ON CALL

USING MOBILE TECHNOLOGIES TO MEASURE LEARNING IN EMERGENCIES

unicef
Office of Research – Innocenti

unicef | for every child

UNICEF OFFICE OF RESEARCH – INNOCENTI

The Office of Research – Innocenti is UNICEF's dedicated research centre. It undertakes research on emerging or current issues in order to inform the strategic direction, policies and programmes of UNICEF and its partners, shape global debates on child rights and development, and inform the global research and policy agenda for all children, particularly for the most vulnerable.

UNICEF Office of Research – Innocenti publications are contributions to a global debate on children and may not necessarily reflect UNICEF policies or approaches.

The Office of Research – Innocenti receives financial support from the Government of Italy, while funding for specific projects is also provided by other governments, international institutions and private sources, including UNICEF National Committees.

The findings, interpretations and conclusions expressed in this paper are those of the authors and do not necessarily reflect the views of UNICEF.

Suggested citation: Valenza, M., Dreesen, T. and S. Kan, *On Call: Using Mobile Technologies to Measure Learning in Emergencies*, UNICEF Office of Research - Innocenti, Florence, Italy, 2022.

ACKNOWLEDGEMENTS

This brief was authored by Marco Valenza, Thomas Dreesen and Sophia Kan (UNICEF Office of Research - Innocenti). The authors prepared this report under the supervision and guidance of Linda Jones and Frank van Cappelle from UNICEF's Education Section and Matt Brossard from UNICEF Office of Research – Innocenti. Additional support and guidance were provided by Aisling Falconer, Juan Pablo Giraldo and Neven Knezevic from UNICEF's Education section.

A special thanks goes to the following experts who provided valuable input and comments throughout the research process: Noam Angrist, Clay Westrope, Rafael Pontuschka, Maria Tubio, Cirenia Chavez Villegas, Xinxin Yu, Lee Crawfurd and Radhika Nagesh.

Amparo Barrera and Lara Stefanizzi provided invaluable administrative support to the team. Thanks also goes to Céline Little, Sarah Marchant and Sabrina Tiffany Giffen Desuasido Gill for their support with the publication and communication of this research.

Cover photo: © UNICEF/UN0393292/Satu
Graphic design: Big Yellow Taxi, Inc.
Illustrations: Rafael Pontuschka

One tool that many families own, across the globe, is a basic mobile phone.[1] The use of low-cost basic mobile phones for educational purposes in humanitarian settings is critical where access to connectivity and higher cost devices is limited. The portability of mobile phones, combined with their communication features, offers multiple uses to support education in emergency (EiE). This report is the second of a two-part series that aims to provide practitioners with guidance and resources on, and lessons learned for, using mobile technology in EiE settings. The series draws from a review of the existing literature with feedback from and interviews with EiE practitioners. The first piece of this series – *On call: The use of mobile phones for learning in emergencies* – explores how basic mobile phones can be used to support EiE learning programmes and teacher training. It emphasizes key practices that education practitioners have adopted to improve the equity and safety of mobile phone-based programmes. The present report, the second in this series, instead provides guidance on planning and delivering learning assessments through mobile technology.

1 In this paper, "basic mobile phones" are phones with the capability for text messaging and phone calls, usually without other advanced features. "Mobile technology" instead designates a broader set of portable devices including audio phones, smartphones and tablets.

© UNICEF/UN0413029/ROMERO

1. WHY MOBILE TECHNOLOGIES CAN BE USED FOR MEASURING LEARNING IN EIE

More and better evidence is urgently needed to tailor learning to the specific needs of children in emergencies. One in five refugee children still have no access to primary education, while almost 70 per cent of refugee children do not participate in secondary education (UNHCR, 2020). Where education systems are weak, even those who can access schooling are not acquiring minimum foundational skills: an estimated 70 per cent of children in low- and middle-income countries are "learning poor", meaning they are unable to read and understand a simple text by the age of 10 (World Bank et al., 2022). Measuring learning outcomes is a critical prerequisite for:

- adjusting pedagogy
- targeting instruction to the cognitive and non-cognitive abilities of children with different cultural and educational backgrounds
- equipping teachers with skills to teach these children
- informing the improvement plans of schools and learning facilities (INEE, 2020a).

More rigorous learning assessments can also be used as a basis for advocating for accreditation - a crucial building block for continued learning and employability that is often unattainable for children in emergency settings (INEE, 2020a).

Yet, children in crisis are often excluded from large-scale learning assessments. National assessment data often do not disaggregate along international protection or displacement status, on grounds of data protection or political considerations (INEE, 2020b). Such summative assessments are one-off exercises; their timing may not align with the decision-making timeline of practitioners. In addition, national learning assessments rarely consider the full spectrum of learning needs. Compared to assessments in normal settings, monitoring learning in emergency requires the additional recognition of social and emotional learning (SEL) and psychosocial support (PSS) alongside academic skills (INEE, 2020a). Children in emergencies are more likely to suffer physically and emotionally from the effects of conflict and crisis, which could disrupt their overall well-being and learning. Specific information on the learning needs and progress of those who have access to the formal school system in their host country is therefore fragmented at best. Similarly, in refugee or internally displaced people camps, where humanitarian or civil society organizations provide education, monitoring learning typically relies on ad hoc or programme-based assessments (INEE, 2020b).[2]

Mobile phones can be an effective complementary tool to traditional tests for assessing the learning needs and progress of children in emergencies.

2 For example, see: Twaweza-Uganda (2018); Piper et al. (2020).

They can make data collection less time-consuming and more frequent in a context of high mobility and where the timing of feedback to educators to act is critical. Phone-based assessments conducted in low- and middle-income countries have also been shown to be cost-effective, while addressing common logistical hurdles and costs related to processing and administering paper-based assessments.[3] Data on learning performance collected through mobile phones can unlock the potential of ability-based instruction. In the context of a recent remote learning experiment in Botswana, primary school children who received targeted modules based on learning assessment data collected through mobile phones showed improved scores in key numeracy skills (Angrist, Bergman and Matsheng, 2022).

Given their ubiquity, mobile phones can reach marginalized children who do not attend either formal or informal schools and learning centres – hence generating much needed and more frequent insights into the learning levels of the most vulnerable children. As a nascent practice, however, the use of mobile phones to conduct assessments requires careful planning, piloting and contextualization. While pick-up rates among sampled households were generally satisfactory (Angrist et al., 2020; UNICEF Nepal 2021), in some cases the phone-administered assessments yielded different results than validated pencil-and-paper tests (Crawfurd et al., 2020). This calls for additional research and adaptation. Once reliability has been rigorously verified, advocacy efforts to formally recognize phone-mediated results should be undertaken; this could be facilitated by involving the relevant departments of a country's Ministry of Education in the assessment process. Besides these issues of reliability and accreditation, other key challenges linked with phone-based assessments that are explored in this brief include mobilizing communities (especially caregivers) for the exercise, getting the research design right (including sampling plans), adapting children's safeguarding and data confidentiality protocols to the phone set up and training and supervising assessors, especially when assessments are administered at a distance.

The next section outlines the key implementation steps that practitioners may consider when planning for a mobile technology-based learning assessment.

3 For instance, an assessment using audio phones in Botswana cost US$4.40 per child, compared with a minimum of US$62.5 per child for in-person assessment (PIRLS) and US$150 for classroom-based assessment in Liberia (Angrist et al., 2020). In Nepal, a UNICEF phone survey with caregivers of children below 5 years old cost US$5 per household (UNICEF Nepal, 2021).

© UNICEF/UNI358638/CRISTOFOLETTI

© UNICEF/UN038175/SINGH

2. PLANNING MOBILE TECHNOLOGY BASED ASSESSMENTS

A scoping phase is critical for practitioners to determine if, and how, the use of mobile phones could be leveraged to track children's learning in EiE settings. This subsection presents key initial steps that practitioners should consider, referencing resources to which they could turn.

1. DETERMINE THE SKILLS TO BE MEASURED AND THE TYPE OF ASSESSMENT

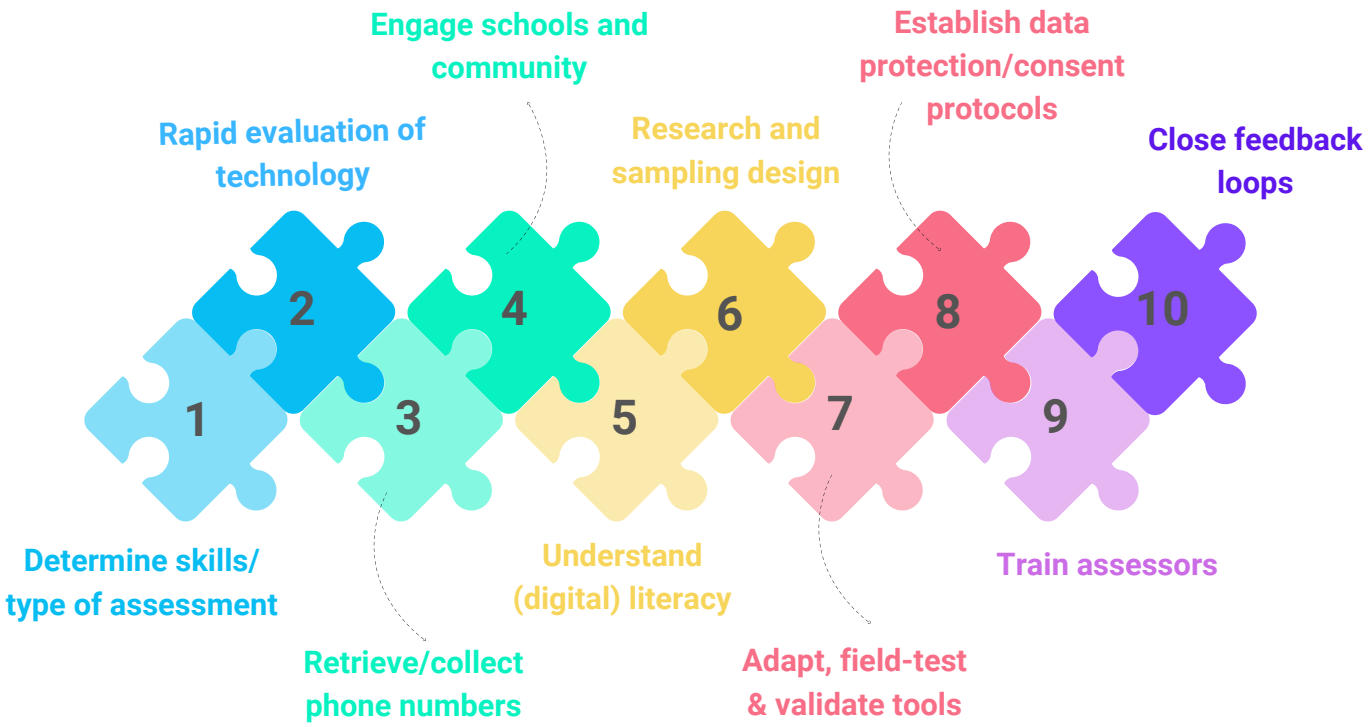
Skills. INEE's minimum standards provide a framework for determining the skills that may be considered for assessment in emergency settings (INEE, 2020b). The minimum standards outline the functional literacy, numeracy and essential life skills that are required to attain a life of dignity and the ability to participate meaningfully in a community (INEE, 2020b). As mentioned above, it is recommended that, besides academic skills, EiE practitioners also consider tracking SEL as well as the needs for psychosocial support (PSS).

The type of skills that practitioners may want to track differ across age groups. Assessment tools for early childhood education have, for instance, focused on measures of school readiness and child development. UNICEF Nepal (2021) has recently conducted a phone-based assessment of early childhood learning, presented in Section 3. Children of different ages may also require differentiated guidance during the assessment, with younger children likely to need more personalized support.

Type of assessment. Assessments are usually categorized as summative (high-stake examinations) or formative (continuous assessment assigned by teachers in the classroom or as homework, to monitor the acquisition of skills and concepts). Summative assessments serve different accountability purposes. Students' progress is assessed across a specific skill set or a comprehensive set of curricular skills to determine grading, progression or certification. Assessment results can also be used by the education system at different administrative levels to verify school/teacher performance and to allocate funding and personnel; as well as by caregivers to keep schools accountable.

Children in emergency settings face multiple barriers in accessing a summative assessment, with the risk of negative repercussions for their progression in school or the achievement of certifications. As mentioned before, many are out-of-school or not taken in by the formal school system of the host country. For those who are enrolled in the formal system, sitting the exams may simply not be feasible on grounds of security threats (e.g., in war-affected zones) or public health grounds (e.g., the Ebola or

FIGURE 1: the 10-step framework to plan learning assessments through mobile technology |



COVID-19 outbreak). Summative assessment can also become a stumbling block if the stakes of the assessment are too high for children whose education has been disrupted due to a crisis or emergency. Another criticism that especially applies to children who could not attend school regularly due to an emergency is that high-stake decisions are based on a single exam, which raises concerns of fairness (UNESCO, 2017). More generally, summative assessments encourage teachers, schools and communities to prepare for a single examination, hence discounting many of the subjects and competencies (such as twenty-first century skills) which standardized assessments do not, or only narrowly, appraise (UNESCO, 2017).

Due to these concerns, international education actors have called for the focus to shift from high-stakes exams to formative assessments, to promote learning and determine progression/certification during a crisis (UNESCO, 2020). Evidence shows that learning loss can be mitigated by maintaining close interactions between teachers and students, including through continuous assessment and especially for disadvantaged groups (UNESCO, 2020). The authentication of the student's work could be an issue, however, and therefore measures must be put in place to minimize cheating (UNESCO, 2020).

Most of the phone assessments conducted to date have been developed to evaluate specific remote learning programmes (Angrist et al., 2020; Save the Children, 2022a). In this respect, mobile phones served formative assessment purposes, as tools for assessors/teachers to assess specific foundational skills that children were learning. Conversely, at this early stage of development and use, phone-based assessments may be less appropriate for high-stake assessments that need in-person guidance and monitoring during the administration of the exercise – and that may be less relevant, fair or feasible in EiE contexts, as explored above.

2. CONDUCT A RAPID EVALUATION OF LOCALLY AVAILABLE TECHNOLOGY

If localized information on technological access is unavailable, field teams should conduct a rapid evaluation in the target sites/communities to determine patterns in ownership of basic phones and smartphones, as well as to assess the reliability of the electrical grid and of the phone network coverage for both audio calls and internet data. Field teams could also retrieve or collect information on household size and the ratio of devices to household members. Identifying service providers, including mobile aggregators, and their costs is also

necessary if practitioners are considering the use of short message service- (SMS) or interactive voice response- (IVR) based assessment platforms (for a checklist of technological requirements here, see Khurana et al., 2022).

The findings of the rapid evaluation of technological access are key to inform strategies on how to distribute phones, smartphones/tablets and/or solar chargers to boost respondents' uptake. For instance, as part of the "Sauti za Wananchi" initiative in East Africa, 2,000 respondents of a longitudinal phone survey received mobile phones and solar chargers based on localized data on phone ownership and access to grid.[4] In Uganda, one of the countries where the initiative was implemented, only 19 per cent of sampled households were connected to the national grid and only 25 per cent in the bottom income quintile had access to mobile phones (82 per cent in the middle, 95 per cent in the top quintile) (Twaweza-Uganda, 2017). Likewise, in Kenya, another implementing country, grid-connected households amounted to 33 per cent of the sample, with phone penetration ranging from 72 per cent (poorest quintile) to 98 per cent (richest quintile) (Twaweza-Kenya, 2016a).[5] Leveraging other existing programmes that use technology as a medium for learning could lead to significant synergies and cost savings (e.g. education programmes distributing tablets/smartphones to learning facilities or households). Mapping such programmes and implementing actors is therefore recommended during this stage.

3. ASSESS THE FEASIBILITY OF RETRIEVING OR COLLECTING PHONE NUMBERS OF CAREGIVERS

It is important to identify whether a database already exists or could be collated from different sources (e.g., schools, NGOs or international organizations) and whether it can be accessed. Alternatively, field teams may need to conduct primary data collection. In East Africa, the baseline round of the "Sauti za Wananchi" longitudinal phone survey was conducted in person. While this approach drives up initial costs, it also offered programme officers and enumerators the possibility of engaging respondents and community leaders face-to-face, with reported benefits on subsequent uptake of the survey, and was an opportunity to distribute phones and solar chargers for the data collection rounds.

4. ENGAGE CAREGIVERS, TEACHERS AND LOCAL LEADERS FROM THE ONSET

Caregivers are crucial allies - not only because obtaining their consent is mandatory but also because they serve as co-facilitators when administration occurs remotely (see Point 5, below). However, practitioners should be aware that getting caregivers to pick up the phone may be a challenge. In some contexts, social and cultural norms may hinder the sharing of information, especially when children are involved (Angrist et al., 2020). Competing priorities and limited educational background can further reduce response rates.

Several strategies exist to promote engagement, including:

- mobilizing caregivers and teachers in the communities from the start, explaining the goals and the practical steps of the assessment, as well as getting approval from local leaders (motivating teachers to participate in the assessment exercise may require a top-down involvement from the MoE, in crisis contexts where this is functioning/operating)
- identifying the right caregiver and assessing their willingness to let their children be interviewed directly (under the caregiver's supervision)
- scheduling phone calls in advance and keeping a flexible schedule, while planning to make 3–5 follow-up calls for each household.
- informing caregivers that they could more closely track their children's learning progress, engaging their curiosity and empowering them in their interaction with the school (see Point 10).

4 While collecting school-related data, this phone-based survey was not specifically designed to measure learning. These statistics on phone coverage and grid connections nevertheless remain a useful benchmark for practitioners to inform strategies on the distribution of devices.

5 The discrepancy between grid-connected households and phone owners can be explained by the use of alternative power sources, such as solar, to charge phones; in some contexts, households may exploit grid/solar-powered neighbors or community venues, such as health centres, as charging stations for their own devices.

5. UNDERSTAND THE (DIGITAL) LITERACY SKILLS OF CAREGIVERS

Determining the contribution that caregivers can provide during the assessment is important, especially if practitioners plan to conduct the test remotely. Rolling out such an assessment requires caregivers to have basic literacy skills to effectively support their children in administering the tasks. For instance, the low-access modality of the Remote Assessment of Learning (ReAL) tool, which relies on a basic feature phone, asks caregivers to copy words and full sentences dictated over the phone by an assessor for the written comprehension modules (Save the Children, 2022a – see Section 3).

Similarly, remote phone testing requires that caregivers possess sufficient digital skills to manipulate a smartphone or an audio phone to follow an assessor's instructions. The stimuli[6] that are part of the learning assessment can be shared via SMS (Angrist et al., 2020) or through low-tech messaging platforms such as WhatsApp (Save the Children, 2022b) – which may reduce the degree of technological skills required. Yet, taking the proficiency of caregivers for granted may lead to the exclusion of vulnerable households incapable of assisting their children, with implications for equity.

6. DEFINE A RESEARCH DESIGN AND SAMPLING PLAN

Practitioners will need to liaise with a partner with expertise in statistics or research, in order to develop a research and sampling plan that equitably represents the population of the target areas. Equity considerations are key when conducting phone-based assessments, since the most marginalized who do not have access to phone/phone network or live in hard-to-reach areas may end up being under-represented in the survey.

The appropriate sample size for the assessment (if sampling is needed, i.e., if not all learners can be assessed) should be carefully calculated; this depends on contingent factors such as the socioeconomic, demographic and ethnic heterogeneity within the target areas. Keeping that in mind, the experience of Uwezo – a large-scale learning assessment in East Africa (see Box 1) – provides a general benchmark of 300 households per district (with populations ranging from 450,000 to 850,000 inhabitants) as a sufficient sample for the measurement of literacy and numeracy skills (Twaweza-Uganda, 2018).

Some phone-based assessments require a preliminary household listing to be carried out, as a basis (sampling frame) from which to (randomly) draw respondents. Practitioners should first verify whether existing sampling frames – such as previous national household surveys or a Cluster-populated database – could be accessed; this would be a cost/time-effective option to consider if recent and solid data is available for the target population.

7. ADAPT, FIELD-TEST AND VALIDATE THE ASSESSMENT TOOLS

Assessing academic and socioemotional skills requires passing rigorous adaptation, field piloting and validation procedures. Testing protocols that have been validated with the general student population may not be appropriate for children who have experienced conflict or crisis; and may cause additional frustration and stress (INEE, 2020b). Practitioners should therefore build on existing assessment tools that have been already validated for EiE, even if most of these have been developed for traditional face-to-face administration. A review conducted by INEE (2020b) documents a few examples of such tools:

In addition, while not specifically conceived for EiE, the Early Grade Reading Assessment (EGRA) and the Early Grade Mathematics Assessment (EGMA) have also been deployed in emergency contexts in different countries and languages.[7]

According to INEE (2020b), the assessment tools that are viable for EiE contexts should fulfil the following requirements:

6 Stimuli are any accompanying audio or visual inputs (such as a reading passage, a drawing or quiz) that is used during the assessment to present a task. Some of these supports can be shared with the respondent through the phone (SMS, WhatsApp, etc.).

7 For instance, EGRA/EGMA were recently deployed in hard-to-reach provinces of Afghanistan (Kan et al., 2022), in conflict-afflicted areas of Niger (Concern Worldwide, 2022) and in South Sudan (Cynosure Evaluation, 2020).

TABLE 1: Assessment tools that have been validated for EiE. Source: INEE, 2020b

INSTRUMENT	SKILLS	AGE OR GRADE(S)
Holistic Assessment of Learning and Development (HALDO)	FLN + SEL + executive functioning	4–12 years old
International Development Early Learning Assessment (IDELA)	Motor development, emergent literacy, emergent numeracy, SEL	3.5–6 years old
UNRWA Monitoring Learning Achievement (MLA)	FLN + classroom practices and school environment	G4–G8
Out-of-School Youth Literacy Assessment (OLA)	Foundational literacy	Youth and young adults

- minimum training required for the assessment administrators
- administrators can be volunteers and community members
- flexible assessment location (household, school, community centre)
- broad age range
- inexpensive administration
- short assessment time: some of the most promising remote phone-based assessments did not exceed 15–20 minutes (see Angrist et al., 2020 and Table 1)
- oral-based assessing for which few physical materials are required.

When introducing technology as the way of administering the assessment, practitioners will have to:

- suit the tool to local settings, by translating its items and protocols into local languages
- adapt the various stimuli packages so that they are relatable to children and are culturally appropriate
- confirm that the assessment is compatible with available phones.

This process requires adequate time and resources, and practitioners should plan accordingly. In practice, this process will require, at a minimum, consideration of:

- **Internal validity:** simplifying or modifying the tool to be compatible with (smart)phones will require an iterative process of field piloting, analysis of pilot results and adaptations. Methodological approaches to assess the tool's internal validity include internal reliability (Cronbach's Alpha), factor structure or item analysis using item response theory models (Angrist et al., 2020). Angrist, Bergman and Matsheng (2022) provide insights on using easily deployed approaches to determine internal validity, including randomizing assessment problems of the same proficiency across respondents, disentangling the measurement of cognitive skills and effort levels (through incorporating an appropriate real-effort task) and so-called "known-groups" validity tests.[8]
- **Concurrent validity:** comparing pilot results of phone-based assessments with validated face-to-face tools administered under similar contexts, in order to check consistency.[9]

Some assessment tools are being specifically developed for phone use. Section 3 presents, for instance, Save the Children's Remote Assessment Tool (ReAL), which has been piloted to measure FLN and SEL skills. ReAL will be tested in at least nine new sites by the end of 2022, including in emergency contexts. Save the Children has partnered with the Child Study Center at Yale University as an external validator, with the goal of launching a fully validated instrument as a global public good in 2023.

The length of the contextualization process may often clash with the urgency of acting, especially during rapid onset emergencies or quickly escalating crises. In these situations, practitioners should rely on already field-tested instruments, prioritizing those that have:

8 Known-groups tests compare expected patterns in learning outcomes (e.g. that upper grade students score better on foundational skills than lower grade counterparts) with observed trends in the sample. A diverging pattern between expected and observed patterns signals validity; no detectable differences may raise validity concerns.

9 Angrist et al. (2020) provides a concrete example of how concurrent validity was accomplished in Botswana.

TO LEARN MORE ABOUT...

Contextualizing learning assessment tools:

- Research Triangle Institute International (RTI), Early Grade Reading Assessment (EGRA) Toolkit: Second Edition - Chapter 6
- Research Triangle Institute International (RTI), Early Grade Mathematics Assessment (EGMA) Toolkit - Chapter 4

Adapting learning assessments for use with phones:

- Angrist, Noam et al., Practical lessons for phone-based assessments of learning.
- Angrist, Noam, Peter Bergman, and Moitshepi Matsheng, (2022). Experimental evidence on learning using low-tech when school is out.
- Luna-Bazaldua, Diego et al., Psychometric Considerations for Implementing Phone-Based Learning Assessments (English).

Adapting learning assessment to EiE contexts:

- Inter-agency Network for Education in Emergencies (INEE), Mapping Exercise: Assessment of Academic Learning Outcomes.

*see References for more details on these publications

i. a simpler structure, and hence require less intensive adaptation (this would come at the cost of assessing skills less comprehensively than more structured tools such as EGRA/EGMA but would still allow to track progress in key FLN skills, in a timelier manner)

ii. been conceived for phone-mediated delivery (e.g., ReAL) or can be easily adapted to phone-mediated delivery

iii. been deployed in similar contexts or with similar target populations (in this respect, practitioners could verify the geography of widely deployed instruments such as EGRA and EGMA, which have been administered through the Tangerine software in more than 60 countries and 100 languages).[10]

8. ESTABLISH DATA PROTECTION AND CONSENT PROTOCOLS

Practitioners should establish protocols for:

- data protection, to ensure confidentiality and data security throughout the data collection, transfer and analysis process
- seeking caregivers' informed consent to let their children participate in the exercise
- seeking children's assent (if applicable, i.e., if children are spoken to directly as part of the assessment).

Switching from face-to-face to remote, phone-based assessments requires some adaptations to the protocols. For instance, consent will have to be sought verbally; calls may be recorded (at least at the beginning) to enable compliance checks by the data collection managers or research team. Since phone-mediated communication is less personal and interactive, the script for seeking caregivers' and children's consent should be extremely clear and concise, including the:

- goals of the assessment (especially that it does not affect school grades)
- confidentiality protocols that are in place, the benefits (including whatever incentive may be extended)
- risks of participating in the exercise.

Save the Children (2022a; page 12) provides an example of a consent form for phone-based learning assessment.

The box below contains guidelines for complying with ethical requirements when conducting learning assessments or phone-mediated surveys.

10 Use Around the World — Tangerine (tangerinecentral.org).

TO LEARN MORE ABOUT...

Ethics in conducting learning assessments:

- Research Triangle Institute (RTI) International, Early Grade Reading Assessment (EGRA) Toolkit: Second Edition - Chapter 2

Adapting ethics protocols to phone set up:

- Angrist, Noam, et al., Practical lessons for phone-based assessments of learning
- Kopper, Sarah, and Anja Sautmann, Best practices for conducting phone surveys

Adapting learning assessment to EiE contexts:

- Save the Children International, Remote Assessment Of Learning - Low Access Administration Guidance – p. 12.

*see References for more details on these publications

9. TRAINING ASSESSORS

Sufficient time should also be allocated for training assessors.[11] The training should also include a pilot exercise where enumerators can practise the protocols and tools and familiarize themselves with the devices.

At a minimum, the training should cover:

- the technology of devices and associated software: smartphone/tablet, computer-assisted telephonic interviewing (CATI) and audio phones, depending on the specific modality of administering the assessment
- mastering the assessment tools, including giving adequate instructions to caregivers/children, timing and sequencing of the tasks, providing prompts, grading and entering data
- informed consent and data protection protocols
- identification, recruitment (in the field or remote) and replacement protocols of respondents.

Pilot data should be carefully analysed, not only to verify the internal reliability of the assessment (Point 7) but also to ensure inter-rater consistency, i.e., that different enumerators grade in a similar manner. In practice, this may entail having two enumerators simultaneously score the same child during the pilot for a restricted sample of children (for instance, Save the Children targeted 10–20 per cent of the sample for paired assessment in the piloting of its ReAL instruments) and compare the assessment scores assigned to the pair to resolve any source of discrepancies.

10. CLOSE FEEDBACK LOOPS

Learning assessment data must translate into action in the classroom and at home. Setting up a user-friendly data management system to analyse and disseminate data to teachers, heads of school and government officers is crucial to enable stakeholders to act upon identified learning gaps. Practitioners will need to prepare the underlying governance system, defining who is responsible for analysis and dissemination, how frequently this is carried out, and to identify with whom they will share evidence with.

Reporting back to government and schools not only allows officials and teachers to take action where most necessary but also incentivizes their participation in the assessment exercise. National and local government officials, school directors and especially teachers, as the primary actors who are called to act upon the assessment results, should be a key audience for this type of data. Mobile phones can be used as a medium to close feedback loops. For example, in Senegal, the directors of 9,500 schools were trained to submit students' standardized test results via SMS using basic mobile phones (Koomar and Blest, 2020). The test results were aggregated by a dedicated dashboard (US$299 per month with an annual subscription) and accessed by regional and district officials for real-time oversight (Koomar and Blest, 2020).

11 Also refers to "enumerators" or "interviewers"; the three are used interchangeably in this brief.

© UNICEF/UN0493975/SANTOS

3. CHOOSING THE RIGHT DELIVERY METHODS AND TOOLS

After completing the initial steps outlined above, practitioners will have gathered the key elements to select the most suitable method for administering the assessment through mobile technology.

This section considers four main categories of delivery modalities of learning assessments for EiE:

1. **Live phone calls**, with stimuli such as reading passages or math quizzes delivered through SMS (or WhatsApp)
2. **Asynchronous**[12] **SMS-based quizzes**
3. **Dedicated software for assessment through tablets**, which requires field teams of trained assessors to visit sampled households for administering the test
4. **Learning platforms with built-in assessment**, which can be accessed through basic mobile phones at affordable costs

Assessing learning through live phone calls requires technology that is commonly available even in emergency settings (mainly basic mobile phones). Some tools require heavier engagement from caregivers, meaning that practitioners will need to ensure that caregivers have sufficient literacy skills to accompany their children during the administration of the test.

Live phone calls enable two-way interaction with the respondents, allowing for direct nudges to complete the tasks (or to reschedule the test for a more convenient time). Testing of different delivery methods in Ghana showed that most participants preferred voice calls to automated forms of testing such as through SMS or IVR (Khurana et al., 2022). The engagement rates of live phone calls were higher than IVR which, in turn, were higher than SMS. The main reasons for the lower take up of automated tests (Khurana et al., 2022) included:

- not being with the child when receiving the SMS- or IVR-broadcast content
- not understanding audio recordings
- inability to navigate through the options or use the keypad and phone-related challenges (network, broken screen, dead battery, etc.)

SMS-based systems require setting up data management and mass distribution systems, typically through mobile aggregators.[13] On the other hand, automating the assessment process can substantially reduce administration costs: the roll out of SMS-based platforms cost less than US$1 per participant in Ghana, compared with US$168 for the IVR system and US$250 for conducting live phone calls (Khurana et al, 2022). The automated marking of quizzes can also provide learners

12 Asynchronous quizzes can be answered at any time and do not require any live interactions between the respondent and the assessor/teacher; answers to the quiz are keyed in on the mobile phone's dial pad.

13 Mobile aggregators provide a platform to mass distribute (and sometimes also receive) SMS to different mobile phone operators.

TO LEARN MORE...

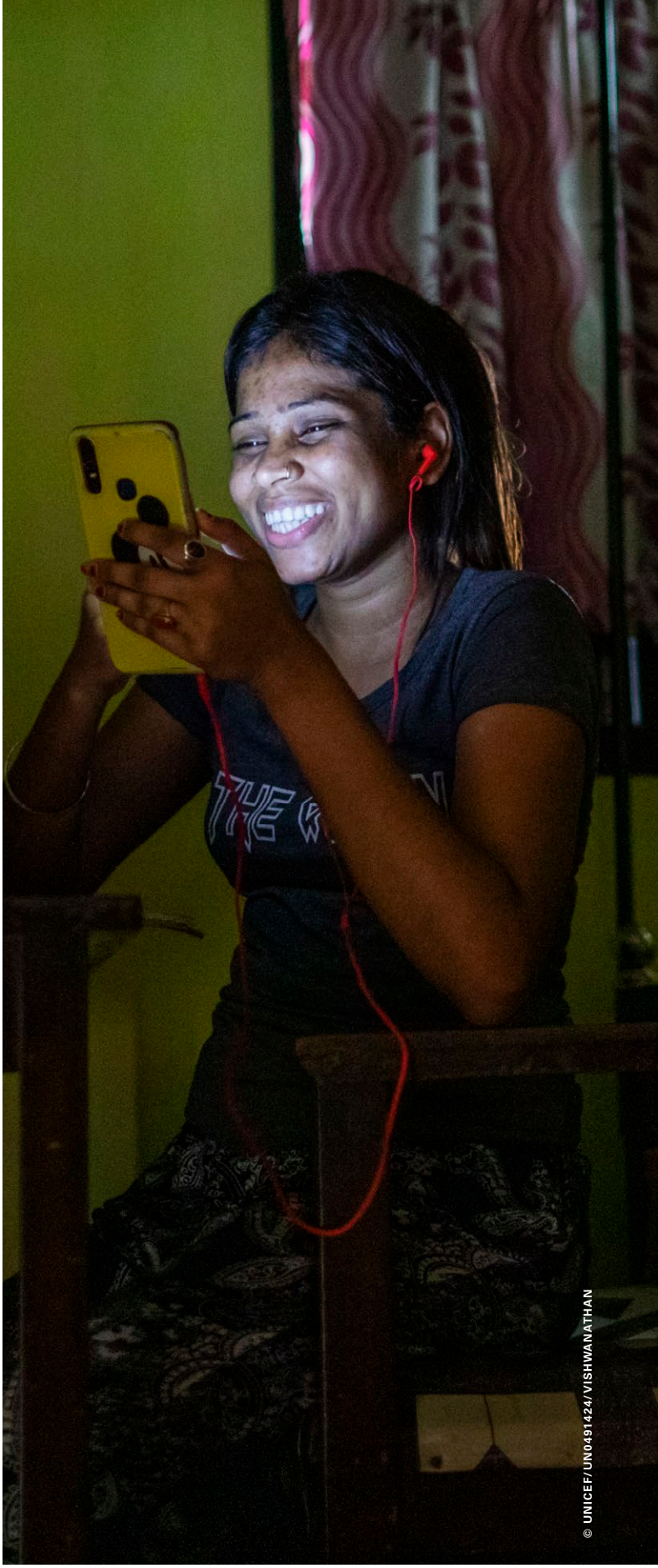
As an additional channel for assessing foundational numeracy tasks, the World Bank and its partners also piloted **IVR-based assessment system** in Ghana. This relied on basic feature phones and could be reached on a toll-free basis by learners. Section 2.2.3 of the Synthesis Report on Piloting of Remote Phone Based Assessment presents the implementation steps required to set the system in place and documented relevant lessons learned.

with instant feedback, while allowing for progressively adjusting the users' tasks, based on ability.

Tablet-based assessments through dedicated software, such as Tangerine, may require less extensive piloting and adjustments than phone-based tests conducted remotely, as they have been field-tested across different contexts, including in emergencies. They also have the advantage of generating user-friendly analytics that can be easily reported back to teachers and programme staff and used for decision-making and tailoring instruction. These tests, nonetheless, are typically implemented on more expensive technology (tablets) which need to be supplied to enumerators, and require in-person administration from trained field teams.

Finally, some learning platforms with built-in formative assessments have been conceived for basic phones and can be accessed at affordable costs. These have the advantage of being engaging (using technology) for end users and of automatically adapting learning content based on in-built assessment. Yet, the use of these learning platforms at home relies on a child's motivation to complete the tasks and requires follow-up from teachers (or teaching aids) and/or caregivers. Being formative in nature, these assessments are not meant to offer a comprehensive and representative overview of children's competencies but rather aim to tailor learning contents based on key abilities.

Table 2 presents methods to track different types of skills (FLN and SEL). It also features related field-tested tools, highlighting the digital and literacy skills and the technology required for conducting the testing, as well as lessons learned for effective implementation.



© UNICEF/UN0491424/VISHWANATHAN

TO LEARN MORE...

This section focuses on relatively low-tech solutions, given the common technological and budget constraints in EiE settings. Mohn et al. (2022) provides snapshot case studies on **higher-tech solutions such as learning apps with built-in exercises**. The analytics generated by these built-in exercises, which can be online or offline, shapes the learning content of users, so that they can progress at their own pace based on their ability. These digital solutions typically require supplying tablets to schools or households.

Another tool being piloted in an EiE setting in Jordan is **War Child's "Gobee"**. Gobee is a gamified learning tool with built-in assessments of foundational literacy and numeracy as well as socioemotional learning skills. It comes with a dashboard that generates data for school and programme staff. The Humanitarian Education Accelerator (2022) presents key lessons learned from the pilot phase of the tool in EiE settings in this blog.

BOX 1. Practical lessons learned from East Africa

Twaweza has been a leading actor in large-scale learning assessments and phone-based longitudinal surveys in East Africa (Kenya, Tanzania and Uganda). Its "Uwezo" initiative aims to collect learning data on foundational literacy (English and Kiswahili) and numeracy for a nationally representative sample of children aged 6–16 in each participating country. Similarly, Twaweza launched a nationally representative phone survey, called "Sauti za Wananchi", in the three countries in which it operates to measure citizen perceptions of public service delivery across a range of policy areas, including education and health. Trained enumerators at designated call centres contact respondents every month using CATI. Although data collection occurs in-person and mainly through paper-and-pencil stimuli for Uwezo[14] and even if Sauti za Wananchi does not include individual learning assessments, practitioners can draw practical lessons from the two initiatives:

- **Involve government entities at different levels from the early stages:** this not only increases officials' buy-in and ownership, but may also reduce the lengthy process of obtaining the necessary research permits (Twaweza-Tanzania, n.d.).
- **Engaging residents from the community in following-up with non-respondents in phone surveys.** Everyone is called at least three times. In every area, participants elect a team leader who helps the callers locate respondents who do not pick up. These monitors were effective in persuading sampled respondents to participate, since they are known to them (Twaweza-Tanzania, 2013).
- **Mobilize the community in conjunction with local leaders prior to starting the survey work.** Accessible information should be provided regarding the goals of the survey, choice of participants, data protection, incentives, use of phones and other hardware provided (including a practical demonstration of solar chargers etc.). In Kenya, booklets in English and Kiswahili were useful in conjunction with community and town hall meetings to explain the goal and the implementation steps of the survey (Twaweza-Kenya, 2016b). The field visit should also be used to identify residents to support the data collection (see point below).
- **Consider distributing devices to phone survey respondents.** Audio phones and solar chargers were distributed to 2,000 participating households. Time and resources were allocated to explaining the role of the devices in the context of the survey, and to avoid tensions in the community between those who received the devices and those who did not. The consent of the head of household should be explicitly sought, as evidence from Tanzania showed this may fuel tensions that may even result in gender-based violence (Twaweza-Tanzania, 2013). The registration process for the associated sim cards should also be agreed upon with the communities, as some participating households got suspicious when offered sim cards registered by the research team (Twaweza-Tanzania, 2013).
- **Consider providing incentives to participants.** Respondents get airtime for every question round completed. The incentives should be clearly explained (timing, amount disbursed) during the preliminary field visit.
- **Put emphasis on closing feedback loops in order to increase participation at the community level.** Uwezo embedded instant feedback mechanisms at the end of the administration of the assessment to inform participating households of the ability level of their children, with a view to increase accountability by the providers of basic education.

14 For an example of the printed assessment pack deployed in Uganda in 2021, see: https://uwezouganda.org/wp-content/uploads/2022/01/Uwezo-2021-Assessment-Pack-Eng-Lusoga-_-Num_final.pdf. Twaweza has also adapted its assessment pack to refugee contexts in Uganda.

TABLE 2: Delivery modalities and tools for implementing learning assessments

A. LIVE PHONE CALLS

1. UNICEF NEPAL, 2021

Target skills	Child development and early learning
Age/grade	2-4 years old
Technology	Audio phone
Delivery	Remote
Delivery of stimuli	Phone (voice)
Respondent	Child's mother
Country	Nepal
Digital/literacy skills required	No specific need for literacy
Instrument	Phone-adapted version of the ECD Index 2030
Cost	US$5 per household (HH)

Key implementation steps and lessons learned

- Involve MoE in the selection and adaptation of the assessment tool.
- Carry out a pilot to confirm the concurrent validity of the phone-adapted ECD Index 2030 vis-à-vis the originally validated face-to-face assessment.
- A suite of practical resources is available online, including instructions for interviewers.

2. ANGRIST ET AL., 2020

Target skills	Numeracy
Age/grade	Grades 3–5
Technology	Audio phone
Delivery	Remote
Delivery of stimuli	SMS
Respondent	Child, with caregiver's supervision
Country	Botswana
Digital/literacy skills required	Minimum literacy to read SMS
Instrument	Phone-adapted version of ASER assessment, with stimuli shared via SMS during the call
Cost	US$5 per HH

Key implementation steps and lessons learned

- Keep questionnaire's length to 15–20 minutes. Logistics (scheduling calls, understanding who's the best-placed household member to assist the child, explaining, etc.) filled up to 50 per cent of call time.
- Keep instructions simple and use practice items to ensure that the target skills – rather than mere language or processing skills – are measured.
- Sending a text message to alert respondents of the upcoming call increases pick-up rates.
- Holding an advance call with a caregiver may increase accuracy, honesty and a willingness to participate.
- Have initial assessment instructions delivered through a caregiver with requests to put the child at ease.
- Minimize sample bias (i.e., participants not responding via phone) by relying on local people (e.g., neighbours who could borrow a phone to the non-respondents who do not own one) and by documenting the bias itself.
- Set a limit of two minutes per question to minimize the likelihood of family members assisting the child; require each child to explain their work, only marking a problem as correct if the child could explain how they solved it (Angrist, Bergman and Matsheng, 2022).

A. LIVE PHONE CALLS

3. CRAWFURD ET AL., 2021

Target skills	Literacy and numeracy
Age/grade	Grades 3–6
Technology	Audio phone
Delivery	Remote
Delivery of stimuli	Phone (voice)
Respondent	Child, with a caregiver's supervision
Country	Sierra Leone
Digital/literacy skills required	Minimum literacy to read SMS
Instrument	Combination of items from EGRA and EGMA, ASER and Banerjee et al. (2017)
Cost	N/A

Key implementation steps and lessons learned

- Sample bias: students who lived outside of the capital or whose parents did not complete school were statistically less likely to be reached by phone – suggesting potential underrepresentation of the most marginalized.
- Consider in-person tracking at the community level to boost participation in the assessment.
- The timing of the call influences accuracy: working parents had to take the calls at noisy and distracting locations (e.g., a marketplace) or in the evening (fatigue). Aim to schedule calls when parents are at home with the child or in a quiet environment.
- Discrepancy between face-to-face and phone assessment scores likely due to students interviewed by phone being helped by caregivers, despite the request of interviewers not to do so (see cell above for mitigating measures).
- Offer children the option of answering the questionnaire in the preferred local language.

4. KHURANA, LEVIN, LUNA-BAZALDUA & LIBERMAN (2022)

Target skills	Foundational math knowledge (place value tasks)
Age/grade	Grades 2–5
Technology	Audio phone
Delivery	Remote
Delivery of stimuli	Phone (voice)
Respondent	Child, with a caregiver's supervision
Country	Ghana
Digital/literacy skills required	Low: no specific need for literacy
Instrument	Adapted from Angrist (2020)
Cost	N/A

Key implementation steps and lessons learned

- The National Council for Curriculum and Assessment collaborated in developing and contextualizing the assessment tools.
- A virtual call centre was set up with toll-free long codes created to make the calls and send/receive SMS.
- The average assessment phone call time was 13 minutes.
- Respondents were recruited through a mix of random digit dialling and referrals from enrolled household (the latter was introduced considering lower-than-expected eligibility rates among randomly called households). First contact call lasted on average 20 minutes.
- The importance of determining disability status was underlined in allocating children to the most appropriate delivery mode: visually-impaired children were assessed through phone calls, and those who were deaf/hard of hearing received SMS quizzes.

A. LIVE PHONE CALLS

5. KHURANA, LEVIN, LUNA-BAZALDUA & LIBERMAN (2022)

Target skills	Foundational math & literacy (reading accuracy & comprehension)
Age/grade	Grades 4-5
Technology	Audio phone
Delivery	Remote
Delivery of stimuli	SMS
Respondent	Child, with a caregiver's supervision
Country	Nepal
Digital/literacy skills required	Minimum literacy to read SMS
Instrument	ASER-adapted tool (literacy) + numeracy tools previously developed by Teach for Nepal
Cost	N/A

Key implementation steps and lessons learned

- The instrument's adaptation process aligned the 12-item questionnaire to the national curriculum.
- Live phone calls and SMS were used in a complementary way to assess different sets of skills (e.g., sentence reading was assessed through calls but not by SMS).
- Numeracy tasks were read aloud and explained over the phone. Literacy tasks were sent in real time during the phone call to minimize cheating.
- Phone numbers were obtained from local governments – 94 per cent were eligible to be assessed over phone.
- Airtime transfer (Rs. 100) incentivized participation.
- Average phone call time was 45–60 minutes.

6. KHURANA, LEVIN, LUNA-BAZALDUA & LIBERMAN (2022)

Target skills	Foundational literacy
Age/grade	Grades 3–5
Technology	Audio phone
Delivery	Remote
Delivery of stimuli	SMS
Respondent	Child, with a caregiver's supervision
Country	Pakistan
Digital/literacy skills required	Minimum literacy to read SMS
Instrument	Adapted from a mix of EGRA and ASER items
Cost	N/A

Key implementation steps and lessons learned

- Multiple SMS with literacy tasks were sent in advance to the participant prior to the live phone call.
- Call lists were provided by provincial education representatives (provincial network) and verified by enumerators (teachers).
- Tabadlab (national implementing partner) set up an audio phone-compatible SMS broadcast application. At the beginning of the assessment, the system sent the assessment items in local language script through the SMS broadcast feature.
- In cases of no, or incomplete, responses, reminders were sent via SMS or live phone calls to prompt the student to complete the quiz. Reminders were also sent to an alternate contact number.

A. LIVE PHONE CALLS

7. REAL TOOLKIT (SAVE THE CHILDREN, 2022A & SAVE THE CHILDREN, 2022B)

Target skills	Literacy, numeracy and SEL
Age/grade	5–14 years old
Technology	Audio phone (low-access version) or basic smartphone (high-access version)
Delivery	Remote
Delivery of stimuli	Low-access version: pre-distributed printed materials and practice items, or stimuli produced on the spot using learning materials in the home with guidance from enumerator High-access version: WhatsApp (or SMS as a second-best option)
Respondent	Child, with caregiver's active participation in the testing
Country	Various
Digital/literacy skills required	Caregivers able to write, read and count (1–20)
Instrument	Remote Learning Assessment (ReAL) Tool – "low-access" version Remote Learning Assessment (ReAL) Tool – "high-access" version
Cost	

Key implementation steps and lessons learned

- Administration protocols and guidelines to adapt the tool are available online.
- Caregivers need to be checked for minimum literacy and numeracy skills to assist in the administration of the test. A certain level of engagement is demanded of caregivers; this is less intensive than for the "low-access" version, as technology (WhatsApp) enables the use of ready-to-use stimuli.
- Plan for the distribution of printed materials and practice items to be used as stimuli during the assessment, or plan for enumerator to guide caregiver to produce stimuli using protocol contained within administration guidance.

B. ASYNCHRONOUS SMS-BASED QUIZZES

8. KHURANA, LEVIN, LUNA-BAZALDUA & LIBERMAN (2022)

Target skills	Foundational math knowledge (place value tasks)
Age/grade	Grades 2–5
Technology	Audio phone
Delivery	Remote
Delivery of stimuli	SMS (asynchronous)
Respondent	Child, with a caregiver's supervision
Country	Ghana
Digital/literacy skills required	Minimum literacy to read SMS
Instrument	Adapted from Angrist (2020)
Cost	N/A

Key implementation steps and lessons learned

- The National Council for Curriculum and Assessment collaborated in developing and contextualizing the assessment tools.
- Innovations for Poverty Action (IPA) was hired as the main implementer; Viamo Technologies was subcontracted to provide technological support such as provision of necessary software, toll-free numbers, short codes and recording of IVR content.
- Two-way SMS-based quizzes used to assess understanding of the place value concept; SMS system included a series of six SMS quizzes delivered to students over two weeks. 74 per cent reported that three quizzes per week was a good amount, while 23 per cent preferred to receive more quizzes per week.
- A mobile aggregator platform was used for sending the weekly SMS and receiving responses.
- Quizzes were incorporated to provide instant feedback for students
- A virtual call centre was set up; a toll-free long code was created to make the calls and send/receive SMS.

C. DEDICATED SOFTWARE FOR ASSESSMENT ON TABLETS

9. TANGERINE (RTI, 2018)

Target skills	Literacy and numeracy
Age/grade	Grades 1 to 6
Technology	Tablet
Delivery	In-person
Delivery of stimuli	Included in the app
Respondent	Children, under the supervision of a trained assessor/teacher
Country	Various
Digital/literacy skills required	Assessors must achieve minimum digital skills to manipulate a tablet; children need to be comfortable using a tablet
Instrument	Tangerine
Cost	The free version includes up to 2,000 surveys stored on the dedicated server; cost plans for larger samples and specific assistance requirements

Key implementation steps and lessons learned

- The assessment can be administered offline, through the dedicated app. Internet is required to upload the data to a dedicated server (cloud).
- Assessor-related data errors during administration and data entry are reduced thanks to skip logic checks, consistency checks and in-built tools (e.g., a timer that supplants traditionally used stopwatches).
- Assessment results can be used immediately: they are instantly aggregated and made available to decision makers through a dashboard on the Tangerine server.
- The app and its protocols have been deployed in more than 70 countries and have content available in more than 100 languages. However, a piloting and validation phase should be envisaged for contextualizing and assess internal validity (see Step 7).

C. DEDICATED SOFTWARE FOR ASSESSMENT ON TABLETS

10. TANGERINE: EF TOUCH (RTI, 2021; RTI, 2018; WILLOUGHBY ET AL., 2018)

Target skills	Executive functions (organization of information, planning, problem-solving, sustaining attention)
Age/grade	Preschool-aged children
Technology	Tablet
Delivery	In-person
Delivery of stimuli	Included in the app
Respondent	Children, with trained assessor-assisted administration
Country	Various
Digital/literacy skills required	Assessors must achieve minimum digital skills to manipulate a tablet; children need to be comfortable engaging with a tablet
Instrument	Tangerine:EF Touch
Cost	N/A

Key implementation steps and lessons learned

- The adaptation process for the instrument followed a four-step process in Kenya:
  1. Content review of all tasks (in English) by experts in early childhood assessment in East Africa
  2. English-to-Kiswahili translation and back-translation
  3. Two-day adaptation workshop with research staff and MoE officials
  4. Assessor training (3 days) and pilot data-collection conducted by research staff and MoE, with in-person evaluation of assessor performance by research leads
- Data collectors should determine the language of task administration (English or Kiswahili) based on initial, rapport-building conversations with children.
- The devices and the testing experience should be standardized: data were collected using Android tablets (9.4 cm × 15.2 cm); tablets were placed in an upright position in tablet stands (approximately US$7 each) in landscape orientation.

D. PHONE-BASED LEARNING PLATFORMS WITH IN-BUILT ASSESSMENT

11. EDOBEST@HOME (EDO STATE UNIVERSAL BASIC EDUCATION BOARD, LAST ACCESSED JUNE 2022; MUNOZ-NAJAR, 2020)

Target skills	Formative assessment against key curriculum-aligned milestones
Age/grade	From preschool to secondary school-age
Technology	Audio phone
Delivery	Remote
Delivery of stimuli	Included in the platform
Respondent	Children, under the supervision of a caregiver
Country	Nigeria
Digital/literacy skills required	Children must have basic literacy and digital skills to use audio phones
Instrument	EdoBest@Home Mobile Interactive Quizzes
Cost	N/A

Key implementation steps and lessons learned

- Quizzes can be accessed daily and are designed to practice, reinforce content learned each day, and receive instant automated feedback through WhatsApp or a text message. All quizzes are also aligned to the Edo state curriculum, the education level and allow students to practice a range of literacy and numeracy skills.
- Mobile interactive quizzes provide real-time data on learning achievement to a dedicated quality assurance team.
- Sustaining student participation was especially challenging in urban areas, where working-class parents have to use their mobile phones for their job, and in larger households, where multiple children had to share available devices.
- Rapid scale-up at the state level was made possible by high phone penetration rates and previous public investments in (i) the digitalization of learning materials and lesson plans and (ii) teacher training on pedagogy for remote learning. In EiE contexts requiring readily deployed solutions, leveraging existing platforms and trained teachers is fundamental.
- Public-Private Partnership (PPP) model as a success factor: the private sector provided expertise in school management, teacher training and educational technology, while the state provided resources to operate the programme in more than 1,500 public schools. The World Bank supported this with technical assistance and results-based financial resources.

4. KEY PRACTICAL LESSONS

- Prioritize the assessment of simple and foundational skills, using practice items/stimuli that are culturally appropriate, at the right ability level and accessible through available technology (SMS, verbally, etc.).
- Involve relevant MoE departments from the onset, including in the selection and validation of the assessment instruments and in the outreach to teachers/schools, where applicable.
- Analysing pilot data and comparing phone-based against face-to-face scores are pivotal steps to determine internal and concurrent validity of the assessment instrument.
- Invest time in first contact call: identify the best-placed member of the household to address the questions/facilitate, identify a suitable time for the call, build rapport and determine disability status/special needs of the child.
- Emerging evidence shows that live phone calls, while comparatively more expensive, tend to yield higher completion rates than automated testing through SMS or IVR.
- Assessments administered remotely through phones should not exceed 15–20 minutes. Sufficient time should be allocated for first contact calls, building rapport with the best-placed member of the household to assist in the call, obtaining informed consent and identifying the right time to hold the assessment. However, practitioners may want to cap the time for each specific assessment item to avoid cheating and minimize survey fatigue and burden on caregivers.
- Leverage community-based agents to follow-up with non-respondents, in order to minimize sample bias (underrepresentation of households with e.g. more limited network or educational background).Sending a text message or holding an advance call to alert respondents of the upcoming call increases pick-up rates.
- Consider the setting up of a virtual or physical call centre with dedicated toll-free numbers; partnerships, including PPPs, with phone service providers and/or tech partners should be sought to drive down costs and facilitate the calling or mass distribution of content/quiz via SMS.

- Consider small incentives to increase pick-up and assessment completion rates, such as airtime transfers.
- Whenever tablets can be easily made available and in-person administering is feasible, software such as Tangerine provides a field-tested and adaptable solution for assessing learning, with in-built automated functionalities for marking and timing tasks that can reduce error margins, as well as user-friendly dashboards for disseminating results to schools and practitioners.
- Apps with in-built assessments are being developed and deployed in emergency or low-resource contexts; assessing the feasibility of online and offline solutions in such contexts may be an alternative or complementary option in contexts where online or offline digital learning solutions can be deployed.

5. CONCLUSION

Mobile technology not only supplements the delivery of instruction, by making educational content accessible, nudging caregivers in engaging in a child's schooling or delivering teacher training (see Part 1 of this series – *On call: The use of mobile phones for learning in emergencies*). It can also effectively serve as a tool for conducting formative assessment of children's learning needs and progress, a domain still characterized by substantial gaps.

Mobile technology can contribute to fill this gap in a cost-effective and timely manner, as it relies on technology that is commonly available even to vulnerable households and does not require assessors to commute to hard-to-reach or insecure areas. Emerging evidence suggests that leveraging the use of basic mobile phones to collect learning data can yield valid results that reflect a child's actual ability and proficiency. As this literature is at its early stages, however, substantial research, piloting and contextualization are necessary to determine mobile phones' concurrent validity in comparison with more traditional assessment methods. In this respect, mobile phone-based testing may be more appropriate for low-stake assessments, including formative and education programme-specific assessments. The use of tablets for measuring learning, on its part, has become a more consolidated practice, with dedicated software such as Tangerine deployed across more than 100 countries, including in emergency settings.

Regardless of the specific type of device being used, planning a mobile technology-based learning assessment requires time and resources. Practitioners can follow the 10-step guide (Section 2) to choose from a suite of modalities and tools for conducting learning assessments through mobile technology (Section 3). First, practitioners will need to determine the type of skills they are interested in measuring; evaluate technological access in target areas; and gain an understanding of the literacy and digital literacy skills of children and caregivers. Mobilizing host communities and/or hosting sites is a crucial prerequisite to the success of the assessment. As mentioned above, equally critical during the preparation phase is the adaptation, validation and piloting of the assessment tools in the specific assessment areas. In order to ensure equity, a research and sampling plan that includes the most vulnerable learners is necessary. Lastly, closing feedback loops is fundamental for translating data into action; a data management and dissemination system that gets the data to government and school personnel, primarily the teachers, will allow for enhanced monitoring and teaching at the right level.

For rapid onset emergencies that require immediate evidence for action, practitioners will have to be strategic in prioritizing already validated assessment instruments that can be more rapidly adapted: tools that are shorter/simpler in their structure, that have been conceived for phone-mediated delivery or that have been deployed in similar contexts or with similar target populations.

REFERENCES

Angrist, Noam et al., 'Practical lessons for phone-based assessments of learning', *BMJ Global Health*, vol. 5, no. 7, 2020, <e003030> accessed 11 September 2022.

Angrist, Noam, Peter Bergman and Moitshepi Matsheng, *School's Out: Experimental Evidence on Limiting Learning Loss Using "Low-Tech" in a Pandemic*, National Bureau Of Economic Research, Cambridge, Mass, 2021, <https://doi.org/10.3386/W28205>, accessed 11 September 2022.

Banerjee, Abhijit, V. et al., *The Untapped Math Skills of Working Children in India: Evidence, Possible Explanations, and Implications*, 2017.

Luna-Bazaldua, Diego et al., *Psychometric Considerations for Implementing Phone-Based Learning Assessments*, World Bank Group, Washington, D.C., 2022.

Chávez, Cirenia, et al., *Continuing learning for the most vulnerable during COVID-19: Lessons from Let Us Learn in Afghanistan, Bangladesh, Liberia, Madagascar and Nepal*, Innocenti Research Brief, United Nations Children's Fund Innocenti Research Centre, Florence, 2021, <www.unicef-irc.org/publications/pdf/RB 2021-02.pdf>, accessed 11 September 2022.

Concern Worldwide, Early Grade Reading Assessment, Concern Niger, Tahoua, Endline 2019–2021, 2022, < https://admin.concern.net/sites/default/files/documents/2022-03/EGRA%202019%202022%20Final%20Report%20Turing.pdf>, accessed 11 September 2022.

Crawfurd, Lee, et al., *Teaching and testing by phone in a pandemic*, Center for Global Development, Working paper 591, September 2021, <www.cgdev.org/sites/default/files/teaching-and-testing-phone-pandemic.pdf>, accessed 11 September 2022.

Cynosure Evaluation, Baseline Evaluation, Education in Emergencies (EiE) in South Sudan | 2017, Cynosure, October 7, 2020. Baseline Evaluation, Education in Emergencies (EiE) in South Sudan | 2017, <https://cynosure-evaluation.com/baseline-evaluation-education-in-emergencies-eie-in-south-sudan>, accessed 28 July 2022.

Edo State Universal Basic Education Board, 'Mobile Interactive Quizzes', <https://edosubeb.org.ng/home-school/mobile-interactive-quizzes>, accessed June 2022.

Humanitarian Education Accelerator, *Piloting Gobee — what we learned from the first pilot of the gamified learning measurement tool*, blog post, <https://medium.com/hea-learning-series/piloting-gobee-what-we-learned-from-the-first-pilot-of-the-gamified-learning-measurement-tool-947402f123ff>, accessed 11 September 2022.

Inter-agency Network for Education in Emergencies, *Investing in Learning: The case for strengthening the collection and use of learning assessment data in EiE contexts*, INEE, New York, 2020a, <https://inee.org/sites/default/files/resources/INEE%20QELO%20Policy%20Paper%20v1.5%20LowRes.pdf>, accessed 11 September 2022.

Inter-agency Network for Education in Emergencies, *Academic Learning Outcomes in Education in Emergencies: Identifying, analyzing, and mapping tools to global guidance documents*, INEE, New York, 2020b, <https://inee.org/sites/default/files/resources/INEE%20-%20Academic%20Learning%20Outcomes%20in%20Education%20in%20Emergencies%20v1.0%20LowRes.pdf>, accessed 11 September 2022.

Kan, Sophia, Mirwais Fahez and Marco Valenza, *Foundational literacy and numeracy in rural Afghanistan: Findings from a baseline learning assessment of accelerated learning centres*, United Nations Children's Fund, *Innocenti Research Report*.

Khurana, Aishwarya et al., *Phone-Based Formative Assessment: Guidance for Evaluating Implementation Feasibility of Phone-Based Formative Assessments in Low-Resource Settings*, World Bank Group, Washington, D.C., April 2022.

Koomar, Saalim and Harriet Blest, Using EdTech to Support Effective Data Monitoring: A Curated Resource List, EdTech Hub, 2020, <https://docs.edtechhub.org/lib/FS4CMYUB/download/RCV8J6A7/Koomar%20and%20Blest%20-%202020%20-%20Using%20EdTech%20to%20Support%20Effective%20Data%20Monitoring.pdf>, accessed 11 September 2022.

Kopper, Sarah and Anja Sautmann, *Best practices for conducting phone surveys*, Abdul Latif Jameel Poverty Action Lab, Cambridge, Mass., 2020

Mohn, Giti. et al., *Resource Pack to Support Remote Learning : Mobile (English)*, World Bank Group, Washington, D.C., 2022, <http://documents.worldbank.org/curated/en/099846403292213739/P1742520988bb10220bc60090a9a4437eaa>, accessed 11 September 2022.

Munoz-Najar, Alberto and Joan Osa Oviawe, *Nigeria: Edo-BEST@Home, Education continuity during the Coronavirus crisis*, World Bank Group, Washington, D.C., November 2020, <https://documents1.worldbank.org/curated/en/371511605243057061/pdf/Nigeria-Edo-BEST-at-the-rate-Home.pdf>.

Piper, Benjamin et al., 'Are Refugee Children Learning? Early Grade Literacy in a Refugee Camp in Kenya', *Journal on Education in Emergencies*, vol. 5, no. 71, March 2020, <10.33682/f1wr-yk6y> accessed 13 September 2022.

Research Triangle Institute International, *Mobile Assessments Made Easy*, RTI International, 2018, <www.rti.org/sites/default/files/tangerine_brochure.pdf>, accessed 13 September 2022.

Research Triangle Institute International, 'Transformational Digital Learning— Insights and Reflections on Student Assessment', Presentation at the 9th International Skills Forum, *Reimagining Education and Skills Development for a New Normal*, 24 August, 2021. RTI, 2021, <https://events.development.asia/system/files/materials/2021/08/202108-transformational-digital-learning-learning-insights-and-reflections-student-assessment.pdf>, accessed 13 September 2022.

Research Triangle Institute International, *Early Grade Mathematics Assessment (EGMA) Toolkit*, RTI International, 2014, <https://ierc-publicfiles.s3.amazonaws.com/public/resources/EGMA%20Toolkit_March2014.pdf>, accessed 13 September 2022.

Research Triangle Institute International, *Early Grade Reading Assessment (EGRA) Toolkit: Second Edition*, RTI International, 2016, <https://shared.rti.org/content/early-grade-reading-assessment-egra-toolkit-second-edition>, accessed 13 September 2022.

Save the Children, *Remote Assessment of Learning - Low Access Administration Guidance*, Save the Children, London, 2022a, <https://resourcecentre.savethechildren.net/pdf/remote_learning_access_2022.pdf>, accessed 13 September 2022.

Save the Children, *Remote Assessment of Learning – High Access Administration Guidance*, Save the Children London, 2022b <https://resourcecentre.savethechildren.net/pdf/remote_learning_high_access_2022.pdf>, accessed 13 September 2022.

Twaweza-Kenya, *Collecting national data through mobile phones - Sauti za Wananchi Methodology*, Twaweza-Kenya, Nairobi, 2016a, <https://twaweza.org/wp-content/uploads/2021/01/Twaweza-MethodologyFactsheet-kenya.pdf>, accessed 13 September 2022.

Twaweza-Kenya, *Sauti za Wananchi, Collecting national data using mobile phones*, Twaweza-Kenya, Nairobi: 2016b, <https://twaweza.org//wp-content/uploads/2021/01/Sauti-za-Wananchi-Technical-Approach-Paper-Kenya.pdf>, accessed 13 September 2022.

Twaweza-Tanzania, *Sauti za Wananchi Collecting national data using mobile phones*, Twaweza-Tanzania, Dar es Salaam, 2013, <https://twaweza.org//wp-content/uploads/2021/01/Sauti-za-Wananchi-Technical-Approach-Paper-Tanzania.pdf>, accessed 13 September 2022.

Twaweza-Tanzania, *Uwezo, a citizen-led learning assessment*, online presentation, n.d., <https://thedocs.worldbank.org/en/doc/115811596531854838-0050022020/original/TemuUwezoExperience.pdf>, accessed 13 September 2022.

Twaweza-Uganda, Collecting national data through mobile phones - Sauti za Wananchi Methodology. Twaweza-Uganda, Kampala, 2017, <https://twaweza.org/wp-content/uploads/2021/01/SzW-UG-MethodologyFactsheet.pdf>, accessed 13 September 2022.

Twaweza-Uganda, *Are Our Children Learning? Uwezo learning assessment in refugee contexts in Uganda*. Twaweza-Uganda, Kampala, 2018, <https://uwezouganda.org/wp-content/uploads/2021/06/Uwezo-Refugee-Learning-Assessment-Report_FinalVersion.pdf>, accessed 13 September 2022.

United Nations Educational, Scientific and Cultural Organization, *Continuous assessment for improved teaching and learning: a critical review to inform policy and practice*, In-Progress Reflection No. 13 on Current and Critical Issues in Curriculum, Learning and Assessment, UNESCO, Paris, 2017, <https://unesdoc.unesco.org/ark:/48223/pf0000255511>, accessed 13 September 2022.

United Nations Educational, Scientific and Cultural Organization, *Managing high-stakes assessments and exams during crisis*, UNESCO's Covid-19 Education issue notes, No 4.3, UNESCO, Paris, 2020, <https://unesdoc.unesco.org/ark:/48223/pf0000373387>, accessed 13 September 2022.

United Nations High Commissioner for Refugees, *Coming together for refugee education*, UNHCR, 2020, <www.unhcr.org/uk/publications/education/5f4f9a2b4/coming-together-refugee-education-education-report-2020.html>, accessed 13 September 2022.

United Nations Children's Fund, A new phone-based learning assessment targets young children. UNICEF Education Case Study, Nepal, 27 October 2021, <www.unicef.org/media/109331/file/A%20new%20phone-based%20learning%20assessment%20targets%20young%20children%20(Nepal).pdf>, accessed 13 September 2022.

United Nations Children's Fund, *Education, children on the move and inclusion in education: Lessons learned and scalable solutions to accelerate inclusion in national education systems and enhance learning outcomes*, UNICEF, 2022, <www.unicef.org/egypt/media/8601/file/Education,%20Children%20on%20the%20move%20and%20Inclusion%20in%20Education.pdf>, accessed 13 September 2022.

Willoughby, Michael. et al., 'Measuring executive function skills in young children in Kenya', *Child Neuropsychology*, 17 June 2018, pp. 425–444, <https://doi.org/10.1080/09297049.2018.1486395>, accessed 13 September 2022.

World Bank, *Synthesis Report on Piloting of Remote Phone Based Assessments (English)*, World Bank, Washington, D.C., <https://documents.worldbank.org/en/publication/documents-reports/documentdetail/099410005022217418/p1742520a76d300510b22a04cb27b098fbf>, accessed 13 September 2022.

World Bank et al., *The State of Global Learning Poverty: 2022 Update*, World Bank, Washington, D.C., <https://www.unicef.org/media/122921/file/State%20of%20Learning%20Poverty%202022.pdf>, accessed 13 September 2022.

for every child, answers

UNICEF Office of Research – Innocenti
Via degli Alfani, 58 – 50121
Florence, Italy
Tel: (+39) 055 20 330
Fax: (+39) 055 2033 220
florence@unicef.org
www.unicef-irc.org
@UNICEFInnocenti on Twitter, LinkedIn,
Facebook, Instagram and YouTube

unicef
Office of Research – Innocenti