprehension at an inferential as well a literal level. This is a highly constrained range of reading skills. Key foundational skills that are not tested in any of the assessments include listening comprehension and word-level comprehension. Reading assessments in which most if not all of the testing tool is dedicated to non-meaning related components of reading (e.g., letter and word recognition and fluency) de-emphasize the importance of reading for meaning, which may result in it being neglected at all levels: policy priorities, the specifications for textbooks, and classroom practice.

Similarly, the math tests assess only a limited set of foundational skills. In the case of ASER, the range is most constrained, in testing only number recognition, subtraction, and division. In some years additional tasks are added such as problem solving with currency and time-telling, but these are not included consistently across years. Beekunko and Uwezo test a slightly broader range: they include addition and multiplication but also lower-level number tasks that assess skills beyond simple recognition, such as counting and identifying which is the greater number. Jängandoo includes tasks that test knowledge of basic geometric shapes and knowledge of units of measurement, in addition to number skills. Understanding of place value should be tested in an assessment of foundational mathematics skills, because developing this understanding is an important part of developing the ability to work mentally with numbers. Beekunko tests children's understanding of place value with a "number ordering 10–99" task and Uwezo tests it with a "which is greater" task; ASER and Jängandoo do not test it at all.

Overall, Jängandoo most comprehensively assesses children's foundational competencies in reading and math. The other citizen-led assessments do not comprehensively assess foundational competencies—but they do not aim or claim to do so, and they are technically sound tools to measure those competencies that they do cover. Notably,

the Jängandoo test takes much longer to administer than the others, and administration time is one of several trade-offs that inform test design.

The leaders of citizen-led assessments have thought carefully about fitness-for-purpose: their tests are not meant to be diagnostic or to inform instructional planning. They argue that the tests are meant to provide enough information to spark interest and concern about learning, and in order to operate at low-cost and be feasible in the context of household administration, they provide just enough information to do so.

Validity

Concurrent validity analyses²³ have been conducted using data from studies using both ASER testing tools (Vagh 2013) and Uwezo Kenya testing tools (ACER 2015, as part of this evaluation).

A quasi-experimental concurrent validity study was conducted in Kenya in December 2014 and January 2015 that explored the relationship between performance on Uwezo and performance on the Early Grade Reading Assessment (EGRA) and the Early Grade Math Assessment (EGMA)—instruments for which there is evidence of reliability and validity. Some exploratory tasks were added to the core EGRA and EGMA tests and analyzed separately to examine the extent to which Uwezo tests, which assess only a limited set of foundational skills, may be able to predict children's performance on a wider range of tasks. Details of the methodology and findings of this study can be found in the [full technical report](#) prepared by ACER (ACER 2015).

The sample for this study was a convenience sample, but it did to some extent reflect the diversity of the Uwezo Kenya target population, since children were tested in five districts, each in a different county that together covered four of the five location contexts that Uwezo uses to classify counties, including "arid/semi-arid, core urban, with large cities, and rural-agricultural (east of Rift Valley)."

Children were tested one-on-one in their households. After cleaning and processing, the final dataset contained 1,207 children, evenly distributed across the five districts in which the study was conducted.

In general, the results of the concurrent validity study revealed that the correlation between the Uwezo tests and the core EGRA/EGMA tests is high, indicating that they are measuring the same construct or very similar

²³ Concurrent validity analyses examine the strength of the association between performance on the new tests (ASER and Uwezo tests in this case) and performance on already established standardized tests. Strong and positive correlations between the two tests indicate strong evidence of concurrent validity. (Vagh 2013)

constructs. Lower correlations between Uwezo and exploratory tasks included in the study (which assessed a broader range of foundational skills) suggest that performance on Uwezo should only be interpreted with reference to the limited range of skills that it directly tests. In other words, performance on Uwezo should not be taken to reflect children’s foundational ability if a broader understanding of this ability is adopted.

The key results from this analysis were as follows for each domain:

English reading domain:

- The correlation between performance on Uwezo and core EGRA was high (0.961), indicating that the tests are measuring the same construct or very similar constructs.
- While the correlation between performance on core EGRA and the exploratory EGRA tasks was high (0.952), indicating that these tests are measuring the same, or very similar, constructs, the correlation between performance on Uwezo and the exploratory EGRA tasks was 0.899. This lower correlation indicates that the exploratory tasks are measuring something that is not captured in Uwezo’s measurement construct.
- The reliability (WLE Person separation reliability²⁴) of Uwezo was 0.653 and the reliability of core EGRA was 0.697, indicating that the two tests are able to explain variations in children’s performance to about the same extent.

Kiswahili reading domain:

- The correlation between performance on Uwezo and core EGRA was high (0.977), indicating that the tests are measuring the same construct or very similar constructs.
- While the correlation between performance on core EGRA and the exploratory EGRA tasks was high (0.942), indicating that these tests are measuring the same, or very similar, constructs, the correlation between performance on Uwezo and the exploratory EGRA tasks was 0.900; this lower correlation indicates that the exploratory tasks are measuring something that is not captured in Uwezo’s measurement construct.
- The reliability (WLE Person separation reliability) of Uwezo was 0.353, and the reliability of core EGRA was 0.651. The low reliability of Uwezo indicates that it can only discriminate between children’s level of skill to a limited extent. This is a result of an issue with the targeting of the Uwezo test. The analysis showed that the Uwezo test did not have items of difficulties that spanned the range of abilities of children in the sample,

but rather that the difficulties of the Uwezo test items were all targeted to the lower end of the ability range.

Mathematics domain:

- The correlation between performance on Uwezo and core EGMA was high (0.954), indicating that the tests are measuring the same construct or very similar constructs.
- The correlation between performance on EGMA and the exploratory tasks was 0.879, and between performance on Uwezo and the exploratory tasks was 0.856. These lower correlations suggest that all three tests are measuring slightly different constructs.
- The reliability (WLE Person separation reliability) of Uwezo was < 0.100, and the reliability of core EGMA was 0.870. The very low reliability of Uwezo indicates that it does a poor job of discriminating between the sampled children’s level of skill. This is a result of an issue with the targeting of the Uwezo test. The analysis showed that the Uwezo test did not have items of difficulties that spanned the range of abilities of children in the sample, but rather that the difficulties of the Uwezo test items were all targeted to the lower end of the ability range.

As part of an evaluation of Pratham’s Read India program conducted by the Abdul Latif Jameel Poverty Action Lab (JPAL , ASER Centre, Pratham 2009), children were assessed using the ASER reading and math tests as well as the following tests:

- The Fluency Battery: a test of early reading ability adapted from the Early Grade Reading Assessment (EGRA) (USAID 2009) and the Dynamic Indicators of Basic early Literacy Skills (University of Oregon Center on Teaching and Learning 2002). It is an oral test that assesses the speed and accuracy with which children can read aloud akshars (Hindi letters/syllables), words, non-words, and short passages, as well as answer comprehension questions.
- Read India literacy and math tests: paper-and-pencil tests assessing basic and advanced math and reading and writing ability, respectively. These tests were drawn from extensively piloted Urdu reading and math tests for use in Pakistan (Andrabi, Das, Khwaja, Farooqi, & Zajonc, 2002) and from the math tests of the Trends in International Mathematics and Science Study (TIMSS).

The items for the Fluency Battery were drawn from ASER reading test items, however there was no overlap in the test content between the two. That is, while the same akshars, words, and passages were used, the tasks were not the same in the two tests. Similarly, the Read India math tests

²⁴ Person separation is the name used to refer to reliability in a Rasch modelling context and it can be interpreted the same way as measures of internal consistency such as KR-20 or Cronbach’s alpha (Adams 2005).

included items from ASER math tests. The concurrent validity analyses therefore show the extent to which children’s performance on ASER math and reading tests correlates to performance on tests that assess fluency and comprehension (the Fluency Battery) and more traditional paper-and-pencil tests that assess a greater range of basic and advanced math, reading, and writing skills.

The concurrent validity analysis showed a high correlation (0.90 to 0.94) between the ASER reading test and the Fluency Battery, but a lower correlation between the ASER reading test and the Read India reading test (0.65 to 0.87), which, unlike ASER and the Fluency Battery, assesses not only basic but also advanced reading and writing ability. The ASER math test and the Read India math test showed a moderate to high correlation (0.74 to 0.90).

Comparability

There are three types of comparability that should be addressed with regard to citizen-led assessments:

- Comparability of findings across different test forms in the same language
- Comparability of findings across assessment tools used in different years
- Comparability of findings across assessment tools in different languages

ASER and Uwezo use type token ratio (TTR) analysis to assess the difficulty of a given form, which is the ratio of the number of unique words in a text to the total number of words in the text. But TTR is a rough measure of the difficulty of a passage, because it addresses only vocabulary and does not account for all other aspects of sentence and text complexity. In addition to TTR analysis, Uwezo also uses the Flesch Kincaid Readability Test for its English tests; this test assesses the level of difficulty of a given passage based on word and sentence length. Uwezo uses it to ensure different test versions have similar levels of difficulty (i.e., similar Flesch Kincaid Readability scores).

In some cases citizen-led assessments indicate that making comparisons is not possible or not desired. But in order to facilitate the exploration of factors that lead to high performance, it is important for citizen-led assessments to support the making of comparisons. The first step in doing so is demonstrating that it is valid to make them in the first place, and this is achieved in part by ensuring that the tools themselves are at an acceptable level of linguistic and psychometric equivalence. Steps that could be taken include:

- An investigation of comparability of tests in the same language can be achieved by a combination of expert judgment and equating studies using common tasks (tasks from different instruments put into a single test) or common cases (children doing two tests).
- An investigation of comparability of tests in different languages can be achieved by a combination of (1) review and verification by bilingual experts and (2) inspection of item statistics, which will show whether the relative difficulties of tasks in two language versions are similar (and thus whether the versions are measuring similar skills). While translation and reconciliation from two parallel sources is generally considered best practice when the priority is comparability of the instruments (and the results), other approaches may be considered more appropriate if different goals, such as alignment of an assessment with diverse curricula, are the priority.

Comparability across languages is less important than the comparability across forms and across years, because citizen-led assessments do not explicitly compare findings between languages. That said, it is likely that states’ results in India, for example, or district results in East Africa, will be compared to each other, and if most children in one state took a test in one language and most children in another state took the test in another language, this comparison would not be valid unless equivalence has been determined.

Comparability across years is particularly important if citizen-led assessments aim to play a role in national monitoring efforts to track learning outcomes over time.

Table 6. Test Characteristics Related to Comparability

	Languages	Forms	Development
ASER	20 languages	Reading: 4 forms per language Math: 4 forms	The same tools are administered for 2 consecutive years
Beekunko	11 languages	Reading: 2 forms per language Math: 2 forms	Tools are revised each year
Jàngandoo	French and Arabic	Reading: 3 forms per language Math: 3 forms	Tools are revised each year
Uwezo	English and Kiswahili In Uganda, 4 additional languages	Reading: 4 forms per language Math: 4 forms	New tools are developed each year

Test administration

The reliability of volunteers as test administrators impacts the credibility of citizen-led assessments. This is particularly important because tests are administered orally with volunteers rating each child's performance and then recording his or her score.

Two inter-rater reliability studies have been conducted, one of the ASER testing tools in 2013 (Vagh), and one of the Uwezo testing tools in 2014-2015 as part of this evaluation (ACER 2015).²⁵ The [full technical report](#) of the latter describes the methodology and findings in detail. The latter study explored the agreement in scores assigned by Uwezo volunteers and by an expert rater to children's responses to the Uwezo tasks. Children were recorded on video as the Uwezo testing tools were administered one-on-one. Later, a group of 20 Uwezo volunteers from ten districts and one expert rater independently scored the children's responses.

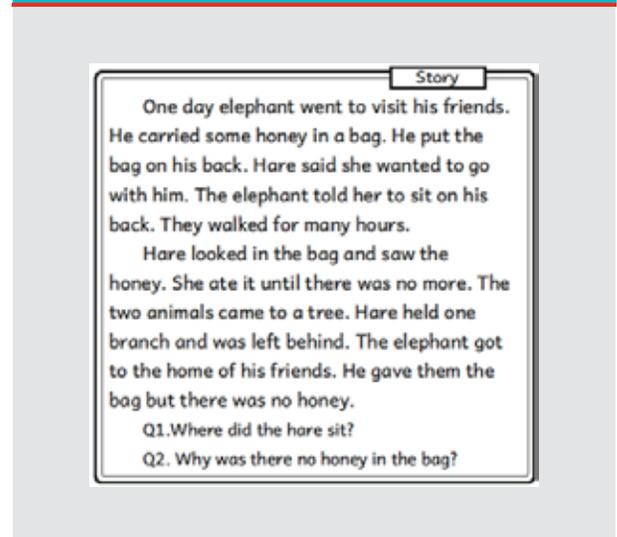
Both inter-rater reliability studies found a high level of agreement in raters' scoring of children's responses.

- The 2013 study of the ASER reading and math testing tools showed "substantial levels of agreement" between repeated test administrations and between raters (Vagh 2013). Agreement was higher for the reading test (0.82) than for the math test (0.79).
- The 2014-2015 study of the Uwezo math and reading testing tools showed high levels of agreement in scores assigned to children's responses both within the Uwezo volunteer group (mean proportions of raters who assigned the same scores to the responses: 0.9412 for English, 0.9617 for Kiswahili, and 0.9561 for Math) and between the Uwezo volunteer group and the expert rater 0.9229 for English, 0.9483 for Kiswahili, and 0.9155 for Math).
 - > In the English reading domain the Story task (read aloud 8-15 sentences) and the second Comprehension task (which requires a minor inference, as opposed to direct recall) were found to be slightly more challenging for volunteers to score reliably than the other tasks (within the volunteer rater group, the mean proportion of raters who assigned the same scores to the responses was 0.8883 for the Story task and 0.8986 for the second Comprehension question, with slightly lower values for the levels of agreement between the volunteers and the expert rater).
 - > In the Kiswahili reading domain the Story task was also found to be more challenging to score reliably, but mean proportions of agreement for scores assigned to responses were still above 0.9.

²⁵ Since the 2013 ASER study and the 2014-2015 Kenya study had different methodological and analytical approaches, reliability values should not be directly compared, but it is reasonable to compare general conclusions about levels of reliability.

²⁶ Each citizen-led assessment has different guidelines for what constitutes successful completion of the reading tasks, including how many mistakes are allowed and how many chances a child is given to read the passage.

Figure 8. Story Task (English)
From Uwezo Reading Test



- These findings were confirmed by observations of survey administration, which revealed that administrators' decisions about a child's competency level did not require subjective judgments and were in most cases agreed upon by other volunteers and observing evaluators. The reading fluency and comprehension tasks were the only ones that were at times difficult for administrators to score.²⁶

What exactly "fluent" reading sounds like, the number of mistakes allowed in a paragraph, and the number of chances given to a child were not always clear to administrators. Assessing fluency is a skill that requires practice, but most volunteers are only able to conduct or observe a handful of children during their training.

- Despite the slightly lower proportion of agreement for the Story and second Comprehension tasks, the inter-rater reliability for these tasks is still relatively high and does not warrant the exclusion of such tasks, which are fundamental measures of early reading competency. Though all proportions of agreement were high, there is room for improvement in scoring reliability for these tasks: offering more comprehensive guidelines and training in scoring them may lead to increased levels of inter-rater reliability.
- The results of the study also revealed that Uwezo volunteers from one location context were no more or less likely than volunteers from another location context to assign scores to responses that disagreed with the scores assigned by the expert rater.

Data analysis and reporting

Citizen-led assessments' raw data was analyzed and ACER was able to replicate the reported findings. Suggested improvements in data cleaning and management, as well as documenting data cleaning processes, have already been taken up by some core teams. Uwezo's publicly-available data cleaning protocol is a good example of documenting and sharing data management processes to increase the credibility of the assessment.

Analyzing children's performance with respect to background factors

The extent to which analyses attempt to link performance and family background varies across the four initiatives.

- Beekunko and Jängandoo conduct frequency analyses on national data disaggregated by indicators such as parents' level of education, displaced status of children, number of meals per day, and household living conditions. These analyses are a commendable effort to investigate how contextual factors might affect children's performance on the tests.
- Uwezo conducts frequency analyses on data disaggregated by socioeconomic status,²⁷ but the relationship between children's performance and background factors is not consistently reported. With respect to the most recent reports (the 2012 survey data), Tanzania conducts frequency analyses of pass rates on data disaggregated by socio-economic status and location (i.e., urban and rural); Uganda conducts frequency analyses of performance on particular tasks on data disaggregated by school type.
- The special "Notes on ASER" section in ASER's annual reports occasionally examine children's performance in the context of their family background.
- Further analysis of the relationships between performance and other factors is enabled through the sharing of the rich datasets generated by citizen-led assessments. Data is made publicly available in the cases of ASER (upon request) and Uwezo (users can download full datasets from its website).

It is important to note that such analysis and reporting requires sophisticated statistical methods to ensure validity; this is why ASER, for example, presents in-depth analytics in separate reports rather than as part of the annual survey report. Even if citizen-led assessments choose

not to report on the relationships between children's performance and background factors, care should be taken when reporting results below the national level to include contextual information alongside the assessment data. All citizen-led assessments report results below the national level (e.g., district, region, *commune*), and this invites comparisons that may be misleading when there are differences in context across the different districts, regions, or *communes*, such as socioeconomic status, which is known to have a relationship with children's performance on learning assessments.

Trend analysis (comparisons across years)

While Beekunko and Jängandoo have not yet conducted trend analyses (comparing learning data over time), ASER and Uwezo have. Their reports caution that observed changes may be due to sampling error or changes in household socioeconomic status, but there is no analysis to confirm that observed changes in performance over time may not also be due to issues relating to comparability of assessment tools.

How performance is reported

Across all four grantee assessments, frequency analyses are conducted which determine the percentages of children successfully completing different tasks in the assessments (e.g., letter level, word level, story level, etc.). Jängandoo and Uwezo conduct frequency analyses for children's performance on the assessments overall in order to report a "pass" rate. Beekunko also calculates average scores in the reading and mathematics assessments for the entire dataset and then disaggregates the data to provide average scores by region, *commune* and community.

It is important to consider the value of repeatedly reporting failure or pass rates when the results are so overwhelmingly negative. Focusing on reporting results with respect to children's specific competencies (e.g., percent of grade 3 students who can read words) as opposed to pass or failure rates allows interested stakeholders and the general public to gain a more informative picture of where children are in their learning, and how contextual factors influence where they are in this learning. It can in theory then support evidence-based interventions at all levels of the education system.

²⁷ The socioeconomic status variable used by Uwezo is obtained with reference to durable assets owned, access to electricity and/or clean water, and mother's formal education level. Based on these indicators, each tested child is categorized into one of three groups: ultra-poor, poor, or non-poor.

How well do citizen-led assessment processes work?

The process evaluation was carried out by R4D in partnership with in-country partners. It assessed the design and implementation of activities conducted by ASER, Beekunko, Jàngandoo, and Uwezo. The process evaluation addressed six key areas:

- Suitability and effectiveness of key players in the survey process
- Training of trainers and volunteers
- Execution of household sampling
- Survey administration
- Quality assurance
- Dissemination and engagement

KEY FINDINGS

The sheer scale of the surveys that are orchestrated each year is commendable, especially considering that such a feat requires mobilizing hundreds of organizations and thousands of volunteers for the cause of measuring learning.

Volunteers are key to the citizen led assessment model.

The evaluation found that volunteers carried out the tasks assigned to them effectively, with the possible exception of household sampling which needs further testing.

Partner organizations are valuable assets, but good ones are hard to find and retain.

Systematic partnership models like partnering with DIETs in India, *Centres d'Animation Pédagogique* in Mali, or a single strong NGO across many districts are promising in that they can be scaled up through existing structures.

Training related to test administration was very strong, but in most cases should include more field practice that mimics volunteers' responsibilities during the actual survey.

Trainers and volunteers were well-equipped to conduct the community survey, household survey, and administer the tests to children. With the exception of Jàngandoo, whose training lasts five days, training time was insufficient. There are of course significant resources implications to increasing the length of training; in all cases citizen-led assessments indicated that training time is one of the key areas affected by resources constraints.

It is important for the quality and credibility of the survey that each initiative develop a systematic approach to quality assurance. Good examples of this are ASER's and Uwezo's quality assurance mechanisms which involve extensive monitoring and re-check activities, including field re-checks for a subset of villages and districts.

Citizen-led assessments have been very successful in using the media to disseminate survey findings, particularly at the national level.

Including key stakeholders at the national level as advisors in the survey design process increases institutional buy-in. Beekunko and Jàngandoo are currently experimenting with pre-survey engagement at lower levels as well (regional, *commune*, and community). This may be a powerful strategy for building support among key constituents (including potential critics); further testing is needed to fully understand the impact of such engagement.

Key players in the survey process

Volunteers are a fundamental part of the citizen-led assessment. Though their use as data collectors has sometimes been used to question the credibility of the assessments, evidence does not support these claims. As described above, inter-rater reliability studies find high levels of agreement in volunteer raters' scoring of children's responses, including when compared with an expert rater (ACER 2015, Vagh 2013). The observations of data collection provided no indication that volunteers lacked the competencies required to perform this task, though no formal comparison was done with paid test administrators. Household sampling conducted by volunteers proved more challenging.

Partner organizations play a critical role in the citizen-led assessment model. All four of the evaluated initiatives rely heavily on these organizations to help them implement the survey. ASER's and Beekunko's partnerships with teacher training institutions present the opportunity for influencing cadres of future teachers, which could be potentially quite impactful. As non-government initiatives, it is important that citizen-led assessments maintain their independence from government, so care should be taken to ensure that partnering with public teacher training institutions does not compromise this status. There is no evidence to date that engagement with pre-service teachers would threaten the initiatives' credibility as independent organizations.

Conducting the survey through partner organizations presents great challenges and also great potential. The task of managing the financing and logistics of hundreds of partner organizations (for ASER and Uwezo) is an enormous undertaking, in addition to the effort that is required to identify and recruit them. Staff of ASER, Beekunko, and Uwezo indicate that recruiting partner organizations and master trainers takes an immense amount of time and is one of the most significant challenges of survey implementation. While many partner organizations and master trainers return for two or more years, replacing the significant number that do not is a time-intensive process. In some states in India and depending on the year, in Kenya, Tanzania, and Uganda, up to half of the district partners need to be replaced every year.

The capacity of master trainers and partner organizations varies greatly. Jàngandoo and Beekunko in particular have struggled to recruit qualified master trainers. With regard to partner organizations, many are small grassroots agencies and struggle to assemble the required number of volunteers.

For ASER and Uwezo, given the effort that is put into identifying new partner organizations, master trainers, and volunteers (through the partners), it is surprising that the network built each year is not utilized in a role that extends beyond the survey administration itself, especially given the partners' location and ongoing engagement in the surveyed communities. The resources required to do this would be significant, given the number of partners in each country. But given that the network is already in place, the efficiency of engaging could be substantial.

Training

Overall, observations of training sessions at multiple levels indicate that they are generally quite strong. Strengths include:

- Comprehensive training manuals have been developed to guide training and serve as a reference for trainers and volunteers in the field.
- Support from the core team of each initiative was strong. For example, ASER national staff attended state-level trainings to support the state teams that were facilitating the training of master trainers; ASER state staff then supported district-level training in the same way. Jàngandoo team members help run all of the seven regional trainings in collaboration with their NGO partners. These practices help ensure consistency and quality control of training.

- The level of engagement of both trainers and volunteers was impressive; both groups showed enthusiasm for the survey training content and processes and put in long hours to get the most value from the limited training time.
- Debriefing between trainers and volunteers during and after field practice sessions was a positive practice observed in the cases of both ASER and Jàngandoo; it appeared to lead to rapid improvement in individuals' survey administration ability.

Challenges associated with training that multiple groups had in common included:

- Most importantly, insufficient time. In almost every country (with the exception of Senegal), nearly every master trainer interviewed indicated that the volunteer trainings (typically 2-3 days) were not long enough to effectively prepare volunteers to conduct the survey. The evaluation team's observations confirm this assertion. Particularly given the importance of field practice sessions, which take significant time, it is critical that the volunteer trainings allow for thorough classroom sessions, at least one field practice session, and debrief and additional practice after the field work. The classroom sessions must cover not only test administration but sampling (perhaps the most complicated component), the village survey, school survey, and household survey as well. The current amount of time allotted to covering these topics in the classroom and field sessions is not sufficient.



ASER National Training in Bihar

- In many cases, limited time led to the reduction of field practice sessions. While they are time intensive, field sessions are critical to volunteers' ability to confidently and effectively conduct the survey. In many field sessions, a group of volunteers went to a community and split up to complete the various tasks: talking to the community leader, mapping the community, conducting the school survey, etc. The result is that many volunteers left the field practice session with no experience having implemented one or more parts of the survey process.
- The majority of the training is focused on the technical aspects of the survey (sampling and test administration). Limited time is spent instructing participants on how to introduce the citizen-led assessment in the communities and households that are surveyed. Observations of field practice sessions and the survey itself indicate that more focus needs to be placed on this task, including each individual practicing the introduction, as trainers and volunteers varied widely in their ability to clearly articulate the purpose of the survey.

Household sampling

In the cases of ASER, Beekunko, and Jàngandoo, several challenges to volunteer-led household sampling were observed during site visits. One commonly faced challenge by volunteers was difficulty in accurately mapping the village or neighborhood (ASER and Beekunko), or difficulties in reading and de-coding the maps provided to them (Jàngandoo). Given the discrepancies observed across the initiatives in the ways the volunteers mapped the village or interpreted a given map, further research is needed to determine the extent to which the volunteer-led sampling technique compromises the randomization of the household sample.

In the case of Jàngandoo and Beekunko, volunteers faced challenges in correctly skipping the right number of households or buildings (*concessions*) each time. This was not observed to pose issues to the ASER volunteers who similarly were instructed to implement the "left-hand rule" (whereby volunteers start in the middle of a village quadrant and select every fifth household on their left-hand side). Part of the explanation for this might be that in ASER's case, the rule is kept very simple, is the same in all contexts, and is given major focus during the trainings. In the case of Jàngandoo and Beekunko, on the other hand, the rules are more complex and differ based on whether volunteers are surveying in a rural area or in an urban area. To the extent that household sampling is left in the hands of volunteers, the degree to which the methodology can be simplified is likely to increase volunteers' ability to implement it reliably.

Linked to this is the degree to which volunteers are prepared for implementing the household sampling step through trainings and volunteer materials. For example, as noted above, ASER spends considerable time teaching and emphasizing the left-hand rule during its trainings, and clear instructions are provided in the volunteer training materials. Perhaps most significantly, the field practice component of ASER's trainings requires volunteers to not only practice administering the household survey, but also village mapping and the implementation of the left-hand rule. Beekunko and Jàngandoo on the other hand do not provide an opportunity for volunteers to systematically practice the household sampling step during the field practice components of their trainings.

While ASER and Beekunko provide clear written documentation on the household sampling steps in the volunteer takeaway materials, Jàngandoo's otherwise comprehensive volunteer manual omits this component. Given the various challenges observed in implementing this step, all three assessments that require volunteers to conduct household sampling should develop and include clear written and pictorial instructions for this step to include in the takeaway training materials provided to volunteers.

Uwezo's household sampling methodology is in theory more robust than the other citizen-led assessments' because a complete household listing is generated before households are sampled. Nevertheless, there were challenges in practice. For example, volunteers sometimes had difficulty finding the households in the sample, particularly in urban areas. Additionally, while District Coordinators interviewed as part of the process evaluation indicated that household sampling was not a challenge for them, according to Uwezo's interal monitoring from 2013, District Coordinators' mastery of the sampling methodology was one of areas where mastery was lowest (Uwezo Assessment Monitoring Report 2013).

Further research is needed to fully understand the trade-offs between volunteer-led sampling without a complete household listing and Uwezo's approach, both of which were found to face significant implementation challenges during the sampling process as described above.

Test administration

While each citizen-led assessment had its own challenges with respect to survey administration, a significant challenge across the board was that children were often surrounded by a crowd while being tested. Even when administrators made efforts to limit distractions (as instructed during training sessions), the testing conditions were not at all standardized. The testing environment varied by child: some are indoors, some outside, mostly sitting but sometimes standing, noise level varied, light varied, whether or not the parent was there varied. Many but not all children were tested while a large group of their peers, family members, and neighbors crowded around them. All of these factors compromise the reliability of the results because testing conditions are not standardized.

Jàngandoo used tablets to collect data for the first time during the 2014 survey. While there were technical issues, as would be expected in a first year of implementation, overall the evaluation team finds the introduction of tablets stands to significantly improve the efficiency and quality of data collection and transmission. The potential for improved data

quality is significant when using a computerized interface because invalid responses are not accepted and the survey questions automatically adapt based on answers to previous questions to avoid contradictory entries. An open-source interface has already been developed for EGRA that could be adapted to citizen-led assessment testing tools. All of this said, the cost of using tablets is significant, and there are other areas where resources should be allocated before electronic data collection, such as deeper dissemination and engagement, as will be discussed.

Quality assurance

Each initiative has some sort of quality assurance system in place, but descriptions of data collection monitoring processes are only publicly available for ASER and Uwezo. In general, databases do not indicate which entries have been part of the quality assurance process.

Dissemination and engagement

National level

Including key stakeholders at the national level as advisors in the survey design process has shown to later increase buy-in of those stakeholders' institutions when presented with survey findings. This can be a powerful strategy for building support among key constituents (including potential critics).

Key informant interviews from the evaluation of impact revealed that the media played a significant role in contributing to broad public awareness about citizen-led assessments and their findings, and that ASER and Uwezo in particular have gained high levels of visibility. For both ASER and Uwezo, this visibility is driven by a media push that gets the survey results into newspapers, on television, and on the radio—though print media was the most-often cited source by key informants. In 2013, over 150 articles on the ASER survey findings were published in print media, five national television broadcasts, and over 20 online articles (ASER Centre 2013). Also in 2013, Uwezo was cited in print and broadcast media 25 times in Tanzania, 47 times in Kenya, and 46 times in Uganda (Twaweza 2013). Jàngandoo and Beekunko, while still at much more nascent stages of their national-level dissemination strategies, have shown promise in terms of leveraging similar channels for disseminating information about the assessments and their results.



Jàngandoo: One volunteer tests the child while another collects the data on a tablet.

Sub-national level

In the case of ASER and Uwezo, the media is also an important channel for dissemination at the sub-national level (state and district) including through print media, radio, and TV (and often in local languages). While Jàngandoo and Beekunko also make use of media at the sub-national level, findings from their respective process evaluations suggest that these channels could be leveraged even more significantly, particularly through the use of local radio.

Beekunko and Jàngandoo conduct sub-national dissemination activities in a much more systematic and centrally organized way than ASER and Uwezo; this is likely facilitated by their significantly smaller size. That said, if Beekunko's and Jàngandoo's lower-level dissemination strategies prove to be effective after initial testing, they can still provide insight to the larger initiatives.

By partnering with NGO organizations that are already embedded within communities, ASER and Jàngandoo are able to gain significant economies of scale in terms of their ability to conduct outreach and dissemination activities. ASER is able to leverage its relationship with Pratham to facilitate dissemination at the state and district levels. This is one advantage that ASER has in that it is tied to a large-scale implementing network that operates throughout the country. Jàngandoo is developing a similar operating structure through its partnerships with NGO organizations in each of the 14 regions of Senegal. Nonetheless, several district partner organizations of both ASER and Uwezo that were interviewed as part of the evaluation of impact indicated that they wished resources were available to include dissemination in their role.²⁸

Community and household level

In the observed communities in Kenya, Tanzania, and Uganda, the instant feedback process was not implemented consistently. Some parents were simply told if their child could or could not read, while others were told nothing at all. This is consistent with findings from a previous study conducted by Lieberman et al. (2014) which found that a majority of parent respondents whose children took part in the 2013 Uwezo assessment claim that Uwezo "did not provide them with any information about their child's performance." In observed communities in Senegal, Jàngandoo volunteers were effective at providing concrete suggestions to parents for ways in which to increase their engagement and support of their children's learning. In Mali,

this was observed as an area of improvement for volunteers during the instant feedback process. Existing literature on the role of information in generating action shows that, in theory, providing concrete suggestions is critical to generating action and that the provision of information alone is limited in terms of the ultimate impact it can engender (Pandey et al. 2008). That said, the evaluation findings on the difficulty of implementing instant feedback in practice indicates that new approaches to training and implementation may be needed in order for this activity to result in the intended outcome of increased awareness and action by parents.



Uwezo: Volunteer in Tanzania giving instant feedback to a parent.

For ASER, the lack of instant feedback at the household level is not a weakness in execution; during the trainings volunteers are not instructed to share the results of the reading and math assessment with parents. The ASER Centre explained that this was due to concerns over the parent blaming or punishing the child if he or she performed below expectations. It also requires significant time to train trainers and volunteers in how to conduct this sort of information sharing effectively—as demonstrated by the difficulty of consistent implementation in other countries. Parents sometimes do observe the assessment and are therefore able to witness their child's performance, and volunteers tell parents what the ASER survey is, but do not share results or suggestions for what parents can do to improve learning.

²⁸ The ASER Centre recently received funding for 2014 district partners to be visited during 2015 by the national and state ASER Centre teams to both disseminate results and discuss possible future collaborations.

At first, this appears to be a lost opportunity to increase awareness of learning outcomes at the village level, especially as one of the aims of the ASER survey is to stimulate awareness and action at the grassroots level. However, Uwezo has tried to provide instant feedback during the survey process and randomized studies have shown that it is largely ineffective at influencing parents' awareness about learning outcomes (Lieberman et al. 2012). The evaluation finds, therefore, that the impact of instant feedback is not at all clear, and much more testing is needed of ways to engage at the community level before, during, and after the survey. It is very difficult to execute this sort of feedback through volunteers with limited training, and even if execution can be improved, there is no evidence that a single visit can stimulate the kind of action that is intended. An example of a variation that could be tested is an intervention in which information and suggestions are provided repeatedly over a period of time (including materials not unlike those that Uwezo has delivered), which has proven effective in raising learning outcomes (see Pandey et al. 2008).

While it is important that information shared is not presented as statistically representative at the village level, sharing findings of previous year's national or regional results, for example, could potentially increase the community's awareness of low learning outcomes. Findings from Beekunko's process evaluation suggest that its day-of-survey community-level meetings were well-conducted and well-received, and achieved positive buy-in and support from local authorities for the initiative. This is an example of an attempt to generate collective action in addition to individual action at the community level. ASER, Beekunko, and Uwezo also share the testing tools and a letter explaining the previous year's findings with both the village leader and the head teacher of the surveyed school in each village.

While none of the citizen-led assessments conduct post-survey dissemination at the community level systematically, examples of very small scale dissemination activities exist, where volunteers, district partners, or even staff of core teams will go to communities and conduct local engagement, but often such activities are sporadic and not formally organized or structured. Examples from India include the following:

- In Maharashtra, the state team uses some of its budget to create pamphlets containing the state level survey findings. They distribute these pamphlets to their district partners and encourage them to disseminate the pamphlets at the grassroots level.
- In Tamil Nadu, where AID India manages the ASER survey statewide, village-level dissemination occurs through AID India's network of partner NGOs. Activities include role plays and evening tuition centers.
- In Karnataka and Rajasthan, ASER Centre state team members did community level engagement activities such as developing "village report cards" with ten and four villages, respectively.

Overall, core teams across the four initiatives do not have the manpower needed to constitute systematic dissemination of findings at the community level. Unfortunately, the survey infrastructure which deploys volunteers to communities is not capitalized on to use the community/volunteer pairings for subsequent community-level dissemination.

While engagement at the community level is desirable, post-survey engagement at this level is also expensive. This suggests that engagement with community-level actors and parents at the time of the assessment and through existing partner organization/volunteer/community pairings should be strongly considered across all four initiatives, given the marginal cost this would incur. Critically, additional and different types of training focused on stimulating community action would be required for such engagement to be effective.

Costs of dissemination

It is important to consider the financial trade-offs of conducting dissemination activities at various levels: dissemination and engagement at the community level, for example, will inevitably require significantly more resources than at higher levels simply because the number of units is so high. That said, the evaluation finds that small-scale, rigorous testing of dissemination and engagement at all levels is necessary in order for citizen-led assessments to make informed decisions about these trade-offs. If specific forms of lower-level engagement prove to be effective at generating awareness and action, they should be pursued and more resources should be sought to make them feasible, even if initially only at small scale.

How well do citizen-led assessments stimulate awareness and action?

Summary findings

This summary describes the impact that ASER and Uwezo have had on awareness and action at the international, national, state, district, and community levels.

International

Both ASER and, more recently, Uwezo, have contributed to an increased focus on learning outcomes in global discourse and agenda-setting. Their contribution has included both providing evidence of the seriousness of the learning crisis (i.e., revealing major deficiencies in even the most basic competencies) and demonstrating how a low-resource model can be used to assess learning on a national scale. ASER and Uwezo did not single-handedly cause this shift toward learning outcomes, but they did contribute to it by raising the visibility of the crisis by quantifying it in very simple, stark terms.

National assessments have for many years revealed low learning levels in many countries, but very few assess children in the early grades or use a population-based sample to ensure that out-of-school and rarely-in-school children are included. The lack of even the most basic reading and math skills that citizen-led assessments reveal is in some ways more powerful in stimulating debate than similarly dismal results of more formal assessments of older children because it reflects such widespread failure of the system to deliver even the most basic education. The evaluation finds that ASER and Uwezo have been quite successful at generating awareness of low learning levels at the international level.

National

Increasing awareness of the learning crisis at the national level is one of the main successes of both ASER and Uwezo. Success in generating awareness at the national level has been achieved largely by a media-centric approach to dissemination of results. Each year's survey results are prominently featured in multiple languages in print media and, to a lesser extent, in radio and television broadcasts as well.

ASER and Uwezo have also stimulated debate at the national level. While national-level stakeholders reported to have already been generally aware of low learning levels in their countries, they indicated that ASER and Uwezo have made low learning levels highly visible in the public sphere.

There is some evidence that ASER and Uwezo have contributed to the prioritization of learning in national education policy documents. For example, India's 2012-2017 national planning document and 2014 resource allocations indicate an increased focus on measuring and improving learning. Similarly, Uwezo results have been cited in government reports and strategy documents, noted as supporting evidence for renewed government focus on learning outcomes.

Generating concrete action to improve learning outcomes has proven challenging. Sporadic but powerful examples of direct action do exist (e.g., programs focused on measuring learning at the state level in India), and they can provide insight into the types of action that are possible and what factors support such action.

State

The level of impact that ASER has had at the state level varies by state, but in the four states included in this evaluation, there was evidence that ASER contributed to stimulating debate, engagement, and in some cases direct action for the improvement of learning outcomes. The potential for ASER's impact at the state level was found to be greater than at the national level, largely because education implementation and practice are state-level responsibilities. Examples of state-level action to which ASER may have contributed include the development of state-wide learning assessments and the use of ASER survey results by officials to identify weak areas (both geographic and issue-based) around which state programs can be designed.

District

Both ASER and Uwezo aim to influence two key groups at the district level: their district partner organizations and district-level government officials. For both initiatives, only sporadic evidence of impact at the district level was found. This can largely be attributed to the lack of resources available for systematic involvement of the network of district partner organizations in dissemination activities, and, relatedly, to the limited capacity of these organizations.

Outside of the government, ASER and Uwezo have triggered the uptake of the testing tools in education programs run by NGOs and CSOs. The evaluation identified dozens of these groups use the testing tools in various ways including initial assessment of children's learning levels, tracking children's progress over time, and monitoring the impact of their work as an organization. As these tools are easy to use and adapt, they can be used by a wide range of organizations with varying capacities.

Community

The evaluation uncovered only limited anecdotal evidence that participation in the survey stimulates awareness or action at the community level. Much more testing is needed of ways to engage at the community level before, during, and after the survey if citizen-led assessments aim to close the feedback loop between data collection and inciting action.

KEY LESSONS FROM SPORADIC EXAMPLES OF SUCCESS

From the set of success stories where action was generated, two key lessons emerge that shed light on what is needed, in addition to the provision of information, to generate action:

- Those few who are motivated to act by information alone, or who have already demonstrated commitment to the issue of learning outcomes, are critical partners. While limited in number, these champions can be very powerful.
- In order for action to take place, there needed to be an entity available to provide suggestions for what those actions might look like. In India, Pratham often played this role by collaborating with government officials to design and implement interventions, but other NGOs, CSOs, and agencies contributed as well to augment the capacity of a champion that wanted to act, but did not know what the first step might look like. In Kenya, Tanzania, and Uganda, the lack of an obvious partner or set of partners to play this role limited Uwezo's ability to generate action.

These lessons align with the recently emerging school of thought in the transparency and accountability space that strategic engagement with government, and particularly with champions within the government, as opposed to (or in combination with) a whistle-blowing approach, may be most effective at achieving common goals of governments and civil society (Fox 2014). But these lessons do not just apply to influence of government officials; they also apply to district partners, CSOs, NGOs, community leaders, and community members. Champions exist within these groups, and if linked with each other and provided with suggestions and support for actions they can lead, they can be powerful forces in bringing about broader action. If more systematically encouraged, the impact of currently small-scale actions (such as the chain of libraries started by a district partner in Bihar, the incorporation of ASER's and Uwezo's testing tools into NGOs' education programs, or the letter from the community leader in Kenya to his District Education Officer demanding an overdue election for a school committee) could be exponentially increased.

Contextual considerations for the success of citizen-led assessments

The evaluation sheds light on the contexts in which citizen-led assessments are likely to have the most traction. Broadly defined, these contexts are ones in which target audiences have autonomy to make decisions about policy and practice. This manifests in different ways at different levels. For example, in India, where state officials have significant control over policy and district officials over planning, citizen-led assessments are likely to have most traction at those levels. In countries whose political systems instead of political economies, the national level is a critical target audience. Targeting district officials in such a context is less likely to lead to systematic reform.

With regard to the community level, citizen-led assessments should, in theory, be able to gain traction in various contexts, as long as the type of information provided and the engagement surrounding that information is designed to educate, instruct, and empower citizens. In a decentralized context, engagement might be designed to generate collective action to put pressure on local policymakers and service providers. Even in a centralized context, individual or collective action at the community level could impact the performance of service providers. In practice, the amount of traction is dependent not just on contextual factors, but on how citizen-led assessments design their engagement activities to respond to those contextual factors appropriately.

Can assessments be expected to change learning outcomes?

One of the goals of this evaluation was to calibrate expectations of the type of impacts that citizen-led assessments can achieve. One of the ultimate intended outcomes of these initiatives is *improved learning outcomes*: ASER and Uwezo survey data clearly shows that this outcome has not been achieved. An analysis of the ASER survey data between the years 2006 and 2012 shows that the number of states showing a declining trend in learning outcomes is increasing while the number of states showing improvement is declining. Uwezo is also unlikely to achieve its goal of increased learning outcomes; they aim for a ten percent increase in the literacy and numeracy of 6-16 year olds in Kenya, Tanzania, and Uganda after four years of the Uwezo assessment — by 2015. Attainment of this goal seems unlikely: Uwezo's own findings indicate that learning outcomes have remained largely consistent across the three countries over the last three years of surveying.

It may not be reasonable to expect that learning outcomes would go up, or stop declining, as a result of actions motivated by citizen-led assessment results. Even flawless implementation of a sample-based survey and related engagement activities, even at the large scale that the ASER and Uwezo initiatives have achieved and should be commended for, leaves hundreds of thousands of communities (the places where education actually occurs) untouched. This is not a criticism of the citizen-led assessment model, but a check on the theory that such a model could cause a significant shift in national learning outcomes.

To fully understand the potential for impact of citizen-led assessments, it is important to frame the broader realm of interventions to which they belong: information-for-accountability interventions. While citizen-led assessments are unique from many such interventions in that they involve citizens not only as consumers of information but in the collection of information, many of the findings are

relevant to their potential for impact. What follows is a brief summary of lessons to be drawn from existing literature on these types of interventions, and their ability to foster action through increased awareness.²⁹ These findings on the existing body of evidence relating to information-for-accountability interventions supplemented the evaluation of the four citizen-led assessments and inform the findings and recommendations described in this report.

²⁹ These are excerpted from a literature review R4D conducted which investigates the impact of interventions where information is leveraged to promote citizen, government, and/or service provider action with the ultimate goal of improving learning outcomes. Citizen-led assessments are of course distinct from many learning-focused information-for-accountability interventions in that they not only provide citizens with information, but also involve citizens in the collection of information. A summary of the main findings of the literature review is provided here; the full review is also available separately in the Annexes to this report.

KEY FINDINGS FROM LITERATURE ON INFORMATION-FOR-ACCOUNTABILITY INTERVENTIONS

1) Information dissemination alone is often insufficient to generate sustained and meaningful improvements in learning outcomes, or even to bring about action meant to improve learning outcomes.

Information-based initiatives seeking to foster improvements in learning outcomes have often proven more effective when implemented alongside a direct intervention. An implementation of the EGRA Plus initiative in Liberia, for example, illustrates the value of providing more than assessment results to parents and teachers in order to have an impact on learning outcomes. The intervention consisted of the administration of EGRA in three groups of schools through three distinct variations: in the first (control) group, results from the assessment were not publicized or disseminated; in the second group of schools, referred to as the “light” intervention, results of the assessment were disseminated to parents through school report cards and student reading report cards prepared by teachers; in the third group of schools, known as the “full” intervention, teachers underwent an intensive training in reading instructional methodologies, in addition to the administration of EGRA and the dissemination of results described in the “light” intervention. An evaluation of the interventions (Piper and Korda 2010) found that the “light” intervention (with information dissemination only and no teacher training) increased student outcomes in just one out of seven EGRA performance categories. Meanwhile, the more involved “full” intervention improved student performance along all seven of the indicators measured.

This intervention has important lessons that can be closely applied to citizen-led assessments. The divergent outcomes of the “light” and “full” intervention reflect the added positive impact of combining information dissemination with a capacity-building intervention such as teacher training. The takeaway for citizen-led assessments is not that they should change their approach to include such an intervention, but that without stimulating such direct action, they are unlikely through information dissemination alone to engender a meaningful improvement in learning outcomes.

Successful interventions conducted with a target audience of policymakers (from national to local authorities) also tend to provide more than information alone. In Tanzania, a civil society group named HakiElimu engaged in advocacy about education policy and budgeting with positive outcomes. HakiElimu’s work not only involved public awareness campaigns, but also directly trained local government leaders, legislators, and journalists, in addition to engaging in high-level lobbying and advocacy. Its impact suggests that increased information plus capacity building support can have a direct impact on policymakers’ decisions, regardless of whether large-scale collective citizen action at the grassroots level takes place.

The effectiveness of interventions that combine information provision to citizens with a coordinated push for increased public sector responsiveness is consistent with prominent thinking in the broader social accountability literature. Described by Jonathan Fox (2014) as “strategic approaches,” these types of multi-pronged initiatives have shown more promising results than tactical (information-only) approaches that rely heavily on assumptions about the role information alone can play in motivating citizen-based action to influence public sector performance.

2) Citizen and policymaker action relies on a critical set of assumptions

Recent research has revealed the weaknesses of certain assumptions about the links in the information-to-action chain with respect to information-based interventions. An evaluation of Uwezo’s impact on levels of citizen action in households and villages in Kenya by Lieberman, Posner, and Tsai (2012) finds that, in order for the information-to-action link to come to fruition, those who receive information must not only understand it, see it as actionable, and care about the topic at hand, they must also believe that their action will lead to improved outcomes. Unfortunately, a seamless pathway between information, accountability, and citizen empowerment is in reality difficult to achieve, particularly in contexts characterized by “entrenched relationships, power structures, and institutions” (Harvard Kennedy School and Results for Development Institute 2012).

For accountability to be successful, citizens must feel that the cost to monitoring and accountability is low and that its value is high; similarly, policymakers must feel that there is low cost and high value to changing their behavior (or conversely, a high cost to not changing their behavior) (Lieberman et al. 2012). If government policymakers feel that there will be no negative repercussions for failing to act, they are unlikely to implement reform or be influenced by citizen attempts to advocate for change.

3) Clearly presented, accessible channels for action are crucial to the success of information-based interventions

Strategic approaches to social accountability, as described by Fox (2014) should “focus on information that is clearly perceived by users as actionable.” Interventions that provide users with a direct pathway to action tend to be more effective than those that do not. A Village Information Campaign conducted in three Indian states provides a useful illustration: by delivering communications materials to each village on a repeated basis (including a short film, a set of posters, wall paintings, take-home calendars, and learning assessment booklets) the Campaign sought to generate increased awareness about the roles and responsibilities of pre-existing, but dormant, school oversight committees, as well as each community’s duty to engage with them. Pandey et al. (2008) found that this information campaign not only increased levels of awareness of citizens’ roles and responsibilities related to the committees, but also had a measurable impact on learning outcomes in all three states where the intervention took place. Two states experienced a positive impact on reading levels (14-27%) in one of the three grades tested; the third state saw an improvement in writing outcomes in one grade (15%) and in mathematics results in the other grade. Levels of teacher effort also increased in two of the three states.

In contrast to citizen-led assessments, this intervention did not publicize information on learning outcomes, but rather on mechanisms through which parents could become involved in their children’s education and hold schools accountable. Thus the information campaign directly armed parents with the knowledge necessary to take action and participate. In this case parents did not necessarily require information on education outcomes or poor learning levels in order to take action—instead, they required instructions or suggestions on how to take action and what concrete steps would look like. While this example does not provide evidence against sharing data on learning outcomes, it demonstrates that this information should be accompanied by actionable suggestions and tools.

Of great relevance to the citizen-led assessment model, Fox (2014) notes that promising information-for-accountability interventions often go beyond providing information at the citizen-level by attempting to “bolster governmental capacity to respond” to voiced concerns from those same citizens. By presenting public service providers with concrete, actionable tools or guidelines on how to best enact changes and take action to remediate poor service provision, information-led interventions can build pressure for change from the grassroots level as well as policy levels, and attempt to create what Fox describes as “state-society synergy.”

Definitions of awareness and action used in this analysis

Increased awareness of learning outcomes is one of the key pillars of citizen-led assessments' theory of change: by providing key education stakeholders with quality information regarding the state of learning within their country, district, and community, these initiatives hope to incite broad-ranging action targeted at improving literacy and numeracy skills. **Awareness of learning outcomes is a necessary prerequisite to action. Action, then, is required for the concern generated by increased awareness to translate through various mechanisms into improvements in learning outcomes.**

The evaluation team has deliberately adopted a broad definition of **action** as any shift in behavior, policy, or practice by an education stakeholder related to the improvement of learning outcomes. To better understand the range of potential actions contributed to by ASER and Uwezo, it is useful to think of a spectrum along which actions are organized according to their potential for impact on learning outcomes (the ultimate goal of citizen-led assessments) (Figure 9). While actions with potential for direct impact on learning exist at the far right of the spectrum, there are important pre-actions to direct action (such as commitments to take direct action) that also exist. It is important to capture these pre-actions and commitments to action in order to understand what influence ASER and Uwezo have had, even when that influence does not manifest in an overt intervention to address learning outcomes. ASER's and Uwezo's contributions to actions all along this spectrum are acknowledged and described throughout this evaluation. The ultimate aim of improved learning outcomes, however, is not likely to be achieved without actions on the far right of the spectrum.

Different stakeholders have varying capacities to engage in direct action. For example, a parent arguably faces fewer barriers to successfully engaging in direct action to improve a child's learning outcomes (e.g., visiting a child's teacher to check on the child's progress) than a state-level policymaker, who can mandate a learning-focused intervention, but relies on lower level actors (district officials, school heads, teachers, etc.) for implementation. Further, the political economy of a country to some extent determines the limits of stakeholders' actions. For example, in India, national-level officials have limited capacity to take direct action in the form of learning interventions, as education planning is not within their purview. Therefore, the allocation of funding to a learning-focused initiative, for example, might be among the most direct types of action possible at the national level, even though this allocation is only an early step that in and of itself will not improve children's literacy and numeracy. The expectations for action at each level need to be calibrated to the tools each actor has at his or her disposal.

International level

Awareness

While concern about learning outcomes existed in the international education space before the ASER survey, and, later, Uwezo were initiated, the data they generate has enabled the quantification of children's competencies and has communicated them in a non-technical format which has allowed for widespread international awareness about just how low learning levels are. The survey findings of both ASER and Uwezo are cited widely in global discourse, and although the ASER survey (initiated in 2004) was not the first initiative to suggest that learning outcomes in developing countries were low, it and Uwezo are unique in

Figure 9. Action Spectrum for Citizen-led Assessments



their national coverage, large scale, easy-to-comprehend findings, and coverage of children across a range of ages. Beekunko and Jàngandoo are emerging as additional examples of large scale national data on learning.

Action

Well before citizen-led assessments came about, the international education community was concerned with learning outcomes. Both the 1990 World Declaration on Education for All (EFA) made in Jomtien, Thailand and its update in 2000 in the Dakar Framework for Action included the prioritization of learning outcomes (UNESCO 1990, UNESCO 2000). During the same period, in its 1995 Priorities and Strategies for Education, the World Bank prioritized supporting institutional capacity development for measuring learning outcomes (World Bank 1995). But this stated importance of learning outcomes did not result in significant action at the international level in the early 2000s (Nielsen 2006). For example, the United Nations' Millennium Development Goals and EFA Fast Track Initiative initiated in 2002 focus on universal enrollment, not learning (United Nations 2000, World Bank 2004).

Today, it is almost certain that the post-2015 Millennium Development Goals related to education will include a learning component. ASER and Uwezo survey results are often used as evidence to support the need for such a global focus on measuring learning, including but not limited to several years' EFA Global Monitoring Reports, the UK Department for International Development's (DFID) Education Strategy for 2010-2020, and the World Bank Education Strategy 2020 (UNESCO 2011, 2013, 2014; DFID 2010; World Bank 2011). Though education quality had been a longstanding concern, findings such as the ones generated by citizen-led assessments provided evidence that such concern was well-founded.

There have of course been national assessments for many years which reveal low learning levels in many countries, but very few assess children in the early grades or use a population-based sample to ensure that out-of-school and rarely-in-school children are included. The lack of even the most basic reading and math skills that citizen-led assessments reveal is in some ways more powerful in stimulating debate than similarly dismal results of more formal assessments of older children because it reflects such widespread failure of the system to deliver even the most basic education.

In this way, citizen-led assessments have contributed to the emergence of learning outcome data at the forefront of the global debate on education development by

providing (1) population-based evidence of the seriousness of the learning crisis and (2) a model of how commitments to improve learning could be held to account by measuring learning outcomes in a low-resource way. Furthermore, the leaders of citizen-led assessments have played important high-visibility roles as international advocates for the measurement of learning outcomes.

Beyond contributing to action in the international policy

sphere, the inspiration that ASER and Uwezo provided for other citizen-led assessments is perhaps their biggest contribution to international action to address learning outcomes. These include initiatives in Mali, Mexico, Nigeria, and Senegal (ASER can also be credited with inspiring ASER Pakistan and Uwezo). Civil society groups in Bangladesh, Ghana, Malawi, Rwanda, and South Sudan have also expressed interest in the model. The ASER Centre in particular has not only served as a model for these initiatives, but has directly supported their development. The newly-formed People's Action for Learning Network (PAL Network) will also contribute to supporting existing and new-citizen led assessments.

Further, ASER has worked with international organizations to either share or adapt ASER or ASER-like tools for their work in various countries. Examples include the widespread use of ASER tools by JPAL, Effective Interventions, the International Rescue Committee, and others. The Peace Corps also has shown interest in the use of ASER tools.

Within East Africa, Uwezo has fostered a broad regional awareness of the learning crisis. Its annual East Africa report brings a regional perspective to the learning crisis, leveraging individual country findings to pressure governments to act together in addressing quality of education. Through regional report launches, Uwezo has also engaged key regional stakeholders, including the East African legislative assembly, parliamentary leaders of each country, and many other regional leaders within and outside of government.

National level

Awareness

Increasing the visibility of the learning crisis at the national level is one of the main successes of both ASER and Uwezo. Informants at the national level in all countries, within the government and outside of it, were highly aware of the assessments and their results. For both ASER and Uwezo, this level of awareness about the initiatives themselves is driven by an intensive media push as previously described.

In addition to the media push, both ASER and Uwezo reports are distributed among government officials, elected representatives, multilateral organizations, universities, think tanks, and NGOs at the national level. Staff of both initiatives lead this outreach by meeting with key stakeholders to share the findings and sharing the report through their national networks. Additionally, for both ASER and Uwezo, large national release events are held to announce each year's findings. This consistent annual event has contributed to the institutionalization of a debate over learning outcomes each year, according to many informants.

Importantly, this increase in visibility should not necessarily be equated to the discovery of low learning outcomes. In other words, national-level stakeholders already knew that learning levels were low—but, critically, ASER and Uwezo have made low learning levels highly visible in the public sphere. To quote a senior official in India's Ministry of Human Resource Development (MHRD), "Government always knew that learning levels are poor in public schools. We did not need ASER to tell us this fact which to us has always been self-evident." In Uganda, a senior Ministry of Education official stated, "Our own assessments had already shown this and therefore for Uwezo to come up with their findings is not news."

Although refuting any influence of ASER on their knowledge of learning levels, representatives from India's MHRD and the Planning Commission did agree that the ASER survey has "shown them the mirror" by forcing public acknowledgment of the learning crisis. In fact, the Minister of the MHRD has released the ASER report at the national launch event three times and the Deputy Chairman of the Planning Commission has released it five times. In East Africa, 83% of national-level stakeholders who were aware of Uwezo agreed or strongly agreed that it contributed to their awareness about learning outcomes.

For both ASER and Uwezo, survey findings are often portrayed in the media in an adversarial manner. In neither case is there evidence that ASER or Uwezo design press releases and other communications materials with the explicit intention of placing blame, but, in many media outlets, the findings are quickly turned into accusations and

criticisms of the public education system. For ASER, this results in questioning of government officials at the national level, who are then forced to respond to survey findings and explain what they plan to do about the low learning outcomes being reported. According to the ASER Center's own analysis, during the 2013 winter parliament session, 15 members of parliament asked questions that included references to the ASER 2012 findings, which showed learning levels lower than in previous years (ASER Centre 2013).

For Uwezo, the media also presents findings in a manner that seeks to allocate blame for low learning levels. In Kenya and Tanzania, teachers' unions have taken a defensive stance against Uwezo's results since they were first released in 2010. They react with sharp criticism of the survey methodology when results are released each year. Teachers often cite the fact that out-of-school students are tested as a reason why the results are not reflective of the education system or their own performance—how can they be held accountable for children who are not enrolled in school? While out-of-school children's results are separated from those who are enrolled in some sections of the report, they feel that they are blamed for the overall findings of low performance. Uwezo's intent is to show the *entire picture* of learning outcomes in the country, a topic of joint responsibility of students, parents, teachers, and government. However, the media at times distorts this approach, leading to defensiveness on the part of audiences that could be key collaborators.

Early relationship-building proves valuable. In both India and East Africa, some who perceive that they are being blamed for the low learning levels exposed by ASER and Uwezo often react by criticizing the assessment methodology and questioning its validity. Individuals and institutions that were engaged in the development of the survey design and its implementation tended not to be critical of the findings.

The evaluation of Uwezo revealed that institutions that were engaged by Uwezo in the development of the survey design and the testing tools did not respond defensively or question the validity of the survey findings. The Uwezo teams in each country seek the advice of national testing agencies and representatives of the Ministry of Education when developing and refining the survey methodology. In Kenya and Tanzania, the teachers' unions are not engaged by Uwezo, but in Uganda, where the Secretary General of the Uganda National Teachers Union (UNATU) is part of Uwezo's Advisory Committee, the antagonism felt by unions in Kenya and Tanzania has not occurred, nor has the methodology been criticized. Similarly, Uwezo Kenya has an ongoing relationship with the head teachers' union in Kenya (Kenya Primary Schools Head Teachers Association), which has been accepting of the survey results.

These findings reveal the power of even limited engagement with key stakeholders in gaining their support and setting the foundation for collaboration. For example,

Uwezo staff members have been asked to take part in several government review processes, including:

Table 7. Uwezo Participation in National Dialogue on Learning	
Kenya	<ul style="list-style-type: none"> Participated in development of the Basic Education Bill, which became law in 2013 Contributions to policy development more broadly, including Kenya's nationwide development blueprint, Vision 2030 Membership in Elimu Yetu, a national coalition of CSOs, research institutions, and other non-state actors who advocate for improved provision of quality education
Tanzania	<ul style="list-style-type: none"> Invited by the Ministry of Education to participate in education reviews through the Tanzania Education Network (TENMET) a national network of non-governmental and community-based organizations similarly focused on advocating for education reform
Uganda	<ul style="list-style-type: none"> Uwezo has been cited as a significant influence in the establishment of a Parliamentary Forum on Quality Education (PFQE) by several members of parliament (Twaweza 2013)

In summary, national-level key informants, even those that have been put on the defensive by the ASER and Uwezo findings, tend to agree that ASER and Uwezo initiatives have already achieved the goal of increasing the visibility of low learning outcomes. However, many informants in all countries believe that **highlighting low learning levels year after year is now, after ten years in India and even after only four years in East Africa, having less impact than it did in the initiatives' first years. That is, new results (especially since the results are not significantly different than previous years') are not further increasing already high levels of awareness.**

Action

In India, while government informants indicate that the debate surrounding issues of learning was not stimulated by ASER, non-government informants disagree. They insist that ASER's role in raising the visibility of the learning crisis has stimulated greater debate within the government. In East Africa, the majority of informants agreed that the Uwezo assessment has increased debate on improving learning outcomes in their respective countries, 91% in Kenya and Tanzania, and 77% in Uganda.

In addition to stimulating debate, ASER and Uwezo have also indirectly contributed to commitments to action at the national level. Evidence does not indicate that ASER

and Uwezo directly caused these commitments, but they have likely played a contributory role. Examples include:

- Uwezo results have been cited in government reports and strategy documents, noted as supporting evidence for renewed government focus on learning outcomes. In Kenya, for example, Uwezo findings have been mentioned in the Kenya National Education Sector Support Plan (NESSP 2014-2018) which cites education quality, including learning outcomes, as one of the key pillars of education to be addressed in the next five years.
- In the Government of India's 12th Five Year Plan (2012-2017), the first priority for primary education is monitoring and improving learning outcomes, and the ASER survey is cited. For the first time, measurement of learning outcomes in a continuous manner was stated as a methodology that would contribute toward improving the quality of education. While no informants claim that the ASER survey data alone led to the prioritization of learning outcomes, most non-government informants believed that ASER's role in raising the visibility of learning issues did contribute.
- Informants from the Government of India's Planning Commission reported that it seeks to build an incentive mechanism to improve the quality of education, and acknowledges that, by increasing the visibility of learning outcomes, ASER has influenced its pursuit of such quality-focused solutions.

Action can also be demonstrated by allocation of resources. In 2014, 1% of India's 2014 national Sarva Shiksha Aabyhiyan³⁰ (SSA) budget, or INR 250 million (USD 4 million), was allocated to conducting learning assessments. The Finance Ministry, which develops the Union Budget of India, has cited ASER data in its annual Economic Surveys, though there is no public citation of ASER as an influence on the 2014 SSA funding for assessments. That said, representatives of the Planning Commission that were interviewed indicated that, while the increase in expenditure on education and its quality has largely been driven by the vast improvement in the economic situation in the country, they believe that the efforts of the ASER Centre have also contributed to a large extent.

The evaluation uncovered several reasons why direct action at the national level has been limited in India:

- Since most education policy and planning happens at the state and district levels in India, direct action to improve learning outcomes is not necessarily a goal at the national level. The assignment of priorities and allocation of resources may be the limit of national-level action, reserving direct action for the state, district, and community levels.

³⁰ Sarva Shiksha Aabyhiyan is the national government's flagship program to achieve universal primary education throughout India.

- Both MHRD and its research and training wing, the National Council for Educational Research and Training (NCERT), were clear that ASER has not impacted their measurement practice. They believe strongly in assessment but are exclusively focused on the National Achievement Survey (NAS) for assessments and denied any influence of ASER on their strategy. They believe that the NAS is a more comprehensive tool to carry out scientific assessments as compared to ASER. Additionally, they felt that the strategy adopted by ASER is more geared towards getting media coverage and raising the level of awareness; aspects in which it has done very well.
- An entitlement and rights-based approach to education over the past several years, as demonstrated by the passing of the Right to Education Act in 2009, constrained the space for policy reform related to quality at the national level. The government was seeking to push rights and entitlements through legislation as a key instrument to social development. This diminished the scope for policy focus on quality.

In contrast to India's decentralized system, in Kenya, Tanzania, and Uganda, education policy and planning happens at the national level, making this level a critical target of Uwezo's influence. Twaweza has identified two cases in which Uwezo contributed to direct actions at the national level:

- According to Twaweza's own monitoring activities, one informant from the Ugandan Ministry of Education indicated that Uwezo's findings were part of the impetus behind a Ministry project to "support teachers in improving their pedagogical skills and teaching literacy" (Twaweza 2013).
- Similarly, Twaweza (2013) notes that Uwezo has contributed to the Big Results Now (BRN) initiative in Tanzania, which takes a variety of approaches (including heightened accountability measures and early learning assessments) to produce quick changes to learning outcomes (Tanzania Ministry of Education).

While these examples provide a useful illustration of what influence at the national level might look like, they do not amount to the level of influence that Uwezo hoped to have. Informants in East Africa indicated that, while Uwezo has indeed played a strong role in pushing national level policymakers to commit to action with the aim of raising learning outcomes, it cannot be said to have contributed to the development of concrete programs.

One key reason why stimulating direct action at the national level has proven difficult to achieve for Uwezo is the lack of concrete suggestions for actions that could be taken. Many national-level stakeholders expressed that though "sounding the alarm" is important, Uwezo should also offer policy suggestions, intervention ideas, and "success stories" if it expects to play a role in improving

outcomes. Representative of a very large proportion of informants, two in Kenya said the following:

"Uwezo should have done more than just communicate its findings. Uwezo should have used the good finds to mount a practical R&D project targeting improving teaching-learning processes and monitoring learning achievement at the same time."

—RESEARCHER, KENYATTA UNIVERSITY, KENYA

"It is not just enough to keep on year after year telling us that children are not learning. Help us to do something to address the problem. What has Uwezo done to rectify the problem and address the bottlenecks and challenges affecting learning?"

—SENIOR GOVERNMENT OFFICIAL IN KENYA

These informants are promoting an approach also supported by evidence in the transparency and accountability field: the provision of information alone will not lead to action (Kosak and Fung 2014, Bjorkman Nyqvist et al. 2014, Bruns et al. 2011, and many others). Twaweza's new "What Works in Education" program will address this limitation by generating evidence, ideas, data, and stories of effective education interventions that can then inform suggestions offered to key stakeholders.

State level (India)

The state level evaluation is only applicable to ASER, but the findings of ASER's significant success at this level are crucial to understanding the most effective strategies that citizen-led assessments can use to generate action, particularly in collaboration with government.

Awareness

The ASER Centre has been successful in increasing the visibility of learning outcomes at the state level. Print and broadcast media are the driving force in the dissemination of findings. Many state teams that were interviewed indicated that state officials did not refute ASER survey results, in contrast to national officials. In addition, across the four states where key informants were interviewed, government officials indicated that the ASER survey gained credibility by being an independent civil society initiative, largely free from bias. Both government and non-government informants indicated that, as a civil society initiative, the ASER survey is independent and therefore performs the vital function of serving as a credible counter-balance to government-led surveys like the NAS.

There are of course exceptions to this acceptance at the state level, but interviews with ASER state teams and government and non-government stakeholders external to ASER showed a general trend toward acceptance.

Action

The potential for impact on action at the state level is greater than at the national level in India for two main reasons: first, state officials have significant authority when it comes to implementation and practice. Second, states in general are more accepting of the ASER survey findings than are the MHRD national policymakers.

The level of impact at the state level varies widely. Several factors inform this variation: the interest of state policymakers in education quality and assessments, the baseline status of a state with regard to learning levels, and the depth of Pratham's engagement with the government seem to be determinants of state level impact on action. Overall, ASER has been successful at inciting debate, engagement, and in some cases direct action for the

improvement of learning outcomes at the state level. State-level action to which ASER may have contributed takes on three particular forms:

- **Several states in India took on the development of their own learning assessments over the past several years, some in partnership with Pratham.** In some cases these were motivated by the desire to refute or validate ASER survey findings. In others, states sought to collect data that was representative below the district level in order to be able to use it to make practical decisions, according to many informants.
- **Another way in which the ASER survey influences direct actions by states is by identifying weak areas around which the state can then design programs.**
- Aiming to improve the quality of education in the state, **some states initiated non-assessment-based programs in recent years.**

Examples of these types of action taken by states to address learning are described below. Informants' perception of the extent to which the ASER survey

Table 8. Examples of Actions Taken by States in India

Bihar	<ul style="list-style-type: none"> • CSOs reported using ASER data in their advocacy efforts; the data allows them to highlight learning gaps within the state and argue for action on the part of the government. • Non-government informants indicate that the high visibility of the ASER survey has motivated the government to internalize assessments into their own system by setting up an assessment unit. • The government has conducted school-based assessments of students in grades 2, 4, and 6 in 1,300 public schools. Pratham is facilitating the assessment and is using volunteers, DIET students, as well as teachers from Cluster Resource Centres (CRCs) to conduct the assessments. They are similar to the ASER testing tools. • ASER survey data was used to identify high-need blocks (where children's learning levels were low) to implement its Mission Gunvatta initiative, a large state-level initiative launched in 2013 and based off of a pilot project conducted in the district of Jehanabad by district officials in partnership with Pratham. The basic approach is to group students by ability level and offer remedial instruction based on their starting competencies.
Karnataka	<ul style="list-style-type: none"> • Informants at the state level in Karnataka reported that the ASER survey findings had very little impact on state level policy or resource allocation. The Karnataka School Quality Assessment and Accreditation Council (KSQAAC), which monitors quality of schools through external evaluations and support in developing school initiatives and community engagement, reports that it has been focused on quality for many years and its practices have not been influenced by the ASER survey. • However, district officials believed that the ASER survey findings and related advocacy likely influenced the recent instruction by the state Lokayukta (Ombudsman) for teachers to spend a set amount of time strengthening children's basic language and mathematics skill.
Madhya Pradesh	<ul style="list-style-type: none"> • The government started a state-wide assessment initiative in 2011-2012 called Pratibha Parva, but informants reported that the ASER survey findings had very little impact on state level policy or resource allocation.
Uttarakhand	<ul style="list-style-type: none"> • Similar to Bihar, CSOs reported using ASER data in their advocacy efforts. For example, Room to Read reported using ASER data to inform its reading program design and develop targets for the expected outcomes of its programs. • Some state-level officials stated that the ASER survey findings have influenced the state's annual plan and the state budget (in which more than half of the state education budget is allocated to addressing quality issues). • Government officials reported that the ASER survey influenced them to work with the State Council of Educational Research and Training (SCERT) to develop assessment tools. The tools were developed in partnership with Azim Premji Foundation and are more detailed than the ASER tool. • ASER survey data on pupil:teacher ratio revealed a shortage of teachers. The Education Department indicated that this realization led to the hiring and training of 2,000 new teachers. • The State Education Department removed non-teaching activities from teachers' responsibilities so they could focus on classroom instruction. Representatives from the department indicated that this change was made in response to the ASER survey data, among other efforts to address quality. • A new early literacy program was designed in collaboration with SCERT, Room to Read, Pratham, and the Azim Premji Foundation. It tracks early literacy using a progression chart. The design of the program draws from the ASER tools.

contributed to the launching of these initiatives is provided in each case (Table 8).

As demonstrated in Table 8, direct actions have been taken by the four state governments aiming to improve the quality of education. With a few exceptions in which key informants explicitly stated ASER's contribution to state efforts, it is very difficult to assess the extent to which the ASER survey contributed to these actions, particularly because states are unlikely to credit a civil society initiative with motivating their efforts, even if it did.

Informants throughout the states, including several ASER state teams from states not included in the evaluation sample, gave many examples of state level partnerships between government and Pratham in response to the ASER survey findings. In effect, the relationship of the ASER survey and Pratham seems, in many states, to manifest as a complementary approach, where the ASER survey exposes the gaps in learning outcomes and Pratham works with governments to develop and test solutions to those gaps. In states where Pratham has significant coverage, like Bihar (where it works with district education offices in all 37 districts), this approach works well. The state-wide Pratham team is able to provide technical support, compensating for the limited capacity of the government education structure, and implement a large scale quality improvement program. This is a quintessential example of assessment driving action. However, in states that do not have such deep Pratham presence, it would be unlikely for Pratham to be able to "pick up" with a solution where the ASER survey left off by identifying the problem, except at very limited scale. While there may be some states where this could happen, the powerful combination of Bihar's committed government leaders and strong Pratham network does not exist in most states.

District level

Awareness

Both ASER and Uwezo aim to influence two key groups at the district level: their district partner organizations and district-level government officials. For both initiatives, the contribution to awareness about low learning levels varies hugely across districts, and is more limited than at the national and, in the case of India, state levels.

For Uwezo, much of the awareness that was generated at the district level was through print media, radio, and television. In contrast to the national level, there was very little evidence of district-level stakeholders being directly engaged.

For ASER, district-level awareness varies by district, and by state. In Bihar, district-level government officials indicated that there was only a modest level of knowledge about the ASER survey among education stakeholders in their districts. In Madhya Pradesh, the district-level officials and officials of the Education Department in the sampled districts had no knowledge of the ASER survey. In Uttarakhand, officials in the sampled districts were aware of the ASER survey and reported discussing the report in their internal meetings, especially immediately following its launch. Notably, in Karnataka, most of the district-level officials were well aware of the ASER survey including the details of the methodology, and reported discussing the ASER report in their meetings.

For both ASER and Uwezo, sporadic impact on awareness at the district level can largely be attributed to the lack of systematic involvement of the network of district partner organizations in dissemination activities. Findings from the ASER and Uwezo process evaluations indicate that dissemination of results at the district level is much less consistent than at the national level (and state level in India), where dissemination is organized by ASER and Uwezo staff. Both initiatives share the annual report and district findings with district partner organizations, but any follow-up activity is dependent on these organizations' capacity and interest. According to ASER state teams, in many cases, ASER district partners are small grassroots agencies that are not able to carry out dissemination at the district level. Examples of district level dissemination through partners and ASER state teams exist but do not occur at scale and are not part of a holistic strategic approach. Karnataka represents an exception to this trend.

The relatively high district-level awareness of ASER in Karnataka may be related to the focus the state ASER team puts on dissemination. The team is composed of Pratham Mysore staff as well as staff from the Akshara Foundation. In Karnataka, district partners are responsible for dissemination as well as survey administration. Active dissemination through partners occurs in approximately half of the districts in the state.

In East Africa, there was an assumption by Uwezo that district partners would organically move to disseminate Uwezo findings and facilitate discussions on poor learning achievement at the district level. However, in the majority of districts in each country, this has not occurred: Uwezo country coordinators in Kenya, Uganda, and Tanzania,

report that only between 10-15%, 35%, and 30%, of partners respectively have made efforts to distribute Uwezo reports and facilitate debates at the district level. While these efforts do demonstrate some activity, district level dissemination does not occur systematically or at scale, so it is difficult to track its impact or reach the majority of districts.

For both ASER and Uwezo, the lack of engagement of district partners in dissemination and other activities is not surprising given the lack of resources allocated for this purpose. Without funding for such activities, but, perhaps more importantly, without guidance and support from ASER

and Uwezo about what those activities could look like, it is not reasonable to expect that district partners could drive forward district-level engagement entirely on their own. Most of these partners are small organizations that do not have the capacity in terms of people, time, or, critically, skills, to engage in dissemination to district-level officials.

For citizen-led assessments, methods to engage the district partner network around dissemination would need to be low-resource—for example, leveraging the existing training structure, which includes district partners. During the 2014 survey, Uwezo Kenya experimented with inviting district and county level officials to observe the survey and even participate themselves in administering the household survey. This is an example of a low-resource way to build engagement of target audiences in to the existing survey structure. Another example is ASER Pakistan’s annual training for district partners after the survey is over and before the findings have been released, which focuses on how partners can disseminate findings in their districts.

Action

Compared to ASER’s contribution to actions at the state level, which have been significant, the ASER survey has had little impact on actions at the district level. For both ASER and Uwezo, there are examples of district-level action being inspired by the citizen-led assessment, but these occur sporadically and therefore do not lead to the type of widespread improvements in learning that ASER and Uwezo aim for. That said, such examples may be able to shed light on effective mechanisms for district-level engagement.

ASER-INSPIRED ACTION AT THE DISTRICT LEVEL IN BIHAR

- A district-level official in Bihar cited the ASER survey findings as a partial influence on the Samijho Seekho program initiated by the government in Nalanda district, where elementary-level students are given extra classes in language and mathematics to improve their learning levels.
- Mission Gunvatta in Bihar is based off of a pilot project conducted in the district of Jehanabad by district officials in partnership with Pratham. The central principle of the initiative is that children are grouped by ability level and then receive remedial instruction based on their competency. The genesis of the program is that a Bihar district official came to Pratham asking for help addressing low attendance levels. Pratham pointed to the ASER survey results to show that children in the same classroom had vastly different ability levels, and suggested sorting children by ability level rather than age for a portion of the day. In addition to improving learning, the idea was that by meeting children where they were, they would be more engaged in school and attendance would go up. In this case, the ASER testing tools are used to assess and sort students.

For both ASER and Uwezo, one small-scale but significant way in which citizen-led assessments have triggered learning-focused action outside of the government is through the uptake of the testing tools in education programs run by NGOs and CSOs. Dozens of these groups use the testing tools in various ways including initial assessment of children’s learning levels, tracking children’s progress over time, and monitoring the impact of their work as an organization. The design of the testing tools enables widespread uptake by organizations with varying approaches and capacities because it is easy to use and adapt.

It is important to note that there are obstacles to district-level action that are to a large extent outside the control of ASER and Uwezo. A few examples of such obstacles include:

- Ideally, as has been observed at the state level in India, district-level officials and NGOs/CSOs could use the survey data to identify locations of high need within their district and design interventions or allocate resources around this information. This is not possible because survey findings are not disaggregated below the district level. In India, several non-government experts at the state level, district officials, and district partners believed that that this made the survey data “less actionable” because they could not use the data to identify locations of critical need within a district. This is a challenge for both ASER and Uwezo because (1) they need to preserve the anonymity of communities and (2) their sample is in most cases not large enough to have findings be representative below the district level.
- In India, structural and capacity constraints in the public education system also limited impact. Awareness and acceptance of results among state leadership did not lead to sustained awareness and action more widely throughout the state. Stakeholders agree that district-level activities are necessary to see impact at that level; it will not “trickle down” from the state level.
- Importantly, the lack of action at the district level limits the potential for impact on teachers, as the district level (or block level, in India) is typically where they receive training and supervision. The recent increase in the number of DIETs serving as ASER district partners may result in increased impact of the ASER survey at the district level and below if, for example, DIETs’ capacity to deliver effective teacher training and incorporate assessment into its curriculum is developed. More time is required before the impact of participation in the ASER survey on DIETs and their students can be assessed.

In summary, while there are examples of ASER’s and Uwezo’s contributions at the district level, the sheer number of districts requires the use of partners beyond their own teams; the lack of resources for strategic engagement with these partners, beyond survey administration, render their overall district-level impact minimal.

Community level

Overall, the evaluation uncovered only limited anecdotal evidence of ASER and Uwezo stimulating awareness and action at the community level.³¹

Awareness

Across the four citizen-led assessments included in this evaluation, three types of community level engagement were identified: before the survey, on the day of the survey, and dissemination to the general public through media.

Interviews with ASER state teams and district partners; findings from Lieberman et al. in Kenya; and FGDs conducted in Kenya, Tanzania, and Uganda indicate that these activities (participation in the survey, instant feedback, materials left behind) did not generate significant awareness about learning outcomes at the community level. Lieberman et al. found that most parents who were present for the assessment were able to recall it and remembered receiving materials, though they generally had a mixed understanding of the initiative, and often had to be reminded about the details of the assessment. In one focus group discussion, a parent in Tanzania said,

We do not remember the organization but there were people passing in our houses. They gave our children work to read and work with simple maths. They did not tell us who they were. They gave us fliers and calendars.

—PARENT, MOROGORO DISTRICT, TANZANIA

Further, Lieberman et al. (2014) found that a majority of parent respondents whose children took part in the 2013 Uwezo assessment claim that Uwezo “did not provide them with any information about their child’s performance.” A study of households in two Kenyan districts (Rongo and Kirinyaga), including a control group (whose members had not been part of the Uwezo assessment), found that Uwezo’s provision of information at the household level had “no discernable impact on either private or collective action” (ibid).

ASER district partners and ASER Centre state teams report that awareness at this level is very low. Given that ASER’s only systematic mechanism for dissemination at the

community level on the day of the survey is participation in the survey and, when possible, observation of the assessment, it is not likely that ASER has been able to generate significant awareness either. The ASER Centre state and national teams agreed with this conclusion—they have not focused on generating awareness or action at the community level.

District partners for both ASER and Uwezo reported that, while volunteers’ awareness of low learning levels was enhanced by participating in the survey, they in general have not seen evidence of that awareness being passed on at the community level. This is to be expected, given that volunteers’ role and training are limited to data collection; they are not prepared to serve as messengers or advocates for learning.

In general, teachers in Kenya, Tanzania, and Uganda were more aware of Uwezo than parents, but they reported that their knowledge of low learning levels was based on their own experience as teachers, not information from Uwezo. The only engagement of teachers in the survey process is that one teacher in each community (the grade 2 teacher) is asked questions about enrollment, attendance, and teaching resources as part of the school survey. Teachers generally did not feel that they could take action in response to the Uwezo survey because they did not know the results of the survey in their village. One teacher in Kenya said,

If they test our children and tell me their performance, then I will know. But you see they never even tell us what they have found out about our children. This thing you are calling “instant feedback” is never given by these volunteers.

—TEACHER, POKOT COUNTY, KENYA

Action

While limited awareness has impeded the stimulation of action at the community level, there are some examples of action at this level driven by district partners, District Coordinators, individual volunteers, or community leaders. Though not found to be representative of typical behavior by these groups, it is useful to describe these examples as they may provide insight into the types of actions that could be encouraged or supported by district partners, volunteers, or other community-level champions.

³¹ The ASER Centre requested that the community level not be included in the evaluation of ASER as it has not focused its efforts on this level and does not expect to have made impact. Therefore, primary data was not collected at the community level in India. Key informant interviews with district partners and ASER core team members, as well as a review of ASER activities, allows for general discussion of the level of impact that might be expected at the community level.

EXAMPLES OF SPORADIC COMMUNITY-LEVEL ACTION

- In one of the eight communities observed in Kenya, a local chief contacted his District Education Officer after the Uwezo assessment to complain about the poor performance of the school serving his community. As explained by the chief himself, this seems to have had some positive impact:

“When I heard of the report on poor performance, I got concerned and wrote a letter to the District Education Officer (DEO) to complain about the poor performance of our school these days. The school used to be so good and it used to produce good results compared to what it is now. Among the issues I raised in the letter were the elections of a new committee which at the time were long overdue and (I explained) how previous elections had not been conducted properly. The DEO called the Head Teacher and instructed him to call elections. I ensured that the elections were done as per the law. From last year when this took place, there has been some change in the way the school is run.”

—ASSISTANT VILLAGE CHIEF, RONGO DISTRICT, KENYA

- One Uwezo volunteer and local school teacher in Kayunga District, Uganda has become known as “Madam Uwezo” for her attempts to teach all parents in her village about the need to become more involved in their children’s education, encouraging even illiterate parents to sit with their children while reading.
- Other anecdotes of volunteer-led activities implemented after the assessment were also found in Uganda, including a reading competition, scholarship funds, and parent engagement sessions.
- Although parents in the focus group discussions did not report such actions, District Coordinators reported that some parents in their districts were motivated by the Uwezo results to take their children to school themselves every day, in order to increase their interaction with teachers. They also reported that some parents were checking in with their children’s teachers in order to monitor their children’s progress.
- A long-term district partner in Bihar, India reported that it had started a chain of libraries for children and initiated a new program called Chalo Padhe to develop reading skills of children in the elementary grade.

These examples demonstrate the types of action that citizen-led assessments intend to motivate at the community level. Due to their sporadic occurrence, they do not constitute achievement of the widespread and sustained community level engagement alluded to in both ASER’s and Uwezo’s theories of change. In order to achieve that, it will be necessary for to invest in community engagement beyond data collection.

Chapter 3. Recommendations

Recommendations for existing and new citizen-led assessments

Overall

- 1 Consider reducing the frequency of citizen-led assessments, moving from an annual to a biannual or triennial assessment.** Given that significant changes in learning outcomes over the course of one year are not expected, the massive annual investment in data collection may not generate enough marginal impact to justify its cost. According to the majority of stakeholders interviewed for the evaluation of impact for both ASER and Uwezo (including education experts, policymakers, and CSOs), the impact of the annual survey has diminished over time, as anticipation of results has waned. Even for Beekunko and Jängandoo, which have not released data as many times, the frequency could still be reduced.

There are undoubtedly tradeoffs to implementing this recommendation. Surely, the level of awareness and debate about learning that ASER and Uwezo have generated at the national level has been partially due to the momentum its yearly assessments have generated. But there is no evidence that new data (especially if it is not notably different data) is required to launch a new awareness-building and action-generating campaign. **The important rhythm that the current annual survey provides could be maintained with annual campaigns or events, without the need for annual data collection.** Distinct changes in learning outcomes could be anticipated over the course of two or three years, giving stakeholders a reasonable timeline for pursuing goals aligned to the survey findings. Most importantly, a biannual or triennial survey would free-up resources to use on dissemination and engagement activities during the off-year(s)—activities that have so far been limited due in part to resource constraints.

Roles of key players in the survey process

- 2 Continue to use volunteers as surveyors, but explore ways to extend their role beyond that of data collectors** in order to increase their potential to serve as champions within their communities. The survey “infrastructure,” which identifies partner organizations in every district/region and deploys volunteers to communities, is not currently leveraged. Citizen-led assessments could consider training partners and volunteers in survey follow-up engagement efforts, and then use community/volunteer pairings to implement community engagement activities. This could be supplemented with automated communication to volunteers (via SMS, for example) providing suggestions for follow up at the community level. This recommendation would be difficult to implement within the current training and funding structure: volunteers in many cases already feel stretched by their existing responsibilities. Reducing the frequency of the survey could again be a useful strategy to free up time and resources for more in-depth volunteer engagement.
- 3 Capitalize on the large networks of partner organizations that have been built. Engage partners beyond data collection,** and identify networks of organizations that could become partners in multiple locations. Systematic partnership models like partnering with DIETs in India, *Centres d’Animation Pédagogique* in Mali, or a single strong NGO across many districts are promising in that they can be scaled up through existing structures. Such engagement must be tailored to align with the capacity levels of partner organizations, or to build their capacity through strategic engagement.

Pre-survey engagement

- 4 **Consider testing new ways to engage in pre-survey awareness-building at all levels**, which has been shown to increase buy-in by target audiences once the survey takes place and its findings are shared. This could include mobilizing master trainers and district partners before the survey. Beekunko and Jàngandoo are experimenting with pre-survey engagement at lower levels (regional, *commune*, and community levels); more testing is needed to determine the impact of this engagement.

Sampling

- 5 Further inquiry is needed to determine the extent to which household sampling methodologies and implementation challenges compromise the randomization of the household sample.

Training

- 6 **Ensure extensive and realistic field practice is included in volunteer training.** In every country, master trainers, partner organizations, and evaluators felt that the volunteer training did not allow for sufficient mastery of the process. In particular, **field practice sessions that allow every volunteer to practice leading all key elements of the survey are critical.** New training models should be rigorously tested to ensure that any increased investment results in stronger volunteer performance. Uwezo began some of this experimentation in September 2014.

Testing tools

- 7 **Consider broadening the range of skills that is tested by citizen-led assessments in order to better inform policy and practice.** While it is important to maintain the simplicity of the testing tools, foundational concepts that are not currently tested could be added without greatly increasing the complexity or length of the test. For example, the reading test could assess listening comprehension. This would allow for distinguishing between children who have inadequate decoding skills and children who have deficits in the ability to make meaning in the test language. It would also give the assessment more power to inform policy, resourcing, and classroom practice. All citizen-led assessments should also test foundational mathematic competencies such as place value, which Beekunko and Uwezo already do, and measurement and shape, which Jàngandoo already does.
- 8 **Take steps to investigate equivalence of testing tools, in order to facilitate the comparison of results.** It is important for citizen-led assessments to do so in order to enable the exploration of factors that lead to high performance. The first step in supporting the making of comparisons is to demonstrate that it is valid to make comparisons in the first place, and this is achieved in part by ensuring that the assessment tools are at an acceptable level of linguistic and psychometric equivalence. Linguistic and psychometric equivalence can be investigated by drawing on expert judgment, by conducting equating studies, and by analysis of assessment data. Investigating and confirming equivalence across years is particularly important if citizen-led assessments aim to play a role in efforts to track learning outcomes over time.

Quality assurance

- 9 ASER and Uwezo have developed publicly-available descriptions of their quality assurance processes. These documents not only guide internal processes but also increase the credibility of the assessment. All citizen-led assessments should have formalized internal quality assurance processes in place, and ensure that monitoring data is used not only to identify entries that may pose quality concerns but that can also be used after the survey is administered to identify areas for improvement for subsequent surveys.

Dissemination and engagement

Many of the recommendations below build on two key findings: First, engagement at every level should involve more than just dissemination of findings. Evidence shows that simply providing information will not generate action (Kosak and Fung 2014, Bjorkman Nyqvist et al. 2014, Lieberman et al. 2012, Bruns et al. 2011, and many others).

Second, an emerging body of evidence demonstrates that improvements in governance and, relatedly, service delivery only happen with a combined effort of top-down and bottom-up accountability, ideally with a high amount of vertical integration (Fox 2014). Due to the structure of the citizen-led assessment process, ASER, Beekunko, Jàngandoo, and Uwezo are well-positioned to achieve this type of vertical integration, but it will take concerted focus on the “bottom”—namely, the district and village levels.

The recommendations below focus on ways to close the feedback loop, a gap that has to date prevented citizen-led assessments from becoming citizen-led movements to improve learning.

10 The network of citizen-led assessments should consider experimenting with systematic ways to develop awareness and action about low learning outcomes at the lower levels (district, community, etc.). At the community level, impact evaluations of transparency and accountability interventions indicate that a common feature of successful ones is facilitation of dialogue and action development between citizens and providers (Bjorkman Nyqvist et al. 2014 and others). This is an example of the type of evidence-based engagement activity that could be tested.

Such experimentations should include testing of approaches that promote various types of action: individual action (e.g., individual parents changing what they do) collective self-help action (e.g., a community organizing a tutoring program) or collective accountability action (e.g., a group of people in a community deciding to go to their district education office to demand better services). **Not only parents but community leaders and, critically, schools and teachers should be involved in this experimentation.** The literature review developed for this evaluation provides a good starting point for identifying interventions that have proven to be promising, including repeated provision of materials and information over time, facilitated dialogue between community members and service providers, etc. (Bjorkman Nyqvist et al. 2014, Pandey 2008).

11 In addition to aiming for general awareness, **citizen-led assessments should systematically work to identify champions** in each of the audiences that they wish to influence. Committed individuals have been a key part of many of the successes achieved to date in generating action. Strategies for identifying and engaging such champions should be experimented with. For example, the initiatives could consider building a network of individual champions (politicians, government officials, journalists, community leaders, teachers, etc.) to mutually reinforce each other’s actions at different levels and in different geographies.

12 Beekunko, Jàngandoo, and Uwezo could consider **developing partnerships with providers of solutions.** This is intentionally distinct from providing solutions themselves, as it is important for citizen-led assessments to retain their neutrality. The ASER and Pratham combination has proven to be effective at enabling action.

13 Build on the observed uptake of citizen-led assessment testing tools by NGOs to share tools more systematically with NGOs, private schools, and teachers.

14 Consider an evolutionary approach: The evaluation finds that, over time, the annual reporting of results increases awareness only marginally while demand for solutions increases. **Citizen-led assessments could consider testing a multi-year trajectory in which the first few years focus on generating awareness and raising the visibility of learning outcomes, and subsequent years focus on building partnerships to promote direct action from the bottom up and the top down.** The assessment could still be carried out biannually or triennially, but the focus in off-years could shift from raising visibility in the early years to facilitating specific actions in later years. In order to identify which types of actions they should focus their efforts on stimulating, after a few years of implementation, citizen-led assessments should take stock of what types of actions they have contributed to and are therefore realistic for them to expect to incite at each level. They could then focus their dissemination and engagement activities on those levels where they are most likely to have success.

Recommendations for supporters and researchers of citizen-led assessments

- 1 Support citizen-led assessments' dissemination and engagement activities, so they have enough resources left over after data collection to conduct meaningful engagement.** This could manifest as a continued level of support but for a biannual or triennial survey, with the off years' support going toward engagement activities as described above, especially at the district and community levels.
- 2 Support core teams to build their capacity in two key areas: technical expertise in assessment, analysis, and evaluation, and practical experience in the field of transparency and accountability—**lessons from interventions in this space may be a driving force behind grantees' impact going forward.
- 3 Support citizen-led assessments to experiment with techniques that transcend the boundaries of "dissemination" and move toward strategic engagement, especially with government officials.** For example, in addition to aiming for general awareness nationally, work with implementers to identify narrow geographies (states, counties, or regions) where early indications of commitment exist, and test approaches to "vertical integration" including simultaneous community, district, and state level activities. This depth-over-breadth approach is likely to deliver more learning on effective approaches than a diffuse approach. It will be easier to assess impact and citizen-led assessments' contribution to it. Both success and failure will be easier to spot.
- 4 Support citizen-led assessments to test specific tactics for identifying and engaging champions at community, sub-national, and national levels.** For example, at the national level, build on anecdotal evidence of the power of person-to-person relationships and challenge stereotypes that put government and civil society at odds with each other by investing in shared experiences between the sectors like joint training or workshops related to assessment and learning. Train volunteers and partner organizations in ways to engage with individuals who show interest during the survey—follow up with them shortly after the survey to provide suggestions on facilitating dialogue between parents and teachers, equip them with test booklets and encourage them to share them widely, etc. At district or regional levels, approach champions with current assessment results, a goal for the district's next citizen-led assessment results, and, critically, a presentation of concrete ideas for individual and collective actions that the champion could feasibly push forward.
- 5 Disseminate lessons from the above experimentation widely.** These will inform citizen-led assessments, but the transparency and accountability field more broadly, as well. **Coordinate such experimentation across the growing network of citizen-led assessments,** so a variety of approaches can be tested and lessons shared across the group. Include experimentation with similar approaches in different contexts to enable identification of contextual factors that may enable or hinder effectiveness.

Annexes

Linked Annexes: Portfolio of Related Research

Kenya field studies

Two quasi-experimental studies of the Uwezo Kenya testing tools were conducted as part of this evaluation: a concurrent validity study and an inter-rater reliability study (ACER 2015). The full technical report of these studies was prepared by the Australian Council for Educational Research (ACER) and is available [here](#).

The **concurrent validity study** was conducted in Kenya in December 2014 and January 2015. It explored the relationship between performance on Uwezo and performance on the Early Grade Reading Assessment (EGRA) and the Early Grade Math Assessment (EGMA)—instruments for which there is evidence of reliability and validity. Some exploratory tasks were added to the core EGRA and EGMA tests and analyzed separately to examine the extent to which Uwezo tests, which assess only a limited set of foundational skills, may be able to predict children’s performance on a wider range of tasks. Details of the methodology and findings of this study can be found in the full technical report prepared by ACER.

The **inter-rater reliability study** was also conducted in Kenya in December 2014 and January 2015. It explored the agreement in scores assigned by Uwezo volunteers and by an expert rater to children’s responses to the Uwezo tasks. Children were recorded on video as the Uwezo testing tools were administered one-on-one. Later, a group of 20 Uwezo volunteers from ten districts and one expert rater independently scored the children’s responses. Details of the methodology and findings of this study can be found in the full technical report prepared by ACER.

Literature review

R4D conducted a literature review exploring the key dimensions of effective learning assessments and synthesizing the available evidence related to interventions aiming to build awareness and stimulate action by gathering and disseminating information. The literature review is available in [English](#) and [French](#).

Annexes

Annex 1: Documents developed as part of this evaluation

Several reports were produced over the course of the evaluation.

- **Literature review:** A review of key dimensions of effective learning assessments and evidence related to interventions aiming to build awareness and stimulate action
- **Pathways to Impact report:** A mapping of each initiative's activities, including the target audiences and intended outcomes of each. (The report also includes a comparison of activities across initiatives and the relative focus each puts on various audiences and types of engagement.)
- Desk review of each initiative's testing tools and processes (provided to each initiative separately)
- Overall synthesis of desk reviews of testing tools and processes
- Evaluation frameworks for Beekunko and Jàngandoo
- Evaluation report for each initiative (provided to each initiative separately)
- Overall evaluation report synthesizing four individual evaluations (this document)
- **Quasi-experimental studies:** A report presenting findings of two studies conducted in Kenya during December 2014 and January 2015 to assess inter-rater reliability and concurrent validity of Uwezo Kenya math and reading tests (English and Kiswahili).

Annex 2: Process evaluation observation sites and interviews

ASER OBSERVATION SITES

2013 National Training

- 3 days classroom sessions, 2 days field sessions

Rajasthan State Training 2013

- 3 days classroom sessions, 2.5 days field sessions

Andhra Pradesh State Training 2013

- 2 days classroom sessions, 2 days field sessions

District training in Ananthapur, Andhra Pradesh 2013

- 1 day classroom sessions, 1 day field sessions

District Training in Medak, Andhra Pradesh 2013

- 2 days classroom sessions, 1 day field sessions

ASER Survey 2013 in Ananthapur District and Medak District, Andhra Pradesh

- Survey observed in 6 villages in Ananthapur District over two days: Anumapalli, Pamidi Narpala, Narpala Korikonda, Gooty, Bukarayasamudram
- Survey observed in 6 villages in Medak District over two days: Perur, Rajapeta, Sangaipeta, Narsapur, Bonthapalli, Peddgotimukula

In total, the survey process was observed in 24 villages for field practice sessions and the actual survey, including approximately 180 households.

ASER PROCESS EVALUATION INTERVIEWS

- In person interviews with six ASER Centre national team members
- In person interviews with 14 state teams
Andhra Pradesh, Assam, Bihar, Chattisgahr, Haryana, Himachal Pradesh, Karnataka, Madhya Pradesh, Maharashtra, Punjab, Rajasthan, Tamil Nadu, Uttarakhand, West Bengal
- In person interviews with Master Trainers (MTs)

BEEKUNKO OBSERVATION SITES

Region	Commune	Village/neighborhood
District of Bamako	IV	Lafiabougou, Taliko, Djicoroni-para, Sébénicoro
	V	Sabalibougou, Kalabancoura
	I	Korofina nord, Fadjiguila
Kayes	Kayes	Sountoukoulé, Lafiabougou, Légal Ségou, Plateau, Khasso
Koulikoro	Kati	Malibougou, Kati Farada
	Mandé	Samaya, Coursalé, Ouezzindougou
	Sanankoroba	Sanankoroba
	Dialakoroba	Dialakoroba
Sikasso	Bougouni	N/A – observation of regional training only

BEEKUNKO PROCESS EVALUATION INTERVIEWS

- In person interviews with four OMAES core team members
- Bamako
 - CAP representative
 - Pedagogy advisor
 - Representative from the mayor’s office (3)
 - Director of Academie Rive Gauche
 - Representative from Academie Rive Droite
 - Advisors to local chiefs/heads (8)
 - Village/neighborhood heads (7)
 - Advisor to the CGS
- Kayes
 - Pedagogy advisor (2)
 - CGS president
 - Advisor to the CGS
 - Advisor to the mayor’s office
 - Advisors to local chiefs/heads
 - Village/neighborhood head (4)
 - Teacher
 - Head teacher
- Koulikouro (Kati)
 - Pedagogy advisor
 - resident of the Cercle Council
 - Secretary General of the Cercle Council
 - Head teacher (5)
 - CGS President
 - CGS representative
 - Village/neighborhood head (5)
 - Pedagogy advisor (3)
 - Secretary General of the Mayor’s Office

JÀNGANDOO OBSERVATION SITES

Region	Training sites	Survey administration zones observed	Partner NGO responsible for the region
Dakar	Dakar Rufisque	CAREF-supervised: Rufisque (neighborhood of Dangou Miname, plan Jaxxay, and in the zone of Zac Mbao (DR Medina Zac Mbao). COSYDEP supervised: Parcelles Assainies U18, Biscuiterie, and Ngor/Yoff.	CAREF and COSYDEP
Ziguinchor	Ziguinchor Kolda Sédhiou	Mpack, Grand Kandé, Kandialan Ouest, Bafikan/ Dialang, département de Bignona	PACTE
Kaolack	Kaolack Fatick Kaffrine		APROFES

JÀNGANDOO PROCESS EVALUATION INTERVIEWS

- In person interviews with four LARTES core team members
- Academy Inspectors (IA) and Inspectors of Education and Training (IEF) from Ziguinchor, Kaolack, and Dakar
- School directors (2) from each zone
- Partner NGO staff
- Secretary General of the Academy Inspector in Kaolack
- IEF representative from Kaolack
- IEF of Pikine
- School directors of Niaguis and Francisco Carvalho Schools in Ziguinchor
- School director in Kaolack and President of the School Directors' Collective
- Members of the Pedagogical Task Force
- One member of the Quantitative Team
- The director of Communications for Jàngandoo
- Animateurs and superviseurs
- Parents of children surveyed in the neighborhoods visited by the evaluation team

UWEZO OBSERVATION SITES

Kenya	<ul style="list-style-type: none"> • Training-of-master trainers • National conference • Zonal training • Volunteer training (Westlands district) • Survey administration (Westlands district)
Tanzania	<ul style="list-style-type: none"> • Zonal training (Kibaha district) • Pilot testing of tools (Morogoro district) • National conference
Uganda	<ul style="list-style-type: none"> • Zonal training • National conference • Volunteer training • Survey administration (Kiruhura district)

UWEZO PROCESS EVALUATION INTERVIEWS

- Regional office
 - Regional Manager
 - Data Management Consultant
 - Regional Trainer
- Kenya
 - Kenya Country Coordinator
 - Regional Coordinators (2)
 - National Trainers (6: 4 interviewed in a group, 2 interviewed individually)
- Tanzania
 - Tanzania Country Coordinator
 - Program Officer
 - Regional Coordinators (2)
 - National Trainers (2)
 - Master trainers (2)
- Uganda
 - Uganda Country Coordinator
 - Regional Coordinators (2)
 - National Trainers (2)
 - Volunteers (3)

Annex 3: Key informants interviewed for evaluation of impact

INSTITUTIONS REPRESENTED AMONG KEY INFORMANTS IN INDIA

<p>National</p>	<ul style="list-style-type: none"> • Ministry of Human Resource Development (MHRD)³³ • National Council for Educational Research and Training (NCERT), which runs the National Achievement Survey (NAS), India's primary government assessment • Planning Commission • National University of Education Planning and Administration (NUEPA) • Education sections of multi-lateral agencies • Research and consulting institutions (both public and private) • Foundation officers • Individual education experts • Civil society organizations (CSOs) • ASER Centre national team
<p>State</p>	<ul style="list-style-type: none"> • State Department of Education (including Sarva Sikhsha Abhiyan) • Planning Commission • Karnataka School Quality Assessment and Accreditation Council (KSQAAC) • State Council of Education Research and Training (SCERT) • Foundation officers • Education sections of multi-lateral agencies • CSOs • Pratham state staff • ASER state team
<p>District</p>	<ul style="list-style-type: none"> • District Education Officers (DEO) • Academics/Professors • ASER district partners • District Institutes of Education and Training (DIETs) • National Service Scheme (NSS) coordinating organizations • CSOs/NGOs • Private colleges/universities

³³ For this evaluation, the MHRD assigned a senior official to articulate the views of the agencies within MHRD on the ASER survey.

INSTITUTIONS REPRESENTED AMONG KEY INFORMANTS IN EAST AFRICA

<p>Kenya</p>	<ul style="list-style-type: none"> • Catholic University of East Africa • Kenya National Examination Council • Kenya National Union of Teachers (KNUT) • Women Educational Researchers of Kenya • ACTION AID • KEPSHA (Kenya Primary School Head Teachers Association) • NOKEPDA (Northern Kenya Pastoral Development Agency) • Kenyatta University • University of Nairobi • Elimu Yetu Coalition
<p>Tanzania</p>	<ul style="list-style-type: none"> • University of Dar es Salaam, School of Education • Dar es Salaam University College of Education • Ministry of Education and Vocational Training • Institute of Adult Education • Plan International Tanzania • Tanzania Institute of Education • EA RADIO • Mwananchi • Tanzania Standard Newspapers
<p>Uganda</p>	<ul style="list-style-type: none"> • Makerere University • Kyambogo Primary School, Kampala • Kampala City Council, Kawempe Division • Kampala District Local Government, Rubaga Division • Kampala District Local Government • St. Paul's Primary School, Banda -Kampala • Kyambogo Primary School, Kampala • Faculty of Education, Kyambogo University • Save the Children-UG • Justice for Children Programme, Kampala • KIBAALE Education Foundation, Kampala • Kampala District Local Government, Nakawa • Wakiso District Local Government HQs • Red-Cross-Kampala • Mbogo Schools, Kawempe- Kampala • Kampala District, Nakawa Division

References

- Abdul Latif Jameel Poverty Action Lab (JPAL), Pratham, and ASER Centre. (2009). Evaluating Read India: The development of tools for assessing Hindi reading and writing ability and math skills of rural Indian children in grades 1-5. Unpublished.
- Alassan, A. 2013. Analyse Psychométrique des Items du Pre-Test de Jàngandoo.
- Australian Council for Educational Research (ACER), Centre for Global Education Monitoring. 2014. "Desk review of four citizen-led assessments of learning outcomes." Unpublished.
- Australian Council for Educational Research, Centre for Global Education Monitoring. 2015. "[Report on the concurrent validity and the inter-rater reliability of Uwezo.](#)" Results for Development Institute.
- ASER Centre. 2013. ASER Centre Strategic Plan 2013-2016. Private submission the William and Flora Hewlett Foundation.
- ASER Centre. 2013. "[Annual Status of Education Report 2012.](#)" ASER Centre.
- ASER Centre. 2014. [National Read India Report.](#) ASER Centre Measurement, Monitoring, and Evaluation Unit.
- ASER Centre. 2014. Sampling and Data Comparison Across JUBA Countries [PowerPoint].
- ASER Centre. 2014. Tests, Tools and Testing - presentation from ASER India [PowerPoint].
- Björkman Nyqvist, M., de Walque, Damien, Svensson, J. 2014. "[Information is Power: Experimental Evidence on the Long-Run Impact of Community Based Monitoring.](#)" The World Bank.
- Bruns, B., Filmer, D., and Patrinos, H. 2011. "[Making Schools Work: New Evidence on Accountability Reform.](#)" The World Bank.
- Department for International Development (DFID). 2010. "[Learning For All: DFID's Education Strategy 2010–2015.](#)" London, UK.
- Fox, J. 2014. "Social Accountability: What Does the Evidence Really Say?" The Global Partnership for Social Accountability.
- International Budget Partnership. 2013. "[Raising the Stakes: The Impact of HakiElimu's Advocacy Work on Education Policy and Budget in Tanzania.](#)"
- Kenya National Education Sector Support Plan (NESSP 2014-2018). Forthcoming.
- Kosack, S., Fung, A. 2014. "Does Transparency Improve Governance?" The Annual Review of Political Science.
- LARTES-IFAN. 2012. Jàngandoo: Baromètre de la qualité des apprentissages des enfants au Senegal: Rapport Statistique. Dakar, Senegal.
- LARTES-IFAN. 2013. Jàngandoo: Baromètre de la qualité des apprentissages des enfants au Senegal: Principaux résultats. Dakar, Senegal.
- LARTES-IFAN. 2014a. Jàngandoo: Baromètre de la qualité des apprentissages des enfants au Senegal: Principaux résultats. Dakar, Senegal.
- LARTES-IFAN. 2014b. RAPPORT DESCRIPTIF (A) Baromètre de la qualité des apprentissages. Dakar, Senegal.
- LARTES-IFAN. 2014c. Baromètre de la qualité des apprentissages des enfants au Senegal: Résultats de l'Évaluation 2013. Policy Brief.
- LARTES-IFAN. (2014d). Jàngandoo : A barometer for the Quality of Children's Learning in Senegal. Evaluation Results 2013. Policy Brief.
- LARTES-IFAN. (Not yet published). Les Facteurs Explicatifs de la Performance des Enfants au Senegal. Rokhaya Cisse, Fatou Ndiaye, Khadidiatou Diagne, Latif Dramani, Abdou Salam Fall.
- Lieberman, E., Posner, D., and Tsai, L. 2012. "[Does information lead to more active citizenship? An Evaluation of the Impact of the Uwezo Initiative in Kenya.](#)"
- Lieberman, E., Posner, D., and Tsai, L. 2014. "Citizens, Education, and the Uwezo Initiative in Kenya."
- Lieberman, E., Posner, D., and Tsai, L. 2014b. "Does Information Lead to More Active Citizenship? Evidence from an Education Intervention in Rural Kenya." World Development 60: 69-83.

- Ministry of Finance, Government of India. 2014. [Union Budget and Economic Survey 2013-14.](#)
- Nielsen, H. Dean. 2006. ["From Schooling Access to Learning Outcomes: An Unfinished Agenda. An Evaluation of World Bank Support to Primary Education."](#) Independent Evaluation Group, World Bank. Washington, DC.
- OMAES. 2013a. Beekunko 2012: Nos enfants sont-ils en train d'apprendre réellement à l'école ? Rapport de synthèse – Evaluation 2012 des apprentissages scolaires au Mali.
- OMAES. 2013. Beekunko 2013: Ce que nos enfants savent lire et calculer - Rapport annuel d'évaluation des apprentissages scolaires au Mali.
- Omollo, Kevine. 2014. ["Knut Trashes Literacy Report"](#) Standard Digital News. The Standard.
- Pandey, P., Goyal, S., Sundararman, V. 2008. ["Community Participation in Public Schools: The Impact of Information Campaigns in Three Indian States."](#) The World Bank Policy Research Working Paper 4776.
- Piper, B. and Korda, M. 2010. ["EGRA Plus: Liberia."](#) Program Evaluation Report, RTI International.
- Planning Commission, Government of India. 2013. ["12th Five Year Plan \(2012-2017\): Social Sectors."](#)
- Salmon, L. 2013. Présentation des principaux résultats: Evaluation des apprentissages Jàngandoo [presentation]. Dakar, Senegal.
- Tanzania Ministry of Education and Vocational Training, 2013 ["Big Result Now in Education : Sector Summary."](#)
- Twaweza. 2013. ["Has Uwezo Influenced Key Actors in 2013?"](#) Twaweza Monitoring Series, Brief 3.
- United Nations. 2000. ["United Nations Millennium Declaration."](#) New York, NY.
- UNESCO. 1990. ["World Declaration on Education For All: Meeting Basic Learning Needs."](#) Jomtien, Thailand.
- UNESCO. 2000. ["The Dakar Framework for Action: Meeting Our Collective Commitments."](#) Dakar, Senegal.
- UNESCO. 2011. "EFA Global Monitoring Report: The Hidden Crisis: Armed Conflict and Education."
- UNESCO. 2011. "Tanzania Education Sector Analysis: Beyond Primary Education, the Quest for Balanced and Efficient Policy Choices for Human Development and Economic Growth."
- UNESCO. 2013. "EFA Global Monitoring Report: Teaching and Learning, achieving quality for all."
- UNESCO. 2014. ["Education for All Global Monitoring Report: Teaching and Learning: Achieving Quality for All."](#)
- Uwezo. 2014. ["Annual Plan."](#)
- Uwezo. 2011. ["Improving Learning Outcomes in East Africa: 2009-2013."](#)
- Uwezo. 2013. ["Are Our Children Learning? Literacy and Numeracy Across East Africa."](#)
- Vagh, S. B. 2013. ["Validating the ASER Testing Tools: Comparisons with Reading Fluency Measures and the Read India Measures."](#)
- Wagner, D. 2011. ["Smaller, Quicker, Cheaper: Improving Learning Assessments for Developing Countries."](#) UNESCO: International Institute of Educational Planning.
- World Bank. 1995. "Priorities and Strategies for Education: A World Bank Review." The World Bank, Washington, DC.
- World Bank. 2003. ["World Development Report 2004: Making Services Work for Poor People."](#) The World Bank and The Oxford University Press, Washington, DC.
- World Bank. 2004. ["Education for All – Fast Track Initiative: Framework."](#) The World Bank, Washington, DC.
- World Bank. 2011. ["Learning for All: Investing in People's Knowledge and Skills to Promote Development. World Bank Group Education Strategy 2020."](#) The World Bank, Washington, DC.
- World Bank. 2011. "World Bank Monitoring Report: Improving the Odds of Achieving MDGs." The World Bank, Washington, DC.



1100 15th Street, N.W., Suite #400
Washington, DC 20005
Tel: (202) 470.5711 | Fax: (202) 470.5712
info@r4d.org | www.r4d.org



**RESULTS FOR
DEVELOPMENT**