Development and Evaluation of the Lao Child Literacy Development (CLD) Pilot

NOVEMBER 2018

Kaitlin Carson | Jamie McPike | Kevin Kamto
Adria Molotsky| Pooja Nakamura

MAKING RESEARCH **RELEVANT**

AMERICAN INSTITUTES FOR RESEARCH | AIR.ORG

Development and Evaluation of the Lao Child Literacy Development (CLD) Pilot

NOVEMBER 2018

Kaitlin Carson | Jamie McPike | Kevin Kamto
Adria Molotsky| Pooja Nakamura

1000 Thomas Jefferson Street NW
Washington, DC 20007-3835
202.403.5000

www.air.org

11/18

Contents

Operational Definitions of Key Terms

Formative assessment: The use of assessments to understand a student's reading level better and to teach according to the student's skill level.

Remedial activities: Specific teaching strategies and activities designed to help students improve in a particular area on the basis of the data from formative assessments.

Reading readiness: A range of skills and abilities that prepare a child to benefit from formal literacy instruction at an elementary school level.

Concept of print: One aspect of reading readiness—namely, a child's ability to know and recognize the ways in which print "works" for the purposes of reading, particularly with regard to books.

Oral language readiness: One aspect of reading readiness—namely, the necessary oral vocabulary and oral language skills needed to be able to learn to read in a particular language.

Decoding: The ability to see a word and make the sounds of each letter or syllable in order to sound out the word accurately.

Oral vocabulary knowledge: The ability to understand a spoken word's meaning.

Phonological awareness: The ability to identify and manipulate sounds in spoken language, including subskills, such as phonemic awareness and syllabic awareness.

Reading comprehension: The ability to read and comprehend both explicit and implicit information presented in multiple phrases or sentences of text.

Executive Summary

Background

Since the establishment of the Millennium Development Goals (MDGs) in 2000, there have been rapid increases in school enrollment around the world, but despite these advances, many countries still face a learning crisis (UNESCO, 2013). Approximately 250 million children worldwide are unable to read a single word, even after attending school for several years (World Development Report, 2018). In Laos in particular, only 8% of students in Grade 3 and about one-half the children in Grade 5 read with high levels of comprehension, and the percentage of nonreaders was significantly higher for non-Lao-speaking children.

While there are several reasons for this learning crisis, one is that millions of children are educated in languages they do not understand (UNICEF, 2016; World Development Report, 2018), presenting a major challenge to effective learning. If children are taught in language(s) they speak and understand, and their linguistic skills are appropriately harnessed, language is one of the most significant resources for effective learning (August & Shanahan, 2006; Alidou, Boly, Brock-Utne, Diallo, Heugh, & Wolff, 2006). Another reason is the lack of effective, yet simple, pedagogies and teacher-learner interactions that can used in contexts of limited resources and training. Studies have shown that programs that most effectively address these learning issues are those that focus on the "teacher-learner interaction" (Evans & Popova, 2016; World Development Report, 2018) and teaching students at their level (World Development Report, 2017).

To address these critical gaps, Catholic Relief Services (CRS) and the American Institutes for Research (AIR) developed, implemented, and evaluated the Child Literacy Development (CLD) pilot project in two districts in Laos for the 2017–18 school year. This report describes both the conceptual development of the toolkit and the results of a qualitative evaluation of the toolkit's usability, implementation fidelity, and perceived impact.

Development of the CLD Model

Developed and evaluated by AIR, and implemented by CRS, the CRS-AIR CLD package includes two formative assessment toolkits—a Reading Readiness (RR) toolkit and a Formative Assessment (FA) toolkit—and supplementary literacy aids tied directly to the results of the RR and FA toolkits that teachers can use to provide targeted remediation to bolster students' skills in particular areas. Three main features distinguish this package from other reading packages:

(a) it is designed to enhance reading readiness, (b) it includes assessments and remediation activities that are tailored for the South and Southeast Asian orthographies and the cognitive and linguistic subskills and mechanisms that are required to learn to read them, and (c) it focuses on the facilitative transfer of skills from one language to another in second-language readers.

Evaluation Design

AIR conducted an evaluation of the CLD program to better understand three main questions: (a) how teachers perceived the usability of the toolkits, (b) whether teachers implemented the toolkits with fidelity, and (c) the perceived effectiveness of the toolkits on student reading results and teacher performance. The evaluation design included a rigorous qualitative study (with qualitative data conducted through interviews, focus groups, and classroom observations) paired with descriptive results from student assessment data, teacher survey data, and coaches' observation data. Teachers implemented the program over the 2018 school year which began in September 2017 and ended in May 2018. We collected qualitative data on the three themes of interest at two stages in the program's life cycle in January/February 2018 (midline) and a second time in April/May 2018 (endline). This report focuses primarily on the findings from the endline round of qualitative data collection but also draws comparisons between these findings and those from the midline data.

Evaluation Findings

Toolkit Usability

When investigating toolkit usability, AIR specifically examined the effectiveness of the teacher training, the coaching provided through the program, and the overall ease of use of the toolkits. We found that teachers and principals found that the training greatly enhanced teachers' ability to implement the toolkits because it gave them an opportunity to experiment with more engaging teaching materials like the Lao symbol cards, and it helped them to learn new instructional practices for literacy instruction. Although teachers and principals largely found the coaching visits to be helpful for the implementation of the toolkits, several teachers remarked that they needed additional coaching specifically on the remedial activities. When asked about the overall ease of use of the toolkits, teachers agreed that the toolkit instructions were clear and the toolkits were easy to use. When asked about challenges related to toolkit usability, teachers often cited toolkit activities that were challenging for students as opposed to discussing challenges related to teachers' use of the toolkits.

Implementation Fidelity

The CLD program was intended to be implemented every Lao language subject hour. Based on AIR observations, this lasted between 1-2 hours per day. To understand one aspect of the CLD implementation fidelity better, AIR examined how closely program stakeholders followed the official program processes for communication between stakeholders, implementation of assessments and remedial activities, and coaching. Regarding communication processes, AIR found that program staff followed official government processes for working in and communicating with participating communities and that government officials indicated that effective communication greatly improved the overall program implementation. Regarding the fidelity of toolkit implementation, AIR found that teachers implemented the assessments with a high degree of fidelity and that, over the course of the program, coaches and CRS staff were able to identify and correct several errors teachers made while assessing. Data from interviews, as well as from coaches monitoring data, demonstrated that coaches had a high level of knowledge about the proper implementation of the toolkits and provided a range of feedback to teachers to support the proper implementation of program components.

However, coaches, principals, and teachers indicated that a minority of teachers consistently (on a weekly basis) implemented remedial activities, and teachers reported struggling with remedial activities due to the complexity of some of the toolkit topics, certain text-heavy sections, lack of time, and lack of ability to implement activities given large class sizes.

Perceived Effects

We found that nearly all interviewees perceived some positive effects on students' reading skills and, in particular, believed that students' motivation to learn and overall reading ability improved as a result of the program. When respondents were asked whether they perceived differences in outcomes between male and female students, qualitative interview data were mixed, but assessment scores indicated that male and female students performed similarly on the assessments over time. When asked about the perceived effects of the CLD pilot program on teachers' performance, teachers and principals agreed that the program improved teachers' teaching methods, enhanced teachers' motivation and excitement for teaching, and enabled teachers to assess and know their students' performance incrementally.

Recommendations

When asked about how the program could be improved, respondents noted that the toolkits should be tailored further to the local context by (a) incorporating additional information on

how teachers can support non-Lao-speaking children, (b) reducing the amount of text and adding more visuals to the toolkits, (c) simplifying the toolkit information where possible so that it is easier for teachers to understand, and (d) changing the physical setup of the toolkits so teachers more easily can remove and use parts of the toolkit.

On the basis of findings from the CLD evaluation, AIR also has several recommendations for improving future programming. We suggest the inclusion of more refresher trainings for teachers, simplifying the toolkit text, and providing more visuals in the toolkits. AIR also recommends providing additional training to coaches that emphasizes how to coach different types of teachers and how better to facilitate behavior change. Finally, to measure the impact of the program further and increase its long-term impact, we recommend aligning and integrating the CLD toolkits and trainings with the government curriculum and national teacher trainings and implementing a rigorous mixed-methods impact evaluation to capture whether program inputs contribute to student reading outcomes.

Conclusions

The CRS-AIR CLD model presents a unique reading package that addresses a need for an evidence-based, tailored, yet globally-relevant reading package applicable to low and middle-income countries. By tailored we mean the program is designed for the various linguistic, orthographic, and sociolinguistic characteristics in which it is being implemented. In this case, that means the programs is designed specifically for alphasyllabic scripts and is applicable to second language learners who may not have the opportunity to learn to read in their own home or community language. By globally-relevant we mean that the approach focuses on educational contexts with limited print exposure, varying scripts and languages, multilingualism – factors that are widespread globally making this model applicable across countries that are facing a learning crisis.

The CLD model centers on the notion of seamlessly linking classroom-level data and teaching methods to ensure that students are taught at the specific reading or language level they are at. Within this framework, it has three main distinguishing features: (a) it is designed to enhance reading readiness, (b) it includes assessments and remediation activities that are tailored for the South and Southeast Asian orthographies and the cognitive and linguistic subskills and mechanisms that are required to learn to read them, and (c) it focuses on the facilitative transfer of skills from one language to another in second-language readers. All these factors – limited print exposure, varying scripts and languages, and learning in multilingual contexts – are

widespread globally making this model applicable across countries that are facing a learning crisis.

Introduction

This report presents the development and evaluation of the Child Literacy Development (CLD) pilot program in Khammouane Province in the Lao People's Democratic Republic (Lao PDR). Funded by Catholic Relief Services (CRS) and developed by the American Institutes for Research (AIR), the CRS-AIR CLD pilot program is a literacy package to support Grades 1 and 2 (G1 and G2) teachers to teach Lao reading better to their Lao-speaking and non-Lao-speaking students. The program focuses on the use of (a) classroom-based formative assessments to group students at their correct level and (b) simple, targeted remedial teaching activities to support children at the level where they are. The assessments and the remedial teaching strategies are unique in three critical ways:

1. They focus on preliteracy (emergent) literacy skills.
2. They are tailored for the writing system of South and Southeast Asian orthographies, called "alphasyllabaries."
3. They focus on the second (or later acquired) language learners (i.e., non-Lao-speaking children learning to read Lao in schools).

The program was developed and evaluated by AIR and implemented by CRS.

Background and Rationale for Development

The Global (and Lao PDR) Learning Crisis

In Lao PDR (or Laos) and across the world, there have been rapid increases in school enrollment. In Laos in particular, children are eligible for 5 years of free and compulsory basic education, and this system has been in place for more than 20 years. Despite these advances, many countries around the world, including Laos, are still facing a learning crisis. Approximately 250 million children worldwide are unable to read a single word, even after many of them have been in school for up to 4 years (UNESCO, 2013; World Development Report, 2018). In Lao PDR, the national Assessment of Student Learning Outcomes (ASLO, 2012) and the Early Grade Reading Assessment (EGRA) showed that only 8% of students in Grade 3 and only about one-half the children in Grade 5 read with at least 80% comprehension. These reports also showed that the percentage of nonreaders was significantly higher in the non-Lao-speaking groups.

Although a multitude of problems contribute to the low quality of learning, large-scale meta-analyses and other studies have highlighted one factor that is present in most high-impact

learning programs: effective pedagogies that focus on the "teacher-learner interaction" (Evans & Popova, 2016; World Development Report, 2017). The precise details of the effective pedagogies may vary, but teaching students at their level (skill-based teaching also called "Teaching at the Right Level" [TaRL]) has been shown to have significant impact on learning outcomes in varied contexts (World Development Report, 2018).

A second and often overlooked issue is that millions of children are educated in languages they do not understand (UNICEF, 2016; World Development Report, 2018). Despite studies that highlight this barrier to learning, there is still limited evidence on how better to support children who may be educated in a language they do not speak or understand.

To address these two key issues, CRS and AIR developed a teacher training package focused on classroom-based formative assessments, remedial teaching strategies tailored for students at the right level, and focused instruction for second-language learners in Lao PDR.

Lao PDR

Lao PDR is a multilingual, multicultural, and multiethnic country in mainland Southeast Asia, bordering Thailand, Myanmar, China, Vietnam, and Cambodia. It has approximately 6.8 million people. Lao PDR has about 50 different ethnic groups (Simmons & Fennig, 2018) and more than 200 ethnic minority subgroups (Australian Council for Education Research, 2015), representing all four major mainland Southeast Asian ethnic groups: Tai-Kadai (~24%), Austroasiatic (~57%), Hmong-Mien (~5%), and Tibeto-Burman (~ 13%) (Simons & Fennig, 2018). About 1% of the population use sign language.

The estimated number of languages spoken in Lao PDR is 84. At present, the official language of instruction at all education levels in the country is Lao. Many remote communities do not speak Lao, and children in these areas have very limited exposure to print prior to schooling (Australian Council for Education Research, 2015). In addition, about 67% of Laotians live in rural areas (Lew, 2014). Ethnic minority groups usually settle at different elevations on hills and mountains, whereas the Lao-Tai traditionally grow paddy rice in the valleys. Non-Lao-Tai ethnic groups in remote areas with difficult agricultural conditions form a large part of the 55% of the country that is poor (Chamberlain, 2007). Children from non-Lao-Tai ethnic groups continue to have significantly lower educational achievements than do their more urban Lao-Tai ethnic group counterparts (United Nations International Children's Emergency Fund [UNICEF], 2015). Although urban Lao-Tai and non-Lao-Tai women have a similar rate of school completion, non-Lao-Tai women in rural areas have, on average, 6.6 fewer years of schooling than do Lao-Tai men (World Bank, 2018).

The World Bank identifies several challenges facing teachers in Lao PDR, including poor-quality teacher training, limited in-service training opportunities, and limited support from government pedagogical advisors (PAs) (World Bank, 2017). A mere 3% of all Lao teachers meet the passing rate of 80% on the Systems Approach for Better Education Results (SABER)[1]-Teacher Average scores (World Bank, 2017). In addition, 53% of teachers report receiving their salaries late at least once a year (World Bank, 2017). These data make it clear that Lao teachers face many obstacles that make teaching a challenge and that support for teachers is needed urgently.

The Role of Language in Designing Effective Reading Programs

There are several facets of the education context that need to be taken into consideration for designing successful literacy programs, including language policy, teacher knowledge, capacity, and motivation, student language proficiency levels, and stakeholder perceptions about language issues. Furthermore, many reading programs in low- and middle-income countries (LMICs) draw on research from monolingual, alphabetic (mostly English) models and theories of reading development (E.g. EGRA toolkit, 2016), but research has shown that there are several dimensions on which various languages and writing systems differ (Share,2008; Nag & Perfetti, 2014). These differences have significant implications for how children learn and how to design effective reading programs. Here we outline some of these to explicate the various linguistic and orthographic dimensions that are taken into consideration when designing the CLD assessments and remediation activities.

We use the Simple View of Reading (SVR) as our main theoretical framework. This is a widely accepted and empirically based reading model whose main premise is that both decoding and oral language comprehension skills are required for successful reading comprehension (Gough & Tunmer, 1986; Hoover & Gough, 1990). This framework has been validated for several languages and writing systems (Florit & Cain, 2011; Joshi, Tao, Aaron, & Quiroz, 2012) and in second language (L2) reading acquisition (Lervag & Aukrust, 2010; Proctor, Carlo, August, & Snow, 2005; Verhoeven, van Leeuwe, & Vermeer, 2011). Decoding and oral language comprehension are both made up of further subskills, such as concepts about print, phonological awareness, and orthographic knowledge. The relative importance of each of these subskills depends on the specific language that the child is learning to read, whether it is a first language (L1) or an L2, and the reading level of the child (Florit & Cain, 2011; Francis et al., 2005; Tilstra, McMaster, van den Broek, Kendeou, & Rapp, 2009; Verhoeven et al., 2011).

[1] The Systems Approach for Better Education Results (SABER) are World Bank tools used to collect, analyze, and compare data on teachers and education systems worldwide.

The Language Specifics

Three main reading components that have language-specific constraints are phonology, orthography, and semantics.

1. *Phonology:* Phonology is the way that sounds are organized in a particular language. Phonological awareness is the ability to hear, identify, and manipulate those sounds. It is one of the most significant predictors of reading achievement in all languages (McBride-Chang, 2004; Ziegler & Goswami, 2005). However, the level of phonological awareness, which is the most important predictor of reading comprehension, varies across languages. Phonological awareness levels range from *syllable* (mora; e.g., "cat" or "sprint" or "I") to onset-rime (e.g., "m-at" or "p-ool"), to *phoneme + schwa* (e.g., ಕ /cə/ in Kannada, which is mostly relevant to certain alphasyllabic languages), to *phoneme* (e.g., /c/ or /ch/). The implication for teaching is that the type of phonological awareness that is taught and measured must mirror the phonological systems of the languages.
2. *Orthography:* Orthography is the rules of how a language encodes (writes) its spoken language sounds. There are four aspects of orthography that need to be considered when developing strong reading programs in any language.
    a. **The size of the phonological unit that is encoded.** For example, in alphabetic languages such as English, Spanish, Arabic, and Hebrew, the phoneme is encoded as a single symbol (grapheme). For example, the letter "c" carries the sound /c/ or the letter "m" carries the sound /m/. In alphasyllabic languages such as Lao (and Thai, Hindi, and Amharic as well), both phonemes and syllables are encoded, but they always are represented in a syllable cluster. For example, in Lao, the symbol ຂາ represents /khaa/ denoting a cluster of the phonemes /kh/ and /aa/ blended into a syllabic cluster. In morphosyllabic languages like Chinese and Japanese kanji, only syllables are encoded, and they are always a morphological unit. For example, in Japanese, the symbol 木 represents /ki/ a syllable which has no separable graphemic components within it to represent the separate phonemes /k/ and /i/, and the whole syllabic symbol also carries meaning (tree), making it a "morpho-syllable".
    b. **The degree of orthographic depth.** Depth (transparency) is the degree to which there is a one-to-one correspondence between sounds and symbols. For instance, the letter "c" in English may be /k/ as in /cat/, or /s/ as in /city/, or /ch/ as in chicken, and therefore there is no one-to-one correspondence between all symbols and sounds in English. This is called a "deep" or "opaque" orthography. In contrast, in Spanish, "g" is always sounded out as /g/ and "a" as /a/. This is called a "shallow" or "transparent"

orthography. This has implications for the way we teach, the way we assess, and the length of time that we teach decoding versus comprehension.

c. **Symbol set size.** This refers to the number of graphemes and graphemic rules that need to be acquired. In Chinese and Japanese, students need to acquire around 3,000 and 2,100 characters, respectively, to be considered functionally literate at the end of secondary school (Taylor & Taylor, 2014). Many alphasyllabic languages, such as Lao, have more than 400 graphemic combination forms (including tonal markers) that need to be acquired. Several Roman alphabetic languages, such as Spanish, have about 29 graphemes (letters) to be learned. Therefore, learning trajectories and length of acquisition will vary across different languages, even if they have transparent or shallow scripts.

d. **Orthographic complexity.** Similar-looking graphemes (Chang, Chen, & Perfetti, 2018), nonlinearity (Kandhadai & Sproat, 2010), uppercase and lowercase, and font types all influence evidence-based teaching approaches, length of time needed for reading acquisition, and reading assessment design and implementation, so they need to be factored into reading program development.

3. *Semantics:* Semantics, or the meaning component of reading, is the foundation on which all reading occurs. Reading, in essence, is understanding text. Thus, "sounding out" or decoding or oral reading fluency are all either stepping stones or mediators of the final goal of reading: comprehension. Reading comprehension requires oral language proficiency, which in turn requires proficiency in components such as morphological complexity, vocabulary, grammar, pragmatics of language use, and background knowledge. Semantics is even more important in contexts in which children are learning to read a language they may not speak at home because the literature clearly shows that a child will not learn to read a language he or she does not understand (Ball, 2010; Kosonen, 2002; UNESCO, 1953)

Therefore, to be successful in bilingual/multilingual contexts, a reading program must promote effectively how to develop oral language skills in the language the child needs to learn to read and determine how much L1 is necessary for a child to transition to (or add) a new language effectively (Nakamura & de Hoop, 2014).

Program Description

On the basis of the theoretical framework and linguistic descriptions presented in the foregoing paragraphs, CRS and AIR developed and implemented the CLD pilot project for the 2017–18 school year. The program is a straightforward yet comprehensive literacy teaching package that provides G1 and G2 teachers with a set of classroom-based assessments and teaching tools to

enhance student reading (and prereading) outcomes. This program was implemented in 34 schools in Xebangfai and Xaibouathong districts in Laos, including schools in urban and rural areas and schools in areas from the three dominant ethnolinguistic groups in Laos—Makong, Lao, and Phu Thai.

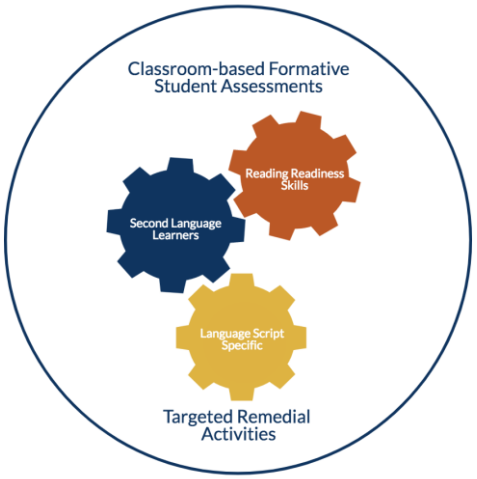
The program model has three main features that distinguish it from other reading packages: (a) It is designed to enhance reading readiness, (b) it includes language-specific assessments and teaching strategies that are tailored to the South and Southeast Asian orthographies, and (c) the program is designed to provide support for L2 learners and the facilitative transfer from one language to another for second-language readers. Each of these targeted pedagogical practices is embedded within a framework of classroom-based data intended to be used to improve teaching practices. We discuss each of these three main features below.

1 **Reading readiness development:** Reading readiness includes the development of necessary oral language skills, specifically semantic fluency – the ability to tap into a rich vocabulary network with fluency and ease; and an understanding of "concept of print" (i.e., a very basic knowledge of how print works). Lack of access to print materials (in any language) prior to the start of schooling, coupled with limited oral language skills in Lao leads to the assumption that reading readiness may be low. For this reason, the program emphasizes building reading readiness in addition to improving basic early literacy skills for students in G1 and G2.

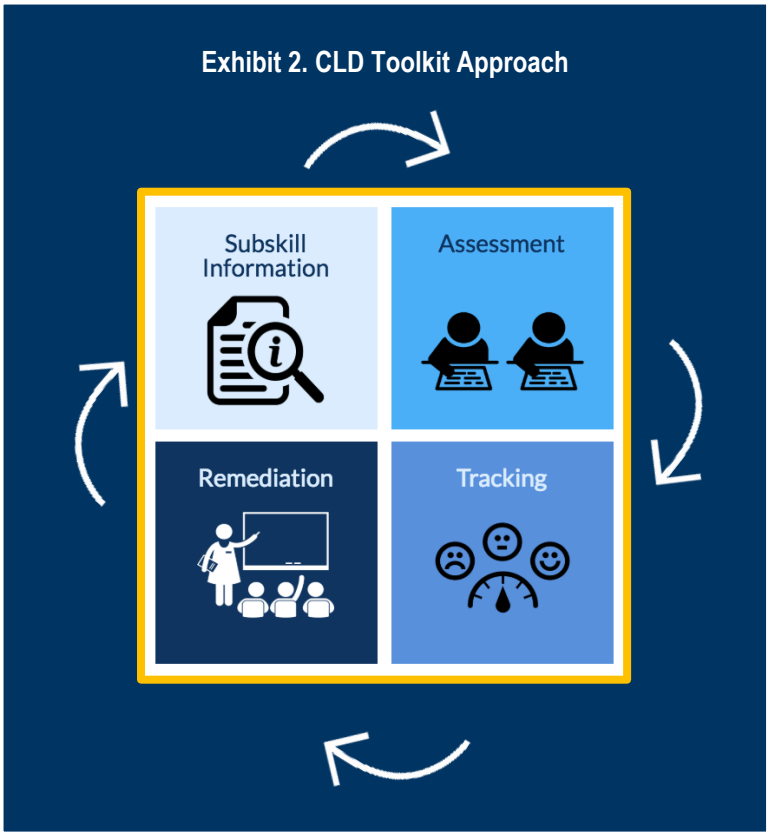
2 **Language-specific factors:** The project focuses on factors that are specific to the Lao language and writing system, as opposed to borrowing from research on literacy acquisition processes in English and other alphabetic languages. Two specific factors are incorporated into the program design: (a) phonological awareness (i.e., awareness of the sound structure of a language), which is at the dual-syllable and phoneme level in Lao and includes tonal awareness, and (b) orthographic knowledge (i.e., knowledge of the rules of the written language and how they represent sounds), which in Lao is the ability to read a primarily transparent script with a large symbol set size that is visually complex and spatially nonlinear. We also take into consideration the linguistic distance between Lao and the child's home language.

3 **Second (or later) language literacy acquisition:** The project is tailored specifically to children who may be learning to read a language that they do not speak or understand well at the start of literacy instruction. For this reason, the program dedicates more time to the development of oral language skills and taps into transferrable metalinguistic skills (see Exhibit 1).

Exhibit 1. Theoretical Framework



To provide the most effective content for a literacy acquisition package for Lao children, the CLD pilot project provides teachers with a set of pedagogical tools to integrate with their instruction, along with training on effective use of these tools. This package includes two toolkits—a Reading Readiness (RR) toolkit and a Formative Assessment (FA) toolkit—and supplementary literacy aids that teachers can use to assess students' literacy subskills and provide targeted remediation to bolster students' subskills in particular areas. Exhibit 2 provides a visual illustration of the iterative cycle between data gathering and teaching, rooted in the three main programmatic features.



Both toolkits include four iterative steps in the pedagogical process: (a) the teacher learns theoretical information about each subskill, (b) the teacher conducts formative assessment for each subskill, (c) the teacher records the students' subskills scores in the tracker, and (d) the teacher conducts the targeted remedial activities for the precise reading subskill that the child may be struggling on. Exhibit 2 shows this iterative cycle.

Training and support for teachers on the use of the toolkits included a 5-day comprehensive training at the beginning of the program and ongoing monitoring of and support for teachers from pedagogical coaches throughout the project life cycle. The program used cascading trainings to roll out the program toolkits. AIR led a training of trainers (TOT) with the District Education and Sports Bureau (DESB) pedagogical advisors and CRS staff community mobilizers (CMs), who later would act as coaches. These government and CRS staff were trained on the forthcoming cascading trainings through a facilitator's guide developed by AIR. The TOT and five cascading trainings with teachers occurred in August 2017 (see Exhibit 3 for training photos).

Through this approach, the program aimed to enhance teachers' knowledge of the subskills that are most challenging for their students to learn and to provide teachers with the tools necessary for focused remediation. We expected the program to improve reading instruction and student reading scores in the CLD pilot schools for both Lao-speaking and non-Lao-speaking children.

Exhibit 3. Photos from CLD Trainings

Training of trainers (above). Teacher training (led by coaches) in August 2017 (below).

Initially, AIR planned to conduct a mixed-methods study that included a randomized controlled trial and a rigorous qualitative evaluation. Because of unforeseen logistical challenges, the evaluation was redesigned into a rigorous qualitative evaluation paired with descriptive results from quantitative data. AIR designed and executed a qualitative evaluation—which included key informant interviews with program stakeholders, focus group discussions with students, and observations of G1 and G2 classes—to examine the overall usability of the toolkits for teachers, coaches, and principals; the implementation of the CLD toolkits (fidelity of implementation); and the perceived effects of this project on students' reading outcomes. We collected qualitative data on these themes at two stages in the program's life cycle: once in January/February 2018 (midline) and a second time in April/May 2018 (endline). This report focuses primarily on the findings from the endline round of qualitative data collection but also draws comparisons between these findings and those from the midline data. Throughout the report, we triangulate qualitative findings with descriptive quantitative analysis.

CLD Theory of Change

In this section, we present a graphical representation (Exhibit 4) of the theory of change for the CLD program on the basis of the theoretical framework and program description provided earlier. This theory of change presents the hypothesized linkages between the activities and the anticipated outputs, outcomes, and perceived impact. Because our research design did not allow us to attribute causal impact to the program, the research questions and theory of change focus on perceived impacts.

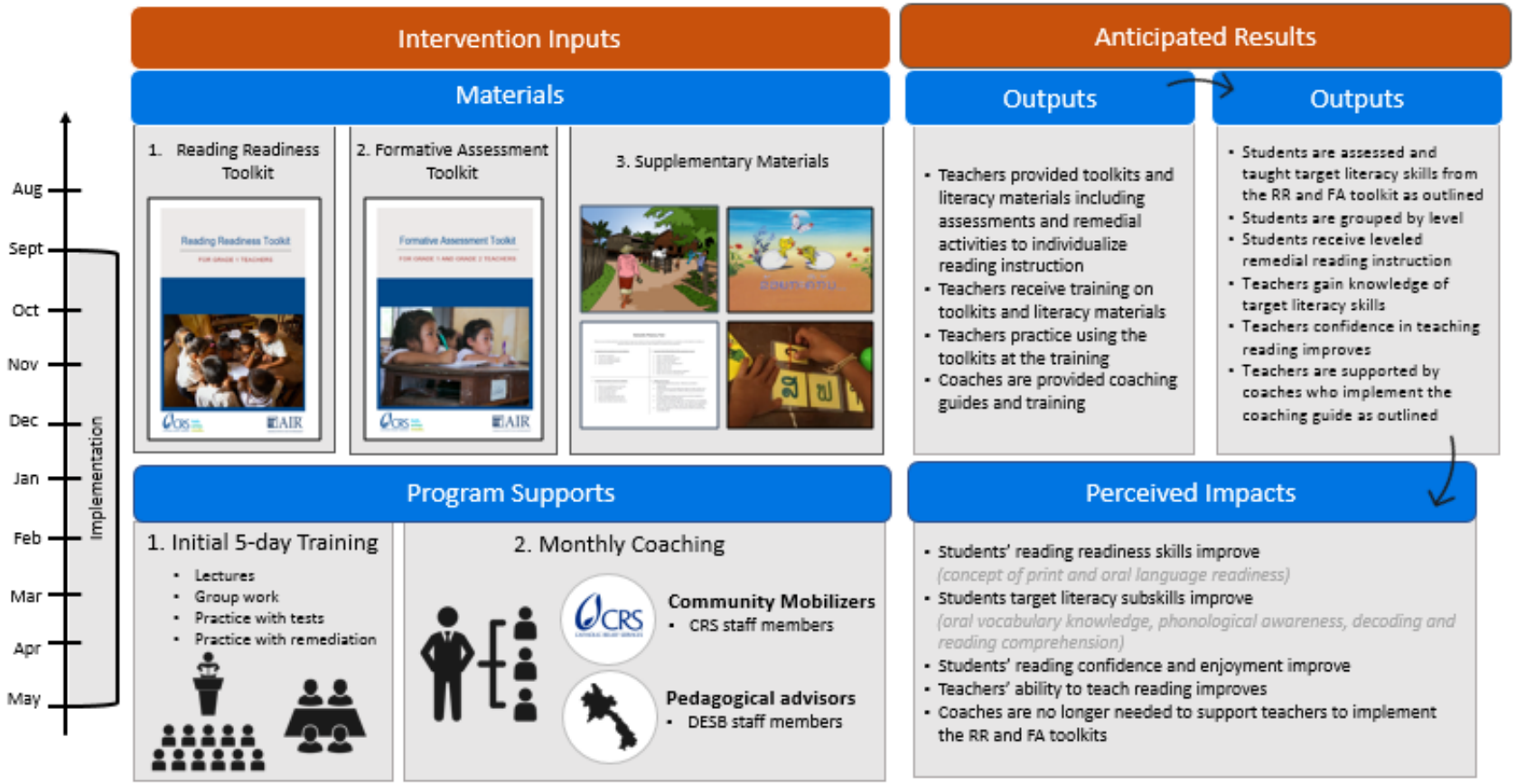


Exhibit 4. Overall Theory of Change

Assumptions

Several assumptions underlie the theory of change. First, this theory of change assumes that education exists in a politically stable environment and that school infrastructure is adequate and safe. Second, we assume that, to learn, children must have access to schools and their nutritional needs must be met. Third, teachers must be available and regularly attend school. Finally, parents and other stakeholders must understand and support children's reading acquisition.

Initial Conditions

Lao PDR is a multiethnolinguistic context with 84 languages spoken across 49 ethnic groups, and, according to the current government-mandated policy, the language of instruction at all grade levels is Lao. A report by UNESCO (2015) found that students in G1 through Grade 3 (G3) demonstrated low literacy scores, with only 23% of G3 children able to read at an independently proficient level. Ongoing obstacles to student learning in this context include poverty, lack of access to schools, low school attendance, inadequate school infrastructure and resources, poor teacher training services, and poor teacher quality (including lack of content knowledge to teach reading, especially in large, multiage/grade classrooms with children learning to read in a language they may not speak or understand).

Intervention Components

The CLD pilot is a new approach to improve the way teachers teach reading. The program uses classroom-based assessments to group students at the right level and pairs these assessment results with targeted remedial teaching activities to support children at their level. The CLD program includes the following three components: the RR toolkit, the FA toolkit, and supplementary materials. Teachers learn about these three components during an initial 5-day training and are supported throughout the program through monthly follow-up coaching visits.

Stakeholder Engagement and Approval

To implement the intervention components and translate these activities into outputs, the program needed approval from several key stakeholders, including the Ministry of Education and Sports (MoES), teacher training institutions, local school principals, teachers, CLD coaches, and village education development committees (VEDCs).

Importantly, the program was designed after CRS and AIR conducted a situation analysis of the literacy needs and demands of the Laotian context. The analysis included an examination of existing reading

programs and interventions; the evolving (at the time) BEQUAL curriculum; key informant interviews of the major stakeholders, including the Ministry (especially Research Institute for Educational Sciences),, VEDC members, teachers, and principals; as well as classroom observations. There was significant positive reaction towards the development of a reading program that uses data to support teaching in simple and clear ways, as well as the need for a program that is tailored to the Laotian language and educational context. As such, the program was developed to reflect the process of acquiring Lao literacy – especially for children who are L2 learners and learners in print-strapped environments – and thus can be integrated into not only the BEQUAL curriculum, but any curriculum that effectively aligns with alphasyllabic literacy acquisition.

Outputs

If the intervention components are implemented with fidelity, and with the involvement and engagement of key stakeholders, the theory of change expects that the program will provide teachers with increased materials and support. Classroom-based outputs include (a) providing teachers with the RR and FA toolkits and supplementary reading materials to teach key literacy subskills better, (b) offering teachers reading assessments (in the toolkits) to group students appropriately by literacy skill level, and (c) giving teachers activities (in the toolkits) to individualize students' remedial reading instruction. We speculate that teacher training outputs would include (a) teachers trained on toolkits and supplementary literacy materials, (b) teachers using the toolkits and literacy materials, and (c) coaches using coaching guides to support teachers.

Outcomes

The final link in the theory of change posits that if teachers use the toolkits, literacy materials, and training and ongoing coaching occurs in the manner anticipated, the program should lead to a set of outcomes for both students and teachers. Providing teachers with and training them on the toolkits and literacy materials should lead to teachers assessing each student with the RR toolkit and the FA toolkit. The results of the RR and FA toolkit assessments should help the teacher place each student in an appropriate, skill-based group. Teachers then would be able to translate assessment results into specific remedial activities for the different skill levels of students so that students receive individualized support for their literacy acquisition. We posited that these changes would lead to students more actively engaging in classroom activities and students receiving teaching methods and reading materials that are tailored to their linguistic and educational context. Although this is what we anticipated given our theory of change, we were not able to measure the causal impact of the program on students' reading abilities in this study.

We assumed teachers would gain knowledge of the target literacy skills outlined in the RR and FA toolkits through the CLD training. This training and additional practice applying the toolkits could lead teachers to

implement the toolkits and supplementary materials, which could result in teachers' improved confidence in their teaching skills. Finally, we posited that coaches would implement the coaching guides as outlined, thereby providing support to implementing teachers.

Impacts

We hypothesized that, if sustained, these teachers' effective use of the RR and FA toolkits would lead to a series of impacts on both teachers and students. If teachers assessed and provided targeted remediation of literacy skills as outlined in the toolkits, we expected students to improve their RR skills (i.e., concept of print and oral language readiness) and target literacy skills from the FA toolkit (i.e., through oral vocabulary knowledge [OVK], phonological awareness, decoding, and reading comprehension) over time. However, a rigorous impact evaluation is needed to test the expected impacts of these toolkits on student reading outcomes. In this study, we provided in-depth information on the use of the toolkits and measured perceived effects of this program. As a result of the improved RR and literacy skills, students' reading confidence and enjoyment also may improve. In addition, we expected teachers' motivation to teach reading would improve, given the usability of the toolkits. This motivation may improve their ability to teach reading without the additional support from coaches.

Research Questions

In this evaluation, we examined how the program was implemented and perceived by stakeholders by looking at three main themes: (a) how teachers perceived the usability of the toolkits, (b) whether teachers implemented the toolkits with fidelity, and (c) the perceived effectiveness of the toolkits on student reading results. In this study, we examined three overarching research questions:

1. How do teachers perceive the RR and FA toolkits in terms of the teacher training, the coaching provided to teachers, the overall user-friendliness of the toolkits and tracker, and the applicability of the toolkits to students' needs?

    a. How does this vary across ethnolinguistic groups and urban/rural locations?

2. Are teachers and coaches implementing the RR and FA toolkits with fidelity (i.e., as outlined in the toolkits and training), which includes applying the assessments according to the given schedule, properly conducting and recording the assessments, and properly conducting remediation activities with groups of students according to their assessment results?

    a. How does this vary across ethnolinguistic groups and urban/rural locations?

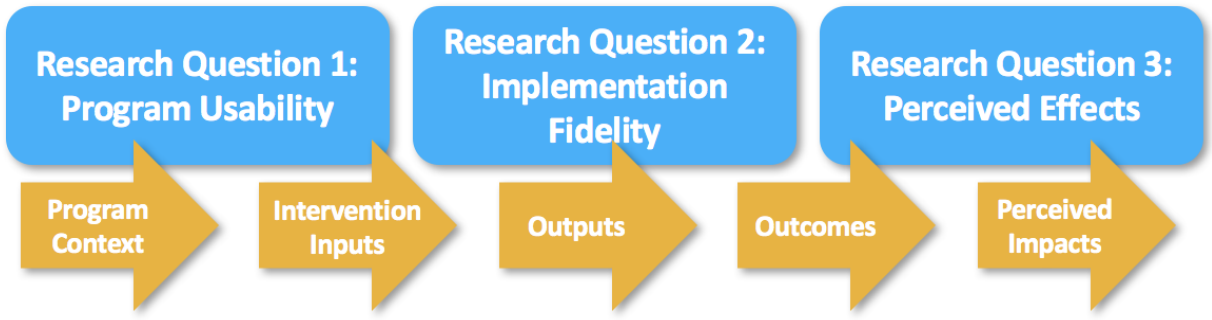
3. How do teachers perceive the effectiveness of the RR and FA toolkits in improving students' reading ability?

    a. How does this vary across ethnolinguistic groups and urban/rural locations?

We explore each research question on the basis of the hypothesized linkages outlined in the theory of change. Research Question 1 explores the way intervention components responded to initial conditions and potential moderators. Research Question 2 analyzes how the intervention components translated into outputs and outcomes. For this research question, we defined "fidelity of implementation" as the implementation of the program (including the toolkits, learning materials, and coaching) as intended and in accordance with the instructions and descriptions in the RR and FA toolkits. Finally, Research Question 3 examines the perceived effectiveness of the program on students' reading ability and teachers' ability to implement skills-based instruction by using the toolkits. For this research question, we defined effectiveness as how well the expected outcomes translated into the perceived impacts (which we defined as those impacts that the informants observed in students and teachers as a result of the program). This information is critical for a program's success because stakeholder buy-in

is central to sustaining and scaling any intervention. This research question does not, however, allow us to attribute any causal impact directly to the program because an impact study with a valid counterfactual would be necessary to make those claims. Our approach is summarized in Exhibit 5.

Exhibit 5. Mapping Research Questions to the Theory of Change



Methodology

To evaluate the CLD program, we used a mixed-methods approach. Qualitative methods were the focus of this evaluation, and quantitative methods were used to triangulate qualitative findings. In this section, we first present our qualitative methodology and then the quantitative methodology for this evaluation.

Qualitative Methodology

The study included two rounds of qualitative data collection: one at midline in January/February 2018 and one at endline in late April/May 2018. We collected data at two time points to better understand the implementation process at different stages in the program's life cycle. In this report, we focus primarily on the results from the endline data but also draw comparisons between these findings and those from midline when relevant.

Data Collection Sites

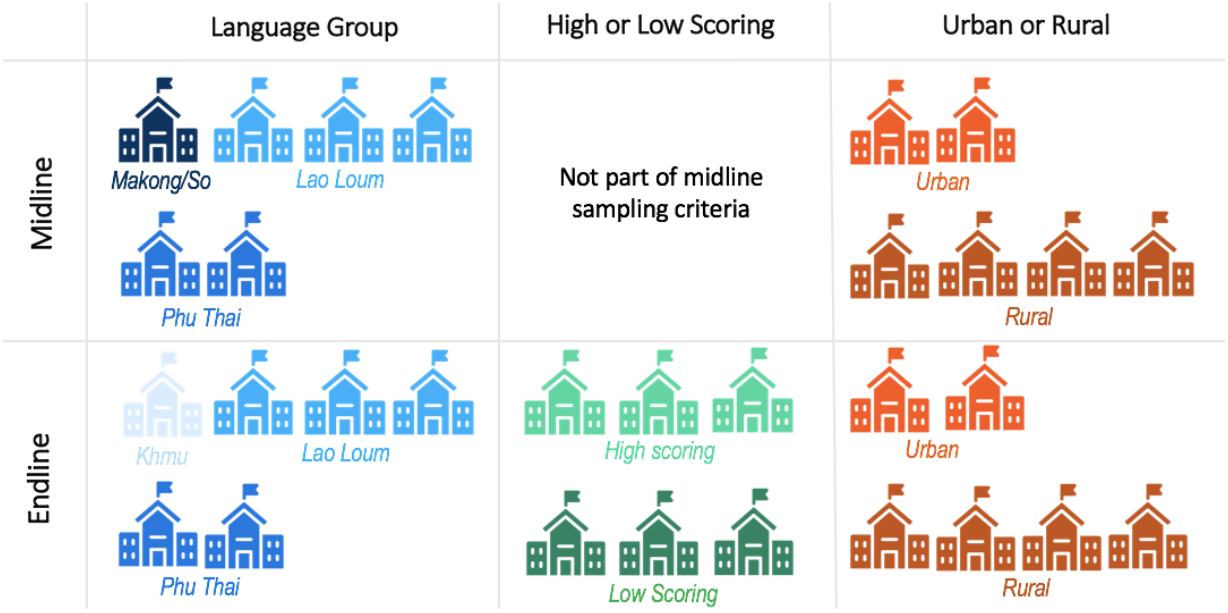
To select schools for qualitative data collection, we employed a purposive sampling strategy in which study participants were selected based on key theoretical characteristics of interest (Bernard, 2017). Given our interest in understanding differences across urban/rural areas, as well as across the three ethnolinguistic groups in the project's target areas, at midline we purposively sampled one urban and one rural school from each of the three ethnolinguistic groups included in the project—namely, Makong, Phu Thai, and Lao. We also divided our sampled schools evenly between the Xebangfai and Xaibouathong districts (see Exhibit 6) so that we could capture program processes across the two implementation districts.

Exhibit 6. Map of Implementation Areas

On the basis of the findings from midline, the research team modified our sampling for endline to investigate further the differences between high- and low-performing schools involved in the program (as opposed to focusing more

heavily on urban/rural differences).[2] To do this, we selected schools from the highest- and lowest-performing schools from each of the three ethnolinguistic groups (see Exhibit 7). Notably, although we aimed to include two Makong schools in our endline sample, the school information provided by the program misclassified a Lao-area school and a Khmu-area school as Makong schools. Because of this misclassification, our final endline sample includes three Lao schools and one Khmu school, which proved advantageous because the Khmu school provided important information about toolkit implementation with speakers of languages that are more linguistically distant from Lao(see findings below in "Fidelity of Implementation" section).

Exhibit 7. Qualitative Sampling



Qualitative Instruments

We used four types of qualitative methods (interviews, observations, focus groups, and structured notes) and sampled diverse stakeholders who played key roles in program implementation (e.g., teachers, principals, government officials, and so on). In this section, we describe each of these methods and the sampled participants for each method. We summarize the methods used for each research question in Exhibit 8.

[2] Based on the quantitative student assessment data from CLD pilot schools.

Exhibit 8. Summary of Qualitative Instruments

Research Questions	Qualitative Instruments							
	Classroom Observations	FGDs Students	KIIs Teachers	KIIs Principals	KIIs Literacy Coaches	KIIs CRS Staff	KIIs PESS and DESB Officials	Coaches' Observation Notes
Research Question 1: Program Usability	●		●	●	●			●
Research Question 2: Fidelity of Implementation	●	●	●	●	●	●	●	●
Research Question 3: Perceived Effects		●	●	●	●	●	●	●

Key Informant Interviews. A key informant is a person who possesses expert knowledge about the program or a topic related to the program. Key informant interviews (KIIs) are particularly useful for obtaining in-depth information about program processes, perceptions, and experiences (Morgan, 1996). For endline we conducted a total of 25 KIIs with diverse stakeholders involved in the CLD project, including G1 and G2 teachers at CLD schools (n= 9[3]); school principals (*n* = 6); CLD literacy coaches, including DESB pedagogical advisors and CRS community mobilizers (*n* = 4); CRS staff (*n* = 2); and officials from the Provincial Education and Sports Services (PESS; *n* = 2) and the DESB (*n* = 2). Interviews with teachers, principals, and coaches focused on the usability and implementation of the toolkits, as well as the perceived impact of the CLD program on students. KIIs with CRS staff and PESS and DESB officials focused on implementation fidelity and the perceived effects of the program.

It was common for sampled schools to have a combined G1/G2 class or only one G1 or G2 class per school. For this reason, we conducted KIIs with available G1 and G2 teachers at these schools (or the combined G1/G2 teacher at the school), as well as with the principal and the relevant literacy coach at each sampled school. We used the same interview protocol for G1 and G2 classes because teachers implemented the same FA toolkit because of low literacy scores in both grades. For the KIIs with DESB and PESS officials and CRS staff, we sampled individuals who were working directly with the project. The interview protocols for these stakeholder groups included similar lines of questioning related to program implementation and perceived effectiveness and also included more targeted questions related to the unique expertise of each stakeholder group (see Annex A for the KII protocols).

[3] Please note: While we originally intended to sample one G1 and one G2 teacher per school for a total sample of 12 teachers, the figure is lower than originally planned due to the fact that there were some mixed G1/G2 classes and several school principals also served as school teachers.

Guided Classroom Observations. During classroom observations, researchers recorded data by using a simple observational protocol that was adapted for the Lao context. Observations help triangulate data obtained through other qualitative methods by capturing practices and processes as they occur (Morgan, 1996). For this study, we conducted a total of 12 guided classroom observations at midline and 11 observations at endline (the sample was smaller at endline due to a combined G1/G2 class). Classroom observations focused largely on teachers' classroom practices, including the classroom setup, teacher-student interaction (i.e., the way students were divided into groups and how teachers conducted activities), the use and availability of teaching tools and materials, the implementation of CLD toolkits (i.e., which remedial activities the teacher implemented from the toolkits), other in-class exercises, and any difficulties the teacher experienced implementing the remedial activities. The same observation protocol was used for both G1 and G2 classroom observations (see Annex A for qualitative research protocols, including observation protocols).

Focus Group Discussions with Students. Focus group discussions (FGDs) involve guiding a diverse group of participants through a discussion on a particular topic. This qualitative method is well suited to obtaining diverse perspectives on particular issues and also offers the possibility of observing intragroup dynamics and norms during the discussion (Morgan, 1996).

FGDs with Students

Generally speaking, focus groups include anywhere from five to eight participants who are guided through various discussion topics by a trained facilitator. For this study, we conducted one FGD in each of the sampled schools, for a total of six student FGDs at midline and six student FDGs at endline (see Exhibit 9). The FGDs focused on students' general feelings about reading, their perceptions of reading instruction, and their experience with activities from the CLD toolkits. The FGD protocol included age-appropriate activities used to elicit information from student respondents effectively. To capture the widest range of responses from students, we randomly sampled students from selected classrooms, stratifying for gender, to ensure a gender balance (see Annex A for qualitative research protocols, including FGD protocols).

Notes from Monthly Coaching Visits. As part of the CLD project, trained literacy coaches provided ongoing support and guidance for teachers in implementing the toolkits. In this role, coaches conducted monthly classroom observations and debriefing sessions with teachers to go over teachers' strengths and weaknesses. Coaches completed an observation checklist and

documented (qualitatively) other notes from their observations. For the midline report, we analyzed monthly coaching observation notes (with a total of 263 data points[4]) on the implementation of student assessments and remedial activities. These data were incredibly valuable for capturing more comprehensive information on implementation processes, given that guided classroom observations conducted at midline included observations only of remedial activities.

Exhibit 10 presents the number and type of interviews, observations, and FGDs conducted for the qualitative portion of the evaluation over midline and endline.

Exhibit 10. Qualitative Sampling

Data Collection per School			Data Collection Across Schools
Observations	**FGDs**	**KIIs**	**KIIs**
• 1 G1 observation • 1 G2 observation	• 1 FGD with G2 students	• 2 KIIs with implementing G1 and G2 teachers • 1 KII with school principal	• 2 KIIs with CRS staff • 2 KIIs with PESS officials • 2 KIIs with DESB officials • 4 KIIs with CM/PA
23 observations of G1/G2 classes (for midline and endline)	12 FGDs with G2 students (for midline and endline)	52 KIIs (for midline and endline)	

Qualitative Data Collection. CRS and AIR partnered with a Lao data collection firm, Enterprise and Development Consultants (EDC), for the qualitative research. AIR first led a 5-day training for EDC data collectors in late January 2018, prior to the midline data collection, to familiarize them with the study's aims and objectives, qualitative methods, ethical procedures, and data collection questionnaires. Participants were given ample time to practice administering the interview and focus group questionnaires as well as use the observation form during the training. AIR staff accompanied the researchers into the field to oversee 2 days of data

[4] There was space at the end of the coaches' checklist where coaches could include qualitative notes about the positive and negative aspects of a teacher's lesson. One data point is defined as one observation note taken by a coach during a classroom observation. A data point could include several comments from a coach on a teacher's performance. For example, despite the fact that the following note includes several comments, we would count the following note as one data point because it was from one classroom observation: "the teacher notes all scores for each student; the teacher makes examples before the assessment; the teacher explains the method of doing the test to children very clearly; the teacher should speak loudly so children can hear clearly; the teacher should provide more activity for waiting students; the teacher should use paper cover on the decoding tool." Classroom observations took place monthly between September 2017 and January 2018. Not all teachers were observed every month, however, because of various factors. For some teachers, we have data for only two or three of the four observation months.

collection (February 5 and 6, 2018) to ensure quality and compliance with data collection procedures and provide additional support to data collection teams when needed.

Exhibit 11. Refresher training with enumerators

For endline data collection, AIR led a 3-day refresher training with EDC researchers in April 2018. Six of the eight researchers who participated in the midline training also attended the endline refresher training. The training covered many of the same topics discussed at the midline training (e.g., the study's aims and objectives, ethical procedures, and data collection questionnaires) and also included an update on findings from midline, lessons learned on the data collection process from midline, and modifications to the overall research design for the endline data collection round (see Exhibit 11).

During both rounds of qualitative data collection, researchers worked in two-person teams to conduct KIIs and student FGDs, with one researcher leading the KII or FGD and the other documenting the KII or FGD via handwritten notes. Researchers also recorded all KIIs and FGDs (after obtaining oral consent from adults and written consent from school principals for the student FGDs) on portable digital audio recorders. For the classroom observations, researchers took detailed observation notes in response to several prompts on the guided observation forms created by AIR (see Annex A). CRS field supervisors debriefed field staff daily to review what worked well during data collection that day, what did not work well, and what could be improved for the remaining days of data collection. Following the completion of field research, data collectors typed up all handwritten notes and transcribed all audio files. These data subsequently were translated into English by EDC data collection supervisors, with translations verified by CRS field supervisors.

Qualitative Analysis. For this evaluation, AIR used a collaborative coding and analysis process to compare research findings and build on emerging themes. First, the AIR research team imported all qualitative data into the qualitative data analysis software NVivo, with case classifications for each transcript. Case classifications provided researchers with important demographic and background information for each interviewee, including position or title, location (e.g., urban or rural), gender, mother tongue, and the dominant language spoken in his or her area, among other classifications. These classifications were used to examine differences in qualitative

responses between different subgroups of respondents (e.g., urban teachers versus rural teachers).

Next, our qualitative team reviewed each transcript and developed thematic summaries for each research question. Once these preliminary summaries were completed, we then discussed contradictions in our findings, emerging hypotheses, open questions, and areas of convergence in our analysis. After finalizing these themes, qualitative researchers independently coded a selection of transcripts and compared researchers' coding patterns by using an interrater reliability test within NVivo. For areas in which coding patterns strongly diverged, we came to a shared understanding of code definitions to ensure consistency with our analysis.

Quantitative Methodology

We triangulated data obtained through qualitative methods with descriptive quantitative analyses. Quantitative data from student assessment scores, teacher surveys, and coaches' classroom observations were taken from the intervention itself and thus were neither from tools developed for research purposes nor collected by trained data collectors (i.e., the data were collected by program staff, not by AIR or EDC). Furthermore, these data were not guided by any particular research questions; rather, they will be used to augment the interpretation of the qualitative evaluation, which will be submitted subsequently. The aim of this report is to respond to the request of exploring the RR and FA data to detect any kind of quantitative patterns.

Quantitative Data

As mentioned earlier, three types of quantitative data were used in these endline analyses: survey data from teachers collected in November 2017 and May 2018, student assessment scores, and classroom observation data collected during monthly coaching visits from September 2017 to May 2018. The student assessment scores were collected through the RR and FA toolkits. Although these data were not intended to serve as program evaluation data, they are used for descriptive analysis in this report to describe some student outcomes as well as teachers' implementation processes.

Exhibit 12 presents how the three main quantitative data sources were used to answer each research question.

Exhibit 12. Summary of Quantitative Instruments

Research Questions	Quantitative Instruments		
	Teacher Survey	Classroom Observations	Student Assessment Scores
Research Question 1: Program Usability	●		
Research Question 2: Fidelity of Implementation		●	
Research Question 3: Perceived Effects	●	●	●

Teacher Survey

Teacher surveys were conducted during the TOT and teacher training for the CLD pilot in Lao in November 2017 and again during the covalidation of the midline findings workshop in May 2018. The surveys, administered to 58 teachers at baseline and 40 teachers at endline, captured teachers' demographic characteristics, classroom demographic characteristics, professional development activities, and reported pedagogical and assessment practices.

Classroom Observations

Classroom observations were conducted by coaches through monthly visits from September 2017 to May 2018. During classroom observations, program coaches used the coaching guide, developed by the program, which included a monitoring checklist, space for written notes, and guidance on providing feedback to teachers. Coaches observed a small cohort of G1 teachers (n = 25) in September 2017, at the beginning of implementation of the toolkits and assessments, whereas a larger group of G1 and G2 teachers (n = 55) were observed consistently from November 2017 to May 2018. The observations focused on gathering information related to the fidelity of implementation of the program toolkits—for example, whether teachers adequately explained the toolkit activities to students, how they occupied students who were not being assessed during one-on-one assessments, and whether teachers were prepared adequately for the assessments.

Sample for Student Assessment Scores

The sample for the student assessment scores included all G1 and G2 classrooms, teachers, and students in the 34 schools participating in the program. Four rounds of assessments took place in November 2017, January 2018, March 2018, and May 2018. In all four assessment rounds, the sample included 1,098 children (568 male students and 530 female students) from Xebangfai and Xaibouathong districts. The number of children in G1 and G2 were approximately the same: 51% of children were in G1, and 49% of children were in G2. Within our sample,

three teachers and 52 students were in mixed-grade classes. Sampled children were between 5 and 15 years old, with an average age of 7. Only eight of the sample were more than 8 years old, and only one student was more than 13 years old. In terms of ethnolinguistic characteristics, the sample included 393 Lao students, 301 Makong students, and 404 Phu Thai students. The majority of children (62.6%) lived in rural areas, and 37.4% lived in urban areas.

At the school and teacher levels, the sample covered 66 classrooms and 64 teachers[5] in 34 schools. There were equal numbers of male and female teachers in the sample (32 male and 32 female). Of the schools included, 14 were in the Xaibouathong district and 20 were in the Xebangfai district. Fifteen schools were in urban areas, and 19 were in rural areas.

Quantitative Analysis

Our main analytical objective was to compare assessment scores across subgroups of students to examine possible patterns of performance within and across subgroups, concurrently and over time. To this end, we performed *t*-tests to compare two averages (means) and determine whether there were any differences in scores by gender (male and female), language dominance (Lao-language dominant and minority-language dominant), geographic area (urban and rural), and teacher pay status (paid teachers and volunteer teachers).

To observe how students' scores changed over time, we compared scores received by students from the first and last assessment rounds conducted in November 2017 and May 2018. We considered a change to be positive if a significantly larger proportion of children scored higher in May for a given subskill than they did in November (as evidenced by an associated *p* value of greater than .05).

[5] If each school had one G1 and one G2 class, we would expect there to be 68 classes and 68 teachers. However, two schools did not have a G2 class, reducing the classroom number from 68 to 66. The number of teachers reflects this reduced classroom number and the fact that two teachers taught two grades, bringing the total number of teachers to 64.

Evaluation Findings

Descriptive Analysis of the Context

In this brief section we highlight two main contextual findings about teachers' background and the linguistic context that help interpret the main findings that are presented in the following sections.

Key Findings

Teachers: In our sample, teachers had an average of 10 years of teaching experience, with 6 of those years at the G1 or G2 level. Teachers in urban areas had almost 5 more years of teaching experience than rural teachers.

Linguistic Context: According to self-reports by the respondents Phu Thai and Lao were considered most linguistically close, whereas Lao and Khmu or Makong were reported as the most linguistically distant from each other. Therefore, respondents perceived more teaching and learning difficulties for Khmu or Makong ethnic minority speakers. The main challenge that was reported by teachers, principals, coaches, and DESB officials was that ethnic minority students and teachers did not have enough proficiency in Lao language to learn and teach in Lao language classrooms.

Teachers

Within our sample, teachers, on average, had a little more than 10 years of teaching experience, with 6 of those years at the G1 or G2 level. There were significant differences between urban and rural and between volunteer and paid teachers' years of experience. Teachers in urban areas had almost 5 more years of teaching experience than did their counterparts in rural areas ($p < .05$). Most teachers in our sample completed secondary education, with some continuing to vocational education programs. On average, teachers attended professional development (PD) activities two to three times per year, for 3 to 5 days each time, with most of these PD activities focused on reading or literacy.

Classrooms

Most classrooms in the program were equipped with fixed seats all in a row or moveable seats arranged in a row, educational posters or charts on the walls, informal written materials, and displays of children's work (see Exhibit 10). We found no evidence of statistically significant differences in teachers' classroom environments between teachers of Lao-dominant and non-Lao-dominant students. We did, however, find a significant difference in whether prescribed textbooks were available, with classes of teachers of non-Lao-dominant students being 0.34 percentage points more likely to have them available. We found only a small, marginally significant difference between teachers in rural and urban areas with regard to the classroom

environment and availability of materials; teachers in rural areas were more likely to believe they had the tools to address students' learning needs.

Linguistic Context

Respondents reported that certain languages seem more linguistically related to each other than to others. This information helped us better understand how perceived effects of program usability differed across ethnolinguistic groups. Phu Thai and Lao were reportedly the most similar to each other, with informants suggesting there were only slight syntactic and morphological differences. One government official offered the following example: *"For example, in Lao the word is 'Pak ka **don**,' but in Phu Thai the word is 'Pak ka **done**.'"* Interviewees in our intervention districts stated that Phu Thai and Lao speakers can understand each other with only slight accentual differences. A CRS staff member ventured that the two languages are *"80 to 90% similar,"* with a provincial minister clarifying that there remain some areas in which the two groups cannot understand one another, stating, *"Some of their words we have to guess what they want to say and what it means."*

Interviewees reported that key differences between Lao and Makong or Khmu make the understandability of Lao more difficult for them (than for native speakers of Phu Thai). Interviewees stated that, in general, Makong and Khmu are very distinct languages from Lao and Phu Thai. Makong and Lao were reported to be "totally different." A program coach explained that *"[the] difference here is that the vocabulary is not the same as well as the names of things, accent, tones, and of course the meaning."* Lao L1 speakers also note that Makong L1 speakers have an accent in Lao that makes their pronunciation, at times, difficult to understand. A CRS staff member stated:

> *Makong language is different from Lao, and Lao could not understand if Makong people speak in their language. When Makong people speak Lao, they also have a different accent from Lao language, and sometimes we could not understand what they mean due to the accent.*

There were both instances where students did not speak and understand Lao, the language of the classroom; as well as instances where teachers did not speak and understand Lao enough to teach it comfortably. The student language issue was typical in areas where the language of the community was either Phu Thai or Khmu. DESB officials, teachers, and principals reported that in non-Lao areas students often enter the classroom with limited to no previous exposure to Lao and very low levels in any Lao literacy subskill. A principal explained the situation in a Khmu area as follows:

> *At first, when students who never attend school before came to school, they could not speak any Lao. For example, eating rice they speak their local language, and the teachers do not understand.*

Teachers stated that in these scenarios, they often relied on translating Lao information into the language of the community so that students could understand. One principal stated, *"I have to speak in local language to explain lessons."* It often takes students who are non-Lao speakers or in non-Lao areas time to adjust to Lao as the language of instruction. One principal in a Khmu area described how it takes children time to feel comfortable speaking in Lao because they often are *"not brave to speak out or communicate with people in Lao."*

For the teacher language issue, respondents stated that teachers who speak Lao as an L2 commit errors in pronunciation and spelling while teaching in Lao. Our data showed that 35% of the teachers were not Lao-language speakers themselves. CLD program coaches and district officials noted that many Lao L2 speakers struggled to match native-level pronunciation. Interviewees specifically stated that non-Lao teachers had moments during which they made *"mistakes writing Lao words,"* had *"incorrect pronunciation,"* or had *"problems with their tones and accent."* However, a minority of respondents stated that all teachers can speak and teach in Lao. One CM reported, *"All teachers are speaking in Lao; even though teachers are different ethnicities, they all speak Lao."*

Respondents stated that situations of language mismatches in schools led teachers to use the L1 in the classroom. Despite the strict national language of instruction policy, a surprising number of teachers, principals, coaches, and district officials explained that teachers used both the mother tongue and Lao in the classroom. Government officials, coaches, principals, and teachers stated that mother tongues were used mainly to help students understand what was happening in the classroom. One coach explained in the following way how this typically occurs:

> *Teachers have to teach/talk with students in two languages. For example, Makong (Bo) or Phu Thai speakers will teach or explain the lesson in their language and translate into Lao language because students in Grade 1 and 2 do not speak Lao clearly and don't know some words in Lao language that are needed to be explained in their languages.*

Many of the teachers in the sample from non-Lao areas reported also using mother tongues specifically to help students transition to a Lao-only language of instruction.

Clearly, both these issues are sources of serious communication and comprehension issues, which, unsurprisingly, are likely to have a significant impact on any kind of learning, including the implementation of the program toolkits.

Overall Program Usability

Key Findings

Research Question 1: How do teachers perceive the RR and FA toolkits in terms of the teacher training, the coaching provided to teachers, the overall user-friendliness of the toolkits and tracker, and the applicability of the toolkits to students' needs? How does this vary across ethnolinguistic groups and urban/rural locations?

Teacher Training: Teachers and principals found that the training greatly enhanced teachers' ability to implement the toolkits for several reasons: It provided a chance for teachers to learn the toolkits' instructions, it gave them an opportunity to experiment with more engaging teaching materials like the Lao symbol cards, and it helped them to learn new instructional practices for literacy instruction. When asked how the training could have been improved, respondents recommended more frequent and ongoing training of teachers, more training on particular components of the toolkits, additional instruction on how to create engaging teaching materials, and inclusion of parents and VEDCs in the training so that they are aware of the program and the possible benefits the program brings to students.

Coaching: Although teachers and principals largely found the coaching visits to be helpful for the implementation of the toolkits, several teachers remarked that they needed additional coaching specifically on the remedial activities (this was particularly true for teachers from rural, low-performing schools). This may be because, despite following the remedial activities for several months, some teachers may not see the progress that they expect for their students and may not know exactly how to address this challenge by using the existing materials.

User-Friendliness of the Toolkits: When asked about the overall ease of use of the toolkits, teachers agreed that the toolkit instructions were clear and the toolkits were easy to use. When asked about challenges related to toolkit usability, teachers often cited toolkit activities that were challenging for students. For example, several teachers responded to the question *"How easy or difficult is it to use the Reading Readiness and Formative Assessment toolkit assessments?"* by noting skills that students commonly struggled with (e.g., decoding) as opposed to discussing challenges related to teachers' use of the toolkits.

Teacher Training

Respondents were asked—at both midline and endline—how well the training prepared teachers to use the RR and FA toolkits, and responses during both rounds of data collection were largely positive. Findings from midline showed that respondents found the training primarily beneficial because it was a PD opportunity that allowed them to engage with other teachers and to learn new teaching methods, something that many teachers did not receive on

a regular basis.[6] Although this theme was mentioned less frequently at endline, there were a few notable overlaps between the findings from the two rounds of data collection. Specifically, teachers and principals found that the training greatly enhanced teachers' ability to implement the toolkits because it provided a chance for them to practice the toolkit instructions, it gave them the chance to experiment with using more engaging teaching materials, and it helped them learn about new instructional practices.

Practicing the Toolkit Instructions

When asked whether the training adequately prepared teachers to implement the toolkits, teachers and principals described how the training helped them learn the steps necessary to complete the assessments and to implement certain remedial activities effectively (e.g., the Lao symbol cards.) For example, one teacher mentioned how *"yes, it [the training] prepared me because now I can read and follow the instructions in the toolkit and then put [them] into action."* Another teacher echoed this sentiment: *"The training was useful because we had the opportunity to do actual practices on the toolkits, and that gives a real lesson learnt."* The step-by-step approach to learning the toolkit instructions was noted as very helpful for a third teacher, who said, *"The previous training was already good. In the training, the trainer advised every step of instruction in the toolkits, and now I put in[to] practice."*

Using New and Engaging Classroom Materials

For many teachers, the ability to learn about and practice using new supplementary classroom materials was a major added benefit of the training and often their favorite part of the training. According to one teacher:

> *Yes, [the training] prepared me because, previously, we mainly followed the textbook. Using the word cards and applying pictures in teaching was very limited. Now after the training, I know more about using word cards in teaching and can help to explain the lessons to students.*

Another teacher provided a similar comment when she said:

> *I feel that my teaching is more comfortable because the project has provided many kind of teaching materials such as pictures, word cards, and other things that we can use—*

[6] Findings from the midline data collection round, for example, indicated that 62% of teachers participating in the CLD pilot program said that they did not take part in any PD activities in the past year.

> *these materials—to facilitate teaching. It is very good for us to have those teaching materials.*

This ability to use different and more engaging materials is seen as a major added value of the program and is discussed in greater detail in the "Fidelity of Implementation" and "Perceived Effects" sections that follow.

Learning Effective Strategies to Teach Reading

Teachers also noted that the training provided them with the knowledge necessary to teach effectively the skills required for reading. Specifically, it helped teachers learn how to integrate the new toolkit activities into their existing pedagogical practices. One teacher remarked:

> *It [the training] helps to prepare to use the toolkits … all the toolkits, I can apply in my teaching plan. For example, the toolkit has the alphabet[7] together with picture. … I can apply this by showing pictures to students before we go into learning how to connect the alphabet with the vowel.*

This training was very helpful for teachers because it enabled them to teach reading skills more effectively and enhance their own pedagogical content knowledge, something one teacher said she felt ill prepared to do prior to the start of the program:

> *[During the training] I learned a practical reading, writing, speaking and how to pronounce and delete words. All these are very clear to me now. Before [when] I taught, it was up to my understanding, but now this project helped make it clearer to me how to teach.*

Teachers and principals recommended that the training could have been improved in several ways, including more ongoing training of teachers because of the lack of refresher trainings (see the "Fidelity of Implementation" section later in this report), more training on particular components of the toolkits, additional instruction on how to produce teaching materials, and training for parents and VEDCs (see Exhibit 15). Of these themes, the two most common recommendations were additional teacher training and training on how to produce classroom materials. These two themes were quite common across respondents, and one respondent suggested that this could be solved through more on-the-job training by CMs and PAs. Specifically:

> *Many teachers do not understand clearly about how to use the variety of toolkits, so it would be better if CM/PA prepare further training or conduct on-the-job training to*

[7] It is likely that this reference is to Graphemes

> *teachers and spend about 2–3 hours to train them. In addition, the PA should demonstrate on how to produce the teaching materials and lead teachers to participate in producing teaching materials together.*

Despite the fact that teachers received ongoing coaching support, there still appeared to be an eagerness for more frequent teacher training. This could be due to a desire for a more formal, group training environment that emphasizes on-the-spot learning in a risk-free environment, as opposed to the one-on-one coaching during which teachers typically receive feedback only after an observation period (see "Coaching" section that follows). Data from teacher interviews at midline and endline indicate that some teachers were uncomfortable with the coaching visits, possibly because they may have felt like these visits were an assessment of their skills as instructors. For example, one teacher described how she felt *"a bit of pressure"* when the coaches visited her class. Another teacher noted how the coaching visits were like a *"guest coming to visit our house; we have to welcome them even though we dislike [them] sometime, but we have to accept."* Although these findings were not consistent across teachers in our study sample, and most respondents were appreciative of the coaching visits, it may be worth investigating the amount of time coaches spend coaching as opposed to observing teachers during a particular classroom visit and whether the observation process feels, to teachers, like an assessment of their performance. If this is the case, teachers may feel a reluctance to admit when they do not understand particular toolkit content, which would have implications for their ability to capitalize fully on the knowledge of the coaches and implement the toolkits with full fidelity (we discuss the coaches' perceptions of their role and the skills required to build trust with teachers in the "Fidelity of Implementation" section below).

Furthermore, teachers' desire for additional training on producing teaching materials likely is due to two main factors. First, because many teachers already have seen some benefits from using more engaging materials provided in the toolkits (see the "Perceived Effects" section), they may be eager to provide students continually with new and updated materials to keep them engaged and motivated. Second, several teachers and principals described their desire to continue using the toolkits even after the program ends, yet some teachers noted that they did not know how to replicate the materials if they were lost or damaged (or if there were not enough materials for all the students in their classes). According to one teacher, *"The problem is the wording card is not enough; sometimes vowel card or alphabet card are not enough for students because there are many words. If I create it by myself, I am not capable of doing it."* For this reason, training on how to produce materials may help improve the overall sustainability of the initiative, especially given respondents' eagerness to continue to use them.

Exhibit 15. Respondents' Recommendations for Improving the Teacher Training

Respondents' Recommendation 1: More Ongoing Training	*"Yes, **I want more training** in order to learn new things and learn more how to use the toolkit clearer than previously."* *"From my opinion, I think **all teachers, including the Directors, need to have a refresher training**. The training has to ensure that all teachers have broadened [their] knowledge. … For following up, it could be continuously; we could [have a] follow up [training] once in a month. It is possible. Follow-up should take more frequently and continuously."* *"If it is possible, this project should have more training for teachers. Only one training is not enough; the teachers forget the instruction and do not implement the activity correctly. **There should be three to four trainings**."*
Respondents' Recommendation 2: Training on Specific Components of the Toolkits	*"I want this project to provide **more advice on assessment** and [my] general teaching where it is not good and where it is already good. Then I can improve."* *"If talking about training, I would like the project to train on how to use the alphabet and vowel. I don't understand yet. … **The third skill in FA about connecting the alphabet with vowel … I really don't know to use this skill**."* *"**From the toolkit, the skill number three**—the reading of vowel with two alphabets together. The students do not understand, and it is also difficult for the teacher to read and [to know] how to teach [the] students to understand."*
Respondents' Recommendation 3: Training on How to Produce Engaging Teaching Materials	*"I want [training] on the method of **how to produce the teaching materials** in order to have variety of teaching tools and develop for using in the future."* *"There will be more effective use of the toolkits if the program supports **training for or activities of how to produce the innovative equipment**. [This is] so the teachers will be able to produce, by themselves, [materials] such as innovative equipment that are made from wood, stone, and pictures."*
Respondents' Recommendation 4: Training for Parents and VEDCs	*"We have to find a topic and **do more training for school and village administrative level** to see that what are the benefits to the school and students. … The previous training is not enough yet. When working with people, it takes time."* *"In my opinion, there should **be more training for VEDC and school administrative level** because they need more training. If they do not get training, they will not understand and realize the importance of this work."*

Coaching

Monthly coaching sessions aimed to improve teachers' capacity to implement the toolkits effectively. As they did in the findings from midline, teachers noted that coaching visits ranged anywhere from 30 minutes to 2 hours, and visits occurred either during the Lao language hour or during a general study hour. Data from KIIs with coaches, principals, and teachers, as well as notes from coaches' classroom observations, indicated that the majority of feedback provided to teachers focused on correcting behavior when a teacher was not implementing the toolkit

with fidelity (e.g., not providing enough instructions to students or not providing activities for the remaining students) and providing advice on how to deal with implementation challenges. According to one teacher:

> *They [the coaches] gives advice/coaching about how to deal with a problem. For example, if my teaching is not good enough or an activity is not properly implemented, then they will give advice after the class. During the observation, they do not give advice/coaching.*

Another teacher echoed this idea:

> *They [the coach] observes Lao language class and after finished, they would come to talk with the teacher and then give advices where things should be improved or an added activity.*

Although teachers and principals largely found the coaching visits to be helpful for implementing the toolkits,[8] several teachers remarked that they needed additional coaching specifically on the remedial activities. One teacher mentioned how she needed *"coaching about the method on how to use the toolkits—especially the remedial activities—for students who already conducted remedial activities, but they still have problems with their reading and writing in Lao language."* Another teacher provided a similar remark: *"I need more support to solve problem for the students that can't read. What remedial activity should I organize? Particularly for decoding, students cannot do or read. What remedial activity should I organize for this?"* A third teacher mentioned:

> *I want this project to observe when I am doing remedial activities for students, because this project might have a new methodology or new way to support students to learn better, and I am not sure what I am doing for students is the correct way.*

Despite the fact that teachers interviewed at midline largely reported that the remedial activities were straightforward and easy to implement, the findings from endline indicated that some teachers wanted more assistance implementing these activities. The teachers who mentioned a need for additional coaching on remedial activities were from diverse ethnolinguistic groups but primarily from low-performing schools in rural areas. This change in

[8] The question *"Has the Community Mobilizer/Pedagogical Advisor provided you enough support to implement the program toolkits as recommended in the program guidelines?"* elicited largely positive feedback. Responses included the following: *"Yes, I am very satisfied with what they are doing here at this school"*; *"Yes, they give quite enough support. I do agree with them, and I am happy that they come to help"*; *"We enjoy working with CM/PA because we can get the knowledge and we can ask them when we don't understand";* and *"Yes, I like because it is good that someone come to do coaching—coaching about something that I am not good at yet; then I improve myself. If no one does the coaching, I would not be able to improve myself."*

teacher perceptions over time could be because teachers now have more data on their students' progress. For some teachers who may not see progress for their students, this presents an important challenge: what to do if students are still struggling, despite the teacher's regular implementation of the remedial activities in the toolkits. This may be why several teachers expressed a desire for additional coaching on teaching strategies and remedial activities to help struggling students in particular.

User-Friendliness of the Toolkits

Overall Ease of Use of the Toolkits

Similar to findings at midline, teachers largely agreed that the toolkit instructions were clear, that the toolkits were easy to use, and that ease of use increased over time. Specifically, the proportion of teachers who were seen by coaches to use the toolkit with ease (according to coaches' observation data) increased from 81% in November to 94% in May. One teacher responded to the question *"How easy or difficult is it to use the Reading Readiness and Formative Assessment toolkit assessments?"* by stating, *"It is not difficult; we already pass the training, and of course we can do it."* Another teacher noted, *"It is not so difficult; I am able to implement it."* Teachers from high-performing schools and from areas where Lao is the predominant language most often said that the toolkit instructions were clear and the toolkit activities were easy to implement.

Regarding the most useful activities from the toolkits, teachers commonly cited the Lao symbol cards and the images from the OVK assessment—specifically, the *"alphabet together with the picture. This is in the toolkit given by this project. I can apply it showing pictures to students before we go to learn connecting alphabet with vowel."* Another teacher mentioned:

> *The activity that is useful for students is pictures. For example, when playing game using wording card, the teacher would draw a picture for them and ask them to connect the vowel with alphabet to a sentence or a word such as a frog eats the moon. I would distribute vowel and alphabet to them and then ask them to do at the blackboard. Students who could not do can learn from the one who can do this activity.*

These data are consistent with findings discussed in greater detail in the "Fidelity of Implementation" section, later in the report, that highlight how teachers most often used the more tactile remedial activities (e.g., Lao symbol cards and the OVK assessment pictures) in their classes because they found them to be more engaging for their students.

Challenges With Implementing the Toolkits

When asked what aspects of the toolkit, if any, were challenging to implement, teachers' responses indicated that there may be a conflation (on the part of teachers) between the overall toolkit usability (e.g., instructions, readability) and the ability of students to complete the assessments and toolkit activities (i.e., student skills). For example, when asked about challenges related to toolkit usability, teachers often would cite toolkit activities that were challenging for students to complete. Several teachers responded to the question, *"How easy or difficult is it to use the Reading Readiness and Formative Assessment toolkit assessments?"* by noting skills that students commonly struggled with (e.g., decoding) as opposed to discussing the challenges with teachers' use of the toolkits.

Teachers' responses to the question on usability challenges, therefore, do not necessarily indicate a lack of toolkit usability but likely reflect teachers' perceptions of the challenges that students faced with the toolkits. This hypothesis is bolstered not only by the responses given previously regarding toolkit usability (i.e., teachers noting that the toolkits are easy to use) but also by data that show that teachers from schools with low student assessment scores tended to cite difficulties with toolkit usability more often than did teachers from higher-performing schools. This distinction between toolkit usability for teachers and students also was mentioned by teachers during midline and illustrates some of the challenges that teachers face not necessarily with administering the activities but with ensuring that students progress in their learning. Although teachers may perceive that the toolkits are user-friendly for their own needs, if students struggle with completing the activities or progressing in their learning, teachers may interpret students' learning progress as connected to the overall toolkit usability.

Teachers noted that students commonly struggled with the script-processing subskills like decoding (i.e., subskills that require the mapping of sounds and symbols) and reading comprehension. Some teachers noted challenges that they faced teaching decoding, in particular, given the language differences between the students and the teacher. According to one teacher, *"The decoding [is difficult] because of different languages that make the students and teacher are not understand to each other … so the decoding is more difficult for students."* Another teacher provided a similar remark:

> *The reading is difficult, as the students always face problems because they are ethnic. For example, the tones of sound, they would pronounce every word in the same tone. Now I make a material for them to learn how to pronounce things in the different tones.*

> *They know that the sound should be different in tones, but they could not pronounce properly.*

When discussing the overall applicability of the assessments to students' skill levels, another teacher remarked that the reading comprehension test was too difficult for students, given that students are still unable to read at G2: *"The assessment with the picture test is appropriate. As I said, the fourth assessment is not appropriate. It is too difficult for this level of students. They cannot read well yet at Grade 2."* However, because the assessments are designed to assess students' skills in particular areas, they therefore will be challenging to varying degrees for different students. For this reason, it is challenging to interpret findings on which assessments were perceived by teachers as easy or difficult for students.

Fidelity of Implementation

Key Findings

Research Question 2: Are teachers and coaches implementing the RR and FA toolkits with fidelity (i.e., as outlined in the toolkits and training), which includes applying the assessments according to the given schedule, properly conducting and recording the assessments, and properly conducting remediation activities with groups of students according to their assessment results? How does this vary across ethnolinguistic groups and urban/rural locations?

Communication: Interviews with coaches, government officials, and principals showed that program staff followed official government processes for working in and communicating with communities that were part of the program, including obtaining approvals from provincial officials and district-level officials for program implementation and implementing the coaching program along with district pedagogical advisors. Government officials were pleased with the project's communication and indicated that effective communication improved the overall program.

Implementing Assessments: Teachers' correct descriptions of their implementation of the toolkits, along with students' and observers' confirmations of teachers' practices, showed that teachers could implement the assessments with a high degree of fidelity at the end of the program. Over the course of the program, coaches and CRS staff were able to identify and correct several errors teachers made while assessing. Compared with midline data, endline data from interviews with coaches and teachers indicated an increase in teachers' ability to provide literacy activities to occupy the remaining students during one-on-one assessments.

Remedial Activities: Coaches, principals, and teachers indicated that a minority of teachers consistently (on a weekly basis) implemented remedial activities. The most common remedial strategies teachers used included (a) mixing high- and low-performing students so that high-performing students could support lower-performing students and (b) using the program's Lao symbol cards. Teachers reported struggling with remedial activities because of the complexity of the toolkits (large amount of text and level of complexity), lack of time, and lack of ability to implement activities given large class sizes. Coaches and principals supported these claims, stating that the toolkits are too complicated for many teachers.

Coaching: Data from interviews and from coaches' monitoring data demonstrated that coaches had a high level of knowledge on the proper implementation of the toolkits. Coaches also provided a range of feedback to teachers to support the proper implementation of program components. Both CMs and PAs reported that it was challenging to shift teachers' existing practices around teaching reading.

Communication

Research Question 2 examined whether teachers and coaches were implementing the program as outlined and, if they were not, why not. We investigated this question by examining how well the program's communication, assessments, remedial activities, and coaching, as well as the integration of the CLD program into the existing curriculum, matched with the program's anticipated outputs. This was particularly important for implementation in the Laotian context in which there are several stakeholders – including the Ministry of Education and Sports – that needs to be involved and sign off on each stage of implementation.

The project staff maintained regular coordination and communication (consistent with government regulations) with all actors involved in the implementation process throughout the period of performance. Within Lao PDR, communication processes must follow the national structures for accessing and working with local communities. Findings from midline and endline showed that program staff followed official government processes for working in and communicating with communities that were part of the CLD program, including obtaining approvals from provincial officials and district-level officials for implementing the overall program and for implementing the coaching program along with district Pedagogical Advisors.

Government officials were pleased with the project's communication, stating that effective communication improved the overall program. At midline, a PESS (provincial-level) official cited the *"good cooperation between project, provincial, district, school, and VEDC"* levels as the main reason the program was able to *"make the work run as planned and achieve the set goals."* At endline, district-level officials echoed how good coordination improved the overall success of the program. The DESB Director stated:

> *The facilitation, coordination, solidarity, and participation of all stakeholders [helped] for implementing the program activities. I believed that it increased success. For example, if we go to work in the school but nobody coordinated [this work], the work could not go smoothly. And, in contrast, if the partners have good coordination, we can do that work easily.*

Notably, informants from PESS, DESB, and CRS, as well as coaches, unanimously agreed that the program's processes for communication and coordination of coaching visits, PA participation in the program, and programmatic updates enabled the program to operate smoothly and meet the expectations of key local stakeholders.

Implementing Assessments

High Degree of Fidelity When Implementing Assessments

After the CLD pilot program, the majority of program teachers could implement and score the toolkit assessments as intended by the program guidelines. In qualitative interviews, most teachers described implementing the assessments on schedule, following the directions for the assessments, and adhering to the scoring guidelines as outlined in the RR and FA toolkits. As they did at midline, teachers interviewed at endline correctly described the basic implementation steps for several of the assessments included in the toolkits. Notes from a guided classroom observation at a CLD pilot school (conducted by an EDC field researcher) confirmed one teacher's correct implementation of the assessments:

> *Then teacher calls the student's name for the student to go out to the teacher's desk to do assessment. The teacher shows the pictures in the Formative Assessment toolkit to the student and says the name of a picture. The teacher then asks the student to point out the picture that she says the name [of].*

In this observation, the teacher is implementing the OVK test with her students correctly (see corresponding prescribed directions in Exhibit 16). Interviews with G2 students provided further evidence of teachers' proper implementation of the assessments. One student stated that during testing experiences, *"the teacher asks me to come out and sit then point out things. If anyone can read a lot, then she would ask to read more like Ka Pi delete Ka out, only Pi remains. If I delete Pi, then Ka remains."* In terms of scoring the assessments, many teachers described both the correct scoring procedures for the assessments and the correct stop rules.[9] Because teachers' descriptions of how they used the toolkits had a high degree of similarity to the toolkit guidelines, and these descriptions were supported by students' and observers' confirmations of teachers' practices, we found that teachers could implement the assessments with a high degree of fidelity at the end of the program.

Improved Assessment Implementation Over Time

With the support of coaches, teachers were able to reduce the number of errors committed while assessing students over the course of the program. In coaching notes from September and October, coaches reported that teachers consistently committed the following errors while assessing students:

- Lack of materials preparation (strips of paper for the concept of print test, timer)
- Not recording students' scores correctly on the score sheet
- Not timing the semantic fluency test
- Not providing remaining students with activities
- Not explaining instructions to students clearly
- Not providing the student enough time to answer test questions
- Not supporting students to be relaxed during test administration

[9] Stop rules are included in the toolkits where teachers should stop assessing students if they do not pass a particular test.

Midline qualitative interviews suggested, and endline interviews and coaching data confirmed, that many of these errors were corrected in teachers' assessment practices by the end of the project. Errors frequently reported in coaching notes during the first coaching visit with all teachers in November were reduced drastically in coaching notes from frequent descriptions of implementation errors in January to frequent descriptions of proper implementation in May. At endline, coaches reported that most teachers could explain the assessment instructions successfully, implement the assessments confidently, and help students remain relaxed during testing; they also provided students more time to respond to assessment questions. In the final set of classroom observations in May, teachers, on average, exhibited a high degree of fidelity, with 100% of teachers recording assessment results in the student tracker. However, throughout program implementation, coaches noted that teachers struggled to occupy remaining students with activities.

Teachers Spent a Range of Time on Assessments

To implement the assessments with fidelity, teachers needed to spend time (a) familiarizing themselves with the assessments prior to implementation, (b) locating the book required for the concept of print test, and (c) preparing a timer before beginning the assessments with students. Several teachers confirmed that they were able to prepare sufficiently to implement the assessments. One teacher stated before conducting the assessments:

> *Prepare—[we] prepare the toolkit and pictures that would be used to test students. For example, this page, what it is about? I read its instructions and [the information on] what the teacher should do and what [the] students should do and the score record form.*

Qualitative data showed that the time teachers spent assessing students varied significantly. Observations conducted by the field research team indicated that teachers spent anywhere from 1 minute to 2 minutes, 5 minutes, and 6 minutes per student conducting the OVK assessment. Self-reported data from teachers also showed that some teachers spent on average 3 to 5 minutes per student, whereas others spent anywhere from 5 to 10 minutes per assessment.

A minority of teachers did not spend enough time preparing for assessments and, as a result, demonstrated lower knowledge on the process required to prepare for assessments. One coach highlighted this problem, stating, *"Sometimes the teachers do not follow the steps of assessment. Some teachers have not prepared well for the tools of assessment such as A4 paper*

to hide the picture because in a page, since there are four pictures." Thus, there was still an identified need at endline for teachers to prepare for the assessments.

Less-prepared teachers and teachers in non-Lao areas typically took more time to conduct the assessments. One teacher explained that this tends to happen because students who do not speak Lao as their L1 need more time to identify the correct answers for the test items because they need to process the information in their L1 and then respond in Lao, their L2. This teacher explained, *"For one word, the students think twice; they think in Khmu and then translate into Lao. They know the name of the object, but they do not know name in Lao."* Another reason for the variation in assessment times could be overall teacher preparedness. Low familiarity with the assessments could cause the teachers to spend additional time finding the instructions, locating materials, and efficiently finishing the scoring. However, more research would be needed to find the exact cause of the variation in test duration.

Occupying Remaining Students

Compared with the findings from midline, data from interviews conducted at endline indicated an increase in teachers' ability to provide literacy activities to occupy the remaining students during one-on-one assessments. At midline, the main activities used by teachers to engage students during assessment included drawing, reading, writing, and looking at pictures, and only two interviewees mentioned occupying students who were not being assessed with literacy activities. At endline, however, teachers reported using more literacy-relevant activities to occupy students, such as reading, using the symbol cards, and writing, which suggests that teachers could be improving their classroom management skills (specifically, their capacity to have students work independently on relevant group-based or individual tasks). The most consistently used activities the teachers reported using were the following:

- Reading (*n* = 7)
  - Including reading together (*n* = 1), storybooks (*n* = 3), and reading aloud (*n* = 1)
- Practicing with word cards (*n* = 3)
- Drawing pictures (*n* = 3)
- Looking at pictures (*n* = 2)
  - Including a teacher *(n* = 1) asking students to look at pictures in their textbooks
- Writing (*n* = 2)

Managing Assessments in Multigrade Classrooms

Our qualitative sample included two multigrade classrooms. Qualitative data suggest that multigrade teachers faced additional challenges managing assessments. One multigrade teacher discussed how complicated the assessment process could be given the size of the class:

> *I assigned exercises for them, for example drawing pictures. My classroom has multiple grades in one class—Grades 1 and 2. I would put one of the grades in the other room and ask another teacher to look after them for me. If I don't put students in another room, the students would be very noisy. When I want to do the assessments for Grade 1, then I keep only Grade 1 in the classroom. When I am assessing one student, the rest of the students would do the exercise. But Grade 2 students would be in another room because the assessments take a long time to finish.*

In this case, the multigrade teacher must manage the student being assessed plus the remaining G1 and G2 students in the classroom not being assessed. Having another teacher who was available and willing to oversee the G2 class in a separate room was critical to effective implementation of the assessment. This situation undoubtedly puts strain on teachers.

Students' Knowledge of Their Assessment Scores

If teachers followed the program processes, they would not share the assessment scores with students. When asked whether students knew their assessment scores, however, data from student focus groups and interviews with key informants were mixed. On the one hand, student focus group data showed that nearly one-half of the study sample participants were told how they scored on the assessments, which may have contributed to their self-identification as either a "good" or a "bad" reader. On the other hand, data from interviews with several coaches, CRS staff, and one teacher indicated that teachers did not share assessment scores with students. This contradiction either could be due to social desirability bias (i.e., coaches and teachers were unwilling to share that they did not comply with program guidelines during interviews) or simply indicated that some teachers shared scores with students but other teachers did not.

Students from three schools stated that they were informed of their assessment scores by their teachers. One PA and a CRS staff member confirmed that some teachers either told their students their assessment results or informed the entire class of the average assessment results. One PA explained this in greater detail:

> *For the students' score of assessment, the teacher will inform the students in general, such as 'the assessment, in this month, the students are getting better.' The teacher will not*

> *inform them who is weak or got a low score because they are afraid that students will feel bad and disappointed.*

Data from student FGDs indicated that many students knew their scores. Approximately one-half of the sample of students knew their assessment scores, though most students knew only if they scored a frowny face, neutral face, or smiley face, as opposed to their numerical scores.

Teachers' deviation from the program guidelines could be attributed to a gap in the toolkit training. One PA stated, *"The majority [of teachers] might not know [whether to tell students their scores] because in the training, the trainer did not focus on whether to tell the score to students or not."* The possible negative effects of students finding out if they received a low score underline a need to clarify and communicate this point to teachers and coaches.

In contrast to data indicating that students were informed of their scores, many respondents noted that teachers did not communicate assessment scores to their students. Most coaches in the sample (two CMs and two PAs), a CRS staff member, and a teacher reported that teachers did not tell students their assessment scores. A PA stated that teachers did not tell students their scores specifically to avoid conflict: *"Teachers will not tell the score to students to prevent yelling among students and keep it as confidential to avoid comparisons among students."* Others noted that teachers did not tell students their scores out of compliance with the program processes.

Although the data are mixed on this topic, for some students, hearing their scores could have contributed to whether they associated themselves positively with being a "good" reader or negatively with being a "bad" reader. An excerpt from an FGD with students shows how students felt when they learned of their assessment scores:

> ***Interviewer:*** *When you know you got a 7 score, how do you feel about it?*
>
> ***Respondent 1:*** *I feel happy.*
>
> ***Interviewer:*** *Why does a 7 score make you feel happy?*
>
> ***Respondent 2****: If got a 10 score, [I am] super happy.*
>
> ***Interviewer:*** *If [you] got a 2 score, what do you feel?*
>
> ***All Respondents:*** *Not happy.*
>
> ***Interviewer:*** *Why does a 2 score not make you feel happy?*
>
> ***All Respondents:*** *Because it is a low score.*

These data show that the sharing of scores, in some cases, strongly impacts student self-perceptions as readers. It is difficult, however, to say definitively whether this impact is always positive or negative. For example, findings from midline indicated that some teachers shared scores with students to spur competition among students, and these teachers perceived this to be positive because they perceived it to motivate students further. Because this theme was not investigated in great detail in this study, we cannot determine the effects of sharing scores on students' self-perceptions. Given the mixed responses from participants, we also cannot say how many teachers told students their scores and the level of detail that teachers provided to students when discussing scores (e.g., saying that a student received a smiley face as opposed to a specific number).

Remedial Activities

If teachers followed the toolkit instructions with fidelity, after they implemented the assessments, they should have implemented the appropriate remedial activities for the different student skill levels for each of the six literacy subskills. The toolkits included low, medium, and hard remedial activities for each of the score thresholds students could receive—smiley face, neutral face, or unhappy face. The toolkits included assessments and remedial activities, such as singing songs and reciting rhymes, building base symbol-diacritic combinations, reading aloud every day, and teaching students the connection between sounds and symbols, so that teachers could understand and then respond to students' literacy subskill levels. The program coaches were tasked with helping teachers understand students' scores by showing teachers how to connect unique scores to particular remedial activities and working with teachers to plan and prepare for remedial activities.

Weak Fidelity of Implementation for the Remedial Activities

At midline and endline, about one-half of the interviewed teachers (*n* = 9) implemented remedial activities on a semiregular basis (i.e., weekly or monthly) as opposed to daily. Similarly, coaching data revealed that only one-half of observed teachers provided learning activities for all students or put students into groups and assigned activities accordingly. Interviews indicated that the remaining one-half of teachers implemented few to no remedial activities and mainly focused on implementing the toolkit assessments. One coach described the different levels of implementation as follows: *"For those who implement well, they use the project's activity continuously, who are medium, they use sometimes only and for those who are very low, they hardly use the activity of this project."*

We find some differences in how Lao-dominant and non-Lao dominant teachers implemented remedial activities in their classes. In particular, Lao-dominant teachers **were marginally more likely** than non-Lao dominant teachers to be conducting remedial activities from the toolkit, modeling these remedial activities for students, and being patient with students learning at a slower pace. Non-Lao-dominant teachers were observed to be marginally more likely than Lao dominant teachers to follow toolkit guidance for how to set up activities for students, including providing remedial activities for all students and conducting group activities.

We also found that Phu Thai teachers within the sample appeared to have a lower understanding of remedial activities than did Lao- and Khmu-speaking teachers. Three Phu Thai speakers in the sample discussed remedial activities either as "materials" or as grouping high and low scorers together. In contrast, other non-Phu Thai teachers were able to list more examples of remedial activities and offer more complex descriptions of these activities that better aligned with the program activities. However, because of the small sample size, we cannot be certain whether Phu Thai teachers had a lower adoption of remedial activities (see Exhibit 17).

We did not see patterns of how teachers implemented the remedial activities differently across key teacher characteristics, including years of tenure, the urbanicity of the school, and the size of the class.

Exhibit 17. Use of Remedial Activities, Comparisons by Language Subgroups

	Lao Teachers		Non-Lao Teachers		Balance Test		
Coaches' Observation Notes Related to Use of Remedial Activities	**Mean**	***N***	**Mean**	***N***	**Difference**	**Standard Error**	***p* Value**
Teacher conducts remedial activities from the toolkit.	0.89	37	0.74	88	-0.15	0.08	.06
Teacher puts students into groups and assigns activities accordingly.	0.38	37	0.55	88	0.17	0.10	.09
Teacher first models remedial activity for students.	0.92	36	0.76	88	-0.16	0.08	.05
Teacher provides clear instructions about each activity.	0.97	37	0.74	92	-0.23	0.07	.00
Teacher is patient with students learning at a slower pace.	1.00	111	0.97	268	-0.03	0.02	.09

For teachers who struggled to implement remedial activities on a regular basis, there were several key factors that limited implementation, including lack of time to implement the activities effectively, the complexity of the toolkits, and difficulties integrating remedial activities with the formal curriculum.

Time Constraints

Teachers frequently noted that they did not have enough time to implement the remedial activities as they are designed. Most teachers implemented the remedial activities during the Lao language class, which typically lasted for 1 to 2 hours during the school day. Teachers were uncertain whether their time should be used to conduct the remedial activities or whether it was spent better by following the activities in the government curriculum and teaching to the end-of-year tests (see the "Difficulty Integrating Remedial Activities With Ministry Curriculum" section later). Because students could receive one of three possible scores across six possible literacy subskills, teachers reported not having the time to plan or implement remedial activities because of the high number of scores and associated activities. One literacy coach explained how most teachers implemented one remedial activity with the whole class because of the time required to plan and organize separate activities:

> *For example, working in groups, sometimes students are not divided into groups, but the whole class does it [the remedial activity] at the same time. If teachers spend time to divide into groups, then they have to arrange tables and it takes time. So, sometimes not dividing into groups, but with the whole class at the same time. Those remedial activities that teachers hardly use are the ones that the teachers need to divide into groups. They prefer to do for whole class.*

Managing time was particularly challenging for teachers with combined G1 and G2 classrooms or combined G1, G2, and G3 classrooms because these teachers dealt with significantly larger class sizes and classroom management demands.

Complexity of the Toolkits

Interviews with coaches indicated that teachers' limited experience with the toolkit topics was a significant challenge to the uptake of remedial activities. One principal stated that because teachers struggled to read, understand, and follow the toolkit guidance, they did not fully integrate remedial activities into their work:

> *As observed, [teachers] can integrate the remedial activity into their responsibility in the classroom about 70–80% out of 100% because the understanding of teachers is low. They slowly read and follow instruction in the toolkits. They could not remember all of*

> *the instructions, but they will gradually improve. Mostly teachers could not remember how to do remedial activities.*

Another principal supported this claim, requesting additional training on the toolkits because of teachers' "limited understanding":

> *The limitation is that our teachers still do not understand well how to use the toolkits because teachers themselves have a limited understanding, [a] short vision. This project did a training for them, but of course it does not cover everything. This is still a challenge for teachers. If I put in easy word is that teachers are confused with the toolkits—when should use and what is the best appropriate for students. But given materials are very good.*

The perceived complexity of the toolkits and the reported lack of detail in the remedial activities were barriers to teachers' adoption of remedial activities.

Remedial Strategies That Teachers Used

Program teachers had a different understanding of remedial activities than what was intended by the program. CRS and AIR developed the program toolkits as pedagogical guides for teaching reading and providing differentiated support to students. Midline and endline data showed that rather than adopting a different overall approach to teaching (skills based, assessment based, differentiated by level), teachers typically mentioned only using the materials-based remedial activities and changing few of their teaching practices. When teachers mentioned remedial activities, they often discussed how they used the materials-based activities to supplement the learning materials from the official curriculum. One teacher explained this process in greater detail:

> *The project activity and curriculum of the Ministry of Education and Sports is very similar, because the literacy development focuses on the wording card [that has the] picture card, vowel, and alphabet. When I show the card to students, some students may not know. They don't know what the alphabet is, but then I show them the pictures and they would realize what the alphabet is because there is a picture of the alphabet.*

In this example, the teacher used the word cards and picture cards provided by the program as material examples to help students understand the lesson. However, the program's goal was for teachers to use supplementary materials to provide targeted, skill-based instruction for students on the basis of their assessment results. Teachers typically only implemented the toolkit remedial activities that were tactile or visual compared with other activities that may not have been as visual but had a greater focus, improving a specific literacy skill (e.g., highlighting rhyming patterns for children and conducting active storytelling to improve oral

vocabulary knowledge). Teacher reliance on material-based remedial activities also was reflected in coaching notes. Coaches frequently advocated for teachers to use more materials and reported teachers consistently using the Lao symbol cards provided by the program.

The main adjustment that teachers made to their teaching practice (that was not part of what the toolkit explicitly called for) was to focus more attention on the low scorers by mixing high and low scorers in working groups or seating arrangements. In the initial 5-day training with CLD teachers, coaches instructed teachers that after assessing, teachers should give students activities based on their level—sad face, neutral face, or happy face. Teachers' actual practices diverged from the intended program model because the majority of teachers provided the entire class with the same activity, but this activity was tailored to the level of the lowest-scoring students.

Focusing on Low-Scoring Students

The program intended for teachers to provide differentiated support to students by grouping students into low-, medium-, and high-scoring groups and providing specific remedial activities (from the toolkits) for each group. However, qualitative data showed that teachers focused their remedial activities primarily on low-scoring students. Teachers described supporting low-scoring students by pairing these students with high-scoring students for class activities. Seven of the nine teachers interviewed at endline reported mixing high- and low-scoring students to help low-performing students. One coach explained this process:

> *After implementing the assessment, the teacher will know what percentage of students can do well or cannot. Such as, there are 20 students in a class, after the assessment they [the teacher] will know how many students are excellent, medium, and weak in study. Then the teacher will organize the group of students by mixing the excellent and the weak [students] by sitting [them] in a group to let students help each other.*

Teachers, coaches, and principals remarked that mixing students was beneficial to the lower-performing students because higher-performing students could provide additional support to low-scoring students, helping the lower-scoring students improve.

Teachers also described allotting more class time to support low scorers. Although students received high, medium, and low scores, almost all teachers (all but two) reported tailoring their remedial activities to the low-scoring students. Midline and endline findings indicated that some teachers implemented the same, "easy-leveled" remedial activity with all students, regardless of assessment scores. Data from interviews with two coaches (one PA and one CM) confirmed this, and an excerpt from an interview with a coach shows how many teachers did not differentiate activities among students:

> ***Interviewer:*** *Overall, are the teachers able to connect the students' assessment scores (such as sad face, neutral face, and happy face) to a specific skill?*
>
> ***Respondent:*** *Teachers will do the general teaching with all students in the class, even though they know that who is weak in oral vocabulary, but teacher ask the question with all students, no specific asking the student who is weak.*

Coaching notes also showed that several teachers would let low-scoring students participate (typically by answering questions) more often than higher-scoring students. Although a small number of teachers (*n* = 2) said they provided certain activities differentiated for low- and high-scoring students (not neutral face), coaching data, classroom observations, and KII data all showed that teachers struggled to provide differentiated support based on the assessment results.

Integrating the Materials With the Basic Education Quality and Access in Lao PDR Curriculum

Teachers participating in the program were trained on how to implement the CLD toolkits and how to use the toolkits alongside the official government curriculum. The remedial activities are differentiated **pedagogical activities** intended to be easily integrated into the official curriculum and tailored to the curriculum content area. For example, if the official curriculum required students to learn farming vocabulary, the teacher could use specific remedial activities—from the OVK section of the FA toolkit—on the basis of students' scores from the OVK assessment to teach this content in a more tailored manner to students' different skill levels. This integrated approach was created so that teachers could use materials provided by the program to bolster activities and lessons from the government curriculum to better support students' development of particular literacy subskills. If teachers followed program guidelines with fidelity, they would be equipped to use the toolkits to assess students effectively and to design leveled (i.e., "easy," "medium," and "hard") remedial activities on the basis of students' scores.

Data indicated that program participants clearly understood (at both midline and endline) that the official Ministry of Education and Sports curriculum should be the main teaching content in classes, and data from all KIIs and observations showed that the Ministry of Education and Sports' curriculum was still teachers' main focus during the Lao language hour.

Only a minority of teachers integrated the toolkits into the official curriculum as the program intended, however. Both midline and endline data showed that, in practice, teachers had a different understanding of how remedial activities could be paired with the official curriculum. Coaches' monitoring data and teacher interviews show that instead of integrating specific remedial teaching activities from the program into their teaching of the official curriculum, teachers more

often used parts of the toolkits to engage students in lessons from the official curriculum. For example, some teachers reported using pictures from the OVK assessment for instructional activities (outside of assessments) as a way to engage students because the government textbooks lacked interesting visuals that they believe were needed to retain students' interest and attention with the content. When asked which remedial activities they conduct in the classroom, teachers reported using materials-based activities most often. This is echoed in qualitative coaching notes which indicate that many teachers just adopted the program materials or used the program materials to support activities from the curriculum. One principal explained this process:

> *Since the CRS project [provides] the program materials [toolkits], it helps the students understand the lesson better because it has pictures that make students enjoy. Before, teachers didn't have the material with pictures like this, so students were not interested to learn, but now we have pictures and other materials to facilitate their teaching.*

Several teachers also discussed using the OVK images, the concept of print books, and supplementary materials to provide examples for students and help engage them with colorful, interesting objects. For example, one teacher explained:

> *I use it together with the curriculum because the curriculum has only the words, the sentences, the content … but the project, it has word cards that you see, it consists of reading as a word and phrase. We take that content and the lesson to put together.*

Most teachers reported using the toolkits alongside the official curriculum, as this quotation indicates, but the integration of the toolkits focused more on using toolkit materials (e.g., pictures, Lao symbol cards) to engage students in learning even if the content from the lesson did not necessarily relate to the remedial activity used (see Exhibit 18). This differs from the desired implementation in the sense that teachers did not select remedial activities based on students assessment results to integrate into lessons from the government curriculum, they more often used visuals from the remedial activities to engage students in learning overall. These differences in desired compared to actual implementation could be traced back to a number of causes including the education level of teachers, the complexity of the toolkits (noted above), or time constraints.

Exhibit 18. Children completing CLD toolkit remedial activities in pilot classrooms



Coaching

Coaches closely adhered to the program's protocol to provide monthly observations and feedback to participating teachers. Coaches visited teachers on a monthly basis for approximately 30 minutes to 2 hours. After observations, coaches provided feedback to teachers for 15 minutes to 1 hour, depending on the level of support the coach determined that the teacher required. One CM explained the different levels of support in the following way:

> *For those who do well when I visit them, I almost have nothing to advise them. I can give them like 90% of good implementation. I would like to emphasize that for those who do not implement or do not do well yet, I as a CM already give then advice and do demonstrations for them. For those who can do well, I do not advise them much; I only suggest an extra activity that might be appropriate for them. For schools that do not perform well, I demonstrate and propose some activities for them that then they can practice.*

One coach described the length of the coaching sessions as follows: *"It is dependent on how much the teacher could understand. Sometimes it takes more than 30 minutes, but in general it takes about 20 minutes."* This implementation process closely followed the program's intended coaching process. Data from interviews as well as coaches' monitoring data[10] demonstrated that coaches had a high level of knowledge on proper implementation of the toolkits and provided a range of feedback to teachers to support the program's fidelity of implementation.[11]

[10] From midline only.

[11] Coaches were intended to observe the Lao language hour without interrupting the teacher. Coaches then should have found a quiet space to debrief teachers on the basis of the observation checklist, using the "world's simplest feedback model." The feedback included the following: (a) What did you do well? (b) Here's what you did well. (c) What would you do differently? (d) Here's what you can do differently.

Data from KIIs with coaches and teachers showed that coaches' recommendations to teachers focused on improving the program's implementation. This included discussing how teachers could integrate program activities more effectively into their teaching practices (including how to assess students, advocating for teachers to use more remedial activities, discussing arranging students in groups, and familiarizing teachers with the supplementary materials). For example, when asked to describe the last coaching visit, one program coach described helping a program teacher identify students for remediation activities:

> *In the last visit, we talked about the remedial activities for the children who did not pass the assessment test or could not get our standard score. Then we are coaching the teachers on how to conduct the remedial activities in order to develop the child literacy as mentioned in the coaching guide.*

Coaches found it difficult to implement the coaching program because of the number of schools and teachers they were assigned to monitor. Specifically, three coaches mentioned how traveling between schools was tiring and affected their ability to support teachers. One coach explained this challenge:

> ***Coach:*** *For classroom observation, of course it impacts my work because it is too much responsibility for 14 schools, and the level of understanding of teacher is not same and my own understanding as well and sometimes I am tired from traveling. When I do class observation and give support for teacher, I may not do very well. In overall, the impact is only long traveling and that I am tired only. But when doing classroom observation, I don't have problem with that because I follow the appointment schedule and our plan.*
>
> ***Interviewer:*** *You mentioned that you do not do classroom observation fully. What do you mean by that?*
>
> ***Coach:*** *As I am tired from traveling, so I could not perform the duty very well.*

Anther coach noted that it was particularly difficult to travel to rural areas, indicating that more remote schools may face additional challenges with the long-term sustainability of the coaching model. Coaches also cited the additional challenge of providing individualized coaching support to multiple teachers with very different levels of teaching skill. Because site visits were conducted monthly, a significant number of coaches reported that the frequency and number of coaching visits to a large number of teachers was an obstacle to providing effective support to teachers.

Both CMs and PAs reported that they often were unable to shift teachers' existing practices around teaching reading. Coaches frequently reported that teachers listened to their advice during coaching sessions but often failed to follow their advice (as observed by the coaches during later visits). One coach stated, *"If [teachers] are not open-minded, they would not learn*

anything. But I as a coach do my best to transfer knowledge. Sometimes they listen but do not put in action. That is challenging." Another coach echoed this sentiment, stating that certain teachers have *"their own way of teaching"* and often *"do some activities of the project but perform very low."*

Data from CMs and CRS staff highlighted that younger coaches also struggled to provide advice to more senior teachers during coaching sessions. A CRS staff member summarized the issue in this way:

> *The challenge for CMs of this project is they are still too young. We have three CMs and two are young people. In our Lao culture, they are still acceptable, but [it's] about effectiveness of coaching or giving advice, because young people [do not] advise older people. … Older people may think a CM is as young as his son or daughter [and think] "how dare he or she do coaching for me.*

The lack of change on the part of teachers possibly could be the result of teachers' low capacity to apply the recommendations effectively or be due to cultural issues and norms (between coaches and teachers) affecting teachers' willingness to try the new approach.

Coaches' implementation of the program closely aligned with the program guidelines. Coaches provided routine support to teachers that was critical to teachers' proper implementation of the program. Examples of coaches providing support to teachers included correcting minor mistakes in teachers' implementation of the assessments, encouraging teachers to use more remedial activities, and practicing specific remedial activities with teachers. The amount of time coaches spent with teachers during the debrief sessions (20–30 minutes) was shorter than anticipated, but, overall, teachers and coaches seemed satisfied with the coaching process.

Perceived Effects

Key Findings

Research Question 3: How do teachers perceive the effectiveness of the RR and FA toolkits in improving students' reading ability? How does this vary across ethnolinguistic groups and urban/rural locations**?**

Perceived Effects on Students: We found that nearly all interviewees (teachers, principals, government officials, and coaches) perceived some positive effects on students' reading skills and overall reading enjoyment as a result of the program. In particular, respondents perceived two main effects on students as a result of the program: (a) Students showed an increased motivation to learn and greater engagement in the classroom, and (2) students' overall reading ability also was perceived to improve as a result of the program. Although the first theme—on enhanced motivation and engagement—echoes responses provided at midline, findings at endline showed that respondents perceived greater effects on students' overall reading ability. When respondents were asked whether they perceived differences in outcomes between male and female students, qualitative interview data and student assessment data were mixed. Assessment scores, however, indicated that male and female students performed similarly on the assessments over time.

Perceived Effects on Teachers: When asked about the perceived effects of the CLD pilot program on teachers' performance, teachers and principals agreed that the program had several positive effects on teachers. Specifically, the program improved their teaching methods and their ability to use new teaching materials creatively, enhanced teachers' motivation and excitement for teaching, and enabled teachers to assess and know their students' performance incrementally. Some respondents noted that the program increased the workload of teachers but that this additional work was worthwhile given the benefits it brought to students.

Perceived Effects on Students

When asked whether the RR and FA toolkits resulted in changes to students' learning and literacy skills, nearly all interviewees perceived some positive effects on students' literacy skills and overall reading enjoyment as a result of the program. Specifically, respondents perceived two main effects as a result of the program: (a) Students had an increased motivation to learn and greater engagement in the classroom, and (b) students' overall reading ability improved. Although the first theme—enhanced motivation and engagement—echoes responses provided from research participants at midline, findings at endline showed that respondents more commonly cited perceived effects on students' overall reading skills. When respondents were asked whether they perceived differences in outcomes between male and female students, data from qualitative interviews and student assessments were mixed. Several interviewees perceived differences in male and female respondents, but assessment data indicated that male and female students performed similarly on assessments.

Results from the assessment data did not show any significant differences between boys and girls on any of the subskills. This is similarly true for perceived differences across ethnolinguistic groups. Although we found that Lao-dominant students on average scored higher in every subskill, we did not find any truly significant difference between their scores and those of non-Lao speaking students.

Enhanced Student Engagement and Motivation to Learn

When we examined the qualitative data on the perceived effects of the program on students, interviewees frequently mentioned students' increased motivation to learn and engagement (as a result of the program materials and content) in the classroom as major benefits of the program. Several teachers and principals noted how the opportunity to use instructional materials that differed from those of the official government curriculum enhanced student interest and classroom engagement. For example, one principal described how *"students pay more attention when teachers have different tools in teaching. … What is given by this project is very useful."* Another teacher noted how *"students are more motivated and the given teaching materials are useable."* A different teacher offered a similar comment, noting how it was much easier to retain students' interest since she started using the program toolkits: *"The main challenges that we faced [before] was [students were not paying] attention. After the program support materials, it is easier."* Respondents typically referenced the visual nature of the toolkits as key to retaining students' interest and enhancing their motivation to learn. For example, one principal described how the visual elements of the toolkits served to motivate students further to learn:

> *The assessment and remedial activities help to motivate students, because previously none of students would draw pictures of poultry by themselves. They see the pictures and they start drawing by themselves; that is one of motivations. When books have pictures, it is very attractive rather than just blank paper.*

Another teacher noted:

> *When I show them the pictures, it is easy because I have pictures in my hands. For students Grade 1, if we have a teaching media, it is easy to teach and they can remember easily. For example, when I teach them about alphabet and vowel and I have the teaching media to show them, they can remember easy.*

The diversity in teaching and learning tools proved important for keeping students' attention and engagement during instruction.

Perceived Improvement in Students' Reading and Writing

Most respondents perceived improvements in students' reading subskills (and some noted a perceived improvement in writing skills as well) as a result of the CLD pilot program. Although there was widespread perception of student improvement, quantitative data from the toolkits showed that there were no significant differences in scores between the first and last assessment sessions (other than a marginally significant increase of $p < .10$ in phonological awareness scores).[12] It is difficult to rectify these contradictory findings given that teachers and principals may be perceiving smaller-scale improvements in student performance that are not necessarily adequately reflected in the assessment scores. Perceptions of student improvements also may be related to other improvements in student performance outside of assessment scores such as students' self-direction and contributions in the classroom.

When describing these benefits, teachers and principals either cited direct improvements in students' assessment scores or observed improvements in reading and writing abilities. According to one teacher, *"Before, students really can't read and write; now it is much better, because now some students are very good."* Another respondent provided a similar remark when discussing the program's benefits:

> *It is good that this project come to this school to support students' reading skills, because before this project came, it was very difficult for the students. They did not even know alphabet and I have to hold their hands to learn writing. When this project comes and provides the books, the students can learn how to write by themselves.*

Another respondent noted differences between the perceived performance of students at CLD schools versus non-CLD schools:

> *This morning I just visit this school and I saw that students in Grade 1 and 2 can read and write very well. If compared to the schools that not receive support from this project, the students are [on a] totally different level because other schools did not have [the] activities as same as [the] program support learning-teaching materials like [the] target schools. So [students] that are learning based on the normal curriculum will not have much enthusiasm.*

[12] As described in greater detail later, this does not mean that there were no gains in students' progress, nor does it imply that any gains (perceived or actual) that were made can be attributed to the program. This simply shows that there were no significant changes in the subskills measured in the toolkits themselves—not in overall reading outcomes or reading enjoyment—in the short period during which the pilot was implemented.

In addition to the improvements they observed in students' reading abilities, several teachers also mentioned students' assessment scores when discussing the perceived effects of the program. For example, when responding to the question *"Have the toolkits, including the assessments and remedial activities, affected students?"* one teacher remarked:

> *I think so because based on the assessment result, from time to time the students improve and their score get higher than in previous assessment. For instance, I already complete the remedial activity for students and assess them again. Their score is higher, they can read more words.*

Another teacher described how, *"as I mentioned, at the beginning, some students do not dare to speak or express themselves. When we have the results of assessment, we see how to improve students that they can express themselves more."* A third teacher provided a similar comment: *"For example, when I do assessment with a student about the picture test, at first, she may not know to answer, then after she would know. When she comes out to do assessment, she would pay more attention."*

Although many respondents perceived positive benefits to students' reading abilities, some teachers noted that it was unclear whether the program was impacting their students positively, particularly lower-performing students. When asked, *"Have the toolkits, including the assessments and remedial activities, affected students?"* one teacher responded, *"No, it does not affect anything."* More specifically, this respondent explained:

> *Students cannot read yet. It is difficult because we have to manipulate script and vowel; I think it is difficult. To read aloud is not easy because you have to pronounce correctly. In May, this project asked me to do test on reading comprehension, but it is so difficult; accent, pronunciation and sound some students could not even follow it.*

A principal responded to the same question by saying, *"It is getting better, but not 100 percentage."* Whereas the principal cited here was from a higher-performing school (as measured by student assessment scores), the teacher cited immediately before worked at a lower-performing school where the dominant language was Phu Thai, possibly indicating why there were no observed improvements for students in this school. These perceptions that the program did not lead to effects on students' performance, however, were less common. Most respondents agreed that the CLD program had led to improvements in students' reading enjoyment and in their assessment scores, and these responses were found across schools with high and low student assessment scores and across diverse ethnolinguistic groups.

Although the first theme—on enhanced motivation and engagement—echoes responses provided at midline, as mentioned earlier, findings at endline showed that respondents perceived greater effects on students' overall reading skills. Quantitative data from the toolkits (see Exhibit 19) showed that there were no significant differences in scores between the first and last testing time (other than a marginally significant increase of $p < .10$ in phonological awareness scores) between the assessment sessions. This does not mean there were no gains, nor does it imply that any gains that were made can be attributed to the program. This simply shows that there were no significant changes in the subskills measured in the toolkits themselves—not in overall reading outcomes or reading enjoyment—in the short period during which the pilot was implemented.

Exhibit 19. Average Scores for Each Subskill for the First and Last FA Testing Sessions

	First Assessment Session (November 2017)					Last Assessment Session (May 2018)				
Reading Subskill	***N***	**Mean**	**SD**	**Min**	**Max**	***N***	**Mean**	**SD**	**Min**	**Max**
OVK	552	9.71	0.68	4	10	1,063	9.53	0.92	4	10
Phonological Awareness	551	6.83	3.28	0	11	1,063	7.56	2.36	0	11
Decoding	552	3.98	4.77	0	16	1,062	3.73	4.73	0	16
Reading Comprehension	134	2.40	1.52	0	4	290	2.81	1.31	0	4

Perceived Effects on Students by Gender and Ethnolinguistic Group

Although some teachers and principals noted gender differences in the perceived effects of the toolkits—typically that female students outperformed male students—others said that the effects of the program were the same across genders. For example, one teacher described how *"it is a big difference. Boys are not interested in lessons, but girls listen to what [the] teacher explains, and girl students always do better than boys. Boy students are not afraid of female teachers."* Another teacher echoed this remark when stating, *"Most of the girls in the class are excellent, but boys are hardly to understand because … the boys [do] not concentrate of their study because of playing the mobile phone, watching TV."* In contrast, however, other teachers noted no differences between male and female students. For instance, one teacher remarked how her students performed *"very similar"* and that there were *"no differences [between female and male students] because they have same level of understanding."* Although teachers and principals provided mixed responses, quantitative analysis showed no differential effects of the toolkits by gender. Exhibit 20 presents the comparative analysis of boys and girls on all

subtasks. Average scores for boys and girls across all time periods were similar, suggesting that there were no major gender differences in student performance.[13] However, girls did perform slightly better than boys on semantic fluency (p = .05).

Exhibit 20. Average Scores for Each Subskill for All Testing Times Pooled Together (Boys and Girls)

	Boys					Girls				
Reading Subskill	***N***	**Mean**	**SD**	**Min**	**Max**	***N***	**Mean**	**SD**	**Min**	**Max**
Reading Readiness										
Concept of Print	300	6.45	3.14	0	10	255	6.26	3.21	0	10
Semantic Fluency	300	8.99	4.89	0	23	255	9.82	4.83	0	25
Formative Assessment										
OVK	1,919	9.60	0.82	4	10	1,827	9.55	0.88	4	10
Phonological Awareness	1,918	7.42	2.77	0	11	1,827	7.45	2.82	0	11
Decoding	1,918	3.98	4.65	0	16	1,826	4.42	4.92	0	16
Reading Comprehension	507	2.52	1.46	0	4	575	2.51	1.46	0	4

Surprisingly, in terms of outcomes across the different language speakers in the sample, Exhibit 21 shows that although Lao-dominant students, on average, scored higher in every subskill, there were no significant differences in scores between Lao-dominant and non-Lao-dominant groups, other than marginal significance (p < .10) in OVK difference (with Lao-dominant students scoring higher than their non-Lao-dominant counterparts).

Exhibit 21. Average Scores for Each Subskill for All Testing Times Pooled Together (Lao Dominant Versus Non-Lao-Dominant Students)

	Non-Lao-Dominant Students					Lao-Dominant Students				
Reading Subskill	***N***	**Mean**	**SD**	**Min**	**Max**	***N***	**Mean**	**SD**	**Min**	**Max**
Reading Readiness										
Concept of Print	357	6.29	3.24	0	10	198	6.49	3.04	0	10
Semantic Fluency	357	9.15	5.07	0	25	198	9.77	4.49	0	23
Formative Assessment										
OVK	2,393	9.50	0.92	4	10	1,353	9.71	0.70	4	10
Phonological	2,393	7.13	2.75	0	11	1,352	7.96	2.79	0	11

[13] We also sought to understand whether respondents perceived different effects for different language speakers in their classes. However, this question was not asked systematically in qualitative interviews (it was asked in less than one-third of interviews), limiting our ability to make rigorous claims about perceived effects. For those who were asked this question in their interviews, the responses were largely the same: Students in their classes primarily spoke the same language, so respondents could not perceive differences between different language speakers.

Awareness										
Decoding	2,391	3.58	4.38	0	16	1,353	5.28	5.27	0	16
Reading Comprehension	528	2.33	1.51	0	4	554	2.70	1.38	0	4

We also detected a significant difference ($p < .05$) in students' concept of print and a marginally significant difference ($p < .10$) in decoding scores between students in rural versus urban areas, as shown in Exhibit 22. To reiterate, these data do not point to the impact (or lack thereof) of the program because the data are from the intervention and there is no control group for comparison of this data. In addition, and importantly, these data were not intended to measure reading outcomes or reading enjoyment.

Exhibit 22. Average Scores for Each Subskill for All Testing Times Pooled Together (Urban Versus Rural)

	Rural					Urban				
Reading Subskill	***N***	**Mean**	**SD**	**Min**	**Max**	***N***	**Mean**	**SD**	**Min**	**Max**
Reading Readiness										
Concept of Print	327	6.13	3.21	0	10	228	6.70	3.09	0	10
Semantic Fluency	327	9.60	4.95	0	25	228	9.04	4.75	0	23
Formative Assessment										
OVK	2,368	9.54	0.88	4	10	1,378	9.63	0.80	4	10
Phonological Awareness	2,367	7.21	2.82	0	11	1,378	7.80	2.70	0	11
Decoding	2,366	3.37	4.61	0	16	1,378	5.62	4.76	0	16
Reading Comprehension	470	2.52	1.40	0	4	612	2.52	1.50	0	4

Perceived Effects on Teachers

Teachers, principals, and government officials interviewed at endline also noted several perceived effects of the program on teachers—namely, enhanced teacher motivation as a result of new teaching techniques and materials as well as enhanced teacher knowledge of student performance (obtained through the assessment scores). According to one principal, *"As I see, teachers at this school improved a lot after they get the training and have these activities integrated into the existing curriculum."* A government official said something similar: *"Even though this is just pilot, but teachers really learn from this project."* Respondents noted that the toolkits helped to enhance teachers' overall motivation to teach and their instructional practices by providing them with new teaching materials and pedagogical techniques. In

addition, the assessment data provided by the toolkits improved teachers' knowledge of student progress, allowing them to provide more targeted instruction at various intervals throughout the year.

New Teaching Techniques and Materials

Several teachers noted that the program improved their overall motivation to teach by providing them with new teaching techniques and more engaging classroom materials. A quote from one teacher illustrates this point well:

> *This project makes me feel more motivated to teach, because every time when I teach, I think of the activity of this project and think about how to integrate the project's activity to the lesson teaching in the class. For example, pictures, vocabulary, bringing all these things to use in the class and many times that I use pictures to help explaining about the lesson.*

Several other teachers highlighted the importance of the more engaging teaching materials brought by the program. When asked, *"How has the program affected how you feel about teaching reading?"* one teacher noted, *"I'm glad to have this program support, because the program activities can help students improve on their study with using many teaching materials such as word building cards, stories books, etc. All these materials can improve students' reading."* Another teacher responded to this question in a similar manner:

> *The first good result is that project provides teaching materials for teachers. Teachers keep in the plastic box. This is a good point of this project. Not only Grade 1 and 2 gets benefits, but Grade 3 to 5 also use this teaching materials.*

As described earlier in the section on perceived effects on students, the use of more tactile, visual classroom materials was also exciting for students, possibly increasing students' motivation and engagement in the classroom. If students are more engaged in the classroom activities, this is also likely to enhance teachers' enthusiasm and motivation for teaching.

Although several teachers discussed the value of the new teaching materials, teachers and principals also noted that the program helped teachers to learn new teaching techniques and methods. One teacher described how she was *"proud that this program supports us, because we have learned the new teaching method, how to organize groups of student, how to prepare the lesson, and how to produce the teaching materials to make students understand the lesson better."* Other respondents mentioned more specific aspects of the program teaching materials that were most useful for them to learn, including how to teach specific reading skills (e.g., phonological awareness), how to conduct assessments, and how to integrate these activities into the existing curriculum. These points are illustrated in Exhibit 23.

Exhibit 23. Perceived Changes in Teachers' Practices

Teaching Specific Reading Skills	*"**I have been trained on how to teach phonological awareness, etc. Before I never did the phonological activities** with using the word building cards like this, just taught from the normal curriculum that has not many activities like this program."* *"**I learn about teaching, skills of teaching and the new ways to teach early literacy skills**. … new techniques and knowledge."*
Assessing Students' Skills	*"**I learn how to teach early literacy skills for children and how to do the assessment to know their reading skill**. This makes me more aware of how I teach students, how to speak and pronounce. All of these I am learning. [T] more diverse techniques rather than just following textbook."*
Integrating Activities into Existing Curriculum	*"Some teaching techniques that the teachers learn from this project are being applied in their teaching in class. That is how they improve their teaching. **Previously, the teachers never do activity to explain the lessons, but after got training, they integrate these together and they understand more**."* *"As I see, **teachers at this school improved a lot after they get the training and have these activities integrated into the existing curriculum**. Even though this is just pilot, but teachers really learn from this project."* *"**Previously teachers do not integrate the activity to help explaining the lessons for students**. They just read from textbook for students, and they do not explain this is a noun, this is a verb … but now they know and they explain for students."*

Improving Teachers' Knowledge of Students' Skills

Several respondents noted that another major benefit of the program was the ability to know the different skill levels of students and to know students' skills at various intervals throughout the year. According to one teacher:

> *[The program] is worth it because it can help students get better on their study. After the assessment, we will know that this student is weak of reading, and then we can give an appropriate reading topic for their homework to practice reading further. Then the following date they can read better.*

Another teacher provided a similar remark: *"I teach to improve students to be better, and [the program is] worthwhile because I can know each level of students."* This ability to know students' assessment scores, according to one principal, has led to several changes, *"such as*

teachers know how to adjust the way of teaching to fit with the level of students, and before the teacher did not know, but now they do."

This ability to understand different students' skills and, in turn, provide differentiated remedial instruction was a major added value for teachers who, according to one government official, previously relied mostly on summative assessment methods. This official described how the FA approach differed strongly from the existing approach of the government curriculum and was an important added value of the program:

> *No project can do like this [project] in Laos—like designing the formative assessment—because our educational system is a summative system. That means the test or assessment is as a whole and [the] score or result is a conclusion, so we cannot know what skill students have. This toolkit is really seeking [to know] the capacity of students or the skills of students and what skill he has and where [he needs] to improve. This toolkit can answer the need of literacy development in Laos.*

Several teachers believed that information about their students' skills (provided though the assessment scores) throughout the school year improved their ability to teach reading more effectively to each of their students. However, two respondents noted that this more ongoing process led to some potentially negative effects on teachers. Specifically, the ongoing assessment was seen as creating more work for teachers, as described by one interviewee:

> *Of course, teachers have more work to do because they have many tools or materials to implement the activities. But this more work is positive. It is more detail to facilitate and improve teachers' [ability to] teach skills, and teachers can gain a lot of experience from this program activities. Implementing with students like this is better than just teaching as normal. Teachers will compare that before the project came, teaching was different, but [now that we] have this project support, [the teachers] feel comfortable.*

As discussed earlier, in the "Fidelity of Implementation" section, teachers allocated different amounts of time to preparing for the assessments; for some teachers, the time required may be too much. In this sense, the program may be requiring more effort on the part of teachers to implement the activities effectively. However, this additional time may be offset by the time saved by not having to create teaching materials,[14] as well as by the other benefits of the program for teachers and students.

[14] According to one respondent, the materials provided by the program can save the time that teachers previously may have spent creating engaging classroom materials and activities for their students: *"Somebody is lazy to produce the teaching materials, but this project provide them. So if using the this tools, it can facilitate them to teach easier."*

General Limitations

It is important to acknowledge some limitations of this study. First, when interviewing teachers, principals, and literacy coaches (i.e., PAs and CMs), we relied on self-reports regarding teachers' implementation of the toolkits and the process of advising or coaching teachers on use of the toolkits in the classroom. Although self-reports are not always the most reliable, we attempted to remedy this limitation by triangulating data between respondents, between qualitative responses and quantitative data, and through classroom observations of teachers' activities. This process of triangulation allowed AIR either to verify particular self-reported findings or to identify differences between self-reported responses and findings from other data sources. Identifying these differences allowed AIR to consider additional hypotheses and questions, which we investigated during midline data collection and examined during endline data collection.

Second, we asked several questions during this study regarding teachers' and coaches' practices in the past, which introduces potential recall bias. To remedy this bias, in some cases AIR asked respondents to speak about their experiences during a specific time frame. For example, when asking coaches about their experience advising teachers, we asked them to discuss their most recent coaching session with a teacher. In doing so, we aimed to eliminate some of the recall bias associated with asking coaches to describe experiences throughout the life cycle of the project. Although these limitations are important to acknowledge, AIR is confident that our research, analysis, triangulation, and synthesis process minimizes biases associated with our data.

Finally, our results showed a very limited number of significant results throughout the subgroup analyses, suggesting that there is no systematic difference between any key groups. However, rather than being collected by trained data collectors, these data were taken from the intervention itself. Accordingly, these results were not guided by any particular research questions nor were they collected by trained data collected. Instead, meant to supplement the results of the qualitative evaluation.

Recommendations

Here, we summarize the key recommendations and next steps from the study respondents, as well as research recommendations based on the findings.

Recommendations from Study Respondents

When asked how the program could be improved, respondents provided a diverse range of answers. Specifically, teachers, coaches, and CRS staff said that the toolkits should be further tailored to the local context. This could occur in the following ways:

- Incorporating additional information on how teachers can support non-Lao-speaking children, including issues around tones and phonological differences
- Reducing the amount of text and adding more visuals to the toolkits, as text-heavy documents are not easily navigated by teachers
- Simplifying the toolkit information where possible so that it is easier for teachers to understand
- Changing the physical setup of the toolkits by moving all the toolkit information, assessments, score trackers, and remedial activities into binders to make discrete pieces of the toolkits easier for teachers to remove and use

Based on teachers' and coaches' reports about the amount of time required to implement the assessments, respondents (teachers, principals, and CRS staff) also noted that the testing time should be shortened. Options for reducing testing time include:

- More group-based assessments or assessment that teachers can use with small groups of children
- More forced stops or places in the sequence of assessments where, if students score below a certain number, it would prompt the teacher to stop assessing the student (for example, if a student received below a 5 on decoding, then the teacher would not assess the student on reading comprehension)

Most teachers in our sample considered the remedial activities to be primarily supplementary materials that they could implement alongside the Ministry of Education and Sports curriculum in order to effectively engage students and provide students with concrete examples. Few teachers thought of the remedial activities as a way to adjust their pedagogical approaches to literacy, as intended by the program. In other words, teachers felt the program provided a

foundation upon which to modify their teaching practices. To increase support for teachers, the program could provide the following:

- Longer or more frequent coaching sessions, aligned to key pedagogical approaches.
- Refresher trainings, specifically on implementing the remedial activities
- More sets of materials, including the Lao symbol cards (mentioned by several teachers, this would make it easier for teachers to conduct remedial activities with students in smaller groups)
- Different types of activities and more variety in the activities

Researcher Recommendations

On the basis of the qualitative and quantitative findings, AIR presents the following recommendations for updating and further scaling and sustaining the CLD program for the Lao contexts.

It is important to build on the program's observed strengths. Both the RR and FA toolkits were used widely by all teachers in the program, and there were consistent reports of appreciation and high levels of usability, especially on the assessments. We recommend continuing the teacher trainings because they yielded very positive feedback. We also highlight the need to capitalize on the high perceived impacts of the program on reading enjoyment and reading outcomes. Even though we were unable to measure the impact of the CLD program, we were able to obtain evidence of the perceived impact, which is an invaluable base on which to continue to scale the program. It is highly recommended to use teachers who are comfortable with the program as advocates for the program with their peers, through a possible peer mentoring program.

In this section, we suggest possible modifications to the program to increase effectiveness further. First, we suggest the inclusion of more refresher trainings for teachers to be coached and mentored consistently on the use of the program. These trainings could include the following:

- Additional strategies for group activities
- More instructions for remedial activities

- A detailed training on the differences between the subskills and how and why certain skills take longer to acquire (we recognize that this may be challenging given the training and education levels of teachers in this region but still believe it is worthwhile to consider)

Second, the toolkits themselves could be revised by AIR to do the following:

- Simplify the text throughout further, especially for the remedial teaching activities so that instructions convey the critical information without being overly detailed
- Provide a greater emphasis on visuals

Third, coaches reported struggling to help teachers change their classroom practices. According to AIR's assessment, the program could increase support to coaches working with teachers by including additional training that emphasizes the following:

- A module on working with different types of teachers (e.g., more tenured teachers, those in rural or urban contexts, teachers struggling with parts of the program)
- Information on behavior change and tips for facilitating behavior change

Finally, to measure the impact of the program further and increase its long-term impact, we recommend that it is important to do the following:

- Align and officially integrate the CLD toolkits and trainings with the curriculum and national teacher trainings
- Conduct rigorous mixed-methods research on the impact of the program regarding key outcomes, such as the capacity of teachers to teach reading, students' reading enjoyment, and student reading outcomes

Conclusion

This report lays out both the development framework and the evaluation findings of the CRS-AIR CLD pilot in Laos. The qualitative evaluation findings laid out several significant successes in the program, and explicated ways to scale and sustain these to improve Lao reading outcomes for the children of Laos. At the same time, the evaluation also helped uncover several areas for improvement.

In terms of toolkit usability, it was clear that teachers found the trainings helped their understanding of how to use the toolkits. They were particularly appreciative of the opportunity to try new teaching methods, with engaging activities and materials, like the Lao symbol cards. While coaching was considered a very useful scaffold to utilize the toolkits, there was a need for more support on the implementation of the remedial activities' component of the toolkit. In terms of implementation fidelity, there were high levels of adherence between intended usage and actual usage, with the level of fidelity increasing through the program. The clear channels of communication between the program implementers and the stakeholders facilitated the fidelity of implementation. Finally, in terms of perceived impact, nearly all respondents agreed that the CLD package improved teaching ability, teacher motivation to teach, and enabled teachers to assess student performance on Lao reading and language subskills. There was also consensus that the program lead to perceived improvements in student's reading and level of enjoyment to try interacting with print and read.

The CRS-AIR CLD model presents a unique reading package that addresses a need for an evidence-based, tailored, yet globally-relevant reading package applicable to low and middle-income countries. By tailored we mean the program is designed for the various linguistic, orthographic, and sociolinguistic characteristics in which it is being implemented. In this case, that means the programs is designed specifically for alphasyllabic scripts and is applicable to second language learners who may not have the opportunity to learn to read in their own home or community language. By globally-relevant we mean that the approach focuses on educational contexts with limited print exposure, varying scripts and languages, multilingualism – factors that are widespread globally making this model applicable across countries that are facing a learning crisis.

The model centers on the notion of seamlessly linking classroom-level data and teaching methods to ensure that students are taught at the specific reading or language level they are at. Within this framework, it has three main distinguishing features: (a) it is designed to enhance reading readiness, (b) it includes assessments and remediation activities that are tailored for the South and Southeast Asian orthographies and the cognitive and linguistic subskills and

mechanisms that are required to learn to read them, and (c) it focuses on the facilitative transfer of skills from one language to another in second-language readers. All these factors – limited print exposure, varying scripts and languages, and learning in multilingual contexts – are widespread globally making this model applicable across countries that are facing a learning crisis.

References

Alidou, H., Boly, A., Brock-Utne, B., Diallo, Y. S., Heugh, K., & Wolff, H. E. (2006). Optimizing learning and education in Africa–the language factor. Paris: ADEA.

August, D., & Shanahan, T. (Eds.). (2006). *Developing literacy in second-language learners: Report of the National Literacy Panel on language-minority children and youth.* Washington, DC: Center for Applied Linguistics.

Australian Council for Educational Research. (2015). *Situation analysis: Student learning outcomes in primary education in Lao People's Democratic Republic (PDR)*. Camberwell VIC, Australia: Author.

Ball, J. (2010). Enhancing learning of children from diverse language backgrounds: Mother tongue-based bilingual or multilingual education in the early years. Retrieved from http://unesdoc.unesco.org/images/0021/002122/212270e.pdf

Bernard, H. R. (2017). *Research methods in anthropology: Qualitative and quantitative approaches*. Lanham, MD: Rowan & Littlefield.

Chamberlain, J. (2001). *Participatory poverty assessment: Lao PDR*. Retrieved from http://lad.nafri.org.la/fulltext/372-0.pdf

Chamberlain, M. (2007). *Reading literacy in New Zealand: An overview of New Zealand's results from the Progress in International Ready Literacy Study (PIRLS) 2005/2006*. Wellington, New Zealand: Comparative Education Research Unit, Research Division, Ministry of Education.

Chang, L. Y., Chen, Y. C., & Perfetti, C. A. (2018). GraphCom: A multidimensional measure of graphic complexity applied to 131 written languages. *Behavior Research Methods*, *50*(1), 427–449.

Evans, D. K., & Popova, A. (2016). What really works to improve learning in developing countries? An analysis of divergent findings in systematic reviews. *The World Bank Research Observer*, *31*(2), 242–270.

Florit, E., & Cain, K. (2011). The simple view of reading: Is it valid for different types of alphabetic orthographies? *Educational Psychology Review, 23*(4), 553–576.

Anthony, J. L., & Francis, D. J. (2005). Development of phonological awareness. *Current Directions in Psychological Science*, *14*(5), 255–259.

Gough, P. B., & Tunmer, W. E. (1986). Decoding, reading, and reading disability. *Remedial and Special Education*, *7*(1), 6–10.

Hoover, W. A., & Gough, P. B. (1990). The simple view of reading. *Reading and Writing*, *2*(2), 127–160.

Joshi, R. M., Tao, S., Aaron, P. G., & Quiroz, B. (2012). Cognitive component of componential model of reading applied to different orthographies. *Journal of Learning Disabilities, 45*(5), 480–486.

Kandhadai, P., & Sproat, R. (2010). Impact of spatial ordering of graphemes in alphasyllabic scripts on phonemic awareness in Indic languages. *Writing Systems Research*, *2*(2), 105–116.

Kosonen, K. (2002, January). *Literacy in a minority language—The case of Chong of Chantaburi*. In proceedings of the 8th International Conference on Thai Studies, Nakorn Phanom, Thailand.

Lervåg, A., & Aukrust, V. G. (2010). Vocabulary knowledge is a critical determinant of the difference in reading comprehension growth between first and second language learners. *Journal of Child Psychology and Psychiatry, 51*(5), 612–620.

McBride-Chang, C. (2004). *Children's literacy development.* New York City, New York: Routeldge.

Morgan, D. L. (1996). Focus groups. *Annual Review of Sociology, 22*, 129–152. Retrieved from http://www.annualreviews.org/doi/abs/10.1146/annurev.soc.22.1.129

Nakamura, P., & de Hoop, T. (2014). *Facilitating reading acquisition in multilingual environments in India (FRAME-India). Final report*. Washington, DC: American Institutes for Research.

Proctor, C. P., Carlo, M., August, D., & Snow, C. (2005). Native Spanish-speaking children reading in English: Toward a model of comprehension. *Journal of Educational Psychology*, *97*(2), 246.

UNESCO. (2013). *The global learning crisis: Why every child deserves a quality education*. Retrieved from http://unesdoc.unesco.org/images/0022/002238/223826e.pdf

Lew, S. (2014). A linguistic analysis of the Lao writing system and its suitability for minority language orthographies. *Writing Systems Research*, *6*(1), 25–40.

Simons, G. F., & Fennig, C. D. (Eds.). (2018). *Ethnologue: Languages of the world*. Dallas, TX: SIL International.

Tilstra, J., McMaster, K., Van den Broek, P., Kendeou, P., & Rapp, D. (2009). Simple but complex: Components of the simple view of reading across grade levels. *Journal of Research in Reading*, *32*(4), 383–401.

UNESCO. (2015). *Lao People's Democratic Republic: National assessment of student learning outcomes*. Retrieved from http://nada.uis.unesco.org/nada/en/index.php/catalogue/126

UNESCO. (1953). *The use of the vernacular languages in education*. Monographs on Foundations of Education, No. 8. Paris, France: UNESCO.

Verhoeven, L., van Leeuwe, J., & Vermeer, A. (2011). Vocabulary growth and reading development across the elementary school years. *Scientific Studies of Reading*, *15*(1), 8–25.

World Bank. (2018). *Growing smarter: Learning and equitable development in East Asia and the Pacific*. Washington, DC: The World Bank Group. Retrieved from http://www.batukarinfo.com/system/files/9781464812613.pdf

World Bank. (2018). *World Development Report 2018: Learning to realize education's promise.* Washington, DC: Author. Retrieved from http://www.worldbank.org/en/publication/wdr2018.

Ziegler, J. C., & Goswami, U. (2005). Reading acquisition, developmental dyslexia, and skilled reading across languages: A psycholinguistic grain size theory. *Psychological Bulletin*, *131*(1), 3.

Annex A. Qualitative Data Collection Protocols

Classroom Observation Protocol

Time: 1 hour

Introduction

Observer data:

Observer name: _____

Observation date: _____ Time started: _____ Time ended: _____

Background data:

Grade: ❑ 1 ❑ 2

Subject/class observed: ❑ Lao Language ❑ Other (specify): _____

Teacher Name _____ Sex: ❑ Female ❑ Male

Number of students present in the classroom: _____ Total _____ Male _____ Female

Are there multiple grades in one class? Please describe:

Number of students from each grade present in the classroom _____ G1 _____ G2 _____ G3 _____ G4

Classroom Material

1. Describe the physical state of the classroom. (Are there walls? Is there enough light?)

2. Describe any classroom materials hung on the wall.
    a. What language(s) are the materials in?
    b. What types of materials are they—posters, student materials, or other?
    c. What is the quality of the materials? (Are they visible? Torn? Weatherworn?)

3. Describe any learning materials the students are using.
    a. What types of materials are they—books, handouts, or other?
    b. What language(s) are the materials in?
    c. How many students have these materials?
    d. When is each material used during the lesson?

4. Describe any teaching materials the teacher uses during the lesson.
    a. What types of materials are they (e.g., books, prepared lesson plans [book or paper], or other)?
    b. When is each material used during the lesson?
    c. How does the teacher use each material?

Classroom Activities

5. Describe the classroom activities taking place during your observation.
    a. How long is each activity?

6. Does the teacher reference the Formative Assessment toolkits during your observation?
    ❑ Yes ❑ No ❑ Unsure

7. Estimate how long the teacher spends on instruction, assessments, and remedial activities to promote students' literacy skills.

    _____Instruction

    _____Assessments

    _____Remedial Activities

Assessment

8. Check (√) any assessments the teacher conducts from the Reading Readiness or Formative Assessment toolkit: (check all that apply)

Formative Assessment Toolkit

- ☐ Oral Vocabulary (picture test)
- ☐ Phonological Awareness (sound deletion and identification)
- ☐ Decoding (reading real and nonreal words)
- ☐ Reading Comprehension (reading the story and responding to questions)

9. Is the teacher conducting assessments with one student at a time ❑ Yes ❑ No ❑ Unsure
    a. If yes, what are the students not being assessed doing while the teacher assesses other students?
        i. Describe if/how their behavior changes when the teacher leaves the room.
    b. If the teacher is assessing students in groups, please describe what that looks like (e.g., spatial arrangement of students in the classroom, number of groups, and number of students per group).

10. Is the teacher following the assessment instructions from the toolkit?
    ❑ Yes ❑ No ❑ Unsure
    a. If no or unsure, how is the teacher's assessment process different from the process outlined in the toolkits?

11. Is the teacher recording students' exact scores (the number each student receives) in the toolkit tracker ❑ Yes ❑ No ❑ Unsure
    a. If no or unsure, how is the teacher recording students' scores?

12. Were all students assessed during the observation period? ❑ Yes ❑ No ❑ Unsure

13. How long did each assessment approximately take? [Record: (1) the name of the assessment and (2) the time per student/overall for the group]

Additional comments:

Principal—KII Questionnaire

Time: XX

Respondent: Principal

Goal: To understand if teachers are implementing the toolkits with fidelity and if principals perceive the toolkits to be effective.

Introduction: Please introduce yourself by stating your name and your role as a researcher, then explain the purpose of the research and obtain verbal consent from the respondent.

Background

State: *To begin, I'd like to ask you a couple of questions about who you are and what you do.*

1. What is your position at the school?
2. How long have you been a principal? [*If teacher*] How long have you been a teacher?
3. What languages are spoken as mother tongues by children in the schools you oversee (please list all languages by school)?
    a. [Go through each language listed by the interviewee]: In your opinion, how unalike is this language to Lao?
        i. How specifically do they differ (*accent, tones, pronunciation, written language, etc.*)?
4. In the schools you oversee, how many teachers are first-language Lao speakers?
    a. For those teachers whose first language is not Lao, what languages do they speak?
    b. For teachers whose first language is not Lao, what are the challenges to using Lao as the language of instruction?
5. Is it common for teachers to be employed as both teachers and farmers?
    a. If yes, how does the agricultural/farming season affect teachers' ability to teach?
        i. Which months are the most affected?
    b. How do teachers manage the responsibilities associated with farming and teaching at the same time?
6. How does the farming season affect students in your school?
    a. [Prompt]: Do caregivers take their children with them to work in the fields?
        i. If yes, why do caregivers take their children with them to work in the fields?
        ii. How old are children, typically, when they go with their caregiver to the fields?
7. How do children manage their schoolwork with the responsibilities associated with working in the fields?
8. How do teachers manage students who may miss class because they are working in the fields?

Program

State: *Great, now I am going to ask you some questions about the CRS Program XX and the Reading Readiness and Formative Assessment toolkits.*

1. Are you familiar with the XX program?
2. What are the main challenges facing teachers in your school related to teaching children how to read in Lao?
3. What are the main challenges facing children who are struggling to learn Lao?
    a. How, if at all, are teachers in your school attempting to address these challenges?

Assessments

State: *Within the XX program:*

1. Are teachers able to integrate program activities effectively with their existing responsibilities in the classroom? If no, why do you think this is the case?
2. When teachers are conducting assessments with a single student, what do they typically do with students who are not being assessed?
3. Do teachers share the assessment results with you? Why or why not?
4. Have the assessment results led to any changes in your work or the work of teachers at this school?
    a. If yes, please describe the changes in your work and/or the changes in the work of teachers in this school.

Remediation Activities

1. Do teachers at your school adapt their curriculum or teaching practices to account for the diverse reading levels and skills of children?
    a. If yes, how do they adapt their curriculum?
2. Are you familiar with the remedial activities in the XX program?
3. Are teachers involved in the program implementing the remedial activities?
    a. If yes, please describe your knowledge of these activities.
4. How well are teachers able to integrate remedial activities into their existing classroom responsibilities?
5. Describe any remedial activities (provide examples) (as part of XX program) that you have observed teachers implementing.

Coaching

1. In a typical month, how often do program coaches visit participating teachers at this school?
    a. Describe what typically happens when the coach visits.
    b. Do teachers enjoy working with a coach? Why or why not?

Other Perceived Effects

1. What effects do you hope the program has on teachers and students in your school?
2. In your opinion, have the toolkits—including the assessments and remedial activities—met these expectations? Why or why not?
3. How else could the program support teachers to implement the toolkits more effectively?
4. Describe any remaining suggestions you have to improve this program.
5. Which evaluation findings would you be most interested to learn about from this research *(possible prompts: children's scores before and after the evaluation, teachers' perspectives on the program, challenges the program faced, lessons learned from implementing the program, next steps)*?
6. Would you be interested in participating in a workshop to validate (*verify with local stakeholders*) the evaluation findings?

State: *Thank you. That is the end of my questions. Do you have anything else you would like to add or anything you would like to ask me?*

Teacher—KII Questionnaire

Time: 1.5 hours
Respondent: Teacher
Goal: To understand the usability of the toolkits, if teachers are implementing the toolkits with fidelity, and if teachers perceive the toolkits to be effective.
Interviewer name: _____
Note taker name: _____
School code: _____
Date: _____
Start time: _____
End time: _____

Introduction: Hi, my name is [state your first name]. I am here today to understand your role in the XX program and to learn more about your perspectives on the program, including how to make the program better. I have a list of questions I would like to ask you, but before we start, I would like to read you a script explaining the research and your rights as a participant [read consent script]. [If the participant agrees to participate, continue] Great, as I said, the interview will take about an hour. Let's get started.

Background

State: *To begin, I'd like to ask you a couple of questions about who you are and what you do.*

1. How long have you been a teacher?
    a. How long have you been a teacher at this school?
2. What is your first language?
    a. (If not Lao) Do you face any challenges teaching children Lao? If yes, what are these challenges?
3. Do you have any other work (*please include any small commerce activities or farming*) you do in addition to being a teacher?
    a. If yes, describe the challenges you face to do both jobs well.
4. How does the farming season affect your students?
    b. [Prompt]: Do caregivers take their children with them to work in the fields?
        i. If yes, why do caregivers take their children with them to work in the fields?
        ii. If yes, how old are children, typically, when they go with their caregiver to the fields?
5. How do children manage their schoolwork with the responsibilities associated with working in the fields?
6. How do you manage students who may miss class because they are working in the fields?

Program

State: *Great, now I am going to ask you some questions about the CRS Program XX and the Reading Readiness (note: only G1 teachers and combined G1/G2 classrooms use the Reading Readiness toolkit) and Formative Assessment toolkits.*

Training and Materials

1. Did you receive any training as a part of the XX program?
    a. What activities were particularly useful/helpful for implementing the toolkits?
    b. Did the training adequately prepare you to use the [Reading Readiness and] Formative Assessment toolkit[s] in your classroom(s)? Why or why not?
    c. In what ways could the training have prepared you better?

Assessments

1. Do you prepare in advance to assess students using the [Reading Readiness and] Formative Assessment toolkit[s]? Please describe.
    a. How much time do you spend preparing?
2. If you are conducting assessments with a single student, what do you typically do with students who are not being assessed?
3. How easy or difficult is it to use the [Reading Readiness and] Formative Assessment toolkit assessments?
    a. Which assessments do you find most difficult to administer and why?
    b. Which assessments do you find the least difficult to administer and why?
4. Are the assessments at the appropriate level of difficulty for students in your classroom? Why or why not?

Remediation Activities

1. Are you familiar with the term "remediation activities?"
    a. [If yes] In your own words, what does the term "remediation activities" mean?
    b. [If no] **State:** *remediation activities aim to help students to develop the skills they are struggling with by assessing what level they are achieving, then providing exercises to help them improve.*
2. Which of the following literacy skills [***prompt**: semantic fluency, concept of print, oral vocabulary, decoding, phonological awareness, and reading comprehension*] do the majority of your students struggle with understanding?
    a. How would you select remediation activities to improve these skills by using the toolkits?
    b. Can you show me in the toolkit which remediation activities you use most frequently?
    c. Do you group your students differently based on their assessment results?
        i. If yes, please explain how you group students.
    d. Are the remedial activities at the right level of difficulty for students receiving a ☹, 😐, or ☺? Why or why not?
3. How do you integrate remediation activities into your required curriculum?

4. Have any of your program materials become damaged?
    a. If yes, describe how they became damaged.
5. If your classroom materials became damaged or lost, would you be able to reconstruct them?
    a. If yes, describe how.

Coaching

1. Please describe what typically happens when a Community Mobilizer/Pedagogical Advisor visits your classroom.
    a. How much time do they spend in your classroom?
    b. What do they do during that time?
2. Do you enjoy working with your Community Mobilizer/Pedagogical Advisor? Why or why not?
3. Has the Community Mobilizer/Pedagogical Advisor provided you enough support to implement the program toolkits as recommended in the program guidelines? Why or why not?
4. What other support do you need to conduct the assessments?
5. What other support do you need to provide students with individualized remedial activities based on their assessment results?

Other Perceived Effects

1. What have you learned from this program about early literacy skills and how children learn to read?
2. Of the reading skills [***prompt**: oral vocabulary, concept of print, decoding, phonological awareness, and concept of print*] included in the toolkits, which do you feel most comfortable teaching? Why?
    a. Which do you feel least comfortable teaching? Why?
3. How has the program affected how you feel about teaching reading?
4. Have the toolkits, including the assessments and remedial activities, affected students?
    a. If yes, in what ways?
    b. Has the effect differed for boys versus girls?
        ii. If yes, in what ways?
    c. Has the effect differed for different language speakers in your classroom?
        iii. If yes, in what ways?
5. Do you think that the time it takes to implement the assessments is worthwhile given the information you are able to get on students' skill levels? Why or why not?
6. After the program ends, do you plan on continuing to implement any aspects of the program activities (e.g., the remediation activities, the assessments from the Reading Readiness or Formative Assessment toolkits) or continue using any materials? If yes, please describe.
7. Which evaluation findings would you be most interested to learn about from this research *(possible prompts: children's scores before and after the evaluation, teachers' perspectives on the program, challenges the program faced, lessons learned from implementing the program, next steps)*?
8. Would you be interested in participating in a workshop to validate (*verify with local stakeholders*) the evaluation findings?

State: *Thank you. That is the end of my questions. Do you have anything else you would like to add or anything you would like to ask me?*

CRS Staff—KII Questionnaire

Time: 1 hour
Respondent: CRS Staff Member
Goal: To understand if teachers are implementing the toolkits with fidelity/the program was implemented with fidelity.

Introduction: Please introduce yourself by stating your name, then explain the purpose of the research and obtain verbal consent from the respondent.

Background

State: *You may remember that our data collection team conducted an interview with you on the CRS Literacy Project in January/February 2018. I'd like to ask you some additional questions about this project since we last spoke. To begin, I'd like to ask you a couple of questions about the teachers and schools you oversee.*

1. What languages are spoken as mother tongues by children in the schools you oversee (please list all languages by school)?
    a. [Go through each language listed by the interviewee]: In your opinion, how unalike is this language to Lao?
        i. How specifically do they differ (*accent, tones, pronunciation, written language, etc.*)?
2. In the schools you oversee, how many teachers are first-language Lao speakers?
    a. For those teachers whose first language is not Lao, what languages do they speak?
    b. For teachers whose first language is not Lao, what are the challenges to using Lao as the language of instruction?

Program

State: *Great, now I am going to ask you some questions about the XX Program. Some of the questions may be similar to the questions I previously asked, but I would like to know if, to your knowledge, there have been any updates in the program's progress.*

1. Have teachers continued to implement the program assessments as outlined by the program?
    a. What challenges have teachers faced implementing the assessments?
    b. What assessments do teachers seem to have the most difficulty implementing?
        i. How have coaches supported teachers to implement these specific assessments?
2. How do teachers use the results of the assessments?
    a. Are teachers able to provide instruction to students on the basis of their assessment results?
        i. If yes, how do teachers provide instruction to students on the basis of students' assessment results?
        ii. If not, why are teachers not able to provide instruction to students on the basis of students' assessment results?

   b. How do teachers communicate the assessment results to their students?
      i. [Prompt]: Do teachers tell students their exact score?
      ii. Are students aware of other students' scores?
         i. If yes, how does this affect students?
3. Do teachers understand how to use the remedial activities?
   a. If yes, how do they use the remedial activities?
   b. If yes, how often do teachers use the remedial activities?
4. If no, why do you think teachers do not use the remedial activities?

Coaching

1. How have coaches supported teachers to implement assessments and remedial activities?
2. What areas have coaches specifically worked on with teachers to help improve their implementation and use of the toolkits?
3. How would you improve the coaching to provide more support for teachers to implement the toolkits?

Challenges, Strengths, Recommendations

1. Please describe any elements of implementation that have been particularly successful.
2. What have been the biggest challenges to date?
3. If you could redesign the program, what, if anything, would you change about it to make it more effective?
4. Which evaluation findings would you be most interested to learn about from this research *(possible prompts: children's scores before and after the evaluation, teachers' perspectives on the program, challenges the program faced, lessons learned from implementing the program, next steps)*?

State: *Thank you. That is the end of my questions. Do you have anything else you would like to add or anything you would like to ask me?*

Education Officials—KII Questionnaire

Time: 1 hour

Respondent: PESS Official/DESB Official

Goal: To understand implementation context and perceived impact of the toolkit implementation.

Introduction: Please introduce yourself by stating your name, then explain the purpose of the research and obtain verbal consent from the respondent.

Background

State: *You may remember that our data collection team conducted an interview with you on the CRS Literacy Project in January/February 2018. I'd like to ask you some additional questions about this project since we last spoke. To begin, I'd like to ask you a couple of questions about the teachers and schools you oversee.*

1. What languages are spoken as mother tongues by children in the schools you oversee (please list all languages by school)?
    a. [Go through each language listed by the interviewee]: In your opinion, how unalike is this language to Lao?
        i. [Go through each language listed by the interviewee]: How specifically do they differ (*accent, tones, pronunciation, written language, etc.*)?
2. In the schools you oversee, how many teachers are first-language Lao speakers?
    a. For those teachers whose first language is not Lao, what languages do they speak?
    b. For teachers whose first language is not Lao, what are the challenges to using Lao as the language of instruction?
3. Is it common for teachers to be employed as both teachers and farmers?
    a. If yes, how does the agricultural/farming season affect teachers' ability to teach?
        i. Which months are the most affected?
    b. How do teachers manage the responsibilities associated with farming and teaching at the same time?
4. How does the farming season affect students in the schools you oversee?
    a. [Prompt]: Do caregivers take their children with them to work in the fields?
        i. If yes, why do caregivers take their children with them to work in the fields?
        ii. How old are children, typically, when they go with their caregiver to the fields?
5. How do children manage their schoolwork with the responsibilities associated with working in the fields?
6. How do teachers manage students who may miss class because they are working in the fields?

Program

State: *Great, now I am going to ask you some questions about the XX Program. Some of the questions may be similar to the questions I previously asked, but I would like to know if, to your knowledge, there have been any updates in the program's progress.*

1. Since we last spoke, are you aware of any changes to the program implementation? If yes, please describe these changes.
2. What effects has the program had on how teachers teach reading?
3. What effects has the program had on students' reading performance?

Challenges, Strengths, Recommendations

1. Please describe any elements of implementation that have been particularly successful.
2. Please describe any elements of implementation that have been particularly challenging.
3. If you could redesign the program, what, if anything, would you change about it to make it more effective?
4. Which evaluation findings would you be most interested to learn about from this research *(possible prompts: children's scores before and after the evaluation, teachers' perspectives on the program, challenges the program faced, lessons learned from implementing the program, next steps)*?
5. Would you be interested in participating in a workshop to validate (*verify with local stakeholders*) the evaluation findings?

State: *Thank you. That is the end of my questions. Do you have anything else you would like to add or anything you would like to ask me?*

Literacy Coach/PA—KII Questionnaire

Time: 1.5 hours
Respondent: Literacy Coach/PA
Goal: To understand if teachers are implementing the toolkits with fidelity/the program was implemented with fidelity.

Introduction: Please introduce yourself by stating your name, then explain the purpose of the research and obtain verbal consent from the respondent.

Background

State: *You may remember that our data collection team conducted an interview with you on the CRS Literacy Project in January/February 2018. I'd like to ask you some additional questions about this project since we last spoke. To begin, I'd like to ask you a couple of questions about the teachers and schools you oversee.*

1. What languages are spoken as mother tongues by children in the schools you oversee (please list all languages by school)?
    a. [Go through each language listed by the interviewee]: In your opinion, how unalike is this language to Lao?
        i. How specifically do they differ (*accent, tones, pronunciation, written language, etc.*)?
2. In the schools you oversee, how many teachers are first-language Lao speakers?
    a. For those teachers whose first language is not Lao, what languages do they speak?
    b. For teachers whose first language is not Lao, what are the challenges to using Lao as the language of instruction?
3. Is it common for teachers to be employed as both teachers and farmers?
    a. If yes, how does the agricultural/farming season affect teachers' ability to teach?
        i. Which months are the most affected?
    b. How do teachers manage the responsibilities associated with farming and teaching at the same time?
4. How does the farming season affect students in the schools you oversee?
    a. [Prompt]: Do caregivers take their children with them to work in the fields?
        i. If yes, why do caregivers take their children with them to work in the fields?
        ii. How old are children, typically, when they go with their caregiver to the fields?
5. How do children manage their schoolwork with the responsibilities associated with working in the fields?
6. How do teachers manage students who may miss class because they are working in the fields?

Program

1. Please describe your last coaching visit.
    a. What lesson did you observe during your last visit?
    b. Did you use the coaching guide during your observation? If yes, how did you use this guide?
    c. What did the teacher do well during your observation?
    d. What could the teacher have done better during your observation?
    e. What recommendations did you provide to the teacher?

2. Overall, have teachers continued to implement the program assessments as outlined by the program?
    c. What challenges have teachers had implementing the assessments?
    d. Which assessments do teachers seem to have the most difficulty implementing?
        i. How have you supported teachers to implement these assessments?
3. How do teachers use the results of the assessments?
    c. Are teachers able to provide instruction to students on the basis of their assessment results?
        i. If yes, how do teachers provide instruction to students on the basis of students' assessment results?
        ii. If not, why are teachers not able to provide instruction to students on the basis of students' assessment results?
    d. How do teachers communicate the assessment results to their students?
        i. [Prompt]: Do teachers tell students their exact score?
        ii. Are students aware of other students' scores?
            i. If yes, how does this affect students?
4. Do teachers understand how to use the remedial activities?
    a. If yes, how do they use the remedial activities?
    b. If yes, how often do teachers use the remedial activities?
    c. If no, why do you think teachers do not use the remedial activities?
5. Overall, are teachers able to connect students' assessment scores (happy face, neutral face, or sad face) to the appropriate remedial activities (happy face, neutral face, or sad face) for a particular skill? Why or why not?
6. Are teachers able to teach the skills in the toolkit (oral vocabulary knowledge, concept of print, phonological awareness, decoding, and reading comprehension) according to the standards outlined in the official document?
    a. Which remedial activities do teachers have the most success with and why?
    b. Which remedial activities do teachers struggle with and why?

Challenges, Successes, and Recommendations

1. What aspects of your job as a literacy coach are the most challenging and why?
2. What aspects of your job as a literacy coach do you enjoy the most and why?
3. What do you think are the major strengths of the program as a whole?
4. Which evaluation findings would you be most interested to learn about from this research *(possible prompts: children's scores before and after the evaluation, teachers' perspectives on the program, challenges the program faced, lessons learned from implementing the program, next steps)*?
5. Would you be interested in participating in a workshop to validate (*verify with local stakeholders*) the evaluation findings?

State: *Thank you. That is the end of my questions. Do you have anything else you would like to add or anything you would like to ask me?*

Students—Focus Group Discussion (FGD) Guide

Time: 1 hour plus 15 minutes for setup
Respondent: Grade 2 students
Goal: To understand the activities children enjoy and children's perspective on one of the toolkit tests.
Materials: Mats, paper, crayons, copy of the Formative Assessment toolkit in Lao
Interviewer name: _____
Note taker name: _____
School code: _____
Date: _____
Start time: _____
End time: _____

Background Survey for Data Collectors
Please complete as much of this survey as possible prior to administering the Student FGD Guide. The purpose of this survey is to help researchers understand the contents and logistics of the FGD.

1. Total number of children interviewed: _____
    a. Number of children by sex: Male _____ Female _____
    b. Number of children by age: 6 _____ 7 _____ 8 _____ 9 _____
    c. Number of children by mother tongue: Lao Loum _____ Makong _____ Phu Thai _____ Other _____
2. Language(s) used to administer the tool (*see data collector manual for reference if there are multiple languages in the student group*): _____

Setting Up the FGD With Children (15 minutes)

- Use your first name to make students more comfortable
- For the focus group, sit in a circle on the floor to facilitate discussions
- Put any materials required for an activity in the center of the circle where all students can reach them
- Take breaks as often as students need or whenever students look bored or tired
- Treat all activities as games to help students relax
- Provide students with positive encouragement
- Remind students that this activity has no impact on their work and grades in school

Warm-Up Activity (5 minutes)
State: *Hi everyone. My name is [state your first name]. Today we are going to play some games to learn about how you practice reading and which classroom activities you like best. We are going to play games for about an hour. If you have any questions or you want to take a break, you can ask me directly or raise your hand. If you want to stop at any time, you can; please let me know. To start, let's sing a song. Everyone stand up and repeat after me…*

Ee koeng
Lullaby

(Refrain)
Oe hoe oe hoe oe hoe oe

Ee koeng eoy, kho khao, kho kaeng.
Kho waen thong daeng khaen kho nong laa.
Kho ch/saang, kho maa pet kai ngua khway.
Kho ka-buay thong kham tak nam.
Kho khao yam maa pon kon kham.
Kho mo lam hai nong khoi boeng.
Ee koeng eoy, tu khoi fao kho.

(Refrain)
Oe hoe oe hoe oe hoe oe hoe oe.

Activity 1: Classroom Activities (20 minutes)
State: Great! Let's all sit down in a circle. To start, I'd like to ask you some questions about reading.

- Who thinks they are a good reader? Raise your hand.
- Why do you think you are a good reader?
    - *If necessary*: Why don't you think you are a good reader (*remember to be sensitive*)?
- What do you need to do to become a good reader?

Activity 2: Reading Flower Drawing (20 minutes)

- Great job sharing! Now I'd like to ask you some questions about how you practice reading. I am going to draw a flower on this piece of paper. I'll add a new petal for each way that you can think of that you can practice reading in your classroom. For example, does your teacher read stories aloud to you? I'll write that here.
- How else do you practice reading? For each new thing, I will add another petal to our flower, like this [draw another petal].
    - Do you like this activity? Why or why not?
        - [If yes] What do you like about it?
    - Does anyone else like XX? Why? (mark the number of students who agree)

Break: Stretching (5 minutes)

- Now we will do a small group activity to get our energy up. Everyone stand up!
    - I want you to reach up and stretch your fingers to the sky.
    - Then, I want you to bend over and reach your fingers to the floor.
    - Now, reach your arms to the side as far as you can.
- Great! Let's all sit back down again.

Activity 3: Assessment Feedback (15 minutes)

- Great job! Now I'd like to learn what you think about some of the activities and games you have been doing in your classroom.
    - [show students OVK pictures from FA toolkit and other assessments from the toolkit] Have you seen these pictures or other games before?
    - Where have you seen these games?
    - What does your teacher do with these games?
    - After you finish, does your teacher tell you your score on the game?
        - If yes, what does your teacher normally say?
        - How does that make you feel?
    - Do you want to play these games again?
        - If no, why not?
        - If yes, why?

Closing

We are all done. Please raise your hand if:

- You had fun during the session.
- You felt uncomfortable during the session.
- You learned something.
- You felt happy.
- There is something you want to tell us.

Please raise your hand if you have any questions you would like to ask. Otherwise, thank you for participating in our games. Let's all clap for each other for a job well done.

Annex B. Covalidation Workshop

AIR conducted a covalidation workshop May 28 and 29, 2018, with teachers and principals from all 34 implementation schools to engage all of the program teachers and principals in discussing our research findings, providing greater detail and nuance to our findings, and brainstorming recommendations based on the findings. During the workshop, participants were divided into small groups (10 groups of nine participants). Each group was given a list of the program's successes and challenges according to the midline research. Participants then were asked to discuss among themselves and write down whether they agreed or disagreed with each assertion. Groups then ranked their top three findings on the basis of how important they believed them to be to overall program success. Facilitators took the top challenges each group listed and condensed them down into 10 key challenges. Each group then brainstormed recommendations for how to respond to each challenge. In the last session, each participant voted on the recommendations they felt were most important.

The data from the workshop include all the participating groups' five handouts with worksheets, the list of recommendations, and the polling information for the top recommendations. Findings from the workshop are included in our overall findings, presented in the following sections, and our recommendations as an additional layer of quality assurance for our research process.

Exhibit B-1: Covalidation Workshop Activities

Left: Voting on recommendations at the covalidation workshop

Right: Group work at the covalidation workshop

AIR®
AMERICAN INSTITUTES FOR RESEARCH®

CRS
CATHOLIC RELIEF SERVICES

Established in 1946, the American Institutes for Research (AIR) is an independent, nonpartisan, not-for-profit organization that conducts behavioral and social science research on important social issues and delivers technical assistance, both domestically and internationally, in the areas of education, health, and workforce productivity.

MAKING RESEARCH **RELEVANT**

AMERICAN INSTITUTES FOR RESEARCH
1000 Thomas Jefferson Street NW
Washington, DC 20007-3835 | 202.403.5000
www.air.org

LOCATIONS

Domestic: Washington, DC (HQ) | Monterey, Sacramento, and San Mateo, CA | Atlanta, GA | Honolulu, HI | Chicago and Naperville, IL Indianapolis, IN | Metairie, LA | Waltham, MA | Frederick and Rockville, MD | Chapel Hill, NC | New York, NY | Columbus, OH | Cayce, SC Austin, TX | Reston, VA

International: El Salvador | Ethiopia | Haiti | Honduras | Zambia