A Synthesis Report on Piloting of Remote Phone-Based Formative Assessment Solutions in Ghana, Nepal, and Pakistan

WORLD BANK GROUP

LeAP
Learning Assessment Platform

WORLD BANK GROUP

LeAP
Learning Assessment Platform

A Synthesis Report on Piloting of Remote Phone-Based Formative Assessment Solutions in Ghana, Nepal, and Pakistan

Aishwarya Khurana

Victoria Levin

Diego Luna-Bazaldua

Julia Liberman

1818 H Street NW, Washington, DC 20433

Telephone: 202-473-1000; Internet: www.worldbank.org

Table of Contents

Acknowledgments

This report was prepared by Aishwarya Khurana (Consultant), Victoria Levin (Senior Economist), Diego Luna-Bazaldua (Education Specialist), and Julia Liberman (Senior Education Specialist) of the World Bank's Learning Assessment Platform (LeAP) team. This work was sponsored by a Global Partnership for Education's (GPE) grant to support continuity of learning during the pandemic. The team worked under the overall guidance of Omar Arias (Practice Manager, Global Engagement and Knowledge Unit, Education Global Practice). Colleagues who contributed to this document with reviews and feedback include Rabia Ali, Marguerite Clarke, Martin Elias de Simone, and Maria Jose Vargas Mancera.

World Bank team members who supported this work at the country level include:

World Bank Ghana team: Yoko Nagashima (Senior Education Specialist), Eunice Yaa Brimfah Ackwerh (Senior Education Specialist), Chris Burningham (Consultant).

World Bank Pakistan team: Juan Baron (Senior Economist), Mary Bend (Consultant).

World Bank Nepal team: Karthika Radhakrishnan (Senior Education Specialist), Uttam Sharma (Consultant), Aarya Shinde (Consultant), Shwetlena Sabarwal (Senior Economist), Sunil Poudel (Consultant).

This report is part of a broader set of global knowledge products on phone-based learning assessments, including a guidance note on considerations for the implementation of phone-based formative assessments, a note on psychometric considerations for phone-based assessments, a landscape review of existing phone-based assessment interventions and their key features, and a checklist template to assess implementation prerequisites and enabling conditions for phone-based formative assessment solutions. The team is grateful to Sarah Fuller Klyberg for editing the report, and to Danielle Willis for graphic design.

Abbreviations

ASER	Annual Status of Education Report
CATI	computer-assisted telephone interviewing
EGRA	Early Grade Reading Assessments
GES	Ghana Education Services
GL-TV	Ghana Learning Television
GoN	Government of Nepal
GoP	Government of Pakistan
IPA	Innovations for Poverty Action
IVR	interactive voice recording
KPK	Khyber Pakhtunkhwa (Pakistan)
LPC	live phone calls
MoEST	Ministry of Education, Science, and Technology (Nepal)
NaCCA	National Council for Curriculum and Assessment (Ghana)
NDRI	Nepal Development Research Institute
RDD	random digit dialing
SMS	short message service
SBC	Standards-Based Curriculum (Ghana)
SNC	Single National Curriculum (Pakistan)
TaRL	Teaching at the Right Level

Executive Summary

School closures related to COVID-19 not only disrupted education but also impacted teachers' ability to know whether and what their students were learning. This information gap was most challenging in contexts with limited internet connectivity and low access to "smart" digital devices. In such circumstances, conducting formative assessments remotely using basic phones was seen as a means to provide timely information to teachers, parents, and students and support learning continuity outside the classroom. As students return to school, the same solutions can be used to complement in-person instruction to accelerate learning recovery, expand the use of formative assessment in hard-to-reach schools, and improve the resilience of education systems when confronted by future shocks.

This report describes the three pilot studies that aimed to test the logistical and technological feasibility of using short message service (SMS), interactive voice recording (IVR), and live phone calls (LPC) to conduct formative assessment of foundational math and literacy skills for primary-grade students. These pilot studies were conducted in Ghana, Nepal, and Pakistan between October 2020 and January 2022. Table 1 describes the main features of the three pilot studies.

Table 1. Pilot Study Features

Nature of pilot	Duration of pilot	Technology solutions used	Grades assessed	Subject area assessed	Sample size (by solution)	Types of data collected	Cost
Ghana							
Stand-alone pilot	Oct. 2020–Jan. 2022 (16 months)	SMS, IVR, LPC	2–5	Foundational math	Total: 259 (82 for SMS, 84 for IVR, and 93 for LPC)	• Engagement patterns • Reasons for nonengagement • Cognitive demand of assessment tasks • Parental support • Satisfaction with the assessment	SMS (per SMS): US$0.04 IVR (per minute): US$0.03 LPC (per minute): US$0.03
Nepal							
Extension of existing activity	June 2021–Jan. 2022 (8 months)	LPC (with SMS)	4–5	Foundational math and literacy	Total: 2,163	• Engagement patterns • Reasons for nonengagement	SMS (per SMS): US$0.016 LPC (per minute): US$0.021
Pakistan							
Standalone initial pre-pilot	June 2021–Jan. 2022 (8 months)	SMS, LPC	3–5	Foundational literacy	Total: 36 (SMS only: 18 LPC [with SMS]: 18)	• Engagement patterns • Reasons for nonengagement • Cognitive demand of assessment tasks • Satisfaction with the assessment	SMS (per SMS): US$0.001 – US$0.004 LPC (per minute): US$0.02–US$0.30

Note: "LPC (with SMS)" indicates an assessment delivery design in which students were assessed orally through live phone calls and received text messages with written questions and tasks. "LPC" refers to an assessment delivered only over a live phone call with the student.

The main findings from the three pilots include:

- Engagement in remote formative assessment is higher when the participants are contacted using LPC as compared to SMS or IVR technologies. Main reasons for nonengagement include unavailability of child and caregiver at the same time and technical challenges such as connectivity issues or invalid phone numbers.
- To increase acceptance of the phone-based assessment, assessors must establish rapport with the caregiver and the child. During the first contact interviews, assessors must develop trust by mentioning upfront the names of the organizations involved and making sure that caregivers do not incur any costs related to the pilot, are provided with regular reminders, and understand that the assessment is not high stakes.
- Where data on satisfaction were collected (Ghana and Pakistan), most children and caregivers were satisfied with assessment delivery. They found it easy to respond to assessment items on mobile phones, they found the assessment tasks to be at the right level, and caregivers found it useful to assess their child's proficiency levels.

Compared to LPC and IVR, SMS is the least expensive technology to conduct remote formative assessments and can easily be automated. Implementing LPC is more expensive due to longer call duration. Although the three pilots focused on testing the logistical and technological feasibility of phone-based assessments, they revealed certain aspects that should be considered for scaled-up implementation, which would entail working with local teachers and school administrators:

- The use of basic mobile phone technology for the purposes of student assessment does not need to be limited to pandemic-related school closures. It can be used for any kind of disruption, including prolonged teacher strikes, school closures due to violence and insecurity, and disruptions caused by climate-related events.
- Deploying these low-tech solutions for assessments to improve the quality of instruction and achieve meaningful results in terms of learning outcomes requires a significant amount of investment of time for planning and setup. Relative to other education expenditures, however, deploying these assessments can be inexpensive.
- It is important to consider that (1) phone-based assessments may be more suitable for early grades and for assessing simpler concepts, and (2) when designing the assessment, it is necessary to collaborate with the government education authorities, align the assessment with school curriculum and calendar, understand the local context, and choose the language of assessment carefully.
- It is critical to hire the right people (including technical partners, voice artists for IVR, and teachers) to implement various aspects of assessment to ensure success. Provision of necessary technology infrastructure (such as mobile aggregators, provision of long and short code, etc.) within the country is critical for the success of phone-based assessment.
- Training of teachers to use formative assessment tools needs to focus not only on how to administer these tools over the phone. Training also needs to address how to provide useful real-time feedback to students as well as how to interpret the results of these assessments to inform instruction and provide targeted support.

1. Introduction

Since its onset, the COVID-19 pandemic has disrupted education systems globally, forcing many countries around the world to close schools. In April 2020, pandemic-related school closures affected 1.6 billion learners in over 190 countries (UNESCO 2020). Overall, school closures have ranged from no closures in a few countries to closures for more than a full school year.

In response to pandemic-related school closures, use of technology for learning became widespread. To mitigate learning losses due to school closures, countries introduced and expanded remote learning modalities. Countries such as Afghanistan, Brazil, Cambodia, Estonia, Kenya, Haiti, Nigeria, Pakistan, Peru, Sierra Leone, Uruguay, and many others delivered educational content using a combination of online platforms, television, radio, mobile phones, and printed learning materials (Barron Rodriguez et al. 2021).

Advanced (internet-based) technology offered an inadequate substitute for in-person learning, and adoption was hindered by preexisting inequities prevalent in most systems. Children from disadvantaged backgrounds were more likely to be excluded from remote learning due to lack of access to electricity, connectivity, devices, and caregiver support at home. According to UNESCO (n.d.), lack of connectivity and devices at home excluded at least one-third of students from pursuing learning remotely. *In this context, more low-tech and readily available modalities, such as basic phones and technologies that are less reliant on internet connectivity, can facilitate learning continuity.*

Basic mobile phone technology can be used during any kind of disruption to learning or as a complement to in-person instruction. Phone-based technology can be used for school closures due to prolonged teacher strikes, local violence, and insecurity, as well as for disruptions caused by climate-related events. Investing in the design and setup of phone-based technologies for learning can thus improve the resilience of education systems when they are confronted by future shocks and crises. These technologies can also be deployed as a complement to in-person instruction, particularly to accelerate learning recovery after school closures, to provide additional support to struggling students, or to reach students in remote schools.

To achieve meaningful results in terms of learning outcomes, phone-based learning interventions need to include an explicit focus on learning assessment. While internet-based platforms often include interactive assessment as part of content design, low-tech learning interventions (those that use basic mobile phones, social media, or print-based materials) often have been focused on delivering learning content to students rather than on assessing the extent to which students absorb this content and whether additional material is needed to support students' movement along a learning trajectory. Formative learning assessment is a critical ingredient to determine student learning, but to date, few feasibility studies have explored implementing formative assessment using basic mobile phone technology.[1]

Also, **deploying nontraditional low-tech solutions for assessments that will yield meaningful results for policy and improving the quality of instruction in schools requires a significant amount of investment of time for planning and setup**—even if these assessments are inexpensive

[1] For a review of studies using phone-based assessments around the world, please see https://documents.worldbank.org/en/publication/documents-reports/documentdetail/099835004182239311/p1742520dcdd2e0009fa00d6a0176bc1c7.

relative to other education expenditures in the country. Investment in planning becomes even more important in a post-disaster setting, since the capacity to deploy assessments, no matter how low-tech, cannot be built during or in the immediate aftermath of a shock, when authorities are busy dealing with urgent relief operations. The machinery for such assessments needs to be put in place, ideally, as part of a comprehensive disaster response system.

With funding from the Global Partnership for Education's (GPE) Consortium grant for "Learning Continuity at Scale that Reaches the Most Marginalized," the World Bank's Learning Assessment Platform (LeAP) team supported learning assessment activities to help ensure learning continuity during the pandemic. From 2020 to 2022, the LeAP team developed global knowledge products to support remote formative assessment of student learning outside physical classrooms using basic mobile phone technologies. As part of this effort, the LeAP team evaluated the adoption and use of phone-based formative assessment tools (World Bank n.d.), particularly in contexts with low internet connectivity and limited access to more advanced digital devices (for instance, smartphones, tablets, and computers) at home to develop recommendations for greater use and effectiveness.

The work involved identifying pilot countries and planning and piloting three remote formative assessment solutions—SMS; interactive voice response (IVR); and direct, live phone calls (LPC).[2] The design of this pilot study leveraged the experiences of organizations that initially responded to the COVID-19 pandemic with the implementation of low-tech phone-based learning and assessment delivery in low-income countries (Angrist et al. 2020). Initial review of past experiences and projects allowed for identification of key factors that may enable the implementation, sustainability, and potential scale-up of phone-based assessment delivery solutions (Khurana et al. 2022; Luna-Bazaldua et al. 2022; World Bank 2022a, 2022b).

Following consultations with World Bank country teams and in-depth feasibility evaluations of implementation in 11 countries, Ghana and Rwanda were selected for pilots in October 2020.[3] In Ghana, the standalone pilot focused on foundational numeracy using all three phone-based solutions (SMS, IVR, and LPC). In July 2021, Nepal and Pakistan were selected as additional pilot countries. In Nepal, the pilot focused on extending the ongoing Teaching at the Right Level (TaRL) initiative by adding formative assessment in foundational literacy to the existing math assessment as well as adding new regions and localities (Radhakrishnan et al. 2021). In Pakistan, an initial small-scale pre-pilot that focused on foundational literacy was launched in July 2021.[4] In all the countries, the pilot rollout happened in October 2021, and the work, including analysis of results, was completed by January 2022. This report presents the experiences of Ghana, Nepal, and Pakistan in implementing formative assessment solutions using basic mobile phone features (that is, SMS, IVR, and LPC). The report also discusses lessons learned from implementation of the three technology solutions across the countries and considerations for scaled-up implementation of phone-based formative assessment.

1.1. Education Context during the COVID-19 Pandemic

When the COVID-19 pandemic started in March 2020, Ghana responded by closing schools, affecting nearly nine million students enrolled in preprimary, primary, and secondary education (UNICEF

[2] The pilot study is further explained in Section 2.
[3] Due to logistical challenges, the country pilot in Rwanda was closed in July 2021.
[4] The work done in Pakistan was limited in size and scale due to budget and time constraints. The aims were to develop pilot materials, check the feasibility of technology solutions, and propose steps for scale-up.

2021). The school closures in Ghana lasted approximately thirty-nine weeks, and schools were fully reopened by January 2022 (UNESCO n.d.).

Nepal announced school closures in response to the COVID-19 pandemic in March 2020. Since then, schools have opened intermittently, and as of January 2022 schools had reopened after nearly one month of closures due to the third wave of the COVID-19 pandemic.

Pakistan also instituted widespread school closures because of COVID-19. Schools in Sindh were closed starting on February 27, 2020. In the rest of the country, school closures started over the weekend of March 14, 2020. Schools have been open for full attendance since September 2021, with partial closures and 50 percent attendance during weeks of high case reportage.

1.1.1. Ghana

To ensure continuity in learning, a contingency plan was formulated by Ghana Education Services (GES) that aimed to support the implementation of virtual and distance learning through radio, television (Ghana Learning Television, or GL-TV), and online learning platforms,[5] which were made available nationwide to children of all grades (Ghana, Ministry of Education 2020). In addition, studies and surveys were conducted to measure student engagement in remote learning activities.

The RECOVR rapid response panel survey (IPA 2020), carried out in May 2020 with 1,357 respondents, found that since school closures in March 2020, only 60 percent of all students were spending time on educational activities, and the same share of students preferred school textbooks over other educational materials (including GL-TV, internet content, etc.). Moreover, due to lack of motivation and parental and teacher support, children were spending only 5.9 hours per week on educational activities—a number significantly lower than the time spent in school. A nationally representative phone-based survey conducted by UNICEF in June 2020 with 5,667 children aged 0–17 years found similar results. The reported reasons for lack of engagement were children's lack of access to digital devices such as computers and phones, lack of learning materials including books, lack of access to the internet, and lack of interest or motivation (UNICEF 2020b). Another study conducted in October 2020, using phone-based surveys and learning assessments of 1,844 children in the Greater Accra Region of Ghana, as well as their caregivers and teachers, found that students who at home were facing economic hardships "engaged in fewer remote learning activities and had lower literacy and numeracy assessment scores" (Wolf et al. 2021, 1).

1.1.2. Nepal

The Government of Nepal (GoN) has implemented several alternative learning programs—including radio and television-based programs, online learning portals, SMS and phone-based teaching support, and small group community-based teaching—to provide learning continuity opportunities during school closures. Access to these learning programs remains a challenge, however. Two-thirds of school children were unable to access remote learning during school closures (UNICEF 2020a). Fewer than half of households have access to television, and only about one-third own radios. In Nepal, computer and internet access also is low and uneven across provinces, ethnicities, and socioeconomic backgrounds. While mobile penetration is high, with 80 percent of households reporting phone access at home, the usage of phones to access learning platforms consistently remains low nationwide.

[5] This plan included a series of worksheets on assessment tasks for all grade levels (kindergarten through grade 9) on various subjects to complement radio and television instruction.

1.1.3. Pakistan

To support continuity of learning, the Government of Pakistan (GoP) launched various remote learning initiatives utilizing television, including Teleschool (a channel launched by the GoP) and Taleem Ghar (a channel launched by government of Punjab province)[6] (World Bank 2021) and education radio shows such as *Broad Class: Listen to Learn* (Najib and Ranjan 2020). In addition to television, Teleschool and Taleem Ghar were also rolled out through online platforms including a mobile application, a website, and a YouTube channel (Martinez, Barron, and Zacharia 2021). Many schools also offered online and paper-based learning options to children during the school closures. Despite these efforts, results from a phone-based World Bank survey of parents located in Punjab province of Pakistan revealed that their knowledge of available options for remote learning was very low: 84 percent of families interviewed were not aware of Teleschool or Taleem Ghar. Adoption was higher among those families who knew of available remote learning opportunities (Hasan, Geven, and Tahir 2021). World Bank phone-based surveys conducted nationwide revealed that poor children and those living in remote areas were least likely to continue studying, in part because they were unaware that the options to do so existed, due to lack of electricity (needed to power a television or charge an online device) or lack of access to an online device at home. Indeed, paper-based materials were the most widely used remote learning resource for children of all income levels and household localities. Among respondents to the phone surveys, wealthier, urban-based households exhibited the highest levels of ongoing study among children who were kept home from classrooms for their own safety.[7]

[6] For more information about the initiative in Punjab province, see the Taleem Ghar website at https://taleemghar.punjab.gov.pk/.
[7] The World Bank report on phone-based surveys in Pakistan is forthcoming.

2. Phone-Based Formative Assessment and Its Administration in Pilot Countries

Pandemic-related school closures not only disrupted education continuity but also impacted teachers' ability to know what and if students were learning. Assessment of student learning is "the process of gathering and evaluating information on what students know, understand, and can do in order to make informed decisions about the next steps in the educational process" (Clarke 2012, 1). Formative assessment, usually carried out in the classroom as part of the teaching process, encompasses everything from teacher observation to continuous feedback, to homework, and is particularly useful for understanding the learning needs of each student and adjusting instruction accordingly (Luna-Bazaldua et al. 2022). Typically, formative assessment activities have relied on in-person interactions, but when learning takes place outside the physical classroom due to school closures, teachers and parents have an even greater need to understand whether students are absorbing the learning content that is delivered to them remotely.

Conducting formative assessments remotely using basic phones is a solution that can provide timely information to teachers, parents, and students and support learning continuity in times of crisis and beyond. *Phone-based formative assessment* refers to the use of mobile phone technologies to engage students outside the classroom; assess their learning in real time; and provide timely, constructive feedback to promote learning continuity. Mobile phone technologies include feature phones, smart feature phones, and smartphones. Students in low-resource and low-connectivity contexts may typically only have access to basic functionalities even on smartphones, such as SMS and LPC. To address their needs, the phone-based formative assessment modalities piloted in this study focused exclusively on three basic phone functionalities: SMS, IVR, and LPC.

The pilot study (pilot or study) tested the logistical and technological feasibility of using SMS, IVR, and LPC to conduct formative assessment of foundational math and/or literacy for primary-grade students in Ghana, Nepal, and Pakistan. In particular, the study aimed to understand (a) how to design or adapt an assessment instrument so that it is feasible to be implemented via SMS, IVR, or LPC; (b) the technological infrastructure that needs to be set up for implementation; (c) the profile of and training needed for assessors; (d) the engagement patterns with phone-based assessments; and (e) the unit costs involved in implementation. It is important to note that some aspects of scaled-up implementation were not targeted in these short-term pilots, such as the recruitment and training of teachers who would be engaged in phone-based assessment of their (or other teachers') students, provision of detailed real-time feedback to students, and the use of phone-based formative assessment for informing subsequent instruction.

2.1. Organizational Approach

After Ghana, Nepal, and Pakistan were selected to pilot the phone-based formative assessment solutions, the LeAP team collaborated with the World Bank country teams to oversee the pilot implementation, to liaise with different counterparts, and to identify and procure implementing partners. The strategy for recruiting implementing partners included identifying firms that could (1)

develop assessment tools relevant to the country context and (2) implement a small-scale pilot using the assessment materials produced.

The implementing partners needed to be familiar with the country context and have experience working with government agencies, including previous experience in implementing education interventions in the country. Additional desirable characteristics included experience with implementing education interventions using basic phones as a delivery mode and experience with designing and implementing learning assessments aligned to learning standards and learning resources. Implementing partners' access to necessary technological infrastructure was critical to the pilot study, but the team understood that many firms might not have all infrastructure elements. Therefore, attention was paid to their strategy of procuring the necessary technology and resources to implement phone-based assessments, including an option to apply as a consortium of firms.

2.1.1. Ghana

In Ghana, the World Bank teams collaborated with the National Council for Curriculum and Assessment (NaCCA) to develop and contextualize the assessment tools, which were inspired by the tools used for a phone-based math assessment in Botswana (Angrist et al. 2020). Innovations for Poverty Action (IPA) was hired to implement the phone-based assessment pilot. For this work, IPA subcontracted Viamo Technologies to provide technological support such as provision of necessary software, toll-free numbers, and short codes and recording of IVR content to implement the pilot. Biweekly meetings were organized between the World Bank global and country teams, NaCCA counterparts, and IPA and Viamo staff to provide updates and discuss next steps. Continuous engagement with and buy-in from government counterparts from the design stage through the evaluation helped to ensure that the pilot sample reflected regional, linguistic, and socioeconomic diversity, and that the assessment instruments aligned well with education policies and curriculum.

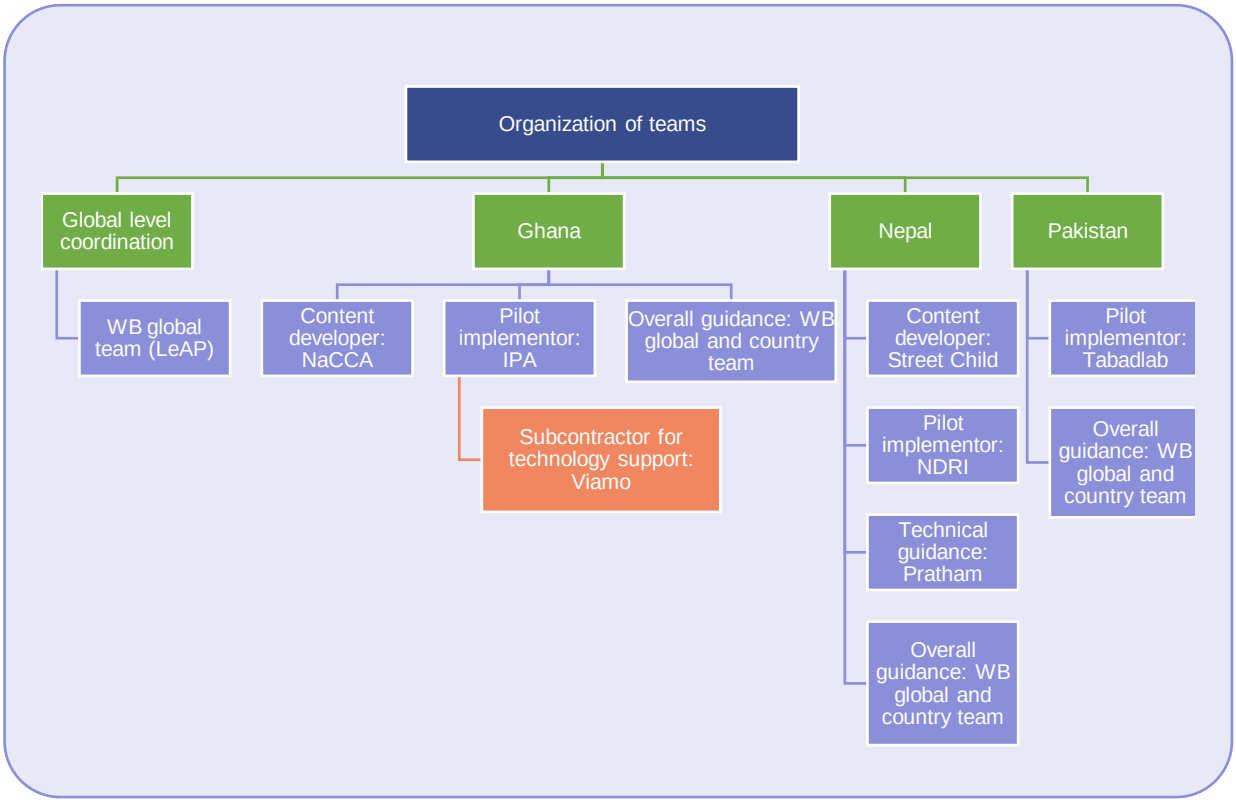
2.1.2. Nepal

In Nepal, the pilot was led by the World Bank in collaboration with the Ministry of Education, Science, and Technology (MoEST) and local governments; Street Child; Nepal Development Research Institute (NDRI); and Pratham USA. Street Child developed the literacy tool and adapted numeracy tools (previously developed by Teach for Nepal with technical guidance from Young 1ove) to be used in the pilot. Pratham USA provided technical guidance in the development of literacy tools. NDRI was responsible for implementing the pilot and collecting the data. The literacy and numeracy component of this pilot was implemented as part of ongoing TaRL work in Nepal. Pratham USA provided technical guidance on the overall implementation of the literacy component.

2.1.3. Pakistan

In Pakistan, the World Bank team hired Tabadlab to design and implement the phone-based assessment pilot. Based in Pakistan with firm presence throughout each of its territories, Tabadlab has collaborated with the Ministry of Education and international donors on policy engagement and numerous education projects. For this work, Tabadlab drew on Early Grade Reading Assessments (EGRA) and Pakistan's national Annual Status of Education Report (ASER) citizen-led assessment to create literacy assessment instruments for SMS, IVR, and LPC delivery through basic mobile phone technology. Weekly meetings were organized between the World Bank country team and Tabadlab to share updates and discuss next steps in the project implementation. figure 1 illustrates the organization of pilot teams.

Figure 1. Organization of Teams



Note: The organigram mentions the overall organization of teams in different pilot countries. The green box mentions the nature of coordination/work (country level or global), the blue box mentions the organizations or firms involved and their role in the pilot, and the orange box mentions the subcontractor(s) hired by the firms to support their work.

2.2. Assessment Solutions

This section describes the types of technology solutions that were piloted in Ghana, Nepal, and Pakistan. The use and implementation of the technology solutions varied slightly in each country, as discussed below.

2.2.1. Short Message Service (SMS)

SMS is by far the most widely used phone-based technology due to its low cost, relative ease to develop and deploy in mobile applications, and compatibility with a wide variety of existing platforms and mobile phones. Formative assessment through text messaging is an asynchronous process during which students respond to questions, sent by SMS, based on content provided either in SMS or in another form.

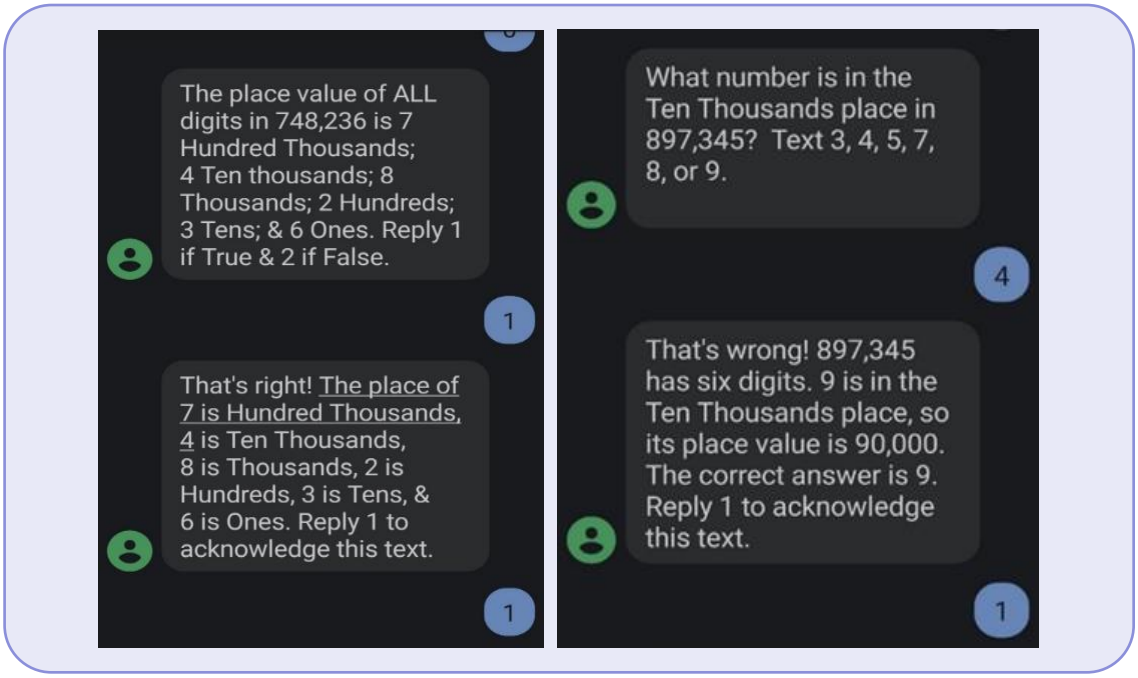
The delivery of assessment content through SMS typically requires using a mobile aggregator[8] platform that sends SMS messages and receives responses. In all three pilot countries, a two-way[9] messaging system was set up to send weekly quizzes and receive each child's responses. Dedicated

[8] Mobile aggregators link different carrier networks with SMS software/platform providers. They simplify the mass distribution of SMS to many carrier networks. Mobile aggregators have agreements with carrier networks to send and receive SMS through these networks.
[9] Two-way SMS is the ability to send and receive SMS.

toll-free short or long codes[10] were set up to support the creation, scheduling, and initiation of both one-way[11] and two-way SMS sessions. The toll-free feature ensured that participants did not bear the cost of replying to the weekly quizzes. Moreover, the quizzes were developed in multiple-choice question (MCQ) format to allow the child to respond without exceeding the 160-character limit. No SMS quiz was repeated during the intervention, and feedback was provided to children based on their answer choice.

In Ghana, two-way SMS-based quizzes were used to check for students' understanding of the place value concept and their ability to identify the place value of all digits in a number. The SMS system included a series of six two-way SMS quizzes delivered to students over two weeks. Each two-way SMS quiz included one place value task, and three tasks were sent per week to students. See figure 2 for an example of the text messages sent to students in Ghana.

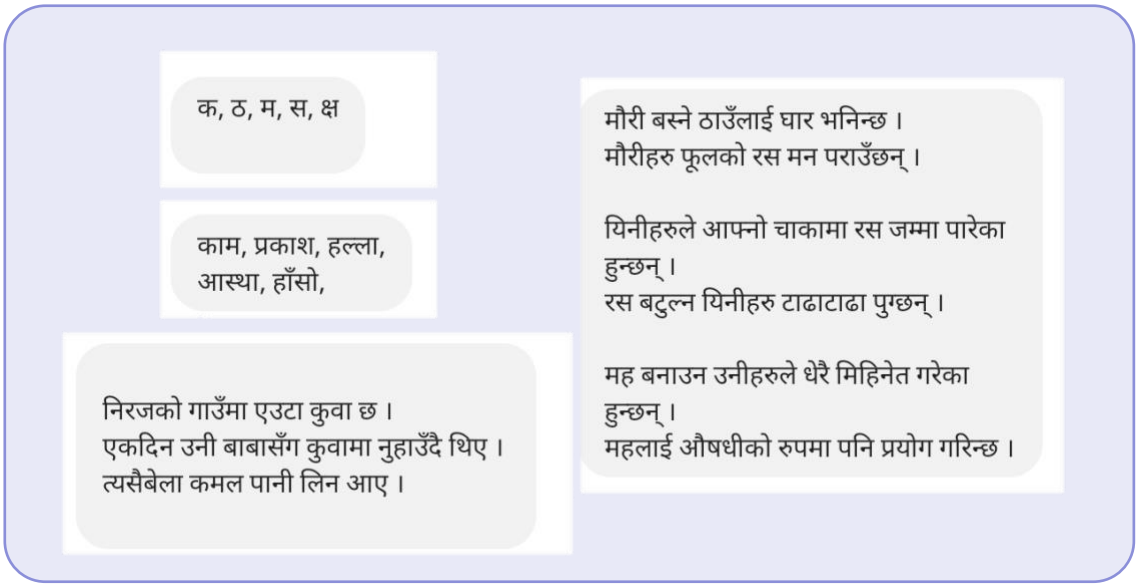
Figure 2. Sample SMS Quizzes from Ghana with Feedback for Correct and Incorrect Responses



Nepal used one-way phone calls along with one-way, real-time, SMS-based quizzes to assess students' reading and comprehension skills. To assess numeracy, each student received a one-way phone call and was asked to take a quiz to check their foundational skills in numeracy. See figure 3 for an example of the text messages (literacy) sent to students in Nepal.

[10] Short codes are 5- or 6-digit numbers used to send and receive texts only (such as 98765). Long codes are 10- or 11-digit numbers that can be used to send or receive SMS and also make LPC (such as 123-456-7890).
[11] One-way SMS is the ability either to send or receive SMS, such as broadcasted SMS.

Figure 3. Sample SMS Quizzes for Nepali Students



In Pakistan, similar SMS-based quizzes were used to check for participants' understanding and mastery of foundational literacy skills, including their ability to discriminate letter sounds, recognize letters, identify nonwords, and answer comprehension questions based on a short text (see figure 4 for an example of a literacy task in Urdu). The SMS assessment included twelve SMS questions (or "items") delivered in separate messages.[12]

Figure 4. Sample Literacy Quiz Sent to Students in Pakistan

[12] Only one round of SMS assessment occurred in Pakistan.

2.2.2. Live Phone Calls (LPC)

Live phone calls between students and teachers or trained assessors can ensure continuity of the learning process. In the context of COVID-19, the primary purpose of direct calls to students is to ensure that students can continue learning through guidance and constructive feedback from teachers or assessors outside the classroom (for instance, when doing homework, reviewing new learning content, or preparing for exams). Direct calls also allow teachers to assess students' engagement with remote learning resources and respond to students' questions.

In Ghana, a virtual call center was set up. To operationalize the call center, a toll-free long code was created, and assessors were trained on how to conduct assessments over the phone. In Nepal and Pakistan, no call center was set up, and trained assessors called potential respondents directly. In the case of literacy assessment, the LPC solution was complemented with an SMS solution in Nepal and Pakistan. The assessors called caregivers and children and read the approved phone script, including demonstrating the learning concept to the students (if needed), and then asked students content-related questions to check understanding and provide support in cases where students were unable to answer a given question. The participants did not incur any costs to receive the LPC.

As a synchronous form of communication and assessment, the LPC required both the caregiver and child to be present at the time of the assessor's call. Caregivers were asked to use their phone's loudspeaker setting during the call session and ensure that the child completed the assessment on their own. Detailed logs of all calls or rescheduled appointments were kept and tracked by the assessors.

In Ghana, four telephone assessors, each with a unique account, worked in the call center. All telephone assessors were fluent in English and in at least one of the four local languages selected for the pilot. On average, most participants completed the LPC in 13 minutes and 8 seconds, with a range from 8 to 40 minutes. The extended duration of some calls was due to participants' pausing the math (place value tasks) assessment and attending to other things before completing the assessment.

In Pakistan, the assessment focused on foundational literacy tasks only, and thus implementation of LPC was complemented with SMS. In other words, multiple SMS with literacy tasks were first sent to the participant and were followed by an LPC from the assessor to the caregiver's phone. The assessor would go over each task, which involved reading the prompt sent earlier by SMS, and check for the child's understanding. In total, six assessors, proficient in Urdu, a local language, and English, were hired for the pilot.

In Nepal, the team included 21 assessors who were proficient in both Nepali and English. For the math assessment, the assessor read the questions aloud over the phone. For literacy assessment, to minimize cheating, an SMS with literacy tasks was sent to the child in real time while they were on the LPC. On average, each call lasted 45–60 minutes.

For more information on team composition, refer to appendix A.

2.2.3. Interactive Voice Response (IVR)

Interactive voice response (IVR) is a tool that can be used to reach populations with limited literacy because it relies on speech rather than text (World Bank 2012). It can be used on any type of landline or mobile phone to communicate information via a voice message system in the form of audio recordings (Diedhiou et al. 2015). IVR works by inviting users to call a toll-free number. An

automated system delivers prerecorded audio messages that may contain educational content or voice-based quizzes to which users respond by pressing a number key or saying the number that corresponds to their answer choice (for example, "Say yes or press 1; say no or press 2"). In addition, IVR can be used to reach individuals with vision impairment, and it facilitates delivery of the same content in multiple languages, including languages with no written form (Diedhiou et al. 2015).

In Ghana, IVR quizzes on place value were delivered through an automated interactive voice response system using prerecorded messages. Professional voice actors were recruited to record equivalent messages in English and in four local languages for students in early grades. Participants received voice-prompted quizzes on place value tasks via a toll-free phone number. The IVR quizzes came with audio content with specific instructions for responding to the content. In the push model,[13] participants received three quiz problems to solve via IVR (one problem for each place value task) per week. To enrich participant engagement, the three IVR quizzes were implemented sequentially based on the participant's response to the first and second IVR quiz problems. Each quiz problem had a corrective feedback response based on the participant's answer. Participants entered answer choices or responses via their mobile phone keypad. For each message or script delivered through IVR, the participant would press the # key on the keypad to repeat the message or quiz and available options. The participant incurred no cost for receiving or responding to the IVR quizzes. Each IVR call took a participant an average of 5 minutes and 9 seconds to complete, with a range from 2 to 9 minutes.

Due to technological and timeline constraints, pilot implementation of IVR technology did not take place in Nepal and Pakistan.

2.3. Pilot Design

The focus of the pilot in all three countries was to check the feasibility of using basic mobile phone technology to conduct remote formative assessments. Overall, the pilot focused on primary grades and assessed foundational literacy and/or numeracy skills using the three technology solutions described above. Specifically, in each country, the pilots varied in scale, assessed skills, technology solutions, and the language of assessment. Table 2 provides information on the grades, content assessed, technology used, and the language of assessment in Ghana, Nepal, and Pakistan.

Table 2: Overall Pilot Design

Country	Grades	Content/subject area	Technology used in pilot	Language of assessment
Ghana	2–5	• Foundational math knowledge (place value tasks)	• SMS • IVR • LPC	• English • Ashante Twi • Dagbaani • Ewe • Ga
Nepal	4–5	• Foundational math • Foundational literacy (reading accuracy and comprehension)	• Math: LPC • Literacy: LPC (with SMS)	• Nepali[14]
Pakistan	3–5	• Foundational literacy skills	• SMS • LPC (with SMS)	• English • Urdu • Punjabi • Sindhi • Pashto

[13] Push models involve automatically sending the content to the participants by calling them using the software.
[14] For numeracy, flexibility was provided to students to use Arabic numerals if they preferred to do so, but the language of communication over the phone was Nepali.

2.3.1. Setting Up Infrastructure for Technology

The implementation of the pilot differed in each country in terms of technology solutions used. In Ghana, equivalent assessment content was delivered using three alternative modalities—SMS, IVR, and LPC. In Nepal and Pakistan, the pilot focused only on SMS and LPC solutions due to technological constraints; both countries used LPC in combination with SMS to assess literacy tasks.

Although each pilot project was slightly different, smooth implementation of the three formative assessment solutions required setup of hardware as well as software.

Ghana and Nepal

In Ghana and Nepal, technology infrastructure included the following components.

1. Data collection software: In both Ghana and Nepal, SurveyCTO was used to collect data. Numeracy and/or literacy items were programmed into the SurveyCTO. Before the pilot rollout, the assessment items were tested rigorously to check the constraints, logic, and entire flow of the questionnaire, and to identify other potential issues. The software allowed storing and processing of users' responses and retrieving of the necessary information based on the chosen option. In addition, SurveyCTO's computer-assisted telephone interviewing (CATI) system was used to place calls to all participants and capture their responses.
2. Internet network and airtime: Adequate data was provided for the assessors to place LPC remotely and to ensure that the assessors had access to reliable internet connectivity to make the calls. Relatedly, the two or three formative assessments required airtime for deployment of learning content.
3. Support infrastructure: Key support infrastructure for implementing the three formative assessment solutions included dedicated computers and tablets, mobile phones, and phone numbers (short codes and long codes in the case of Ghana). In Ghana, the assessors were provided a computer and mobile phone and/or phone numbers as well as quality headsets with microphones to access the web-based application and make the calls. The assessors used dedicated phone numbers to avoid interruptions from personal calls. In Nepal, the assessors used their own devices for making the calls and recording data. The survey firm in Nepal provided devices to assessors who did not own a computer, tablet, or phone. Fees incurred by LPC and SMS were reimbursed to the assessors.

   In Ghana, for the IVR component, a studio was rented, and voice artists (proficient in pilot study languages) were hired to record the audio scripts.

Pakistan

For the SMS solution, Tabadlab used software that was compatible with basic phones to set up an SMS broadcast application. At the beginning of the assessment, the system sent the assessment items in local language script through the SMS broadcast feature, which delivered each item as a separate SMS.

For the LPC solution, the assessors made phone calls using their mobile phones. They were provided adequate phone credit to ensure calls could be made as required. Since the calls were made from each assessor's device rather than through a central telephonic system, a phone application was installed on their phones to record the calls made for the pilot. All recording applications were pretested before each round to ensure that the device could record the student's responses accurately throughout the assessment. Recordings were used as an internal backup for any post-assessment audits, as well as to triangulate results recorded on data sheets, as needed.

2.3.2. Content Development

In consultation with the team that implemented a phone-based learning intervention in Botswana that focused on foundational math skills (number operations and place value tasks), the World Bank country teams adapted the original instruments (received from Young 1ove in Botswana) according to country context with support from either the local implementing partner (in the case of Nepal) or a government agency (in the case of Ghana). A curriculum mapping exercise was undertaken to align the math assessment tools used in the Botswana intervention with the country-specific curriculum in Ghana and Nepal.[15] In Nepal and Pakistan, literacy assessment tools were also developed to assess foundational literacy.

Ghana

In Ghana, the pilot focused on the foundational math skills of students in grades 2 through 5. NaCCA, an independent statutory body focused on developing curriculum and assessment standards in Ghana, collaborated with the World Bank to adapt the assessment materials received from a phone-based learning intervention in Botswana. A remote workshop was organized in two phases, described below.

1. Developing and adapting grade-specific assessment content: Four numeracy experts participated in a one-day orientation and a five-day workshop to develop and adapt grade-specific assessment materials. During this workshop, the mathematics experts (a) studied the content of the package of pilot materials provided by the World Bank team and aligned them with Ghana's Standards-Based Curriculum (SBC); (b) drafted a template and developed a curriculum-based assessment framework indicating the strand, sub-strand, content standards and indicators, and the key behaviors or skills on the number operations content for grades 2 through 5; and (c) used the draft curriculum-based framework to develop grade-specific content and tasks for the materials provided.
2. Translation of materials into four Ghanaian languages: Four Ghanaian language experts fluent in Ashante Twi, Dagbani, Ewe, and Ga participated in a five-day workshop to translate the assessment materials for grades 2 and 3 that were prepared in phase 1. Ghana's language policy mandates the use of a child's first language as the language of instruction from kindergarten through grade 3 and the use of English as the language of instruction starting in grade 4 (USAID 2020). Therefore, assessment materials for grades 4 and 5 were not translated into local languages.

Nepal

In Nepal, Street Child developed the literacy tool (informed by ASER) and adapted numeracy tools (previously developed by Teach for Nepal with technical guidance from Young 1ove) to be used in the pilot. Pratham USA provided technical guidance in the development of literacy tools. The foundational numeracy and literacy materials were first aligned to Nepal's curriculum and then were adapted according to the country context. The language of assessment used was Nepali, but the tools were also translated into English.

Pakistan

In Pakistan, the development of assessment tools also followed two phases: (1) development of assessment tools in English and (2) translation of assessment items into local languages. Phase 1 was informed by a desk review of EGRA and ASER literacy instruments and consultations with experts

[15] In Nepal, a curriculum mapping exercise was undertaken for the literacy instrument as well.

on learning assessment tools. The National Curriculum 2006[16] was used to align assessment items with competencies and student learning outcomes mentioned in the curriculum. Once developed, the assessment instruments were translated from English into five local languages (Urdu, Punjabi, Sindhi, Pashto, and Balochi). The assessment instrument comprised 12 items.

Test items were developed for all three technology solutions (SMS, IVR, and LPC), but ultimately various skills and content were assessed using two different modalities (SMS and LPC). For example, sentence reading was assessed through LPC but not by SMS. The skills assessed through each technology solution in Pakistan are described in table 3.

Table 3: Skills and Content Assessed in Pakistan by Modality

LPC	SMS
Letter recognition	Letter sound discrimination
Word recognition	Alphabet letter recognition[17]
Sentence reading	Non-word reading
Reading comprehension	Reading comprehension
Lexical aspects of language	Lexical aspects of language

2.3.3. Pilot Setting and Sampling

Across all three countries, pilot regions and a predominant local language were identified after consultations with ministry counterparts. First contact interviews (see Section 2.3.4. for more details) were conducted to recruit participants for the pilot study. The region, gender, and grades of children were identified as the main stratification criteria. Within regions, participants were recruited from both rural and urban areas using caregivers' preferred language for communication.

Ghana

In Ghana the pilot was conducted in four regions (Ashanti, Greater Accra, Northern, and Volta), selected to include representation from each of three zones—Northern, Middle, and Southern.

The sample (including alternates) consisted of 259 students across grades 2–5 and their caregiver. Caregiver-child pairs were recruited from the four regions. The alternate caregiver-child pairs were recruited in anticipation of possible nonengagement in the implementation of the formative assessment solutions and the difficulty in meeting the quota for some strata.

Table 4 provides details of the characteristics of the children who were enrolled in the pilot study and randomized into the three technology solutions for assessment delivery. The data showed that 54 percent of children were female, and 46 percent were male. Approximately 3 percent of children had any form of disability.[18] From the total sample, 31 percent were second graders, 21 percent were third graders, 27 percent were fourth graders, and 21 percent were fifth graders. Ashante Twi was the most

[16] The Single National Curriculum (SNC) adopted by two of the provinces (Punjab and Khyber Pakhtunkhwa [KPK]) closely resembles the National Curriculum 2006 language framework. Further, the new curriculum is in the process of being implemented. Consequently, the 2006 framework was deemed most suitable for this pilot.
[17] Letter recognition and alphabet recognition have been differently named to account for the difference in the way the task was implemented in LPC as opposed to over SMS. Letter recognition was used in LPC where students read out the alphabets provided. In SMS this implementation was not possible, so students were asked to respond with a word that begins with the alphabet shared via SMS.
[18] Among children with a disability, the most common categories included limited mobility, including wheelchair use; intellectual disability; and health conditions.

widely spoken language, spoken by 39 percent of the children, followed by English, which was spoken by 34 percent of the student participants.

Table 4: Child Characteristics for Ghana

Variable	Total	SMS	IVR	LPC
Gender				
Male	120 (46%)	40 (49%)	38 (45%)	42 (45%)
Female	139 (54%)	42 (51%)	46 (55%)	51 (55%)
Grade				
Basic 2	80 (31%)	27 (33%)	24 (29%)	29 (31%)
Basic 3	55 (21%)	15 (18%)	17 (20%)	23 (25%)
Basic 4	70 (27%)	24 (29%)	22 (26%)	24 (26%)
Basic 5	54 (21%)	16 (20%)	21 (25%)	17 (18%)
Disability Status				
None	250 (97%)	79 (96%)	81 (96%)	90 (97%)
Yes, visually impaired	3 (1%)	0 (0%)	1 (1%)	2 (2%)
Yes, hearing impaired	1 (0%)	1 (1%)	0 (0%)	0 (0%)
Yes, other types	5 (2%)	2 (2%)	2 (2%)	1 (1%)
Primary Home Language				
English	87 (34%)	25 (30%)	32 (38%)	30 (32%)
Ashanti Twi	100 (39%)	31 (38%)	30 (36%)	39 (42%)
Ewe	37 (14%)	14 (17%)	9 (11%)	14 (15%)
Ga	9 (3%)	2 (2%)	3 (4%)	4 (4%)
Dagbani	26 (10%)	10 (12%)	10 (12%)	6 (6%)
Sample Size	**259**	**82**	**84**	**93**

Note: The variables listed in this table came from the survey that was administered in Ghana.

Nepal

In Nepal, the pilot included 2,163 students in grades 4 and 5 from 64 schools in three local governments (or municipalities): Ganeshman Charnath Municipality, Tripurasundari Rural Municipality, and Siddhakumakh Rural Municipality. Of the 2,163 students, 1,467 (68 percent) of those who were reached either fully completed (1,437) or partially completed (30) the assessment (see table 5). The pilot was part of TaRL intervention support in these local governments and the concept note for TaRL was endorsed by MoEST for implementation.

Table 5: Child Characteristics for Nepal

Variable	LPC (with SMS)
Gender	
Male	686 (48%)
Female	751 (52%)
Grade	
Grade 4	685 (48%)
Grade 5	752 (52%)
Primary Home Language	
Nepali	974 (68%)
English	198 (14%)
Both	180 (13%)
Other	70 (5%)
Missing/Unknown	15 (1%)
Total	**1,437**

Note: Table only includes participants who fully completed the assessment.

Pakistan

In Pakistan, the pilot was conducted in three provinces: KPK, Punjab, and Sindh. Although initially selected, Balochistan was dropped from the list of pilot regions due to delays in obtaining the required permission from authorities. A sample of 36 students was recruited across grades 3 through 5. The sample was distributed equally across the grades and regions.

The characteristics of children enrolled in the pilot study are shown in table 6. The data showed that 48 percent of the children were female, and 52 percent were male. None (0 percent) of the children had any form of disability.

Table 6: Child Characteristics for Pakistan

Variable	Total	SMS	LPC (with SMS)
Gender			
Male	18 (50%)	7 (39%)	11 (61%)
Female	18 (50%)	11 (61%)	7 (39%)
Grade			
Basic 3	13 (36%)	6 (33%)	6 (33%)
Basic 4	11 (31%)	6 (33%)	6 (33%)
Basic 5	12 (33%)	6 (33%)	6 (33%)
Primary Home Language			
English	0 (0%)	0 (0%)	0 (0%)
Pashto	9 (25%)	6 (33%)	3 (17%)
Punjabi	12 (33%)	6 (33%)	6 (33%)
Sindhi	11 (31%)	6 (33%)	5 (28%)
Urdu	4 (19%)	0 (0%)	4 (22%)
Sample Size	**36 (100%)**	**18 (50%)**	**18 (50%)**

2.3.4. First Contact Interview

This section outlines the procedure of contacting potential participants for the first time, how they were screened for eligibility, and how consent was obtained to enroll in the pilot study.

Contacting Potential Participants

The telephone assessors were given a database of phone numbers to contact. These numbers were obtained either using the random digit dialing (RDD) method (in the case of Ghana), or they were provided with support from schoolteachers or local government (in the case of Pakistan and Nepal, respectively) and were verified by the assessors by calling them. Telephone assessors dialed each phone number and screened out inactive phone numbers or rejected calls. The call attempt protocol involved calling the same number multiple times at various times of the weekday and weekend.

In Ghana, as the data collection progressed and the number of inactive calls increased, the recruitment protocol was adjusted by requesting referrals from enrolled households, whereby assessors would ask respondents to provide referrals to recruit relatives, friends, or neighbors whose children attended grades 2–5 in public schools and resided in one of the four pilot regions. A total of 3,432 calls were placed, lasting 20 minutes on average. Overall, 64 percent of caregivers who completed the first contact interview were contacted using RDD, and 36 percent were contacted using the referral system.

In Nepal, the assessors were provided the phone numbers of households with students enrolled in grades 4 and 5. These phone numbers were obtained from the local governments. The assessors called each number to verify the number and schedule the survey. Of the initial list of 2,305 students, 334 did not own mobile phones, so it was explored whether those students' neighbors had phones that could be used to reach the child. Ultimately, 2,163 students were identified as able to be assessed over phone. The survey (including the learning assessment) was either completed on the first call attempt or was scheduled for a time convenient to the student and caregiver. The average number of calls placed per student was five, and phone calls lasted 45–60 minutes when successfully completed.

In Pakistan, call lists were provided by provincial education representatives (provincial network) and verified by enumerators (teachers). A total of 36 calls were placed to potential participants.

Screening for Eligibility

The first contact interview began by determining the language for the interview, introducing the organizations involved (World Bank, ministry of education or other government agency, and local firms), and explaining the purpose of the call. Every active call included a series of screening questions to determine whether the potential respondent (the student's caregiver) met the inclusion criteria for the pilot study. The eligibility criteria required the caregiver to:

1. Own a working mobile phone (basic, feature, or smartphone) not exclusively for business purposes;
2. Be an adult, aged 18+, within an eligible household;
4. Be a resident of one of the pilot regions and speak any of the pilot local languages or English;
5. Be a caregiver of a school-going child who attends a pilot study grade in a public school;[19] and
6. Be willing to participate in the pilot of the formative assessment solutions (that is, implementation of the assigned solution).

[19] The adult caregiver is defined as a person who is responsible or shares responsibility for the care and education of children in the household and can talk about their experiences in school and at home.

Consenting caregivers who were deemed eligible were enrolled in the pilot study.

In Ghana, due to the RDD recruitment method, roughly 45 percent of the people who answered the calls were ineligible to participate. The main reasons for ineligibility were that a caregiver "has no child in the target grade" (40 percent) and was "not a resident in pilot study region" (38 percent).

In Nepal, phone numbers were obtained from the local governments. In general, caregivers who did not have a child in the targeted school grade and did not live in the pilot municipalities were ineligible. A small share of students was also ineligible for the pilot because of difficulty with reading fonts on the phones. Only 6 percent of the children said that they had trouble reading Nepali text on their caregiver's phone, and three-fourths of those children had another phone available at home which they could use for the Nepali language assessment.

In Pakistan, all contacted caregivers had children who were eligible to participate in the pilot study.

Consenting and Enrolling Eligible Participants

After determining eligibility, consent was obtained from the eligible caregiver for themselves and for the selected child to participate in the pilot study and data collection process, including post-pilot interviews and a feedback survey (employed in Ghana and Pakistan). The caregivers who did not agree to participate in the study and data collection were not enrolled to participate in the pilot. Once verbal consent was obtained, eligible research participants were enrolled in the pilot study. Verbal consent to record study communication for the duration of the project was also obtained for quality assurance purposes. During the calls, caregivers were informed that the assessments were not high-stakes exams, and the individual results would not be made public.

In Nepal, the assessors first spoke to the caregiver to seek their consent to talk with their child. Fifty percent of the respondents were mothers, almost 32 percent were fathers, and the remaining 18 percent were other family members. To incentivize participation, they were informed that Rs. 100 would be transferred to their mobile numbers to be used for LPC expenses later.

Communicating to Participants about the Intervention Modalities

Upon consenting to be part of the pilot study, the eligible participants in Ghana and Pakistan were assigned to one of the remote formative assessment solutions. In Ghana, children with disabilities, such as those who had low vision or were blind or partially sighted, were purposively assigned either LPC (or IVR, in the case of Ghana), and those who were deaf or hard of hearing were purposively assigned to receive SMS quizzes. Consent to participate in the pilot was followed by communication to eligible participants regarding the modalities, including when the pilot would begin, how the assessor would reach the caregiver and child, how frequently they will receive the communication, the mode of delivery of assessments and quizzes, and instructions on how to respond to quizzes.[20]

2.3.5. Pilot Implementation

Before the pilot rollout, all the technology systems and the sequence of assessment items were tested internally. The pilot activities in all three countries lasted two weeks. In Ghana, the pilot ran from October 4 to 19, 2021. In Nepal, the pilot took place in the month of October 2021. In Pakistan, the pilot was implemented between October 15 and 30, 2021.

[20] This step did not take place in Nepal because the assessment was implemented as soon as a caregiver's and child's approval were obtained. Moreover, the pilot in Nepal included only the use of LPC for both numeracy and literacy assessment.

Rollout

In Ghana and Pakistan,[21] the pilot rollout followed this script map: a pre-welcome message, a welcome message, quizzes on place value with corresponding corrective feedback, and a concluding message.

1. Pre-welcome SMS nudges and LPC: Before the pilot began, SMS nudges and LPC were used to remind participants about the start of the pilot and the phone number and the medium (SMS/IVR/LPC) through which the learning content would be delivered to the participant.
2. Welcome message: A welcome message—consisting of the introduction of the key organizations that were responsible for implementing the formative assessment, the purpose of the study, an example quiz with answers, and a guide to responding to the learning content—was sent to all participants at the start of the pilot. Participants then received their first quiz with the corresponding answer choices. Pre-welcome and welcome messages were sent to all the participants across the technology solution groups in the pilot study.
3. Assessment quizzes: For SMS (and IVR in case of Ghana), the foundational literacy or numeracy quizzes were sent one at a time. Participants' replies to the quizzes were sent to the server that processed the responses and triggered a follow-up prompt, usually in the form of corrective feedback and subsequent quizzes or content. The corrective feedback clearly indicated whether the child's response was (a) incorrect or (b) correct and included an explanation. After reading or listening to the corrective feedback, the child was asked to acknowledge receipt of the message (in the case of SMS) or to indicate whether the explanation made sense (by pressing a button in the case of the IVR or by saying "yes" or "no" in case of LPC).

   The second and third quizzes were automatically generated once participants replied to the first and second quizzes, respectively.
4. Concluding message: The assessment ended with a concluding message that thanked the participant and communicated the next steps.

In Pakistan, participants assigned to the LPC group received the literacy assessment items via SMS before the assessor called them over phone.[22] In Ghana, participants in the LPC and IVR groups did not receive text messages for numeracy assessment tasks; the assessor or system read out the questions.

As mentioned above, in Nepal the LPC assessment was completed in the first contact interview, or it was scheduled for a time convenient to the student and caregiver. Once the caregiver's consent for the assessment administration was obtained, the child's assent was also obtained, and the child was asked to bring a piece of paper and pencil to the phone. The assessor also requested that the caregiver not help their child with the assessment. The caregiver was asked to put their phone on loudspeaker mode to the extent possible and hand the phone over to their child. Once the caregiver gave the phone to the child, the assessor spent some time building rapport with the child. For instance, assessors were instructed to talk with children politely and in a friendly manner to ensure that each child felt comfortable. The assessment began with foundational math questions, which the assessor asked verbally. The math assessment was followed by the literacy assessment, for which a supporting text message was sent to the caregiver's phone.

[21] In Nepal, pre-welcome and welcome messages were not sent to participants since the pilot primarily focused on LPC.
[22] This approach was used because the literacy assessment included assessing skills such as letter recognition, word recognition, sentence reading, and reading comprehension, which cannot be assessed unless the child sees the text.

Reminders

In cases of nonresponse and incomplete responses, in Ghana and Pakistan,[23] reminders were sent via SMS or LPC to prompt the student to complete the quiz. The number of follow-up contacts depended on the technology used. For SMS, the reminders were sent at varying hours to target the availability of the participants. In the case of LPC, if both the caregiver and the child were not available or did not answer the phone, the call session was rescheduled to a more convenient time. In Ghana, each prerecorded IVR audio content was sent once, followed by a maximum of nine redials at different times over one day and across days for nonresponse. Moreover, all nonengaging participants were called a day before the end of week 1 and week 2 to prompt them to complete the place value quiz and to understand any barriers that were preventing their participation. In Pakistan, reminders to respond to the quiz also were sent to an alternate contact number.

Language of Assessment

The learning content was delivered based on the grade of the child and their preferred language.

In Ghana, the SMS system used the English language across all grades due to character length limitations for implementing SMS. For the IVR and LPC systems, learning content for grades 2 and 3 was delivered in the child's preferred language—English or any of the four local languages. For grades 4 and 5, the assessment was deployed only in English to align with the use of English as the medium of instruction at the upper primary level in Ghana.

In Nepal, the assessment was delivered in Nepali across grades 4 and 5. Assessors allowed children to use Arabic digits for numeracy if they preferred to do so, but the verbal communication was in Nepali for the literacy and numeracy assessments.

In Pakistan, caregivers were allowed to select the language in which their child would take the assessment. The assessment was available in Urdu, English, and three local languages spoken in each province (Pashto, Punjabi, and Sindhi) across all grades for both SMS and LPC.

2.3.6. Post-Pilot Interview/Survey

In Ghana, after the pilot, an in-depth interview with both the caregiver and child collected information on reasons for not engaging in the formative assessment process, the level of satisfaction with the assessment technology and assessment content, cognitive demand of the phone-based assessment, technical challenges encountered with responding, and any suggestions. The assessors in Ghana contacted 64 caregiver-child pairs who (a) never engaged with assessment, (b) engaged with the assessment only once (1 out of 2 times), and (c) engaged with the assessment consistently (2 out of 2 times). The results of in-depth interview are discussed in Section 3.

In Pakistan, all potential participants were sent a short post-pilot survey, which was delivered through LPC, regardless of whether they participated in the LPC or SMS assessment. The purpose of this short survey was to gather information on various aspects of the pilot, including:

- Did they find the exercises useful to assess their child's proficiency levels?
- What was the difficulty level of the assessment exercise in general?
- How did they find the experience of responding to the assessment questions over LPC or SMS?
- Were they able to read the SMS texts easily?
- In which language did they prefer such an assessment?

[23] Reminders were not sent to participants in Nepal since the assessment took place during the first contact interview.

3. Results

This section discusses the results of the piloting three technology solutions and looks at factors such as level of participation, reasons for non-engagement, difficulty of the assessment, nature of parental support provided and the overall satisfaction with the delivery of pilot.

3.1. Level of Participation in Formative Assessment

To measure the level of participation, in Ghana and Pakistan, two key performance metrics were measured: (1) engagement rate and (2) persistence rate.[24] *Engagement rate* is the percentage of all recruited participants who responded to any quizzes. *Persistence rate* is the percentage of those who completed quizzes in both week 1 and week 2.

3.1.1. Ghana

Of 259 participants, 65 percent received the SMS or picked up the calls (IVR and LPC). The overall engagement rate was 46 percent for the two-week deployment of the 6-question place value quizzes, suggesting that almost half the participants not only opened the SMS or picked up the call and read or listened to the place value quizzes, but also answered the quizzes. Engagement patterns differed significantly, however, across the three formative assessment solutions (see table 7): 21 percent of those assigned to the SMS group answered all quizzes in week 1 and week 2, while 27 percent and 33 percent of participants assigned to the IVR and LPC groups, respectively, answered both week 1 and week 2 quizzes. Overall, more participation was seen in week 2 quizzes compared to week 1 quizzes. The results also suggest that most participants seemed to prefer voice calls to text messages. The engagement rates in LPC were higher than IVR, which in turn were higher than SMS.

Table 7: Engagement Patterns by Formative Assessment Solution in Ghana

Engagement	SMS	IVR	LPC	Total
None	57%	51%	54%	54%
Week 1 only	9%	10%	4%	7%
Week 2 only	13%	12%	9%	11%
Week 1 and Week 2	21%	27%	33%	27%
Sample Size (Total)	**82 (100%)**	**84 (100%)**	**93 (100%)**	**259 (100%)**

3.1.2. Pakistan

As shown in table 8, 58 percent of households engaged consistently in both rounds 1 and 2 of the pilot, although the engagement was higher when the child was reached through LPC (89 percent) as compared to only SMS (28 percent).

[24] Only engagement patterns were measured in Nepal since the assessment was conducted at the time of the first contact interview.

Table 8: Engagement Patterns by Formative Assessment Solution in Pakistan

Engagement	SMS	LPC (with SMS)	Total
None	13 (72%)	2 (11%)	15 (42%)
Round 1	5 (28%)	14 (78%)	19 (53%)
Round 2	N/A	10 (56%)	10 (27%)
Round 1 and Round 2	N/A	16 (89%)	21 (58%)
Sample size (Total)	**18 (100%)**	**18 (100%)**	**36 (58%)**

3.1.3. Nepal

In Nepal, the pilot followed a staggered approach wherein participants were contacted in smaller groups, and the assessment was completed during the first contact interview. To begin, each assessor was provided a list of approximately 20 households with the student's name, school grade, and phone number. After receiving the assessment data for group 1, calls to group 2 were initiated, followed by groups 3, 4, and 5. As the calls progressed to subsequent groups, efforts were made to resolve issues wherever possible, such as obtaining correct phone numbers for students from their schools since the schools were open. See table 9 for information on engagement patterns.

Table 9: Engagement Patterns in Nepal

Group	Total participants	Participants who engaged	Engagement rate
Group 1	514	384	75%
Group 2	484	269	56%
Group 3	806	535	66%
Group 4	92	56	61%
Group 5	267	223	84%
Total	**2,163**	**1,467**	**68%**

3.2. Reasons for Not Participating in the Pilot Study

3.2.1. Ghana

In Ghana, in-depth interviews with 32 child-caregiver pairs revealed that 18 percent of caregivers and children who participated in the in-depth interview reported not receiving the math lessons. Among interview respondents, 82 percent of caregivers and children reported receiving the quizzes; 63 percent of this group said that they did not read or listen to the math lessons, and 37 percent of this group indicated that they read or listened to the math lessons (a) alone and (b) with the caregiver/child. It thus appears that some participants received the message but did not read or listen to the content.

Not being with the child (48 percent) and technical reasons (24 percent) were the two main reasons given for non- or partial participation. Technical reasons mentioned included difficulty in understanding audio recordings; inability to navigate through the options or use the keypad; network-related challenges; phone-related issues such as a lost phone, broken screen, or dead battery; and unclear questions. Other reasons cited for not participating or for participating partially in the math

lessons were lack of time to participate or to call back after missed calls, nonreceipt of the lessons, caregiver and/or child sickness, lack of understanding of the instructions, lack of interest in the math lessons, inability to read and write, and forgetting to give the phone to the child. When asked what would make caregivers and children participate in a similar program in the future, responses centered on the availability of the caregiver and/or child, the timing of delivery of the math lessons, preference for other modes such as SMS and LPC, focus on addition lessons, consistency of the lessons, and acquisition of new phones.

3.2.2. Nepal

Of the 2,163 students reached, 1,467 (68 percent) fully or partially completed the calls with assessors. Reasons for nonparticipation are provided in table 10.

Table 10: Categorization of Call Outcomes

Outcome	Number	Share
Fully completed	1,437	66%
Partially completed	30	1%
Child unavailable or did not talk	10	1%
Phone number wrong or invalid number	243	11%
Child's information (grade or name) incorrect	43	2%
Child moved to another school	73	3%
Not reachable	169	8%
Switched off	54	3%
Call not received	31	1%
Call ended automatically	34	2%
Incoming call barred	32	2%
Call rejected and other reasons	7	0%
Total Cases	**2,163**	**100**

3.2.3. Pakistan

In the case of those opting out of the pilot study, the three top reasons—reported in a post-pilot qualitative study of all potential and actual participants—for nonparticipation were:

1. Caregiver did not have time to support participation in the assessment exercise.
2. Connectivity issues prevented participation.
3. SMS sent were deleted, and therefore participation was not possible.

3.3. Cognitive Demands for the Formative Assessment Tasks and Parental Support

This section discusses how difficult or easy the participants found the assessment tasks to be when administered over phone and the kinds of parental support that students received during the pilot.

3.3.1. Cognitive Demand

The interviews with children presented different perspectives on the cognitive demands[25] of place value quizzes. In Ghana, less than a quarter (21 percent) of children found the place value lessons difficult (see table 11). In Pakistan, 10 percent found literacy tasks to be difficult, while 64 percent found them to be of moderate difficulty.[26]

In Ghana, when asked why they found the place value lessons difficult, most of those children indicated their inability to grasp the concept of identifying the place value of a number. Relatedly, almost all children (96 percent) considered the instructions to be precise and clear—all participants viewed the instructions as helping them to know what to do for the tasks. The in-depth interview responses from sampled children further Showed that the corrective feedback that was provided was not only useful in helping the children understand why they got the quiz right or wrong (93 percent); the feedback was also helpful for responding to other questions (86 percent). The vast majority (96 percent) of the children expressed that they would have liked to have had the chance to retry a quiz or question they got wrong.

Table 11: Assessment of Cognitive Demands for Assessment Tasks in Ghana

Statements	Percent agreed
Tasks were difficult	21%
Instructions were precise and clear	96%
Instructions help you know what to do on the tasks	100%
Find the immediate feedback on your answer useful in understanding why you got the quiz right or wrong	93%
Would have liked to have a chance to re-try a question I got wrong	96%
Use the feedback I got for a previous quiz to answer another quiz	86%
Find the questions over the phone more difficult than on paper	29%
Caregiver helped or supported me in any way to participate in the place value lessons	71%
Sample Size	28

In Pakistan, when asked about the usefulness of literacy tasks to assess their child's proficiency levels, 72 percent of caregivers found them to be very useful, and 64 percent of caregivers found the assessment to be moderately difficult. Only 18 percent preferred the assessment to be conducted in a local language. See table 12 for more information.

[25] Cognitive demand is the difficulty of solving a task when it is administered over the phone.
[26] In Nepal, information on cognitive demand was not collected.

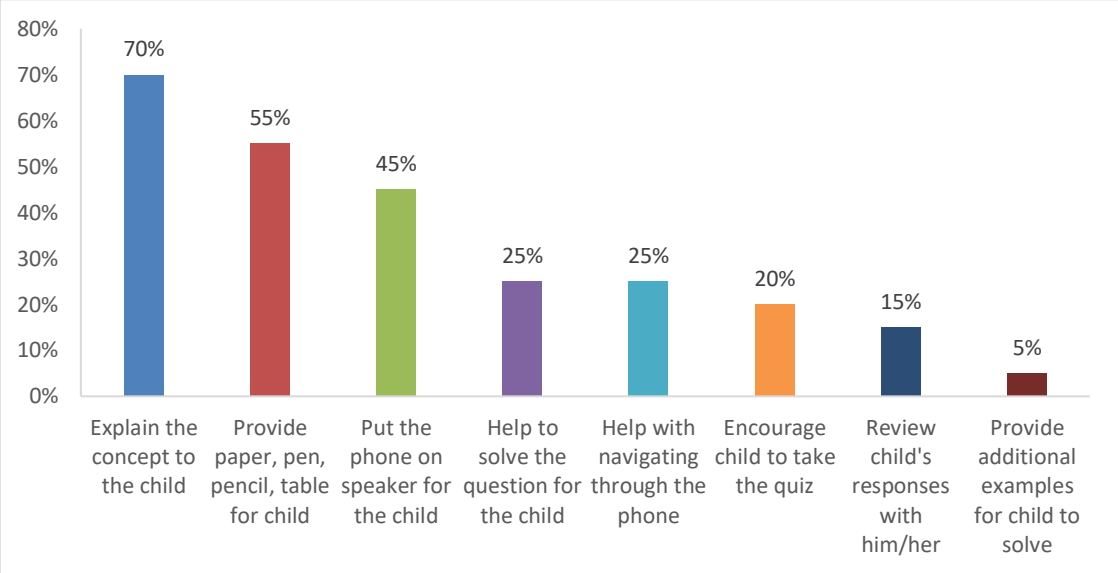
Table 12: Assessment of Cognitive Demands for Assessment Tasks in Pakistan

Statements	Percentage agreed
The exercises were useful to assess their child's proficiency levels.	72%
The assessments were moderately difficult.	64%
It was easy to respond to the assessment questions over LPC/SMS.	91%
Were able to read the SMS texts easily.	60%
Would have preferred the assessment to be conducted in local language.	18%

3.3.2. Parental Support

In Ghana,[27] caregivers helped their children, but mostly in ways that were consistent with instructions. Only 29 percent of children reported that their caregivers did not help or support them in any way to participate in the place value lessons. Of the 71 percent who reported receiving support from caregivers (see figure 5) the most frequent form of support was the caregiver's explaining the place value concept to the child (70 percent).

Figure 5: Nature of Support Provided by Caregivers to Children in Ghana



[27] Information on parental support was not collected in Nepal and Pakistan.

3.4. Satisfaction with the Length of the Pilot and Program Delivery

The section discusses participants' satisfaction with the length of the pilot and delivery of the program.

3.4.1. Ghana

In Ghana,[28] about 97 percent of participants (caregivers and children combined) reported that the learning content was engaging. Almost 74 percent considered the learning content easy to understand. When asked what they thought about the number of quizzes they received per week, about 74 percent reported that three quizzes per week was a good amount, while 23 percent preferred to receive more quizzes per week. On average, most participants who preferred more quizzes indicated about 6 quizzes per week, with a range of 4 to 15 quizzes per week. When asked about their view on the duration of the place value lessons, about 66 percent said that two weeks was an adequate amount, while 26 percent thought it was too short. Only 5 percent were of the view that two weeks of place value lessons were too long.

Furthermore, about 58 percent and 38 percent of participants reported that they were satisfied and very satisfied, respectively, with the math lessons on place value. Among participants assigned to the IVR and SMS groups, 76 percent and 71 percent, respectively, were more likely to be satisfied with the math place value lessons. About 62 percent of participants assigned to the LPC group were very satisfied with the place value lessons (see table 13). One caregiver said, "[The child] was finding it difficult in the beginning. How to go about it wasn't simple, but as the program went on it became easier, and I realize it will help the child."

Table 13: Satisfaction with Math Lessons

	No response	Not satisfied	Satisfied	Very satisfied	Total
IVR	0 (0%)	1 (4%)	19 (76%)	5 (20%)	25 (100%)
LPC	0 (0%)	0 (0%)	13 (38%)	21 (62%)	34 (100%)
SMS	2 (14%)	0 (0%)	10 (71%)	2 (14%)	14 (100%)
Total	**2 (3%)**	**1 (1%)**	**42 (57%)**	**28 (38%)**	**73 (100%)**

When asked about the aspects of the math lessons on place value with which they were satisfied, about 28 percent of participants each reported being satisfied with all aspects and with the place value demonstrations. About 14 percent of participants were satisfied with the place value topic. One participant shared, "I was satisfied with the fact that I was taught what I didn't know." Other participants expressed their satisfaction with being able to answer all questions, with the child's progress in week 2, with the corrective feedback provided, with the professionalism of the call center operatives, with the question format and responses, with the use of phones to learn, with the free nature of the lessons, with the linkage to curriculum, and with government support and the mode of delivery. One caregiver expressed his satisfaction with how the child progressed: "All was okay for me. I was initially worried that he would make mistakes with the way he was answering the questions. But I later bought him drinks given his performance on the math lessons."

[28] Information on satisfaction with the length of the pilot and program delivery was not collected in Nepal.

4. Costs

The three main costs for the SMS, IVR, and LPC systems that were identified during the implementation of the pilot were:

1. Personnel time for the technical design and setup of the SMS and IVR platforms,[29]
2. Setting up of virtual call centers for LPC, and
3. Provision of short code for SMS.

Table 14 provides information on the total cost and the unit cost per technology solution incurred during the pilot in each of the three countries.

Table 14: Total and Unit Cost of Implementing SMS, IVR, and LPC Solutions

	Total cost (US$)			Unit cost (US$)		
	SMS	IVR	LPC	SMS (per text sent)	IVR (per minute)	LPC (per minute)
Ghana	US$44.22	US$17.69	US$17.69	US$0.04	US$0.03	US$0.0330
Nepal	US$191.7	N/A	US$1557.8	US$0.016	N/A	US$0.021
Pakistan	US$10	N/A	US$100	US$0.001–US$0.004	N/A	US$0.02–US$0.3

Note: The costs mentioned only include the cost of sending a message or placing a call. Assessor time and/or fixed costs are mentioned in appendix B.

See appendix B for more details on the costs associated with this pilot study in each country.

[29] These activities included creating or adapting the assessment content for each grade, reviewing and translating the developed content, performing quality control checks, recording audio scripts based on the content, training call center operatives, internal pretesting, finalizing scripts, uploading content to application platforms, and deploying contents in a sequenced manner.
[30] The cost of placing LPC including assessor cost is $0.53 per minute.

5. Considerations for Scaled-Up Implementation

Based on the results of the pilot and feedback from participants, this section considerations for planning for a scale-up implementation.

5.1. Assessment Design

The objective of the assessment needs to be clarified up front, and this objective will vary with context. Formative assessments usually aim to understand the learning level of the student to identify knowledge gaps and inform subsequent instruction. As these assessments are usually carried out by teachers as part of their daily teaching practice, phone-based formative assessments at scale would also involve teachers as implementers of these assessments and/or as users of the assessment results. One vision for the use of these assessments could be as complements to print-based or other low-tech learning content, where phone-based formative assessment is used to gauge students' use of and understanding of this content; provide the student with useful real-time feedback; and inform teachers' decisions on additional or subsequent content that should be given to the student, also known as the "teach at the right level" (TaRL) approach. While the three pilots focused on logistical and technological feasibility rather than on the formative nature of the assessments, the ongoing TaRL intervention in Nepal will yield lessons on the usefulness of phone-based assessments in the context of a TaRL intervention.

Phone-based assessments might be more suitable for early grades and simple concepts. Due to the limited ability of basic phones to share advanced content such as graphs, maps, or long reading passages, phone-based assessments might be more relevant at lower levels of education to assess simple concepts.

The language of assessment must be carefully selected after reviewing the language of instruction policy of the country. This process can also include a more rigorous large-scale exercise for mapping language preferences before conducting assessments in local languages. In addition, technological factors, such as the character limit for SMS and language script compatibility with phones, should be confirmed beforehand. For voiced solutions such as IVR and LPC, use of multiple languages may require the hiring of several voice artists or assessors who are proficient in selected languages.

Understanding local contexts, especially the languages used for the interventions, how the questions are set and asked, and understanding the interventions provided at schools, are crucial for the success and legitimacy of interventions. More time may need to be allowed to improve the content, language of assessment, design of the intervention, training of teachers, etc. Adapting learning resources both to a local context and to SMS, IVR, or LPC format is a real challenge and requires an iterative approach with testing.

It is critical to **work closely with the government education authority** from the very beginning of designing the interventions. It is also important to present the potential benefits of phone-based formative assessments to the education authorities and discuss implications for current and possible future interventions to improve learning outcomes in the country.

Alignment of assessment with the school curriculum and calendar is critical to ensure the assessment tasks are at the right level. Moreover, necessary time should be given such that once implementation starts, it can proceed without delays or breaks.

5.2. Time and Human Resources

Hiring the right people to implement various aspects of assessment is most crucial at the design and piloting stage. A significant amount of personnel time is spent on the technical design and setting up of the platforms. These activities include rigorous curriculum mapping, creating or adapting the assessment content based on each grade, reviewing the content developed, translating assessment content, performing quality control checks, recording audio scripts based on the content, training of call center operatives, internal pretesting, finalizing scripts, uploading the content on the application platform, and deploying the contents in a sequenced manner. Audio recording of assessment content should be done by professional voice actors or artists who are fluent in selected languages.

Sample design at the pilot stage needs to leverage existing school systems. Sampling methods such as RDD may take longer to reach the target group size since calls have to be placed to all phone numbers. This task can be avoided by collaborating with a government agency that can provide a list of schools and grades and that can reach out to school principals through their own channels to obtain caregivers' phone numbers. Using school systems could not only increase the sample, but also increase precision in reaching those whom this intervention aims to target.

Training of teachers during scale-up is critical as formative assessment is ultimately an instrument to inform teachers' instruction. Teachers can be implementers of phone-based formative assessment, as was done in Nepal, or they can just be the end users of formative assessment results, where experienced assessors administer the phone-based formative assessment, but the results are reported, in a timely and actionable manner, to students' teachers to help them identify struggling students or topics that need to be reinforced with additional remote content or in-person instruction.

5.3. Technology Solutions

Necessary technology infrastructure at the country level is another critical aspect for the success of phone-based assessments, including having mobile aggregators that can provide infrastructure to set up a virtual call center using necessary software and web applications and provision of short code for SMS and long code for IVR and LPC.

Use of well-developed learning management systems in activities such as scoring, data collection, and reporting, with some quality-focused post-reviews using teachers and literacy experts, can enhance efficiency.

Setup and efficient use of virtual call centers, in the case of LPC, should be supported by provision of appropriate software that allows for easy recording, placing, and tracking of various operational aspects, such as call attempts and redialling.

Reminders in cases of non- or incomplete response are necessary and should be sent within 24 hours to increase participation in remote formative assessments.

Pretesting all assessment items and technology before the rollout with a focus group can help make the assessment more relevant. This process could result in the identification of items that are too cognitively demanding for phone-based administration, particularly for disadvantaged students, and personnel could either revise the items or remove them from phone-based administration.

Comfort with mobile phone technology is key. Although mobile phones are more widespread than most other digital technologies, not covered in this work, but possibly worth understanding and exploring, is the extent to which poor performance on an assessment is simply due to a participant's low comfort level with the use of phones versus inadequate understanding of concepts covered by the assessment.

Phone-based assessments (despite their low-tech nature) could be more challenging for participants who come from disadvantaged backgrounds, where children may have considerably less access even to basic phones and hence lower proficiency in handling them. In such contexts, with sufficient resources, support can be provided to users, such as leaflets that can be mailed to guide instruction on using even basic phones for the assessment purposes.

5.4. Costing

No additional costs should be incurred by families and assessors for their participation. The cost of sending SMS and receiving and placing LPC should be incurred by the implementing or donor organization. This arrangement could be facilitated by providing airtime or phone credit or mobile data to participants and assessors.

5.5. Participation and Engagement

Trust can be developed by mentioning the names of the organizations involved. To increase participation in the pilot activity, when placing the calls, the assessor should clearly mention the names of the organizations involved in the pilot study. This information would increase the legitimacy of the intervention and may increase not just engagement, but also completion.

To ensure participation of all students, it would be helpful to involve head teachers in providing information to caregivers in advance of the calls. This action will help build trust among caregivers to allow students to participate in the survey.

Assessors should also mention that the assessments are not high-stakes, and the data collected will be kept confidential.

Sending a pre-welcome message before the start of the pilot study can help remind caregivers of their enrollment.

Use of regular reminders in case of nonresponse or incomplete responses and feedback in case of correct or incorrect response should be considered to keep participants engaged. Reminders and feedback also make the assessments more interactive.

6. Conclusion

This report presented the experience of Ghana, Nepal, and Pakistan in testing the feasibility of using SMS, IVR, and LPC to conduct formative assessments of foundational math and/or literacy skills for primary-grade students. The use of basic mobile phone features such as calls and SMS allowed for assessments that, crucially, reached children without smart devices and/or internet connectivity. The approach tested in this pilot is noteworthy not just for its effectiveness but also for its cost-effectiveness.

In the future, the use of basic mobile phone technology for the purposes of student assessment should not be limited to pandemic-related school closures. It can be used for any kind of disruption, including prolonged teacher strikes, school closures due to violence and insecurity, and disruptions caused by climate-related events. Moreover, these assessments can be used even during periods of regular schooling to complement in-classroom assessments. Phone-based remote formative assessments could be used as a complement to the growing availability of digital learning resources, such as the Edmodo Learning Management System in Ghana. Additionally, the Government of Nepal is planning to launch a learning recovery plan, and the approach used in this pilot could be an important component of this plan.

Appendix A: Team Composition in Each Pilot Country

	Assessors	Content developers	Translators	Administrative staff
Ghana	4 (proficient in a local language and English)	4	4	n.a.
Nepal	21 (proficient in Nepali and English)	1	1	1
Pakistan	6 (proficient in Urdu and English)	2	5	2

Appendix B: Costs Associated with Remote Phone-Based Assessments in Ghana, Nepal, and Pakistan

Table B.1: Ghana

Items	SMS	IVR	LPC
Participants who fully or partially engaged in the quizzes	35	41	43
Engagement rate (for reference, not used in calculations)	43%	49%	46%
Marginal cost calculation—relevant for scaling			
Units (per participant)			
Messages sent	34	0	0
Airtime minutes used	0	17	18
Minutes of interviewer time	0	0	18
Unit costs			
Cost per message (total cost/ (messages sent + participants who engaged fully or partially in the quizzes)	US$0.04	n.a.	n.a.
Cost per airtime minute[31] (total cost / (airtime minutes used + participants who engaged fully or partially in the quizzes)	n.a.	US$0.03	US$0.03
Cost per minute of interviewer time (includes a supervisor) ([total cost / (airtime minutes used + participants who engaged fully or partially in the quizzes] – [cost per airtime minute])	n.a.	n.a.	US$0.53
Total costs (variable)	US$44.22	US$17.69	US$432.82
Marginal cost per participant per solution (total cost / participants who fully or partially engaged in the quizzes	US$1.26	US$0.43	US$10.07
Average cost calculation—relevant for implementation on scale of the pre-pilot			
Fixed Costs			
Communication campaigns technical design & setup	US$3,187.50	US$3,187.50	n.a.
Content design	US$637.50	US$637.50	n.a.
Translation and recording[32]	n.a.	US$2,390.63	n.a.
Short code	US$3,100.00	n.a.	n.a.
Platform subscription	US$637.50	US$637.50	US$1,275.00
Call center setup	n.a.	n.a.	US$6,375.00
Interviewer training	n.a.	n.a.	US$2,648.49
Total costs (fixed)	US$7,562.50	US$6,853.13	US$10,298.49
Total costs (fixed + variable)	US$7,606.72	US$6,870.82	US$10,731.31
Average cost per participant	US$21.33	US$167.58	US$249.57

Note: All costs are based on loaded prices, which include taxes and indirect costs.

[31] Cost for airtime is same for IVR and LPC. The difference in total cost is a result of accounting for assessors' time in LPC.

[32] This expense includes hiring voice actors to record the IVR scripts. There was no translation cost for LPC as the translations were done by the field team as part of field staff training.

Table B.2: Nepal

Total cost of SMS + LPC	US$
Cost of SMS + LPC	1,750
Incentive for caregivers (phone charge-back of Rs. 100)	1,198
Content development	1,491
Total	**4,439**

Table B.3: Pakistan

	PKR	US$[33]
For SMS		
Design and development of a sophisticated SMS platform (one-time)	5–10 million	30,000–60,000
Short code approval and setup across all mobile operators (one-time)	1.5–2 million	9,000–12,000
Operations and maintenance to add or update assessment items	0.1–0.25 million per month	600–1,500 per month
Cost per SMS (zero-rated/free for users)	0.25–0.75 per SMS	0.001–0.004
For LPC		
Information System	10–12 million	60,000–70,000
Connectivity setup costs	2–3 million	11,500–18,000
Agent costs/seats	0.1–0.2 million per agent per month	600–1,200
Cost of calls	3–5 per minute	0.02–0.30 per minute

[33] This cost would vary with exchange rate differences and has been rounded up for approximate numbers.

References

Angrist, Noam, Peter Bergman, David K. Evans, Susannah Hares, Matthew C. H. Jukes, and Thato Letsomo. 2020. "Practical Lessons for Phone-Based Assessments of Learning." Working Paper 534, Center for Global Development, Washington, DC. https://www.cgdev.org/publication/principles-phone-based-assessments-learning.

Barron Rodriguez, Maria, Cristobal Cobo, Alberto Muñoz-Najar, and Iñaki Sánchez Ciarrusta. 2021. *Remote Learning During the Global School Lockdown: Multi-Country Lessons.* Washington, DC: World Bank. http://hdl.handle.net/10986/36141.

Clarke, Marguerite. 2012. "What Matters Most for Student Assessment Systems: A Framework Paper." SABER (Systems Approach for Better Education Results) Working Paper 1, World Bank, Washington, DC. http://hdl.handle.net/10986/17471.

Diedhiou, Abdoulaye, Kate E. Gilroy, Carie Muntifering Cox, Luke Duncan, Djimadoum Koumtingue, Sara Pacqué-Margolis, Alfredo Fort, Dykki Settle, and Rebecca Bailey. 2015. "Successful mLearning Pilot in Senegal: Delivering Family Planning Refresher Training Using Interactive Voice Response and SMS." *Global Health: Science and Practice* 3(2): 305–321. https://doi.org/10.9745/GHSP-D-14-00220.

Ghana, Ministry of Education. 2020. *COVID-19 Coordinated Education Response Plan for Ghana.* Accra: Ghana, Ministry of Education. https://ges.gov.gh/wp-content/uploads/2020/04/EDUCATION-RESPONSE-PLAN-TO-COVID-19-IN-GHANA-APRIL-2020-1.pdf.

Hasan, Amer, Koen Geven, and Ayesha Tahir. 2021. "SMS Girl Data Insights: How Has COVID-19 Affected Support for Education at Home in Punjab, Pakistan?" World Bank, Washington, DC. http://hdl.handle.net/10986/35477.

IPA (Innovations for Poverty Action). 2020. "RECOVR Ghana: Tracking the Effects of the COVID-19 Pandemic." IPA, Washington, DC. https://www.poverty-action.org/recovr-study/recovr-ghana-tracking-effects-covid-19-pandemic.

Khurana, Aishwarya, Lucy Kruske, Victoria Levin, Julia Liberman, and Diego Armando Luna-Bazaldua. 2022. "Phone-Based Formative Assessment: Guidance for Evaluating Implementation Feasibility of Phone-Based Formative Assessments in Low-Resource Settings (English)." World Bank, Washington, DC. http://documents.worldbank.org/curated/en/099804404262233170/IDU02609900b01e3c042c409da402dc0f38b0787.

Luna-Bazaldua, Diego Armando, Julia Liberman, Victoria Levin, and Aishwarya Khurana. 2022. "Psychometric Considerations for Implementing Phone-Based Learning Assessments (English)." World Bank, Washington, DC. http://documents.worldbank.org/curated/en/099850003292245698/P1742520b6dfa70e708b0d07ae5a169e654.

Martinez, Mabel, Maria Barron, and Sharon Zacharia. 2021. "How Can Countries Implement Low Tech Remote Learning? Using the Experiences from Edu Radio and Edu TV in Sierra

Leone and Pakistan to Develop Knowledge Packs." *Education for Global Development* (blog), May 18, 2021. https://blogs.worldbank.org/education/how-can-countries-implement-low-tech-remote-learning-using-experiences-edu-radio-and-edu.

Najib, Fakhira, and Pukhraj Ranjan. 2020. "Reaching Remote Students and Families in Pakistan Using Radio as a Response to COVID-19." HundrED.org, March 27, 2020. https://hundred.org/en/articles/reaching-remote-students-and-families-in-pakistan-using-radio-as-a-response-to-covid-19.

Radhakrishnan, Karthika, Shwetlena Sabarwal, Uttam Sharma, Claire Cullen, Colin Crossley, Thato Letsomo, and Noam Angrist. 2021. "Remote Learning: Evidence from Nepal during COVID-19." World Bank, Washington, DC. http://hdl.handle.net/10986/36031.

UNESCO (United Nations Educational, Scientific and Cultural Organization). n.d. "Global Monitoring of School Closures." Education: From Disruption to Recovery. UNESCO, Paris. Accessed April 21, 2022. https://en.unesco.org/covid19/educationresponse#schoolclosures.

UNESCO (United Nations Educational, Scientific and Cultural Organization). 2020. "From COVID-19 Learning Disruption to Recovery: A Snapshot of UNESCO's Work in Education in 2020." UNESCO, Paris, December 16, 2020. https://en.unesco.org/news/covid-19-learning-disruption-recovery-snapshot-unescos-work-education-2020.

UNICEF (United Nations Children's Fund). 2020a. "COVID-19: At Least a Third of the World's and Two-Thirds of Nepal's Schoolchildren Unable to Access Remote Learning During School Closures, New UNICEF Report Says." Press release, August 28, 2020. https://www.unicef.org/nepal/press-releases/covid-19-least-third-worlds-and-two-thirds-nepals-schoolchildren-unable-access.

UNICEF (United Nations Children's Fund). 2020b. "Primary and Secondary Impacts of COVID-19 on Children and Women in Ghana." UNICEF, Accra, Ghana. https://www.unicef.org/ghana/media/3461/file/Effects%20of%20COVID-19%20on%20Women%20and%20Children%20in%20Ghana.pdf.

UNICEF (United Nations Children's Fund). 2021. *Primary and Secondary Impacts of the COVID-19 Pandemic on Children in Ghana.* Accra, UNICEF Ghana. https://www.unicef.org/ghana/media/3486/file/Effects%20of%20COVID-19%20on%20Women%20and%20Children%20in%20Ghana%20(II).pdf.

USAID (United States Agency for International Development). 2020. *Language of Instruction Country Profile – Ghana.* Washington, DC, USAID. https://pdf.usaid.gov/pdf_docs/PA00X9JT.pdf.

Wolf, Sharon, Elisabetta Aurino, Noelle Suntheimer, Esinam Avornyo, Edward Tsinigo, Jasmine Jordan, Soloman Samanhyia, J. Lawrence Aber, and Jere R. Behrman. 2021. "Learning in the Time of a Pandemic and Implications for Returning to School: Effects of COVID-19 in Ghana." CPRE Working Paper, Consortium for Policy Research in Education, University of Pennsylvania Graduate School of Education, Philadelphia. https://repository.upenn.edu/cpre_workingpapers/28.

World Bank. n.d. "Remote Formative Assessment Solutions." World Bank, Washington, DC. Accessed April 21, 2022. https://www.worldbank.org/en/topic/education/brief/remote-formative-assessment-solutions.

World Bank. 2012. 2012 Information and Communications for Development: Maximizing Mobile. https://openknowledge.worldbank.org/handle/10986/11958.

World Bank. 2021. "Thinking Inside the 'Box': Pakistan Turns to Education TV During COVID-19." World Bank, Washington, DC, May 19, 2021. https://www.worldbank.org/en/news/feature/2021/05/19/pakistan-turns-to-education-tv.

World Bank. 2022a. "Supporting Student Learning at Home with Phone-Based Formative Assessments: Feasibility Template (English)." World Bank, Washington, DC. http://documents.worldbank.org/curated/en/099840004182220529/P174252095fdf90509d2800c3a3c96d1e2.

World Bank. 2022b. "Supporting Student Learning at Home with Phone-Based Formative Assessments: Landscape Review (English)." World Bank, Washington, DC. http://documents.worldbank.org/curated/en/099835004182239311/P1742520dcdd2e0009fa00d6a0176bc1c7.

SUPPORTED WITH FUNDING FROM THE
GLOBAL PARTNERSHIP FOR EDUCATION