# Can an Accelerated Intervention Close the School Readiness Gap for Disadvantaged Children? An Evaluation of the Effects of the LEARN Project's Summer Pre-Primary Program on Literacy Outcomes in Northern Lao PDR

Jodie Fonseca

December 2021

A thesis presented to the Institute of Education, University College London
in partial fulfilment of the requirements for the degree of Doctor in Education

***Declaration***

I, Jodie Fonseca, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Signed:

Dated: 21 December 2021

***Acknowledgements***

First and hardiest thanks are due to Dr. Jane Hurry, who shepherded my doctoral career from its inception in Ethiopia to the mountainsides of Vietnam and Laos and most recently to Kathmandu, Nepal. Jane's incisive advice and tireless support propelled me to achieve enormous academic and professional growth – so much that, looking back, I'm happily unable to recognize the researcher I was when I started. Little did she imagine the journey would take nearly 10 years to complete!

Just like Jane, my husband probably never anticipated that this path would start early in one decade and end in the next. From our very first trip to London with our 1-year-old daughter to where we are now, I could never have finished this program without my family providing the space to translate thoughts into words on paper during countless nights and weekends. I also credit the example my mother set of advancing her own higher education as a working parent. My father didn't live to see the end of the dusty, potholed road I began following as a practitioner-researcher in Addis Ababa in 2011. But I know how proud he was.

Finally, I would like to extend thanks to the LEARN Project donors and advisors, without whom this study would not have been possible. LEARN was resourced by Dubai Cares – a philanthropy that prioritizes catalytic programs to inform education sector investments – with contributions from Plan International and Save the Children International. I was fortunate to be part of these organizations' dedication to building evidence-based solutions for some of the most marginalized children in the world. I was also privileged to count on the incredible colleagues at Plan Laos and Plan Canada when management and methodological decisions got real-world tough. This thesis reflects their hard work as much as it does mine.

***Abstract***

Developed against the backdrop of Sustainable Development Goal 4, as well as a global trend towards rigorous assessment of early childhood programs, this thesis answers questions about the effects of an accelerated school readiness intervention for non-Lao children in disadvantaged communities of Lao People's Democratic Republic. Through a longitudinal, cluster randomized control trial, the study employs multi-level regression with an analytical sample of 391 children to examine the outcomes of a summer pre-primary program piloted from 2015-2018 by the Lao government with support from Plan International and Save the Children International in the Dubai-Cares funded Lao Educational Access, Research, and Networking (LEARN) Project.

Research questions are investigated through a design in which the same panel of children are assessed against a control group at three intervals using the Measurement of Development and Early Learning. The thesis identifies significant associations between receiving the treatment and achieving higher gain scores on several emergent literacy tasks between baseline and midline, with effects roughly in line with similar interventions in other contexts. At the same time, the thesis finds that those effects had largely faded by endline. An interaction between treatment and ethnicity was only evident in a few instances, suggesting that the intervention may have boosted school readiness for Khmu children more by the start of grade 1 and for Hmong children more during grade 1.

The thesis raises important recommendations about how to improve the fit between the ultimate objectives of accelerated interventions, the evaluations they undergo, and the needs of the broader education system. New contributions to knowledge are also found by interrogating a global assessment paradigm through a comparative linguistic lens, so that forthcoming evaluations benefit from the lessons learned based on LEARN's attempt to fit a square peg into a unique alpha-syllabic, tonal Southeast Asian language.

***Impact Statement***

For policymakers, this thesis offers substantive evidence to inform decisions on equitable expansion of school readiness services in resource-constrained settings. It suggests that accelerated, foot-in-the-door interventions may help education systems reduce inefficiency in the early grades and support the most disadvantaged children. Nevertheless, it raises important questions about the durability of learning outcomes in summer-time school readiness programs, particularly when children will go on to enter grade 1 classrooms with lower quality instruction and curricular and language mismatches.

For practitioners, the thesis describes how an accelerated pre-primary program might be set up for similar populations, and how it may help to close school readiness gaps for disadvantaged children. It offers critical suggestions about improving the uptake of the intervention and reducing sample attrition for the very children who are most likely to be excluded from formal service delivery. The thesis also recommends that more attention be paid to strengthening the literacy-related content of the summer pre-primary model rolled out through LEARN, in Laos and other locales.

For program evaluators, the thesis appraises the pros and cons of using a global assessment tool to measure programmatic impact. By describing the pitfalls experienced in the LEARN evaluation, the thesis may help others avoid similar limitations in the future. It also provides detailed suggestions regarding how the assessment tool might be further adapted to ensure appropriateness to a relatively unstudied language in a multilingual environment, and to effectively measure persistence of skills into early primary school.

For academic and linguistic audiences wishing to catalyze more contextualized evaluation practice, the thesis offers an abundance of questions. How can we expand and deepen the body of literature on effective literacy assessment in alpha-syllabic languages? How can we contribute to better assessments of children's skills

in lexical tone, and how they learn to read without spaces between words? What steps can we take to build knowledge about the cross-context validity of assessment approaches in disadvantaged, multilingual settings, as well as the predictive value of emergent literacy skills for later progress in primary school in similar environments?

In terms of avenues for further discussion and dissemination, I propose two concrete suggestions to strengthen the impact of the thesis.

- To the evaluators and policymakers working on SDG monitoring through tools like the Measure of Development and Early Learning: Find ways to share lessons learned across stakeholders with experience in similar language systems in South and Southeast Asia. By doing so, you will be providing an enormous service to the millions of second language learners who struggle to build reading skills in the early primary grades in these contexts.

- To the growing number of national governments and development agencies that now have considerable experience with accelerated pre-primary interventions: There is much to be learned from each other, and it will be critical to find opportunities to convene and share insights about the most – and least – effective interventions in pursuit of equitable, high-quality school readiness.

***Reflective Statement***

This statement provides reflections on the content I covered in the different courses in the Doctorate in Education program as well as the linkages between each element. The common threads across all my experiences in the program include a focus on literacy development in the early years for disadvantaged children in lower- and middle-income countries; coupled with a spirit of critiquing research methods that have been 'imported' into a context or that privilege the views of outside experts.

The Foundations of Professionalism (FoP) assignment reflected on the professional status of early childhood development (ECD) teachers in Ethiopia, describing how ECD around the globe has traditionally been seen as a 'caring' profession but has become increasingly professionalized in recent decades, and discussing the implications of this shift for Ethiopia. The paper argued that, to increase the quality and availability of ECD services in Ethiopia, ECD teachers in the country should follow the trend of professionalization while also retaining some of the key caring aspects that make their profession more accessible to remote communities.

The Methods of Enquiry 1 (MoE1) paper presented a proposed research design for a participatory, child-focused situational analysis for a new emergent literacy project in Ethiopia. The assignment discussed the evolution in thinking over the past 50 years about young children's literacy development, from a positivist stance in which literacy is a set of predictable, mechanical acts; to, more recently, a social constructivist approach in which the environment surrounding young children is seen to play a key role in their literacy acquisition. The design of the proposed situational analysis built from the latter perspective, using a participatory, 'mosaic' approach to capture not just traditional literacy skills among pre-school age children, but also perceptions, practices, and priorities related to literacy among children, their families, and their communities in a rural Ethiopian setting.

The Methods of Enquiry 2 (MoE2) paper explored the use of a particular research methodology – the expert interview – in pursuit of information about the sustainability

of Save the Children's literacy programming around the world, which includes a substantial focus on early literacy development. The MoE2 assignment first provided a literature review and conceptual framework about the issue of sustainability in international development with an emphasis on literacy interventions. Then, the paper interrogated the expert interview as a tool for digging deeper into the issues and concepts surfaced by the literature review. In the paper, I concluded that the expert interview can be a useful point of entry into a topic, helping the researcher access the tacit knowledge of experts in a given field. However, it is also critical to expand the circle of inquiry to capture the perceptions and priorities of 'everyday people' engaged in the struggle to provide services to marginalized children rather than relying only on experts.

The first three assignments helped build a foundation for work on the Institution Focused Study (IFS) and thesis by strengthening my understanding of the theoretical underpinnings of education research and my ability to engage critically with the body of literature on my topic of interest. In addition, the three modules helped me understand my own implicit theoretical stance – best described as social constructivism coupled with a positivist perspective about the fundamental skills that children need to develop to become literate – and make it explicit in my research design. Finally, the first three modules helped me uncover my own biases as an insider researcher and strive to understand the power dynamics that can be created in practitioner-led research.

My initial intention in the EdD program was to focus my IFS and thesis research on Save the Children's emergent literacy programming in Ethiopia, but due to new circumstances (the birth of a child and a move to Vietnam and later to Laos), I had to make some changes to that plan. My revised IFS topic focused on helping the Save the Children Vietnam team conduct a situational analysis for a new ECD program, which had a strong emphasis on emergent literacy in the multilingual context among ethnic minority communities in the remote mountain highlands of the country. The situational analysis reflected the research design I laid out in the MoE1 assignment, but with some changes based on the Vietnamese context and the programmatic needs

for Save the Children. This activity also built upon the work I completed in the first three taught modules of the EdD program by allowing me to bring new theoretical perspectives and methodological knowledge to bear on the practical exercise of data collection, analysis, and reporting.

The situational analysis was conducted in the proposed new target areas in Lao Cai Province, Vietnam in September 2013, with the aim of developing an understanding of the context and thus being able to design the most relevant program possible. This qualitative study entailed 36 classroom and school observations, in-depth interviews and focus group discussions with 181 stakeholders, including children, teachers, parents, and educational officials. The situational analysis identified an inter-connected web of early grade reading challenges in Lao Cai. Although there was some evidence of positive instructional practices, for the most part teachers were ill-equipped to deliver effective early grades reading education. These shortcomings stemmed from problems related to the language of instruction; inadequate teacher training; and limited scope to adapt national-level approaches to the local context. An array of background factors in children's home and community lives compounded this picture.

The study was limited by a small sample size as well as 'real-world' constraints such as limited time leading to breadth of data collection at the expense of depth. Nevertheless, it provided a wealth of useful information about early grades reading in the target schools. After completing data collection and reporting for the analysis, I then supported Save the Children to conduct detailed program design based on the study findings.

The first four assignments I completed for the International EdD program represented a logical progression from the more abstract (a hypothetical "professional project" for Ethiopian ECD teachers in FoP) to the more concrete (the design of a situational analysis for a proposed new literacy program in MoE1) to the highly practical (the use and critique of a qualitative research methodology in MoE2, and the actual implementation of the MoE1 situational analysis during my IFS). Each of the

assignments tackled my topic of interest – emergent literacy – from a slightly different angle, looking first at the professionalization of those who teach early literacy; second at a research design that would effectively capture people's perceptions, practices, and priorities about children's emergent literacy; third at the sustainability of literacy interventions over time and after Save the Children support ends; and fourth at the status of early childhood development and early grade reading skills as well as the governance social structures that support children's learning in disadvantaged communities in Vietnam.

The formative feedback that I received on the three assignments for the EdD program moulded my academic thinking in a few key ways and helped shape my approach to the thesis stage. First, the feedback consistently showed that I needed to be more thorough and critical in my engagement with the literature in the fields of ECD and early literacy. This reflects the challenges I faced in shifting from solely a practitioner mindset to that of a practitioner-researcher, who needs to interrogate her own field with a critical eye and from a strong theoretical foundation rather than accepting things at face value. The formative feedback also pointed out that I needed to be clearer and more transparent in describing proposed data analysis methods, and explaining the processes I followed and choices I made in data analysis and interpretation, as well as the shortcomings of those approaches.

My work in the EdD program has been closely linked to my everyday professional practice as an education technical advisor and project manager, where I am required to guide teams implementing early childhood development and early grade reading programs that are grounded in the latest research and evidence. The EdD allowed me to bring my day-to-day work up to a higher standard by becoming more self-reflective and strengthening the links between theory, research, and practice. For instance, I had always implicitly understood that it might be useful for ECD teachers in Africa to become more professionalized, but in FoP, I was able to use the literature about professionalism to engage more critically with the benefits and drawbacks of professionalization of the ECD field. In MoE1, I interrogated a very common research approach in international development (the situational analysis) from a theoretical

perspective, uncovering new elements of ‘everyday’ literacy – as opposed to formal or ‘classic’ literacy – that I had not previously considered. MoE2 allowed me to become more critical and reflective about a methodology that I use frequently in my research, the expert interview, but had never stopped to consider as a methodology per se. In the IFS, I engaged critically with the literature on early learning; participatory, qualitative research methods; and ethno-linguistic marginalization in Southeast Asia.

These prior steps provided a valuable theoretical and critical foundation for the thesis as the culminating experience in the EdD program. After taking up a new job and moving to Laos, I adjusted my intended thesis topic to include an impact evaluation of the accelerated school readiness program for which I had become the project director. The thesis started out with a largely quantitative design that involved examining the effects of the school readiness model in remote communities in Laos. However, prior EdD training on research methods coupled with literature reviews on the unique socio-linguistic features of Southeast Asia led me to question the suitability of the quantitative assessment paradigm in the Lao context. As a result, my thesis retained some of the original quantitative intent while at the same time broadening out to include an interrogation of global assessment tools that were first developed for Western contexts and languages. The thesis also unfurled against the backdrop of the Sustainable Development Goals, which created a need for standardized approaches to measuring children’s learning across a vast array of nations and contexts around the world – an inherently fraught process.

Although the EdD program was a long journey, it was woven with many common threads and each step built upon the previous in a logical progression. In the end, I am pleased with the new contributions to knowledge that I was able to generate, both for my own understanding, and for my professional field more broadly. I type the last sentence in this reflective statement knowing that these skills of questioning and thinking critically about common challenges in the international education field will continue to serve me and my work for years to come.

*Table of Contents*

*Index of Tables*

*Index of Figures*

## 1. Introduction

This thesis aims to answer questions about the effects of an accelerated school readiness intervention for non-Lao children in highly disadvantaged communities of Lao People's Democratic Republic (PDR). Using a longitudinal cluster randomized control trial (RCT), the study examines the outcomes of a summer pre-primary program piloted by the Lao government with support from Plan International and Save the Children International through the Dubai-Cares funded Lao Educational Access, Research, and Networking (LEARN) Project. Implemented and refined over the course of three school years from 2016-2018, the accelerated intervention was purposefully designed to target harder-to-reach communities where children are typically educationally disadvantaged.

The central hypothesis of the thesis is that the accelerated model will effectively promote school readiness and early grades learning for children who would otherwise fall outside the purview of the formal education system due to resource constraints, linguistic marginalization, and the remoteness of their locations. This is investigated through an experimental research design in which children who participated in the accelerated intervention are assessed for school readiness against a control group at three intervals: At baseline prior to the start of activities; at mid-term, just after the accelerated model has concluded and the study population is entering grade 1; and at the end of grade 1.

The study aims to bridge practice, research, and policy by providing actionable evidence to government decision-makers and international development agencies for making choices about where to invest limited education sector resources in the face of competing demands. As the Lao government aspires to achieve Sustainable Development Goal (SDG) 4 related to inclusive and equitable quality education by 2030, this study also aims to build a better understanding of the unique educational constraints facing linguistically, socially, and economically disadvantaged children and the flexible approaches that might be best tailored to their realities. The intervention assessed in the study was designed to address the fact that, without

catch-up support, children in the most marginalized communities were likely to begin grade 1 vastly below curriculum expectations at the start of primary school and remain unlikely to surmount this inherent disadvantage in the first few primary grades – a common threat to achieving educational equity in least developed countries (Willms, 2018).

Theoretically, the thesis is guided by an interactive model of reading that blends developmental, linguistic, and social constructivist perspectives. According to this model, reading acquisition results from the interaction between specific cognitive skills (so-called 'bottom-up' skills); and the ability to make meaning out of text through social factors such as prior knowledge and language exposure, experience, and motivation (so-called 'top-down' skills). The thesis is grounded in a developmental approach to literacy acquisition, in which children progress along a predictable continuum in building reading skills between the ages of five to seven years old as they transition from early childhood into primary school. This includes cognitive skills that previous research on early grades reading have highlighted as critical building blocks for second language learners, such as oral vocabulary, phonological awareness, letter knowledge and listening comprehension; with some adaptations for the alpha-syllabic, tonal nature of the Lao language.

Development of these core skills at the nexus between early childhood and the early primary grades is complicated for ethnic children in remote villages of Lao PDR by their lack of exposure to previous school readiness opportunities and limited prior knowledge of Lao language. Thus, their status as second language learners in a highly challenging context necessitates a strong emphasis on the literature related to second language teaching and learning, combined with a social constructivist lens that addresses the importance of background factors as predictors of learning.

This thesis makes important original contributions to knowledge on several levels, from the local to the global. Within the Lao and regional context, the study represents one of the first ever large-scale, rigorous efforts to measure the effectiveness of an accelerated school readiness intervention for highly marginalized children. As such,

the study contributes to deeper understanding about interventions that can close equity gaps for children who are 'at the bottom of the pyramid' of learning outcomes (Wagner, Wolf, and Boruch, 2018) within and across countries and more effectively measure their learning in a contextually appropriate way. It provides evidence to inform policy decisions to address what has been described as 'weak early childhood development (ECD) syndrome:' Persistent poor performance of primary education systems as a consequence of inadequate ECD and school readiness service provision (Crouch, Olefir, Saeki, and Savrimootoo, 2020).

At a wider level, the study ties into global trends in the development sector towards more rigorous assessment of programmatic impacts. By utilizing internationally developed instruments from the Measuring Early Learning Quality and Outcomes (MELQO) Initiative[1] and the Early Grade Reading Assessment (EGRA), the study contributes to a global effort to strengthen measurement approaches in low to middle income countries, moving beyond historical reliance on an evidence base that was largely developed in wealthier nations. The study provides valuable data to allow proponents of MELQO and EGRA to gauge the reliability and validity of their instruments in assessing school readiness and early grades learning in a Southeast Asian linguistic context and for an accelerated intervention model.

Finally, the study contributes to a more fine-grained understanding of the specific concerns that underpin assessment of emergent literacy and early grade reading for second language learners in the relatively under-studied Lao language. Specifically, the thesis' examination of the cross-language reliability and validity of globally developed assessment approaches in an alpha-syllabic, tonal, *scriptio continua*[2] language and writing system pushes the boundaries of the existing literature that mainly encompasses non-tonal, alphabetic languages. As such, the study findings will be of wider use for others seeking to use similar assessment tools in conducting assessments in Lao language, as well as alpha-syllabic and / or tonal languages in

[1] A collaboration between Brookings Institution, UNESCO, UNICEF, and the World Bank with inputs from a wide group of development partners.
[2] A *scriptio continua* writing system is one which has no spaces or other delineating marks between words.

other parts of Southeast Asia and the world, and in the extensively multi-lingual environments that are often the norm in international development practice.

## 2. Literature Review

### *2.1. Literature Review Purpose and Search Strategy*

The literature review combined elements of historical, conceptual, and methodological reviews (Hart, 2018; Robson, 2011) and by nature was academic but also interventionist – related to the use of knowledge to solve practical challenges (Hart, 2018). The first task was to relate ideas and theory about how young children acquire literacy – and synthesize prior knowledge and experiences from similar contexts – to the circumstances of Laos. This was first achieved by investigating models of reading and early literacy acquisition from the developed world, relying mainly on classic works and meta-analyses to do the job of synthesis for me. Next, I turned to the literature on emergent literacy as well as second language learning and literacy development across cultures and languages, particularly in developing countries, using landmark work on emergent literacy and second language learning, as well as meta-analyses related to literacy in developing countries.

Following this, I queried all the electronic resources available through the University College London (UCL) library – which includes a wide array of electronic databases, electronic journals, e-books, dissertations, etc. – to identify a growing corpus of published books and peer-reviewed journal articles. Since not all the relevant literature on literacy in developing countries was available in peer reviewed publications, I also conducted an extensive internet search (using Google/Google Scholar) and tapped into my network of contacts with practitioners and researchers to identify the significant body of gray literature.

The literature review was also used for establishing the standard of evidence and analysis in my field and comparing findings to those of similar interventions elsewhere. An important result of this process was the identification of critical

independent variables that underpin children's learning achievement in Laos and that therefore needed to be incorporated through the overall analytical approach to the extent possible. This methodologically oriented review was conducted partly through an electronic search of academic journals, but because academic publications are relatively scarce for emergent literacy assessment in developing countries (although increasing in number rapidly), the search had to rely in some measure on gray literature.

To deepen my understanding of the interplay between linguistics, learning, and assessment, I used a snowball search approach to the literature in literacy development in languages that are similar to Lao. To achieve this, I reviewed a set of core publications from contexts where alpha-syllabic, tonal, and/or *scriptio continua* languages exist. I then utilized the citation lists from those publications to identify further relevant documentation through books and journals.

Classic works and meta-analyses on emergent literacy and school readiness tended to appear in the 1980s and 1990s, while I restricted my search of evidence on literacy in developing countries mainly to the 2000s onward. In the case of the literature on alpha-syllabic and tonal languages, there simply was very little rigorous published documentation prior to the mid-2000s.

### *2.2. Theoretical Models of Early Reading Development*

This section describes the theoretical model of reading that guides this thesis, first situating it within historical foundations and trends in thinking about reading acquisition; and continuing with discussion on the theoretical fundamentals of emergent literacy and second language learning for young children. The section concludes with an interrogation of the literature on literacy acquisition in alpha-syllabic, tonal, and *scriptio continua* languages like Lao.

This thesis is grounded in a theoretical framework that blends Piagetian conceptions of child development as a process that unfolds along a predictable continuum, with

constructivist notions of learning as social endeavor embedded within the learner's context and experiences. As such, it draws heavily on an interactive model of reading in which young learners build on cognitive skills as well as previous contextual and linguistic knowledge, with the support of knowledgeable teachers, to become skilled readers.

Although there is no global consensus on a single theoretical model that can describe all aspects of reading, there is clear agreement in the literature that reading is a complex, varied, interdisciplinary process. A plethora of models attempt to describe aspects of reading from a cognitive perspective, such as the skills of word identification, phonological decoding, and syntactic parsing (Rayner et. al, 2010). Indeed, from roughly the 1920s through the 1970s, the dominant discourse around reading acquisition aimed to break down reading into an innate skill that can be built through repetitive drills. Later and more sophisticatedly, reading was deconstructed into a series of cognitive tasks and conceptual understandings that begin with visual inputs to the human eye in the form of text (Alexander and Fox, 2004) that must be achieved with a level of automaticity before comprehension can be reached (Perfetti and Lesgold, 1977). These are often referred to as 'bottom-up' models, in which the learner assembles meaning out of the constituent parts of words and sentences (Rumelhart, 1985).

While bottom-up lenses are informative in attempting to describe the mechanics of reading, the period from the 1970s to the present ushered in an increasing focus on so-called 'top-down' skills, in which learners draw on previous knowledge of the features of written text and spoken language as well as the general context of the text they are reading (Alexander and Fox, 2004). This includes an emphasis on the use of prior knowledge on a given topic as a mental strategy for organizing knowledge and facilitating its later recall (Rumelhart and Ortony, 1977; Rumelhart, 1980).

In the late 1970s, Rumelhart was among the first to put voice to an interactive model of reading, which weaves the bottom-up and top-down threads into an integrated

process that pulls from the reader's knowledge of syntax, semantics, orthography, and lexicon to make sense of written text (Rumelhart, 1985). Subsequently, the field of New Literacy Studies (NLS) began to take the critique of bottom-up models to greater lengths. While "a focus on the acquisition of literacy and the forms of literacy associated with schooling led to the dominance of cognitive science and psychological approaches in education" (Roswell and Pahl, 2015, p. 1), thinkers such as Street and Gee highlighted the contested, political nature of literacy (Ibid.). In contrast to the cognitive (and typically quantitative) models of research and learning assessment often associated with formal schooling, NLS promoted more qualitative, contextualized methods such as socio-linguistic ethnography (Gee, 1998).

Three extensions of these negotiations between cognitive and constructivist paradigms that are most relevant to this thesis came to prominence at around the same time as NLS: Emergent literacy and school readiness, the Simple View of Reading (SVR), and the study of reading acquisition in second languages.

First coined by Teale and Sulzby (1986), emergent literacy drew on the child development work of Piaget and Vygotsky and conceptualized how children begin to build reading and writing skills from birth through entry into formal schooling. Emergent literacy as a sub-discipline propounded several tenets: That literacy acquisition begins long before formal instruction; that skills such as listening, speaking, reading, and writing develop concurrently and not sequentially; and that literacy advances through children's engagement not just with its forms but also its functions in the real world (Ibid.). Further, the link between oral and written language development came to the fore in conceptualizations and studies of children's literacy development prior to schooling (Rhyner, Haebig, and West, 2009). These views had important implications for research, policy, and practice for young children because they required earlier and more home- and community-based interventions that promoted literacy as a multi-faceted endeavor and centered the role of oral language as a springboard.

The practice area of school readiness draws heavily from these developmental conceptions and interactive models of child development, including emergent literacy. The concept of 'readiness' for formal schooling has existed in the literature since at least the late 1800s, often following a 'maturation' paradigm in which children reach – or fail to reach – certain predictable thresholds of readiness during their early life course and this should determine their progression in the education system (Scott-Little, Kagan, and Frelow, 2006; High, 2008; Flewitt and Ang, 2020). Although there is no universal, agreed-upon definition of school readiness, from a research lens, "school readiness refers to the state of child competencies at the time of school entry that are important for later success" (Snow, 2016, p. 8). Evolving from initial conceptualizations that focused predominantly on cognitive skills – and particularly reading readiness – those school readiness competencies are now typically taken to include, in addition to language and cognitive skills, aspects such as physical well-being and motor development, approaches to learning and executive function, and socio-emotional development (Linan-Thompson, 2014; Kagan, Moore, and Bredekamp, 1995).

A landmark meta-analysis in the 1990s of school readiness programs in the United States identified that the language competencies needed for success in school encompass both verbal language – including skills such as listening, speaking, and meaning-making – as well as emergent literacy skills such as print and alphabet awareness, story sense, and familiarity with the conventions of writing (Kagan, Moore, and Bredekamp, 1995). Later, a UNICEF-commissioned conceptual framework set the tone for policies and interventions to promote school readiness in the international development sphere. This framework highlighted not just the importance of children's competency development during a time of important life transition from home into a formal learning setting, but also the need for schools to be ready for learners and parents and communities to be prepared to support their children's success upon and after school entry (Britto, 2012).

An emergent literacy and school readiness perspective tells us that children at different stages of life tend to have the ability to master different reading-related

skills. Before the age of five or six, they are usually pre-readers, imitating reading and playing around with letters and sounds but not able to read by themselves (Snow, Burns, and Griffin, 1998). Then at around five or six – the school readiness window in most formal education systems – children make a life-changing transition from “logographic” or “pre-alphabetic” (perceiving words as similar to pictures) to “alphabetic” (realizing that words are made up of constituent parts such as letters or syllables, and aware of the connections between graphemes and phonemes) (Dombey, 2009, p. 8; Moats, 1998, pp. 1-3; Ehri, 2005).

Based on meta-analyses of evidence primarily on reading acquisition in English, this development of the alphabetic principle allows children to begin the decoding stage at around six to seven years of age (typically grade 1), blending letters and sounding out simple words (Chall, 1996). Fluency is then usually achieved at around seven to eight years old, or roughly between grades 1-3 in most education systems (Ibid. and Snow, Burns, and Griffin, 1998).

The outgrowth of this developmental perspective was captured in seminal meta-analyses of reading acquisition from the late 1990s and early 2000s, which made recommendations for effective reading instruction based on the evidence at the time (Adams, 1990; National Reading Panel, 2000). These analyses identified that, to support children to become strong readers, reading instruction in the early grades should first help them understand that the smaller written units of the language – units that do not have meaning in and of themselves – assemble into larger units that do have meaning. Following this, they must be able to ‘decode’ from the written text they see on the page into coherent speech.

In alphabetic and to a lesser-studied extent in syllabic languages, children must first learn the alphabet of the language (letter knowledge); learn that the language is made up of sounds and sound patterns and that those sounds correspond to letters or groups of letters (phonological and grapho-phonemic awareness); and learn to group those sounds associated with letters into syllables and words (Adams, 1990, p. 3; National Reading Panel, 2000). Children also need to know what many words in

the language mean (lexicon); and be able to read quickly, accurately and with the correct intonation (fluency and accuracy). Mastering these four skills will together contribute to achieving a fifth skill, which is the ultimate goal: Reading with understanding (reading comprehension). These core skills have now come to dominate the discourse around early reading instruction and assessment in the international development community.

In the mid-1980s to early 1990s, the bottom-up and top-down views of reading development were distilled into the SVR. A direct challenge to strictly bottom-up models, the simple view utilized the existing, rigorous evidence at the time to identify two components that both make individual contributions but when taken together are even more consistently and highly predictive of skilled reading: Decoding and language comprehension (Gough and Tunmer, 1986; Hoover and Gough, 1990). In the SVR, decoding is a composite of building block skills such as print concepts, alphabet knowledge, and word reading; while language comprehension, syntactic and semantic knowledge as well as other background knowledge form the foundation of language comprehension (Tunmer and Hoover, 2017). Decoding and language comprehension, when taken together, are powerful predictors of later reading skill and remain robust as children age (Lervag, Hulme, and Melby-Lervag, 2018).

Recognizing the limitations of a reading paradigm such as SVR that was first developed through study of English, subsequent researchers problematized the original view through a comparative linguistic lens. For instance, a 2011 meta-analysis to determine whether the SVR holds up for languages with shallow orthographies (particularly Spanish) as well as it does for opaque orthographies like English “found that the relative influence of decoding and linguistic comprehension on reading comprehension is influenced by the transparency of the orthography of the language that has to be mastered” (Florit, p. 569). As this analysis identified, in the deep orthography of English, decoding was more influential than linguistic comprehension for beginning readers and remained strongly influential for the next few years of instruction; while in shallow orthographies that are considered easier to

master at a younger age, linguistic comprehension was more predictive of reading comprehension than decoding for beginner readers (Ibid.).

Another 2014 meta-analysis (Nag, Chiat, Torgerson, and Snowling) looking into the existing evidence base on literacy acquisition and assessment in developing countries identified the critical importance of oral language skills, particularly vocabulary and sentence composition (syntax) for later success – and especially so for children who do not speak the language of instruction at home. This analysis also found that it was key for children to begin reading instruction with well-established concepts about print; and that there was moderate evidence to support the importance of phonological awareness, or awareness of the sound structures and patterns in the language. In addition, despite the universality of many foundational skills across languages, some predictors are stronger for specific language types and writing systems where different types of psycholinguistic units play different roles. This included syllable recognition in an alpha-syllabic language like Kannada (Ibid.); and the ability to manipulate tones as a phonological unit in Chinese (Ho and Bryant, 1997).

### *2.3. Reading Acquisition in Alpha-syllabic Languages*

The defining feature that differentiates alpha-syllabic writing systems from alphabetic languages is the fact that, “whereas alphabetic scripts represent speech at the level of the phoneme, syllabaries represent it at the level of syllables” (Nag, Treiman, and Snowling, 2010). In alpha-syllabaries, phoneme marks representing consonant and vowel sounds are typically clustered together into a syllabogram, with vowel graphemes often appearing in a non-linear fashion around the central consonant.

Alpha-syllabic writing systems are mainly found in three areas of the world, collectively encompassing large numbers of native speakers: South, Southeast, and East Asia, where the greatest diversity of these types of languages exists; Ethiopia and parts of Eritrea; and pockets of northern North America (Nag, Caravolas, and Snowling, 2011). As interest in cross-linguistic analysis of reading acquisition has

burgeoned in recent decades beyond the initial works on English and other alphabetic languages, the literature on alpha-syllabic writing systems has grown apace, challenging common paradigms of reading acquisition that resulted in treating alpha-syllabic languages as alphabets for the purposes of reading instruction and assessment (Share and Daniels, 2016, p. 27).

Most of the published research on alpha-syllabaries focuses on what are identified as the 'prototype' Indic alpha-syllabic scripts, including Hindi, Kannada, and Urdu, but also Korean Hangul and Japanese Hiragana (Nag, Caravolas, and Snowling, 2011). Much of this literature examines alpha-syllabic languages from a psycholinguistic perspective, investigating how predictors of early reading acquisition may differ in these languages from typical skill progression in alphabetic languages.

A 2014 meta-analysis (Linan-Thompson) identified that in lower- and middle-income countries (LMICs) and across a range of languages, the hierarchy of reading acquisition skills is consistent: "Large unit skills such as blending and segmenting of compound words and blending and segmenting words at the syllable level emerge prior to formal reading instruction while skills at the phoneme level, such as blending and segmenting, develop with formal reading instruction" (p. 5). Even across languages that differ in orthographic depth and phonological structure, syllable awareness universally emerges before phoneme awareness and is predictive of early reading; but "phoneme-level awareness is not a uniformly important factor in reading syllabaries and morpho-syllabaries" (Perfetti and Verhoeven, 2017, p. 460). However, evidence from the Indian alpha-syllabaries indicates that when children learning to read in these writing systems develop a meta-cognitive understanding of the particularities of how sounds map to symbols, they are more effective in transitioning from a mechanistic decoding to a skilled reading phase (Nag, 2017).

A 2011 compilation of research on alpha-syllabic writing systems in India and East Asia also identified that the underlying processes for reading acquisition – such as phonological awareness – are similar to alphabetic languages, but that syllabaries and alpha-syllabaries present additional unique challenges to early readers (Nag,

Caravolas, and Snowling, 2011). In particular, these languages typically have very large consonant-vowel syllable inventories: Up to 1,000 in Korean and 200-500 in the Indic languages, for example. These types of languages also tend to place greater demand on children's orthographic knowledge – their knowledge of the individual symbols in the language and the spatial and sequential rules for combining them – than the more linear and graphemically consistent alphabetic languages (Nag and Snowling, 2012).

Alpha-syllabic writing systems also require specific skill in syllable-level phonological processing as opposed to just letter-sound relationships (Vasanta, 2004; Nakamura, Koda, and Joshi, 2014). In addition, these orthographic systems may place greater demand on visual processing skills because of their non-linear nature and the variation in type and location of diacritics in the scripts. This aligns with the orthographic depth hypothesis, which argues that how quickly and easily children can learn to read in a given language is affected by the "directness and simplicity with which a writing system represents the phonology of the language" (Frost and Katz, 1989, p. 302).

The consequence is that reading instruction in these languages must explicitly teach orthographic knowledge and syllable-related skills, and in most cases requires more instructional time than alphabetic languages for children to achieve mastery (Nag, 2007; Nakamura, Koda, and Joshi, 2014; Nakamura and de Hoop, 2014). Given the longer timeline needed to develop skilled reading, targets for aspects such as decoding, oral reading fluency, and reading comprehension in early primary school in these languages will likely also need to be more modest than in less complex alphabetic scripts (RTI International, 2017).

Other psycholinguistic studies of alpha-syllabic languages have tended to investigate the relative contributions of different foundational skills – such as phonological awareness, syllabic processing, rapid automized naming (RAN), and oral language skills – to children's development of skilled reading. For example, a 2012 study of Kannada examined the extent to which children's syllable and phoneme-level

processing and orthographic knowledge explained variance in their reading accuracy and fluency. The authors found that both orthographic and syllabic knowledge were predictors of later reading skill, with the importance of syllabic knowledge persisting as children aged to a greater extent than for alphabetic languages, but with phonological awareness skills developing more slowly (Nag and Snowling, 2012). This finding was repeated in a study of literacy acquisition in multilingual contexts in India focusing on Kannada, Telugu, and English (Nakamura and de Hoop, 2014). The findings add support to the importance of explicitly teaching both syllable and phoneme-level processing in alpha-syllabic languages.

Nag and Snowling (2012) also identified the consequences of orthographic knowledge, inferring that children in these language types must develop the 'alpha-syllabic' principle much like children in alphabetic languages need to develop the alphabetic principle. Children who have insight into the orthography and can easily access both syllable-level understanding as well as individual phonological properties within syllables tend to have better RAN (Ibid.). In turn, RAN has been shown to be a robust predictor of reading acquisition in the Indian alpha-syllabaries (Nag and Perfetti, 2014; Wijayathilake, Parrila, Inoue, and Nag, 2014).

A 2018 study of Kannada and Telugu (Nakamura, Joshi, and Ji) investigated the relative contributions of syllabic awareness, phonemic awareness, and oral vocabulary knowledge in alpha-syllabic decoding, and looked at changes in the relative predictive power of these variables across primary school grades. The study utilized oral phoneme/syllable deletion, receptive vocabulary, and decoding tasks to assess these skills. The authors found that syllabic awareness was strongly predictive of decoding skills in grade 1-5, and increasingly so as children got older. The converse was true of phonemic awareness skills, which were still a good predictor of decoding skills but lost their prognostic power as children moved up in grades. At the same time, the authors found inconsistent results related to oral language as a predictor of decoding ability.

Nag, Treiman, and Snowling (2010) found that children learning to read and write Kannada are typically taught to read using whole syllables, although they would also benefit from learning to build words using phoneme-level segments, and this skill would potentially improve their ability to spell more complex syllables in the language by tapping into their knowledge of segmentation. This finding has been repeated in other alpha-syllabic languages, including Rumi (Liow and Lee, 2004) and Oriya (Mishra and Stainthorp, 2007).

Nag's (2007) study of Kannada identified that children's skill in decoding consonants with inherent vowels[3] emerges earlier than skill in ligaturing more complex combinations such as consonant-consonant-vowel. This complexity, coupled with an intricate orthography that requires children to master many symbols, means that it may take children up to four years of instruction to achieve reading fluency (Ibid.). Similarly, a study of Malayalam, a South Indian alpha-syllabic language, found that children could master simpler consonants and vowels by grade three but still struggled to decode compound consonants and consonants that carried vowel diacritics (Tiwari, 2011).

Moving from South Asia to East and Southeast Asia, the published literature on the implications of alpha-syllabic orthographies is sparser. A study on predictors of reading comprehension in Korean identified that orthographic awareness was an important factor and recommended that "instruction on orthographic awareness—letter patterns and multi-letter units (e.g., graphemes, phonograms)—should be considered as part of text-reading fluency instruction, particularly during an earlier phase of reading development" (Kim, 2015, p. 477).

A 2015 review of the Cambodian education sector identified that Khmer has features, like Lao, that make children's early reading acquisition more complicated. This includes the fact that some vowels are 'dependent' upon nearby consonants

[3] Inherent vowels are a feature of alpha-syllabic scripts in which each consonant is taken to have a default vowel sound unless a different vowel sound is explicitly marked, or the inherent vowel is cancelled.

and are often written in a non-linear fashion, represented by superscripts, subscripts and diacritic marks placed around a base consonant (RTI International, 2015a). A 2018 EGRA adapted to this reality by including not just a test of consonant identification, but also of dependent vowels. This study in grades 1-3 found that students in the sample, whose average age was approximately 7 years, performed better on consonant identification but could identify only four vowels out of 20 with more than 50% accuracy (RTI International, 2018a).

A limited number of studies of the alpha-syllabic languages in Ethiopia and Eritrea exist in the published literature. A 2009 study (Afsaha et al.) identified that children learning in the alpha-syllabic Ge'ez script had an advantage in developing early reading skills over those reading in a Latin script, even though Ge'ez has a much larger number of symbols to learn because it is an alpha-syllabary that combines consonants and vowels into syllabic blocks. The authors postulate that this is because the larger grain size of the script is easier for children to grasp and supported the idea that teaching decoding of larger grain sizes in these languages (i.e., syllable- rather than letter-based decoding) will lead to better outcomes.

An analysis of the effects of different scripts and writing systems on children's reading skills in Ethiopia looked at children's skills in reading letters, familiar words, non-words, and stories in languages written in syllabic Ge'ez or alphabetic Latin script, controlling for socio-economic status and other background factors. The authors found that script matters – "children acquired the Ge'ez script more accurately earlier in the reading process" (Piper and van Ginkel, 2017 p. 54). They postulate in line with Afsaha et al. that the larger grain size of Ge'ez facilitates children's early reading acquisition.

### *2.4. Reading Acquisition in* Scriptio Continua *Languages*

Another script feature that is unique to a handful of the writing systems of East and Southeast Asia – including Thai, Lao, Burmese, Khmer, and others – but that has largely died out in other systems is the use of *scriptio continua*, or script with no

spaces or other delineating marks between words. A few studies of how this lack of spacing affects the reading process of Thai script, a writing system with similar characteristics to Lao, have been conducted. Kasisopa et al. (2013) found that adults' eye movements when reading Thai were like those when reading spaced scripts like English, but that where possible, readers called on morphological knowledge of characters and sentence-level lexical knowledge and contextual understanding to determine where word boundaries fall.

The authors experimented with adding spaces between words in Thai text, demonstrating that this facilitated individual word recognition but slowed down sentence speed, perhaps because skilled Thai readers were unused to the presence of such spaces. These outcomes broadly mirror a similar 2009 study conducted on adult Thai-English bilinguals and English monolinguals (Winskel, Radach, and Luksaneeyanawin, 2009).

Kasisopa et al. (2016) set out to investigate if the same findings hold true for children aged 12-14 years and found that adding spaces significantly improved their word recognition skills, facilitating their ability to use cues related to word-initial and word-final characters to identify word boundaries. The authors noted that Thai children are taught to read with spaces in primary grade 1, but that textbooks and instruction revert to *scriptio continua* starting in grade 2. They call for further research on younger children to determine the optimal instructional approach.

A similar study of space addition in Chinese, which does not typically have spaces between characters, found that it facilitated children's ability to recognize words more than unspaced text. The authors identified an especially striking effect for new vocabulary words, where the advantage of learning them in spaced format carried over into more rapid recognition of the same words in unspaced format, with strong implications for instructional practice for early readers (Blythe et al, 2012).

### *2.5. Reading Acquisition in Tonal Languages*

In addition to the alpha-syllabic, *scriptio-continua* aspects of writing systems like Lao, early reading development in these languages is also complicated by the occurrence of tonality. Tone is a feature of an estimated 60-70 percent of languages around the world (Yip, 2002), including large swathes of China, mainland Southeast Asia and the Pacific, as well as some languages in Africa and the Americas (Maddieson, 2013).

Tone is a suprasegmental aspect of language - a “vocal effect that extends over more than one sound segment in an utterance” (SIL, 2019, para. 1). Although all languages use variations in intonation – such as the rising melody at the end of a sentence to indicate a question in English – in truly tonal languages, tone is a contrastive feature in which variations in pitch and contour are built into specific aspects of words in a fixed and predictable way and change the meaning of an utterance (Yip, 2002). Tone is often lexical in nature, as is the case in many of the Asian tone languages, where “the distinctive pitch level carried by the syllable of a word … is an essential feature of the meaning of that word” and a change in tone on the same morpheme completely alters the meaning (SIL, 2019, para. 1).

There is evidence for a developmental and neurological basis for tone as a basic human speech characteristic. For instance, infants’ first utterances have been shown to exhibit the specific intonation patterns of their native language, and children can reproduce the tonality of their spoken native language after the first few years of life (Mang, 2001). Infants may in fact be able to perceive the specific tones of their home language prior to six months of age, even before developing the ability to perceive native-language vowel and consonant sounds (Yeung, Chen, and Werker, 2013). Moreover, the auditory ability to recognize tonal pitch and contour appears to be hard-wired in brain circuitry at an early age for speakers of tone languages, but not for speakers of non-tonal languages (Mukari et al., 2015).

Tone generally is taken to have three aspects: Pitch contour, amplitude, and duration (Liu et al., 2011). For example, with pitch contour, a tone may begin with a low, medium, or high pitch and it may then continue steadily, rise, fall, rise and fall, fall

then rise, or waver creakily. Tones that remain level are often referred to as ‘static,’ while those that vary in pitch are ‘dynamic’ (Wayland and Guion, 2003). Many of the tonal languages in Southeast Asia, such as Thai and Hmong, are highly complex, with large numbers of tones and variations in the types of tones that occur (Maddieson, 2013). The existence of multiple tones dramatically expands the inventory of word meanings that are linked to a relatively small set of monosyllabic morphemes (Chen and Pasquarella, 2017; Mang, 2001).

Orthographically, tone can be explicitly marked or implicit and unmarked. When marked, this is typically achieved in three ways: Through word-final silent letters; word-initial and word-final punctuation; or the use of diacritics (Roberts, Walter, and Snider, 2016). When it is unmarked, tone must be inferred by the reader based on prior knowledge of features such as phonology, morphology, and spelling rules, and lexical meaning (Wayland and Guion, 2003; Dooley, 2015).

Tone has increasingly been recognized as a critical feature of language impacting literacy acquisition (see, for example, Graham and van Ginkel, 2014; Abadzi, 2012; and RTI International, 2015b and 2017). Nevertheless, there is little published research on exactly how and why this impact occurs. Much of the research has been conducted with adults, and the bulk has been published on the Chinese tonal languages, with a few studies on other Asian languages such as Thai.

For example, several investigations have shown that adult native speakers of a tonal language tend to have better pitch perception than non-tonal speakers both in language and music (Mukari et al., 2015). Several studies have demonstrated that speakers of tonal languages have a distinct advantage in recognizing tone in a second tonal language than non-tonal speakers (Burnham and Jones, 2002; Stevens, Keller, and Tyler, 2004; Wayland and Guion, 2004).

At the same time, when people who already speak a tonal language attempt distinguish the tones of a second, unfamiliar tonal language, the tone structures of their first language (L1) may interfere with recognition of some tones in the second

language (L2) (Hao, 2012; Qin and Mok, 2011). In addition, the characteristics of tones affect the ability to differentiate them from each other, even for native speakers of tonal languages but more so for non-natives (Mukari et al., 2015; Stevens, Keller, and Tyler, 2004).

In her seminal meta-analysis on literacy instruction in developing countries, Abadzi (2013) recommended that reading instruction in tonal languages must not take tone for granted, but rather should include an explicit emphasis on learning tone sounds. This conclusion is supported by two studies that presented non-tonal speaking adults with minimal pairs[4] of different tones in Mandarin Chinese or Thai and asked them to identify or differentiate the tones from each other, using audio recordings by native speakers. Both studies identified strong and lasting positive gains in tone recognition and differentiation through targeted training, even for tones that are typically difficult to distinguish (Wang, Spence, Jongman, and Sereno, 1999; Wayland and Li, 2008).

A study of tone discrimination in Thai by both Thai-naïve and Thai-exposed adult English speakers in comparison to native Thai speakers also identified noteworthy findings. In particular, the authors found that English speakers who had previously been ambiently exposed to spoken Thai outperformed the Thai-naïve speakers in distinguishing Thai tones, suggesting the importance of prior oral language exposure in developing 'an ear' for tone. What is more, both naïve and experienced L2 learners could achieve near-native tone recognition with practice. Finally, the authors note that the task of recognizing tones places significant load on the working memory of L2 learners in a way that does not occur for L1 speakers, highlighting the importance of automatizing tone recognition (Wayland and Guion, 2003).

A limited set of studies on children's development of tone skills have been conducted, with comparable findings to the studies on adults. Kasisopa et al. (2018) conducted experimental training with groups of six- and eight-year-old children who came from tonal or non-tonal language backgrounds. Children were given six

[4] Minimal pairs are words in the same language that differ in only one phonological (or in this case, tonological) element and have contrastive meanings (Handke, 2017).

training sessions in learning to distinguish foreign Mandarin tones using either audio-only (voice recording) or audio-visual (voice recording plus video recording of the speaker's face) cues. Overall, the eight-year-olds showed more improvement in tone recognition than the six-year-olds, with the authors suggesting that the task may have been difficult for the younger group because it required a significant amount of sustained attention; and because six-year-olds may have less perceptual sensitivity to the sounds of a foreign language than eight-year-olds.

For the children in the study, prior tone language experience facilitated the ability to learn tones in another language more than no prior tone language exposure. In addition, monolingual children improved only when given audio-visual but not audio-only cues, while bilinguals improved most with audio-only cues and to a lesser extent with audio-visual cues, suggesting that tone-naïve children need visual scaffolding to support tone learning (Kasisopa et al., 2018).

A similar study looked at the effects of two weeks of targeted training on tone acquisition in Mandarin Chinese – with six sessions of 40 minutes each – for non-tonal children ages six to fourteen. The authors found that the training significantly improved tone recognition by all age groups, but that the gains were more modest for the younger group, again suggesting that longer periods of sustained attention to L2 tone acquisition is likely to be more developmentally challenging for younger children (Wang and Kuhl, 2003).

Although this theory has not been empirically substantiated, there is speculation that tone affects aspects of early reading development such as fluency, and particularly the optimal oral reading fluency rates at which comprehension can be achieved. For example, Graham and van Ginkel (2014) included Sabaot – a Kenyan language with six marked grammatical tones and an agglutinating morphology in which morphemes can be added to each other to form new words – in their problematization of word-per-minute reading fluency benchmarks across different language types. They found that children achieved 80% comprehension at a rate of 30 words per minute (WPM) in Sabaot, well below typical international benchmarks of around 45 WPM. They

postulate that this may occur because reading tone marks places a greater cognitive load on children as they are decoding and therefore slowed down the pace at which the Sabaot-speaking children in their study could read while still achieving comprehension.

Abadzi (2012) also notes that tonal languages may have lower WPM thresholds at which children can achieve comprehension because they "convey two bits of information in one syllable, and may thus [be] more 'efficient'" at conveying meaning using shorter words (pp. 8-9). She recommends that this be further tested through empirical study and benchmarking of reading skills in tonal languages.

In the next section, the thesis turns to a discussion of the Lao educational and child development context, delving into the unique linguistic features of Lao and other ethnic languages and how they interact with learning and assessment. Subsequent sections then apply the understanding from the broader literature review and from the Lao contextual overview to investigate the effectiveness of LEARN's accelerated school readiness interventions, suggest policy and practice priorities, and identify recommendations for improved assessment approaches.

## *2.6. The Lao Context*

### 2.6.1. An Overview of the Lao Context

Lao PDR is a landlocked, mainland Southeast Asian country bordering China, Vietnam, Cambodia, Thailand, and Myanmar, with a land area of more than 236,000 $km^2$. A population of roughly 6.9 million people lives in its 18 provinces, with nearly 66 percent inhabiting rural areas, although urbanization is occurring at a yearly rate of 4.9 percent (UNDP, 2019).

Laos was conquered as a Siamese vassal state in the late 1700s and became a French colony from 1893 until gaining independence in 1953. This was followed by decades of civil war between the Royal Lao Government, which had resumed power

after the French withdrew, and the communist *Pathet Lao* ('Lao Nation'). Subsequent involvement in the Second Indochina War – referred to in the West as the 'Vietnam War' – left the country impoverished and riddled with unexploded bombs and land mines that are still being cleared today. With the defeat of the royalists, the monarchy was abolished, the communist Lao People's Revolutionary Party instituted a one-party state, and the Lao People's Democratic Republic was established in December 1975 and continues to date (Stuart-Fox, 2005).

The country subsequently experienced more than a decade of financial backing from the Soviet Union, coupled with geopolitical isolation outside the confines of the Communist bloc. Laos has opened up considerably in recent years, however, implementing market-based economic reforms, becoming eligible to graduate from Least Developed Country (LDC) status in 2024 if it sustains current trends (UN Lao PDR, 2018), and actively pursuing integration through the Association of Southeast Asian Nations. Now classified as a lower middle-income country, Laos' gross national income per capita was US$ 6,070 in 2019, and the country had the 20$^{th}$ fastest annual gross domestic product growth rate in the world in 2017 at 6.9% (World Bank, 2019).

### 2.6.2. The Lao Education System and Challenges

The Lao education system currently comprises four levels of instruction: ECD, general education, vocational education, and higher education. This section will focus on the first two levels, which are most relevant to the LEARN study.

The Lao Ministry of Education and Sports (MOES) reported that as of the 2017-2018 school year, the country had 2,597 pre-primary classes with 4,995 teachers and 116,553 students; and 8,558 primary schools with 33,339 teachers and 780,008 students (Ministry of Education and Sports, 2018). The vast majority of these are government schools, as private education is not widespread. The school year runs for approximately nine months, starting in September of one year and ending in May of the following year.

Pre-primary and primary school curricula and materials are developed at the national level by MOES structures, notably the Research Institute for Education Sciences (RIES). Since 2015, RIES has been in the process of revamping the curricula for ECD and general education sub-sectors to move away from rote learning approaches and towards a student-centered, competency based, thematic curriculum. In the early primary grades in particular, reform efforts aim to reduce what is considered an overcrowded curriculum that does not adequately teach basic literacy and numeracy skills or incorporate effective language learning techniques for non-Lao speakers (UNICEF, 2015).

Lao PDR has made impressive strides in recent decades with regard to educational access and participation. Nevertheless, these gains at the national level mask significant ongoing shortcomings in educational access and quality for sub-populations and in sub-sectors related to early childhood development and school readiness. These issues disproportionately affect children living in poverty and in rural areas without a road, as well as non-Lao children, particularly girls (World Bank, 2016b). Importantly, non-Lao populations report that one of the steepest barriers to educational participation is the limited relevance of the education they receive to their daily lives and livelihoods (Ibid; and Ansell et al., 2018).

When available, ECD services are typically delivered through crèches for children ages three months to two years in urban areas; kindergarten for children ages three to five in urban and peri-urban areas; and a one-year pre-primary course for five-year-olds in rural areas. Community-based ECD and school readiness services are also provided through the United Nations Children's Fund (UNICEF) and World Bank support as well as international non-governmental organization (INGO) projects. Nevertheless, despite expansion in recent years, the ECD sub-sector suffers from the lowest levels of learner access to service provision in the education system (MOES, 2018b).

To address persistent access gaps, the government's 2016-2020 Education Sector Development Plan (ESDP) established a target of achieving an 80% gross enrollment ratio for five-year-olds in kindergarten or pre-primary by 2020, an increase from 66% in 2015 (MOES, 2015b). This target aligns with one of the key indicators in SDG 4: *4.2.2 Participation rate in organized learning (one year before the official primary entry age), by sex.* The country was on track to achieve the ESDP goal, with the national gross enrollment ratio (GER) already reaching 77% by 2017.

However, these national-level figures conceal stark disparities between well-off and disadvantaged districts that undermine the country's ability to push the GER forward nationally. According to the 2018 ESDP Mid-term Review (MTR), the district with the lowest enrollment had a GER of only 28%, while in the district with the highest access, the rate was 104%. Progress in reducing disparities between districts over the period covered by the ESDP has not kept pace with progress in raising access overall nationally (MOES, 2018b).

With regard to ECD quality and outcomes, another relevant indicator of SDG 4 is *4.2.1 Proportion of children under 5 years of age who are developmentally on track in health, learning and psychosocial well-being, by sex.* The government has demonstrated strong commitment to measuring this indicator more effectively through a partnership with the World Bank – which was partly funded through the LEARN Project – to adapt the Measurement of Development and Early Learning (MODEL) global assessment tool to the Lao context and conduct the first-ever large-scale diagnostic assessment of child development and school readiness in the country. This diagnostic focused on disadvantaged areas in the north and included a sample of 7,520 children ages 2-5 years old (70% non-Lao) across five provinces, 14 districts and 376 villages (World Bank, 2016c).

The World Bank-supported baseline study found that children scored significantly lower on literacy skills than on numeracy skills, although achievement levels in both domains were discouraging. Focusing on the five-year-olds in the sample, the average score on the literacy domain was only 25%, only .8% of children who

responded could correctly identify the initial sounds in words, and on average they could name only 2.7 letters out of 20. Outcomes in all domains were markedly lower for children from non-Lao ethnic groups, particularly the Hmong (World Bank, 2016c), indicating that the government and its partners face a considerable challenge in achieving the ECD equity goals in SDG4.

After completing the early childhood level, children progress into the general education system, which encompasses five years of primary education, four years of lower secondary, and three years of upper secondary (Ministry of Education and Sports, 2015b). At the primary education level, Laos's ESDP goals are similarly guided by the tenets of SDG 4, particularly SDG target *4.1 By 2030, ensure that all girls and boys complete free, equitable and quality primary and secondary education leading to relevant and effective learning outcomes.* In line with this target, the ESDP set objectives such as reducing grade 1 dropout from 8.5% in 2015 to less than 5% by 2020, improving learning outcomes above levels achieved during a national-level assessment in 2009, and increasing the primary cohort completion rate to 89% (MOES, 2015b).

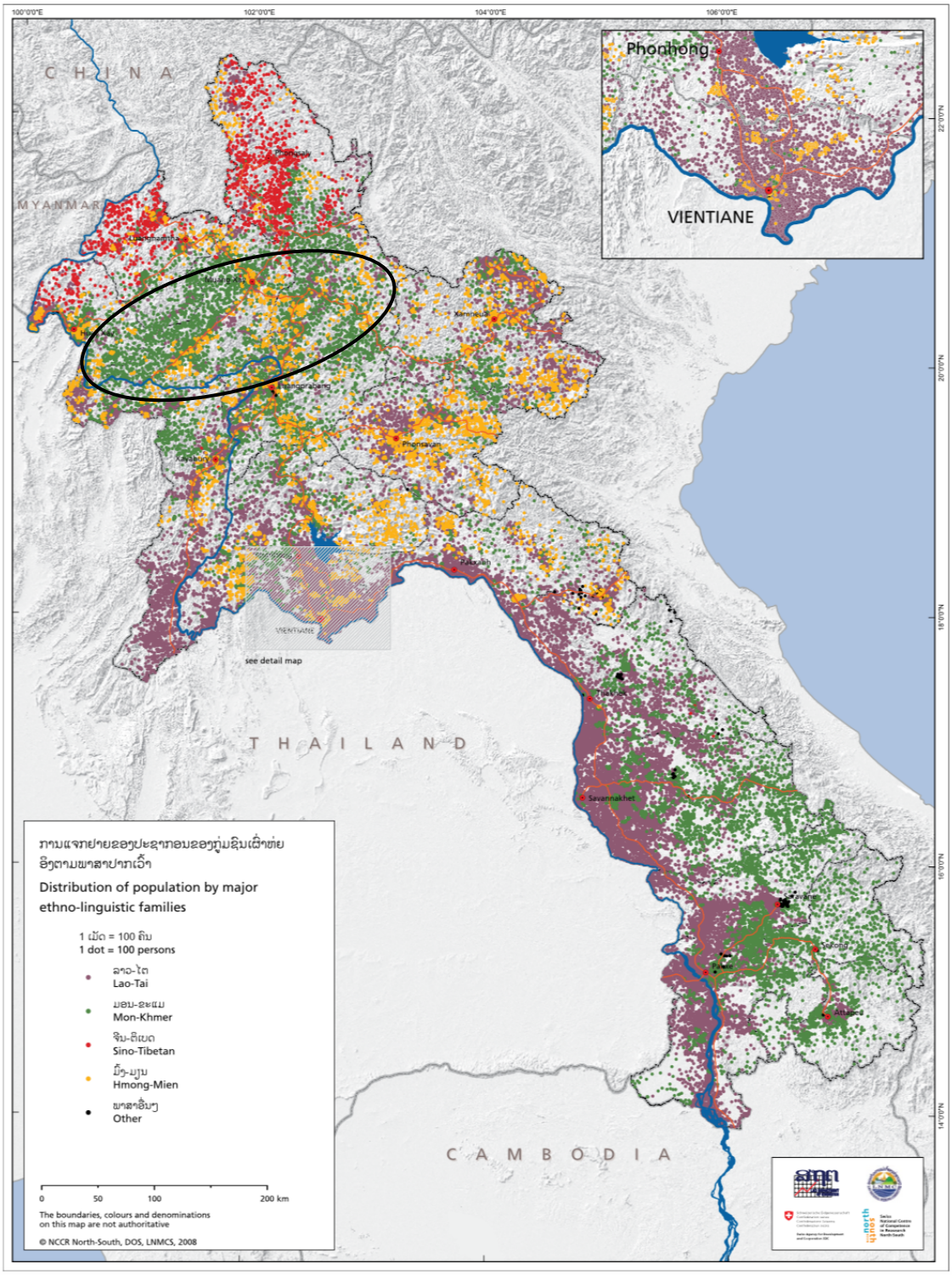
Echoing the findings on disparities in the ECD sub-sector, the ESDP MTR highlights substantial inequalities between districts on the projected cohort completion rate, with the worst-off districts achieving only 43% and the best-off reaching 100% completion (MOES, 2018b). Findings related to learning outcomes in the primary sector – as highlighted in SDG indicator *4.1.1 Proportion of children and young people: (a) in grades 2/3; (b) at the end of primary; and (c) at the end of lower secondary achieving at least a minimum proficiency level in (i) reading and (ii) mathematics, by sex* – are similarly discouraging. For instance, a 2012 EGRA conducted with World Bank technical support found that fully 99 percent of non-native Lao speakers were classified either as non-readers or as poor readers with limited comprehension (World Bank, 2016b).

### *2.7. Ethnicity and Language in Lao PDR*

Given that language and ethnicity are core determinants of children's educational outcomes in Laos, it is critical to conduct an in-depth analysis of these aspects of the country's socio-cultural and educational landscape. Laos is one of Southeast Asia's most ethnically diverse nations, yet intersections between ethnicity and disadvantage abound. Non-Lao ethnic groups are more likely to live in poverty than their Lao counterparts, and socio-economic and educational inequalities have deepened rather than improved in recent years (World Bank, 2015).

Ethnologue (2019) currently lists 73 indigenous languages in the country, and a leading Lao linguist identifies "at least 160 languages spoken in the country by 49 distinct ethno-linguistic groups" (Chamberlain, 2017, p. 2). Each of these languages is assigned by the government into four broad ethno-linguistic families. This includes the Tai-Kadai, typically lowland dwellers along the Mekong plains; the Mon-Khmer or Austro-Asiatic, usually inhabitants of middle altitudes; and the Hmong-Mien and Sino-Tibetan, both of whom tend to reside in the northern highlands and mountains (Schlemmer, 2017). **Figure 1** below provides a map of the distribution of these four ethno-linguistic groups according to the 2005 census. The districts covered in this thesis, demarcated roughly by the black oval, are dominated by the Mon-Khmer (green) and Hmong-Mien (yellow) groups.

## Figure 1: Distribution of the Population by Major Ethno-Linguistic Families




*Source: Messerli et. al., 2008*[5]

[5] This source indicates that the map may be reproduced provided that full acknowledgement is given to the sponsoring agencies: The Swiss National Centre of Competence in Research North-South, the Department of Statistics of the Lao Ministry of Planning and Investment, and the Lao National Mekong Committee Secretariat.

The most recent published census does not provide a map but identifies that 53% of the population is Tai-Kadai (with ethnic Lao only a portion of that figure). 47% come from other ethnicities, and the Khmu (a Mon-Khmer group) and Hmong (a Hmong-Mien group) make up the largest non-Lao ethnicities at 11% and 9% of the country's total population, respectively (Ministry of Planning and Investment, 2015).

The 2003 Constitution establishes Lao as the statutory national language and script (National Assembly, 2003), and the country's education law codifies Lao as the official language and script of instruction (National Assembly, 2007). Although attempts have been made by international development agencies over the past few decades to increase awareness of the benefits of home language instruction and advocate for its use in pilot initiatives in Laos, these ideas have gained limited traction (Kosonen, 2005; Chamberlain, no date). As such, Lao widely retains its role as the language of instruction across all language communities in the nation.

### *2.8. Contrastive Linguistic Analysis of Lao, Khmu, and Hmong Languages*

To shed light on how linguistic features may impact Lao language learning for non-Lao children, this thesis now turns to a contrastive analysis of the language profiles of the three main ethnic groups in the present study: Lao, Khmu, and Hmong. Contrastive analysis facilitates predictions of common challenges and errors that may occur when the native speaker of one language is learning another language, and therefore can guide educators in more effective teaching (Willems, Defrancq, Colleman, and Noël, 2003; Chamberlain, 2017).

Linguistically, Lao belongs to the broader Kra-Dai family of languages spoken across parts of Southern China, northern Vietnam, Myanmar, Thailand, and Laos. Like some of the other languages in Southeast Asia, Lao is a monosyllabic, isolating language: Almost every word is made up of a single morpheme, and the morphemes are typically short syllables of 2-3 letters, except for those borrowed from external languages such as English and Sanskrit. Lao does not use grammatical inflection; instead, grammatical features such as verb tense are indicated through separate

morphemes. Like many monosyllabic languages, Lao uses contrastive lexical tones to increase the number of possible meanings of each morpheme. In the case of Lao, its five tones[6] are contoured: They vary in pitch over the course of a syllable or word and are classified as mid, high rising, low rising, high falling, and low falling (Chamberlain, 2017; Dooley, 2015; Enfield, 2007).

The use of tone in Lao is complex, with a set of rules that are not extensively taught to beginning readers in Lao schools but are largely internalized without realizing it by native speakers of the language. Consonants are first grouped by their inherent tone. From there, the contour of the tone (falling, mid, or rising) depends on vowel length (long or short), the type of final consonant or vowel in the word (stopped or unstopped), and whether there is a modifying tone mark or a class change consonant (Dooley, 2015).

A relatively simple example of how tone markers affect word meaning in Lao is the following three words, which all have the same basic consonant-vowel combination of ປ (‘p’) and າ (‘aa’): The word ປາ has a low rising tone and means ‘fish,’; ປ່າ has a mid-tone indicated by a tone marker, and means forest; and finally, ປ້າ has a high falling tone indicated by a tone marker, and means aunt. An important caveat is that this is a straightforward example where tone is explicitly marked; in most words, tone is not made orthographically explicit. In addition, it is critical to note that this is how tone functions for these words the standard Vientiane pronunciation, and there can be significant regional variations in how tone is pitched and contoured (Dooley, 2015; Chamberlain, 2017; Law, 2014).

Orthographically, Lao evolved over centuries from the Brahmi script of South Asia. It has elements of an alphabetic language, with sounds represented at the phoneme level through both consonants and vowels. It also has characteristics of a syllabary, with the consonant as the base unit of the syllable and vowels grouped around the base, but not necessarily in the same temporal order in which the sound occurs in

[6] There is debate among linguists on whether there are five or six tones, with the majority agreeing on five.

the word. Consonants do not have inherent vowels – vowel sounds are always represented by graphemes – and therefore Lao is considered an alpha-syllabary and not an abugida (Lew, 2014).

Modern Lao has 26[7] consonant graphemes (letters) that represent 19 consonant phonemes (sounds). There are more consonant letters than sounds in the alphabet because some of the same consonant sounds can be represented by two different letters depending on how they appear in a word (Lew, 2014). There are also six consonantal ligatures (combination consonants) that group the letter 'h' with other consonants and affect the tone of the syllable, and there is a placeholder letter that 'holds' the vowel diacritic when a word starts with a vowel sound (Dooley, 2015).

Lao has 38 vowel graphemes that represent 12 vowel phonemes (including 3 diphthongs). The 12 vowels can each have contrastive length (long and short), resulting in 24 different vowel sounds in the language (Dooley, 2015). All consonants can appear as syllable onsets (initial consonants) while only eight can appear as syllable codas (final consonants). Words can start with consonants or vowels, and end in stopped or unstopped consonants or vowel sounds.

Because every grapheme in Lao is associated with a specific sound, Lao is relatively phonologically transparent when it comes to learning the grapho-phonemic rules (Lew, 2014; Dooley, 2015). Tonality, however, is a much more complex matter. The writing system does sometimes indicate tonality through tone markers, as discussed above, but in many cases, tone is not explicitly marked and must be judged using phonological and graphical clues. Incorrect use of tone can radically change the meaning of a spoken word; and incorrect spelling can also considerably alter the tone that is applied. Therefore, with its complex set of rules, "tone expression is a deep or opaque characteristic of the Lao orthography, meaning that readers must use their knowledge of vocabulary and spelling patterns in the language to decode and spell written text" (Dooley, 2015, p. 12).

[7] 27 if you include the letter 'r,' a phoneme that does not occur in Lao but can be used for borrowed words.

When it comes to letter names in the Lao alphabet, consonant names are pronounced as syllables ending in ‘aw,’ and they are almost always taught using a specific set of example words to illustrate the sound of the consonant. For example, the letter ‘k’ is named ‘kaw’ and the example word is ไก่ (‘kaw-kai’ or ‘chicken’). Vowels are typically listed out using an X as a placeholder for the consonant, as in the example in **Figure 2** below, which shows the placeholder with long vowel า (a):

**Figure 2: Example of Written Vowel Representation in the Lao Alphabet**


*Source: Plan International, 2018. Used with permission.*

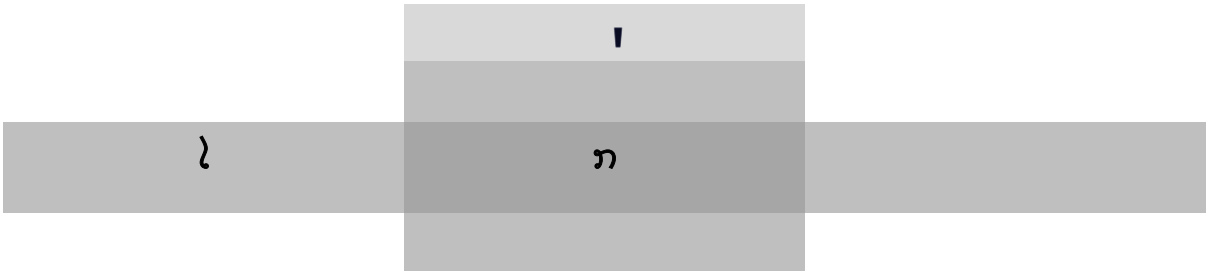
Consonants are written from left to right, while vowel diacritics can appear before, after, above or below the consonant, and tone can be indicated with a modifying mark above (Dooley, 2015). Although the graphically simplest vowels in the alphabet are represented with a single diacritic, as in the word ปา (‘paa’ or ‘fish’), the more graphically complex vowels have two to three diacritics placed around the base consonant, such as in the word ‘tiger’ or เสือ (‘seuah’).

To illustrate how the syllables are formed,

**Figure *3*** below (adapted from Lew, 2014), shows the generic syllable structure of written Lao. **Figure 4** then provides an example of this structure for the word ‘chicken’ or ໄກ່, which includes the base consonant ‘k,’ a prescript ‘long vowel ai’ and the modifying tone mark called ‘mai aek,’ the mid-tone marker.

**Figure 3: Syllable Structure in Lao Script**


**Figure 4: Sample of Syllable Structure for the Word ไก่ ('Chicken') in Lao Script**




As a *scriptio continua* writing system, Lao has no spaces between words, although spaces to mark the end of clauses or sentences are increasingly common, perhaps due to the influence of outside languages (Lew, 2014). There are features that partly compensate for the difficulty that the lack of spaces may cause readers in distinguishing morphemes from each other: For example, most words are predictably monosyllabic and are formed around a single base consonant, and there are often visual cues to delineate words, such as the existence of a prescript vowel in word-initial position to indicate that a new word is starting.

In addition to this in-depth analysis of the structure and functioning of the Lao language, it is useful provide a contrastive analysis with the other two major ethnic languages in the present study population: Khmu and Hmong.

Khmu, the second-most spoken language in the country, belongs to the Mon-Khmer ethnolinguistic family that originated around the Angkor Kingdom in what is modern day Cambodia and spread north through conflict and displacement. Khmu is considered 'less divergent' from Lao than some of the other minority languages in

the country (Chamberlain, 2017, p. 31). The Khmu spoken in the study areas in Oudomxay and Luang Prabang typically does not have tones (Dooley, 2015). However, like Lao, it does have final consonants. Khmu includes more than a dozen consonant sounds that are not found in Lao, while all but one of Khmu's vowel sounds are represented in the Lao phonology, and like Lao, Khmu also contains both long and short vowels (Chamberlain, 2017; Lew, 2014). Khmu is considered sesquisyllabic rather than monosyllabic (as in the case of Lao), with its words typically consisting of a minor syllable followed by a major syllable. Although the Lao government has not given permission for Khmu to be written in its own alphabet, a written form of Khmu does exist in the country using the Lao alphabet (Lew, 2014). However, it is not common in the study areas to find written Khmu, as written Lao predominates.

In summary, children whose native language is Khmu may face challenges when learning Lao due to the lack of tones, the existence of many consonant sounds that do not appear in Lao, and the sesquisyllabic rather than monosyllabic features of Khmu word structure. Conversely, the features that may help Khmu-speaking children learn Lao more easily include the existence of vowel length and final consonants, as well as the geographic proximity to Lao speaking populations that stems from their midland altitude band.

The Hmong language, part of the Hmong-Mien ethnolinguistic family that originated in Southern China, is considered 'very divergent' from Lao (Chamberlain, 2017, p. 31). Hmong is a highly tonal language, with a similar categorization of low/mid/high/falling/rising tones but also a more complex tone system that includes creaky and breathy sounds that do not exist in Lao (Esposito and Khan, 2012). This fact may help Hmong-speaking children when learning Lao in recognizing the tone types that are similar, but it may also cause interference if they attempt to apply Hmong tone rules that do not appear in Lao (Chamberlain, 2017; Lor and Gao, 2020). Hmong also differs from Lao in the very large number of initial consonants that it contains, and the complete lack of final consonants. What is more, Hmong vowels do not have contrasting (short/long) lengths, unlike Lao and Khmu

(Chamberlain, 2017). Morphologically, Hmong is similar to Lao as an analytic, isolating language that does not use inflection and includes predominantly mono-syllabic words (Mortensen, 2017). Hmong has been written in various scripts, but perhaps the most common across national boundaries in Southeast Asia, as well as the Hmong diaspora in the west, is the Romanized Popular Alphabet (RPA) (Ansell et al., 2018).

Hmong speaking children, in summary, may benefit from inherent skill in hearing and using tones but also be hampered by some tonal interference from their native language when learning Lao. They are likely to face significant challenges in mastering Lao spelling, pronunciation, and tone because of the extensive differences in consonant inventory and vowel length, and the lack of final consonants – all of which have significant bearing on hearing tone and mastering Lao spelling rules (Dooley, 2015). In addition, because of the geographically isolated nature of their highland altitude band, Hmong children in general are far less likely than Khmu students to have had exposure to spoken and written Lao language in their everyday life, although they may have some exposure to homegrown reading instruction using Hmong RPA (Chamberlain, 2017). At the same time, aspects of Hmong that may help children learn Lao include the similar sentence syntax, monosyllabic word morphology, and lack of grammatical inflection in both languages.

### *2.9. Section Conclusions*

This literature review has provided an overall theoretical framework for the LEARN study by highlighting the developmental continuum of children's literacy acquisition in the early years and underlining how literacy development is embedded within a broader ecosystem of school, family, and community interactions as well as a web of language factors. The review has problematized received wisdom about emergent literacy and early grade reading interventions and assessment approaches through a comparative linguistic lens, outlining what is known to date about alpha-syllabic, tonal, and *scriptio-continua* languages and writing systems and where the present

study can contribute to deeper understanding of unique linguistic contexts in Southeast Asia and beyond.

The review underscores how alpha-syllabaries such as Lao present more complex orthographic challenges to young readers than purely alphabetic languages due to extensive sets of diacritics and large inventories of consonant-vowel combinations. Learning to decode in these languages may take a few years longer than it typically would in alphabetic languages, and the required orthographic skills must be deliberately taught. In addition, children learning to read in these languages must be specifically instructed in syllabic awareness, to understand how the written language encodes meaning through larger orthographic grain size.

The review also raises questions about how the lack of spaces in a *scriptio continua* language affects early reading acquisition, highlighting the limited literature in this area that nevertheless suggests how adding spaces may facilitate meaning-making for younger learners. Moreover, the review makes it clear that lexical tone is a critical feature of languages like Lao, which must be accounted for in both early reading instruction as well as assessment. Encoding of tone in a written language can be intricately tied to other orthographic and phonological features, such as spelling and vowel length. Perception of tone in one language may facilitate or interfere with tone perception in a second language, and the importance for developing an 'ear' for tones in spoken language should not be underestimated. The presence of tone to convey meaning in a language also has relatively under-studied implications for important aspects of early literacy assessment, including words-per-minute thresholds to measure fluency.

By describing the Lao education and child development context, the literature review has revealed the Lao government's ambitions related to expanding school readiness services and improving the quality of learning in the early primary grades. At the same time, the review has highlighted the persistent challenges the government has faced in achieving those goals, and the need for innovative responses benefitting the most marginalized communities in the country to close equity gaps.

Lastly, the literature review has employed contrastive linguistic analysis to outline how features of the Lao, Khmu, and Hmong languages may affect children's literacy acquisition of Lao as a second language. The review underscores the importance of not taking key differences between the languages for granted in instruction and assessment. The role of geographic proximity and therefore ambient exposure to Lao should also be borne in mind, given the advantage it conveys to Khmu children in contrast to Hmong learners.

Section 3 now turns to a discussion of the study context, including a description of the LEARN summer pre-primary intervention, its theoretical underpinnings, and its fidelity of implementation. The section also covers the broader assessment trends that informed the LEARN evaluation design, and the new contributions to knowledge generated through this thesis.

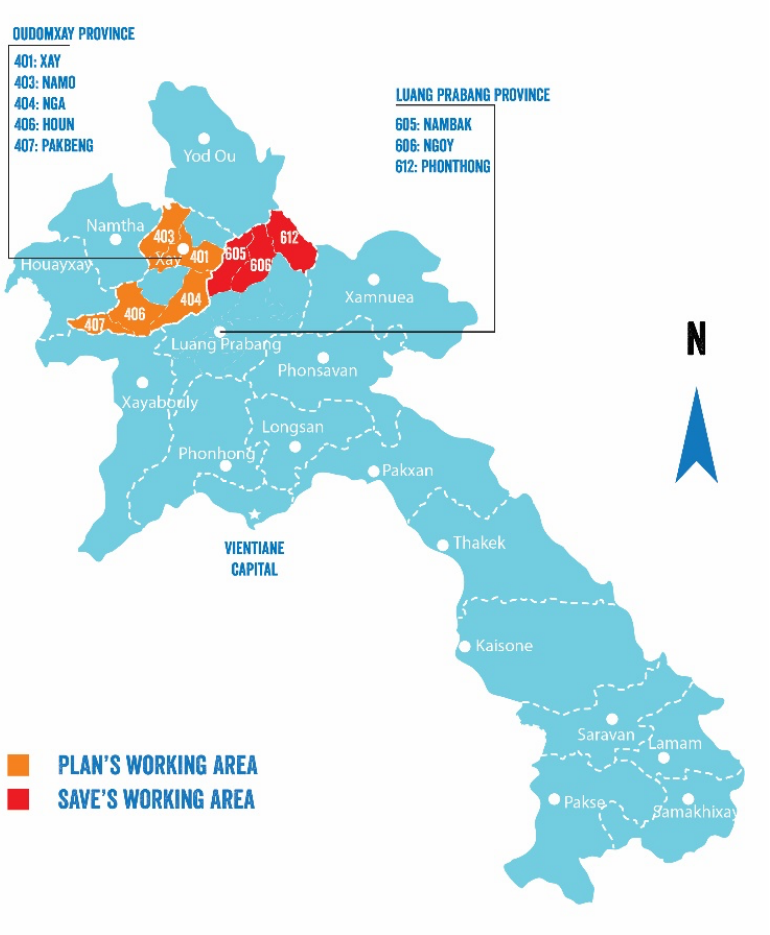
## 3. The Study Context

This research study was embedded within the LEARN Project in Lao PDR. LEARN was a 4¼-year, research-based project implemented by Plan International in Laos and Save the Children International in Laos that began in December 2014 and ran through March 2019. The overall project objective was to improve the learning outcomes and retention rates of boys and girls in grades 1 and 2 of primary school in Oudomxay and Luang Prabang Provinces.

LEARN was implemented across five target districts in Oudomxay and three districts in neighboring Luang Prabang, as indicated in

**Figure *5*** below. The target districts are 450-550 km to the north of the capital city, Vientiane. Most of the schools included in the project intervention and the control groups were located between one and four hours by unimproved road or riverboat from the respective district capitals.

**Figure 5: Map of LEARN Target Districts**




*Source: Plan International in Laos, 2017. Used with permission.*

Target districts were chosen based on a number of factors, including 1) being on the government's list of priority disadvantaged districts due to high poverty levels and poor social indicators, including basic education access and completion rates; 2) having predominantly ethnic (non Lao-Tai) populations; 3) having limited penetration of formal early childhood and school readiness service delivery in rural parts of the district; and 4) benefiting from the existing presence of Plan and Save the Children offices.

LEARN included three intervention models focused on the link between school readiness and the transition and progression through grades 1 and 2. Only the first

model, a summer pre-primary course, is the subject of this study, with the second and third models not covered here.

### *3.1. School Readiness Interventions: Theoretical Underpinnings and Historical Perspectives*

The LEARN summer pre-primary was an accelerated school readiness intervention built on the foundation of decades of research and practice in early education. This evidence base includes landmark programs that improved outcomes for disadvantaged children at large scale, such as Head Start in the United States, which began in 1965 as an eight-week summer pre-school demonstration activity (Office of Child Development, 1970; Office of Head Start, 2021). Head Start has since been joined by a substantial body of similar programs in the United States that have had similar success in promoting school readiness to help break the cycle of poverty and disadvantage (for example, HighScope, 2021; Reynolds and Temple, 1998; Helburn, 1995). Robust evidence from recent decades in the UK also points to the effectiveness of pre-school programs in improving outcomes in that setting, especially for vulnerable children, and particularly when interventions are longer and of high quality (Sammons, 2010a).

Early childhood development and school readiness efforts in the international development context saw a growth in attention in the 1960s as more women entered the workforce in rapidly industrializing, post-colonial nations (Kamerman, 2006). This was followed by an explosion of international attention to ECD in the 1990s that has continued apace to date, in connection with global movements for child rights, Education for All, and nurturing care (World Health Organization, UNICEF, and World Bank, 2018). These efforts culminated with inclusion, for the first time ever in a global education compact, of the SDG target focusing on ECD access and quality (Target 4.2) as well as the two specific indicators (4.2.1 and 4.2.2) to measure progress towards the target.

Since the mid 2000s, the World Bank, UNICEF and other human development-focused organizations have been at the forefront of funding programs and research on ECD and school readiness in LMICs. As of 2016, the World Bank had published at least 28 studies on the feasibility and impacts of such interventions across the globe as part of its focus on building a high-quality knowledge base to drive education reform (World Bank, 2021; Flewitt and Ang, 2020). At the same time, INGOs such as Plan International and Save the Children have strategically prioritized ECD for more than a decade, supporting interventions throughout a wide range of developing nations (Plan International, 2016; Save the Children, 2016).

According to the evidence base generated through these and other programs in the US, the UK, and globally, pre-primary education can provide multiple returns on investment to individual children and families as well as education systems and societies. Children who attend quality pre-primary programs are more likely to enter primary school with stronger literacy, numeracy, and social skills; and are more likely to enroll on time, persist, and succeed in primary school (UNICEF, 2019b). Importantly, these programs help to narrow achievement gaps between disadvantaged children and their better-off peers. Widespread pre-primary participation also contributes to greater educational efficiency by reducing the systemic costs of over/under-age enrollment, dropout, repetition, and remediation of poor learning outcomes in primary school (Ibid.). In addition, longitudinal studies have shown that pre-primary programs can have positive impacts years or even decades later in life in areas such as education, health, and socio-economic status (Bauer and Schanzenbach, 2016; Reynolds and Temple, 1998; Sammons, 2010b; Schweinhart, 2013; Gertler et al., 2014; Shafiq, Devercelli, and Valerio, 2018).

Despite the well-documented benefits of pre-primary education, “services for young children are inadequate and inequitably distributed [and the] burden of children not reaching their developmental potential remains high” across the globe (Black et al., 2017, p. 87). By recent estimates, almost half of pre-primary age children around the world fail to participate in early childhood services, a figure that rises to eight in ten

children in low-income countries – and is likely to be significantly impacted by the coronavirus disease 2019 (COVID-19) pandemic (Muroga et al., 2020).

The push for accelerated approaches to help address these gaps in school readiness service provision for disadvantaged children in LMICs is now more than a decade old. Accelerated education has traditionally been provided in the international humanitarian sector, through age-tailored interventions that help children achieve educational outcomes they missed out on due to poverty, conflict, or natural disaster (Myers and Pinnock, 2017). Run in a compressed time frame, these programs aim to "provide learners with equivalent, certified competencies for basic education using effective teaching and learning approaches that match their level of cognitive maturity" (UNHCR, 2016, p. 2).

Although the global education community generally recognizes that longer-term interventions to promote school readiness would be most desirable, such programs are often out of the reach of national governments struggling with limited resources and large populations of marginalized children. While those governments embark on (or continue) the long-term work of ensuring universal pre-primary access, they may turn to low-cost, accelerated services in the short term "to help better prepare more children now for primary education" (Moscoviz, 2019, para. 5).

The summer pre-primary approach described in this thesis was first developed as this type of stopgap measure for disadvantaged children in Turkey by the Mother Child Education Foundation (AÇEV in its Turkish acronym). The Turkish accelerated model aimed to improve school readiness for disadvantaged 5- 6-year-old children in Turkey who would otherwise not benefit from ECD services in a country where existing programs mainly reached the well-off in urban areas. Aligned with the broad conceptions of school readiness in the literature, the program promoted physical, socio-emotional, and cognitive development through a mix of interactive activities and materials. There was a strong emphasis on the language, literacy, and numeracy skills that the course designers had identified as critical to promoting later

school success in the Turkish primary education system (Bekman, Koç and Taylan, 2011).

Conducted in 2003, the pilot of the program in Turkey provided 200 hours of school readiness services over a 10-week period during the summer break prior to grade 1, along with a parenting intervention in which mothers met once a week for 12 weeks (Bekman, Koç, and Taylan, 2004). The AÇEV summer pre-primary model was later scaled up and tested in additional locations based on the success of the initial pilot (Bekman, Koç, and Taylan, 2011), and was also adapted to an emergency context for Syrian refugee children and children in their host communities in Turkey (Sezen, no date).

Collaborating with AÇEV, Plan International adapted the Turkish summer pre-school model and implemented the course on a 'proof of concept' basis in five villages of Bokeo Province in northern Laos in 2015. This initial adaptation process focused on translating course materials – including a step-by-step guide for teachers, cognitive training workbooks for students that emphasized numeracy concepts and pre-writing, and a daily routine calendar – from Turkish to English and then to Lao. The Plan International team also replaced some of the original course content, such as story books, songs, rhymes, and vocabulary words, with Lao-specific content.

The Bokeo course was well accepted by key stakeholders in the Lao government, who viewed it as an appropriate school readiness approach for hard-to-reach, non-Lao communities. In addition, using a quasi-experimental, pre-post evaluation design with a curriculum-linked assessment tool (refer to **Section 6.2.1** for more details), the pilot found statistically significant improvements in average scores for the treated group compared to the control group in both the literacy and numeracy domains immediately after implementation concluded (Bekman and Diri, 2016). At the same time, feedback from participating government officials and teachers underscored that further corrections and refinements to the course materials and approach would be useful to ensure that it fully aligned with the Lao language and context.

Based on these positive findings, the LEARN Project conducted further adaptation, scale-up, and evaluation of the model. LEARN also specifically set out to increase the robustness of the evaluation design for the model in the Lao context by utilizing experimental methods and globally recognized assessment approaches. This process included a pilot phase in the summer of 2016, an evaluation phase during the summer of 2017 (the subject of this study), and a scale up and government handover phase in summer 2018.

LEARN's main adaptations before the 2017 evaluation phase focused on further strengthening the contextual appropriateness of the course and preparing the model for wider sustainability. For instance, the LEARN team conducted additional text corrections and adjustments to the illustrations in the teacher guide and the student cognitive training workbooks; created new classroom visuals more appropriate to the Lao context; and revised the list of classroom materials to include more locally produced items such as wooden blocks and toys. Lao government counterparts were also prepared as trainers so that they could take ownership of the model and train a larger number of summer pre-primary teachers as the course was expanded to additional locations. Further details about the specifics of the model that was evaluated in the LEARN Project – including the classroom setup, course duration, and the daily routine – are provided in **Section 3.2** below.

Over the past decade, the intervention in Turkey and Laos has been joined by other similar accelerated school readiness programs to address service gaps for disadvantaged children. Together, these efforts have contributed to a growing evidence base for similar interventions – empirical findings that are covered in depth in **Section 6.2.1** to allow for comparisons of outcomes between LEARN and the other programs.

For example, in 2015, Tanzania began piloting a 12-week, accelerated school readiness course with a similar focus on children without formal access to pre-primary. This UK-supported model – run prior to the start of grade 1 by trained

volunteers in community-based centers – was then expanded to 16 weeks and across a larger number of sites from 2016-2018 (EQUIP-Tanzania, no date).

A similar UNICEF-funded activity started in 2015 in Ethiopia provided some of the original inspiration for the implementation and research design of the LEARN Project. The Ethiopia Accelerated School Readiness (ASR) program was intended to fill a gap in service provision for disadvantaged children who are otherwise difficult to reach with the government's existing "0" class or other ECD interventions due to geographic and cost barriers (Spier, 2019). Operating on a pilot basis in Benishangul-Gumuz Region, the program provided eight weeks (roughly 150 hours) of intervention to six-year-old children through a summer pre-primary, delivered by the existing 0 class teacher for children who have a 0 class in their village but may not have attended it. The model later scaled up to additional locations in 2017 (UNICEF, 2019a).

UNICEF and Save the Children partnered to conduct a similar ASR pilot in Zambézia Province, Mozambique, with the goal of improving school readiness for more than 11,000 children ages 5-6 (Bonilla et. al., 2018). Implemented from 2016-2020, this project provided a 120-hour summer course delivered by trained community volunteers plus 12 weeks of parent-to-parent education sessions to help them better support their children's school readiness.

Further replication of these accelerated school readiness approaches is also occurring in India and Côte d'Ivoire. In 2018, the Central Square Foundation tested a 40-day program at the start of grade 1 to promote school readiness for children without prior ECD access in Karnataka and Gujarat (Serikari, 2019). In the same vein, Cote d'Ivoire aims to test an eight-week summer pre-primary program in 2021, targeting five- and six-year-olds who have not participated in any early childhood development interventions (Education Partnerships Group, 2019).

Finally, the World Bank supported the Government of Lao PDR to conduct a multi-year ECD and school readiness project ('ECE Project') that began in 2016. The ECE

Project was not testing accelerated models – rather, it promoted community child development groups (CCDGs) for three to four-year-olds, multi-age teaching (MAT) for three to five-year-olds in the 9-month pre-primary class, and ECE community awareness campaigns, among others. However, LEARN's activities and assessment approach were co-designed with the World Bank in Laos to complement rather than duplicate the approaches being piloted in the ECE Project, particularly to offer a low-cost, accelerated school readiness model best suited for the most geographically marginalized areas in the country.

### *3.2. Summer Pre-Primary Implementation Details*

Building on the experiences with accelerated interventions in Turkey, Bokeo, Ethiopia, and elsewhere described above, the LEARN summer pre-primary was implemented in 46 target villages in both Oudomxay by Plan and Luang Prabang by Save the Children. The intervention targeted children approximately age six years old immediately before they started grade 1 in September, where the minimum age of enrollment is six years and six months. Average class size in the summer pre-primary was roughly 15 students, in line with the low population density typical of the remote target areas. Cumulatively, the summer pre-primary model reached 2,171 children (49% female) from 2016-2018.

The course was delivered for 10 weeks from mid-June through mid-August and included 4¼ hours per day or about 213 instructional hours in total over the 10 weeks of the course. The timetable ran in the morning hours, from approximately 8 am to 12:15 pm, varying slightly based on local conditions.

The LEARN summer pre-primary model retained much of the content and methods of the original program in Turkey as it had been adapted for the pilot in Bokeo, as described in **Section 3.1.** However, the Bokeo program (and subsequently the LEARN summer pre-primary) omitted a supplemental language module that had been included in the Turkish program to build emergent literacy skills in Turkish as a second language (Bekman and Diri, 2016). This occurred because the language

module could not simply be translated from Turkish to Lao. Instead, it would have required a significant amount of time (which was not available to the team at that stage) and linguistic expertise to adapt the module to the unique features of Lao language and the needs of local non-Lao learners. As noted in **Table 1,** the adapted approach in Bokeo still covered core language skills but with less targeted emphasis on learning Lao as a second language due to the lack of the supplemental module.

**Table 1: Language and Literacy Skills Covered in the LEARN Summer Pre-Primary Course**

| |
|---|
| Perception of sounds (2 skill areas)[8] |
| Grammar (6 skill areas) |
| Voice control (6 skill areas) |
| Oral expression (11 skill areas) |
| Oral vocabulary (5 skill areas) |
| Concepts about print and story sense (5 skill areas) |

The summer pre-primary classroom environment was organized into six learning corners to help children develop and extend knowledge and skills across developmental domains: Blocks, house, art, educational games, books, and music. Specific materials were stocked in each corner, including objects such as wooden blocks, dress-up clothes, puzzles, books, and musical instruments. Nutritious snacks, clean drinking water, handwashing, and teeth brushing facilities were also provided in close collaboration with parents and community members.

The course was delivered by the existing grade 1 teacher in the target village plus a pre-primary teacher from a neighboring village who was seconded for the summer, with the condition that at least one of the teachers should speak the same language as the enrolled children. An 11-day teacher training was delivered prior to the start of the course by District Education and Sports Bureau (DESB), Provincial Education

[8] "Perception of sounds" was the term utilized by the original Turkish course developers. This skill area focused mainly on helping children recognize contrasting sounds, such as loud/soft, to assist in their ability to follow oral instructions, participate in activities, and grasp meaning during storytelling.

and Sports Service (PESS), and project staff who were previously trained as trainers during a 6-day training of trainers (ToT). The same staff conducted monitoring and coaching throughout course implementation. One reflection meeting and one follow-up refresher training per year were also provided to teachers.

Classroom activities were complemented by parenting education activities delivered to the parents of children in the course by local non-Lao interns, designed to improve parents' engagement in their children's development. These interns were first trained in Plan's parenting approach, which included 11 sessions of approximately 1.5 hours each on diverse child development, protection, safety, and gender equality topics. They then conducted parenting sessions every few weeks while the summer pre-primary course was running and then approximately once a month afterwards during the remainder of the school year.

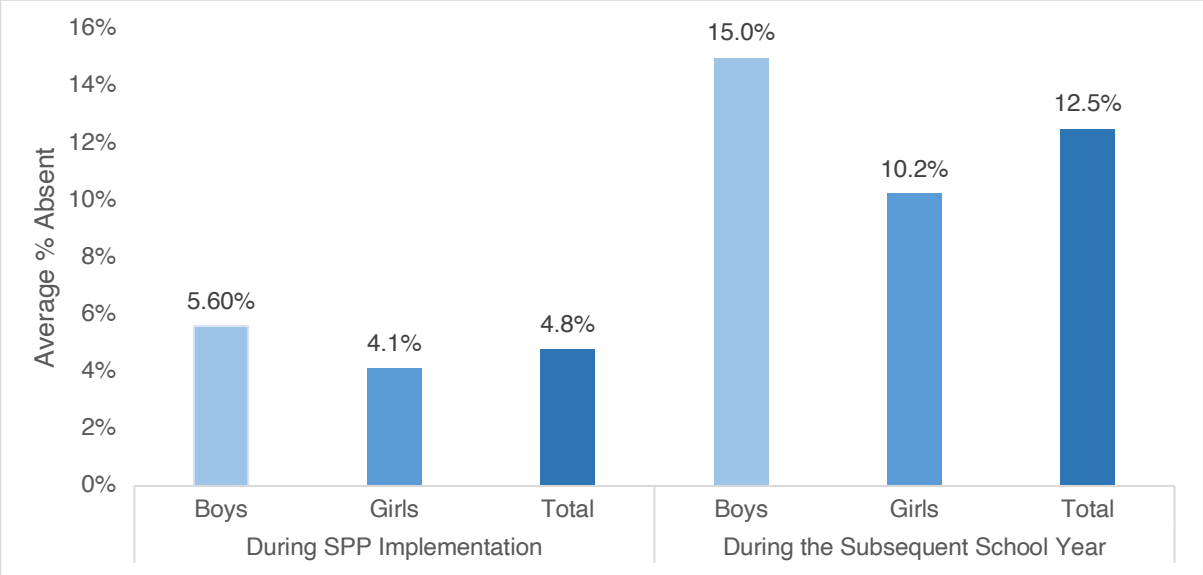
### *3.3. Intervention Dose and Fidelity of Implementation*

The intended dose of the summer pre-primary course was approximately 213 hours of instruction, compared to roughly 1,170 hours of instruction that children would receive through the government's existing 9-month pre-primary program and 3,510 hours through the 3-school-year kindergarten program. To measure fidelity of implementation, DESB and project staff checked attendance through pre-announced spot checks[9] at least three times throughout the 10 weeks of the summer pre-primary course, and then at least three times in grade 1 during the school year after course implementation.

As shown in **Figure 6**, these spot-checks identified that children in the summer pre-primary attended approximately 95% of the intended dose, or around 202 hours. During the subsequent 2017-2018 school year, absentee rates were consistent at around 12% for children who had participated in both courses, and higher for boys (Yang, 2018).

[9] Because permission is required from the Lao government before visiting schools, the spot checks could not be unannounced.

**Figure 6: Absentee Rates for Summer Pre-Primary Participants, Summer 2017 vs. Subsequent School Year, by Sex**




The project monitoring system indicated that the course was implemented with good conformity to the original design: In general, it began and ended on time and covered all the intended content. Classroom observations to assess teacher application of the skills taught through the teacher training found average scores of 77% (Yang, 2018). Implementation research also identified some maintenance of positive teaching practices during the subsequent school year, although teachers reported facing difficulties in sustaining the use of child-centered teaching and learning practices and course materials, because they felt constrained by the existing grade 1 curriculum (St. George and Khoonbarthao, 2018).

### *3.4. Situating This Study within Broader Assessment Trends and Paradigms*

Methodologically, this study is an example of real-world research conducted in a multifaceted context that requires trade-offs between rigor and practicality. As a programmatic evaluation, the study is situated within a growing trend over the past 15 years towards evidence-based decision-making in the international development sector, which has been “marked by a shift in focus from inputs to outcomes and

results" (Gertler et. al., 2011, p. 3). The study is also embedded within a wider global movement that began in the 1990s towards prioritizing and measuring the quality of education and children's learning in addition to the educational access objectives that were previously captured in the Millennium Development Goals. And finally, the research includes elements of an exploratory study in that it attempts to use the data generated by the broader programmatic evaluation to investigate how the unique features of the Lao language affect measurement of child-level outcomes and can be more effectively addressed in global assessment approaches.

The study works within established research paradigms by employing a combination of two globally developed assessment approaches: MODEL and EGRA. MODEL was developed through the MELQO initiative, a collaboration of United Nations Educational, Scientific, and Cultural Organization (UNESCO), the World Bank, the Center for Universal Education at the Brookings Institution, and UNICEF. Begun in 2014, this initiative aimed to "promote feasible, accurate and useful measurement of children's development and learning at the start of primary school, and of the quality of their pre-primary learning environments" (UNESCO, UNICEF, Brookings Institution, and the World Bank, 2017, p. 7).

EGRA, promoted by organizations such as the United States Agency for International Development (USAID) and the World Bank, has a longer but similar history. Emerging in 2006 as part of the response to the global crisis in educational quality and learning outcomes in the Education for All era, EGRA was created for USAID and the World Bank by Research Triangle Institute International (RTI) (Gove and Wettenberg, 2011). It was the product of an extensive review of existing reading assessment approaches and adapted largely from the Dynamic Indicators of Basic Early Literacy Skills (DIBELS) assessment in the United States (Gove and Wettenberg, 2011; Bartlett, Dowd, and Jonason, 2015). The impetus of this effort was to provide a rapid, reliable oral reading assessment that could feasibly be administered across a range of contexts and provide actionable data to policymakers and practitioners (Gove and Wettenberg, 2011).

According to a taxonomy described by Munoz-Chereau et al. in 2021, both MODEL and EGRA fall roughly in the middle of the spectrum of assessment approaches that have been utilized in LMICs. That is, the two assessments were not lifted directly from standard Western tools with no adaptations, but neither were they fully locally developed and culturally specific. Instead, both approaches represent "an amalgamation of a number of translated and/or adapted items from several different western [child development assessment tools]" (Ibid., p. 4). As will be discussed in **Section 6**, this may have contributed to some limitations in the ability of the two tools to adequately measure school readiness and early grade learning in the Lao context.

### *3.5. New Contributions to Knowledge*

This thesis makes vital original contributions to knowledge from policy, practice, and research and assessment perspectives and from the global to the local levels. At a global level, the study is rooted in the wider movement in SDG 4 towards equitable, inclusive education for the most marginalized children. Through its evaluation of a low-cost, accelerated school readiness model for children who would otherwise fall outside the formal education system prior to grade 1, the study joins an emerging body of evidence around how governments can increase equity at the so-called 'bottom of the pyramid' for children who face a double disadvantage as the most marginalized citizens in LMICs that are already below the global curve (Wagner, Wolf, and Boruch, 2018).

As these national governments struggle to raise overall average learning outcomes, they face two broad options: Focusing on "those at the middle of the distribution, or even at the upper end … [so] that the 'right hand' of the distribution can 'pull over' the whole distribution;" or pushing the distribution from left to right by investing more in those at the lower end of the achievement spectrum (Crouch, 2019, p. 61). Indeed, recent analytical work on learning poverty indicates that education systems have a better chance of raising overall average performance by reducing the proportion of children in the lowest learning categories (Crouch, Rolleston, and Gustafsson, 2020).

Nevertheless, findings of prior large-scale interventions in these contexts suggest that “while the poor and minorities have better outcomes as a result of the interventions, persistent interaction effects between treatment and poverty mean that the nonpoor actually benefit more from programmes designed to help the poor” (Piper, 2018a, p. 25). The LEARN study evaluates a deliberate attempt to target interventions to those at the bottom of the pyramid and provide the government with a viable option, in a resource-constrained fiscal space, to understand what works in ensuring that the poor benefit, and to push national average learning outcomes to the right by investing in children at the left of the distribution.

Within the Lao and regional context, the study represents one of the first-ever large-scale, rigorous efforts to measure the effectiveness of an accelerated school readiness intervention. It contributes actionable information to government policymakers and planners around interventions that are effective and feasible in the most difficult-to-reach and disadvantaged areas of the country, as they roll out their next Education Sector Development Plan as well as the related costed action plans that accompany the ESDP; conclude the process of revising national pre-primary curriculum and materials; and make decisions about professional development and allocation of teachers in the school readiness subsector.

The fine-grained findings around areas of strength and weakness in the accelerated intervention, and how it has or has not succeeded in closing the gap for specific ethnic groups, will allow the government to refine the model to produce the most impact. The study also contributes comparative evidence to the small but increasing pool of evaluations of similar accelerated school readiness interventions from Asia to Africa, which have demonstrated feasibility and impact for highly marginalized children. This is particularly timely, because, in 2020, Plan International and its partner organizations AÇEV and American Institutes for Research (AIR) received one of only seven grants worldwide through the Global Partnership for Education (GPE) Knowledge Innovation Exchange (KIX), which will be used to scale up the

summer pre-primary model and test its system-level sustainability in Cambodia, Lao PDR, and Tanzania (Global Partnership for Education, 2021).

Finally, the study takes its place among a rapidly expanding set of assessments that are part of a global movement towards more effective measurement of ECD and school readiness using common evaluation approaches. At the same time, the study problematizes those international paradigms using a comparative linguistic lens.

As of 2019, similar assessments had been conducted in over 30 LMICs across Central and South America, Africa, and Asia. In Laos, the assessment approach has been used at least three times (by the World Bank, the LEARN Project, and UNICEF), and may be used for the KIX innovation described above, as well as others in the future. Based on global experience to date, the original proponents of the assessment approach have acknowledged that the assessment tools may not have responded adequately to local idiosyncrasies and that further work is needed to help national governments make adaptations (Raikes and Sayre, 2019).

This thesis records the experiences, challenges, and lessons learned through adapting and using these assessment approaches in the Lao setting so that others can anticipate and address those challenges proactively in the future. Importantly, the thesis interrogates the assessment approach from a Southeast Asian linguistic perspective for the first time ever, contextually situating it in the highly multilingual environment of Northern Laos and providing recommendations for improved future assessment practice for alpha-syllabic, tonal, and *scriptio continua* languages. As such, the thesis breaks new ground through an in-depth review of critical aspects of assessment validity in this language and country context, including elements such as ecological, content, predictive, cross-language, and systemic validity.

## 4. Research Design and Methodology

### *4.1. Research Questions*

This study was guided by two main research questions around the effectiveness of the summer pre-primary in comparison to no intervention, considering children's linguistic backgrounds as an important determinant of their learning trajectories in the Lao context.

***Research Question 1:*** *What is the effectiveness of the LEARN Project summer pre-primary model in improving children's Lao language school readiness at the start of grade 1 and sustaining their gains through the end of grade 1, in comparison to a control group?*

***Research Question 2:*** *How does the effectiveness of the model vary according to children's ethnicity?*

The study hypothesis is that the model will be more effective than no intervention in the control group, and that children from non-Lao ethnic groups benefit at least as much as Lao children from the intervention.

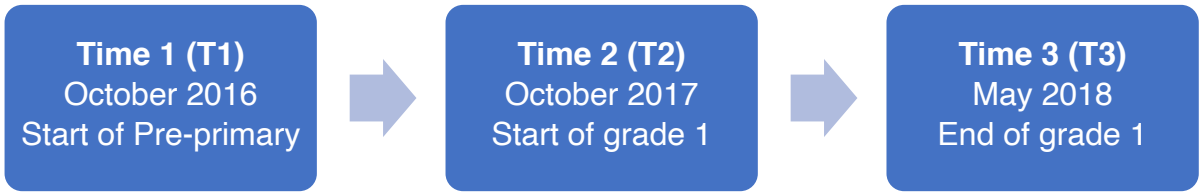
In addition, the study is concerned with extending the state of knowledge and the state of the art around how to conduct this type of learning assessment for an alpha-syllabic, tonal, and *scriptio continua* language like Lao, where there is currently a dearth of documented experiences. The study also aims to generate recommendations to inform decision-making on use of language in these kinds of assessments when children come from diverse, multi-ethnic backgrounds. As such, the study has the following sub-research question:

***Research Question 3:*** *How do the unique features of the Lao language affect measurement of child-level outcomes and how can those features be more effectively addressed in global assessment approaches?*

### *4.2. Research Methods*

This study employs a longitudinal, cluster RCT design to estimate the impacts of the summer pre-primary on key school readiness and early grade reading outcomes. As described in **Figure 7**, this was a repeated measures study, with direct assessment of children at three points in time – baseline (Time 1), mid-term (Time 2), and endline (Time 3) – using largely the same instruments on the same panel of children at each wave.

**Figure 7: Three Waves of Data Collection**




The baseline captured children's status at around age five, prior to their participation in any school readiness intervention. Following baseline data collection in October 2016, the summer pre-primary group was then offered the project's school readiness interventions for 10 weeks during the following summer of 2017.

Subsequently, the midline was conducted in October 2017 when children were approximately six years old and had completed the first month of first grade in the 2017-2018 school year. Endline data were collected at the end of grade 1 of the same school year, in May 2018, when children were around six to seven years old.

The control group received no extra interventions from the project or the government apart from the status quo services they would have eventually received anyway in grade 1.

### *4.3. School and Subject Selection*

#### 4.3.1. Sampling Approach

The sampling frame was developed prior to the study's start, in mid-2016, using administrative data on locations, numbers and types of schools from the Lao Education Management Information System (EMIS) system as an entry point, followed by detailed mapping by project and government staff to fill in EMIS's gaps. Villages were considered eligible to participate in the treatment arms of the study if they met the following conditions:

- Were primarily non-Lao-Tai ethnicity
- Were expected to have at least six children between 4.5 and 6.5 years at baseline
- Did not have other ECD services
- Did not have other development partner investments in areas that would potentially influence the outcomes of interest in the study

The sampling exercise identified an overall sample of 64 eligible schools across six project target districts: Nga and Xay Districts in Oudomxay Province; and Nambak and Ngoi Districts in Luang Prabang.

After establishing this overall frame, randomized assignment of schools to the treatment and control groups was then conducted by LEARN's external research firm, AIR. The sample was stratified by ethnicity to ensure a balance in ethnic groups between the treatment and control groups. Randomization was conducted as close as possible to the start of the actual intervention, in February 2017, with stratification resulting in an acceptable balance in the distribution of ethnicities across groups. 31 schools were allocated to the treatment group (the maximum that was deemed implementable within project staffing structures and financial resources) and 33 schools to the control group.

#### 4.3.2. Sample Power Analysis

Power analysis was conducted by AIR prior to the baseline using the features of the original, randomized sample of 64 total schools. This analysis identified a minimum detectable effect size for child-level outcomes of .244, as outlined in **Table 2**[10] below.

**Table 2: Minimum Detectable Effect Sizes**

| | |
|---|---|
| Alpha level (α) | .05 |
| Power (1–β) | .80 |
| Rho (intraclass correlation) | .15 |
| P (proportion of Level 2 units randomized to treatment) | .48 |
| $R_1^2$ (proportion of variance in Level 1 outcome explained by Level 1 covariates) | .50 |
| $R_2^2$ (proportion of variance in Level 2 outcome explained by Level 2 covariates) | .50 |
| g (number of Level 2 covariates) | 3 |
| n (average cluster size) | 10 |
| J (sample size [# of clusters]) | 64 |
| **Minimum detectable effect size** | **.244** |

#### 4.3.3. Adherence to Sample Randomization

Despite the best efforts to ensure an accurate sampling frame, randomization was compromised in five villages (8% of the intended sample) due to three reasons:

- Prior to baseline data collection, three villages originally assigned to the summer pre-primary treatment group (all in Ngoi district, Luang Prabang) were found to have an existing government pre-primary and thus were removed from the summer pre-primary study.
- One village in Nambak District, Luang Prabang, originally assigned for summer-primary treatment, was dropped by the survey firm, Indochina

[10] Adapted from AIR, 2017.

Research, during baseline data collection on the request of government staff when it was found to already have a government pre-primary class during baseline data collection. It was replaced on the spot with a nearby village – not part of the randomization process – that did not have a government pre-primary.

- One village in Nga District, Oudomxay Province was originally intended to be included in the summer pre-primary treatment group, but the program was not implemented there due to a mix-up with the village name. Implementation took place in a nearby village with an almost identical name instead – also not included in the randomization process – and thus the original village was dropped.

All five of these instances of imperfect randomization were caused by real-world constraints related to incomplete or out-of-date EMIS data or difficulties in verifying correct names and EMIS codes for each school. These challenges are not uncommon in experimental studies conducted in "naturalistic settings where variables are more difficult to control than in laboratory conditions" (Flewitt and Ang, 2020, p. 181) – a constraint that will be further unpacked in **Section 6.**

**Table 3** below captures original randomized assignments versus actual uptake of schools by district. The final sample contained 27 villages in the treatment group from an original total of 31 and remained with 33 in the control. As such, the total of 60 schools in the analytical sample nearly aligned with the original sample of 64 schools needed to achieve the minimum detectable effect size of .244 described above.

**Table 3: Original and Actual Village Assignments, by District**

| | Original Assignment | | Actual Uptake | | Net Change | |
|---|---|---|---|---|---|---|
| **District** | **Treatment** | **Control** | **Treatment** | **Control** | **Treatment** | **Control** |
| Nambak | 7 | 7 | 7 | 7 | 0 | 0 |
| Ngoi | 8 | 8 | **5** | **8** | -3 | 0 |

| | | | | | | |
|---|---|---|---|---|---|---|
| Nga | 7 | 8 | **6** | **8** | -1 | 0 |
| Xay | 9 | 10 | 9 | 10 | 0 | 0 |
| **Sub-total** | **31** | **33** | **27** | **33** | **-4** | **0** |
| **Grand total** | **64** | | **60** | | **-4** | |

#### 4.3.4. Sampling of Children within Schools

Baseline sampling of children within the target schools was conducted by the survey firm using a systematic protocol. First, they explained the purpose of the study to village and school leadership and conducted a census with those stakeholders of all children in the village who were in the stipulated age range of 4.5-6.5 years. In cases where there were more than 20 eligible children in the village, the survey firm then randomly selected 20 from among the eligible list, stratifying by sex, with the aim of including an equal number of girls and boys. In villages with fewer than 20 eligible children, all were included in the sample.

In addition, the survey firm conducted follow up at each wave to find children in the original census-derived sample who were not present in the village at the time of data collection, to help maximize sample size and reduce attrition. For instance, if the child was with his/her parents at family agricultural plots within a few kilometers of the village, the survey team made reasonable efforts to locate and assess the child *in situ* or wait for the family's return to the village.

As each child took part in assessment, the survey firm assigned a unique ID that was used to link all data from the data collection tools in the study, including the parent/caregiver questionnaire and direct child assessment tool. At baseline, a number of children who were listed in the sample were nevertheless not present for data collection due to factors such as agricultural demands on their families. However, at mid-term and endline, the survey firm were instructed to collect data on

all children from the original sample who were present in the village at the time of the mid-term and/or endline visits, regardless of whether those children had been present during previous visits. This resulted in some children being added during at Times 2 and 3 and assigned identifications (IDs) at that time.

### *4.4. Sample Characteristics at Baseline*

Out of the total sample of 673 children in 64 villages in the original sample, there was an average of 10.5 children per village at baseline, with only two children in the smallest village and 30 in the largest. This original sample consisted of 328 male (48.7%) and 342 female (50.8%) children, with 3 children (.4%) missing data on sex. The average age at baseline was 62.7 months (Standard Deviation (SD)=6.01) or nearly 5 years and 3 months, with a range from 48 months, or around 4 years old; up to 79 months, or nearly 6 years and 7 months old. The sample contained children from three main ethnic groups, including 60 Lao (8.9%), 514 Khmu (76.4%), and 98 Hmong (14.5%) children, with one child indicating an 'other' ethnicity.

**Table 4** below breaks down the sample of children according to their original assignment and actual uptake. Out of the original sample of 673 children, 307 were randomized to the control and 366 to the treatment groups. However, verification of uptake resulted in 490 total children, with 243 in the control and 247 in the treatment. As indicated in the table, 106 children who were originally assigned to the treatment group did not participate, while 13 children had no data or contradictory data about their actual treatment status recorded at different treatment waves, making a total of 119 children who were supposed to be part of the original treatment sample but did not remain in that arm and were dropped from this study. Approximately 59 (50%) of those children were missing from the original treatment sample because of the compromised randomization, in which four villages were dropped, as described above. The program team hypothesized that other children who failed to take up the intervention may have done so in part because they were not the correct age to start grade 1 that September, which was a strict criterion for participating in the summer

pre-primary – indicating their concern that the age range for enrolling children in the study may have been set too wide.

In addition, 48 children who were originally assigned to the control group participated in a nearby 9-month pre-primary, while 16 children had no data or contradictory data between different data collection waves about their actual control status. All were also dropped from this study.

**Table 4: Original Child-Level Assignment and Actual Uptake**

| | **Original Assignment** | | **Actual Uptake** | |
|---|---|---|---|---|
| | **N** | **%** | **N** | **%** |
| Summer pre-primary treatment | 366 | 54.4% | 247 | 36.7% |
| Summer pre-primary control | 307 | 45.6% | 243 | 36.1% |
| ***Total Retained in Sample*** | ***673*** | ***100.0%*** | ***490*** | ***72.8%*** |
| Summer pre-primary treatment – did not participate | | | 106 | 15.8% |
| Summer pre-primary control – participated in 9-month pre-primary | | | 48 | 7.1% |
| Summer pre-primary treatment – no data or contradictory data | | | 13 | 1.9% |
| Summer pre-primary control – no data or contradictory data | | | 16 | 2.4% |
| ***Total Dropped from Sample*** | | | ***183*** | ***27.2%*** |
| **Total** | **673** | **100.0%** | **673** | **100.0%** |

In total, 183 children or 27.2% of the original sample, were removed, resulting in a final set of 490 ‘compliers’ to their original assignment. Out of this sample of 490, 50.4% were male and 49% female, with 3 children (.6%) missing data on sex. The average age of children was 62.6 months (SD=6.2) or just over five years and two months, ranging from a minimum of 31 months (about 2 years and 7 months) to a

maximum of 79 months (about 6 years and 7 months). This sample included 45 Lao children (9.2%), 371 Khmu (75.7%), 73 Hmong (14.9%), and 1 'other' ethnicity (.2%).

### *4.5. Instrumentation*

This study used a combination of instrumentation developed through collaboration with the World Bank in Laos to jointly measure ECD impact in the country, as well as other tools that were purpose-built to measure specific aspects of the project, and that had been designed and tested for other similar projects and populations in Laos. The instrumentation used during different waves of data collection included child direct assessment and anthropometrics, a parent/caregiver questionnaire, a school administrative profile, school and classroom observation tools, and a costing data tool. This thesis focuses mainly on analysis of the data from the child direct assessment.

The direct assessment protocol used in this study was adapted from the MODEL tools first developed through MOES's World-Bank supported ECD diagnostic assessment. MODEL itself is an amalgamation of other ECD and school readiness assessment constructs and tasks that have demonstrated feasibility, validity, and reliability in other contexts, and low-resource settings in particular. It draws heavily from Save the Children's International Development and Early Learning Assessment (IDELA) instrument and includes tasks from the Wechsler Memory Scales, the Early Grades Mathematics Assessment and other early math tools, and EGRA, among others (ECD Measure, no date). The IDELA instrument that forms the foundation of a significant portion of MODEL has itself also been used in at least 55 countries across Asia, Africa, Europe, and North and South America (IDELA Network, 2019).

Led by the MELQO Initiative, the MODEL tools were designed through a process of expert consultation to define the critical domains and constructs that the tool should measure; extensive review of existing assessment approaches; and initial testing of draft items in low-resource settings (UNESCO, UNICEF, Brookings Institution, and World Bank, 2017). Once an instrument was arrived at through this process, a four-

country pilot in Lao PDR, Madagascar, Mongolia, and Tanzania was conducted to assess aspects of validity and reliability in real-world settings. Specifically, the pilot assessed internal reliability, concurrent validity, and cross-country validity, "the extent to which MODEL scores demonstrated expected factor structures within and between countries through measures of measurement invariance" (Raikes, no date). The Laos pilot included a sample of 200 children – 49.5% female and with an average age of 4.3 and a range of 2 – 7 years old – and utilized both the direct child assessment and the parent/caregiver report. A range of ethnicities was included (however, no percentage breakdowns are provided) and the assessments were conducted in children's home languages (Ibid.).

To assess internal reliability, the MELQO research team used Item Response Theory. This identified items that were consistently easy (some of the mathematical vocabulary items) or difficult (the phonological awareness items and the backward digit span) across countries and enabled the team to make recommendations about items to leave in or remove from the assessment based on the expected age of the children tested, and to remove items that were considered redundant or did not usefully contribute to the overall score (Ibid.).

To investigate concurrent validity, the MELQO team looked at correlations between outcomes in different domains and between domain scores and background demographic and family characteristics. This analysis identified that the cognitive domains of math and literacy were strongly correlated with each other, but less correlated with the socio-emotional domains. In addition, children's outcomes or the cognitive domains were more strongly correlated with age than the socio-emotional domains.

Confirmatory and exploratory factor analysis were also conducted as part of the four-country pilot to see if items in the direct assessment and teacher/caregiver report confirmed expected factor structures related to child development and school readiness. This analysis found no differentiation between factors indexing mathematics, literacy, and executive function. Instead, "results indicated a one-factor

solution, or a structure in which all items loaded onto one factor of 'school readiness'" (Ibid., p. 3).

And finally, to investigate cross-country validity, a test of measurement invariance was conducted, which indicated some similarity between countries regarding item difficulty and the strengths of correlations, but the picture was not consistent. Based on this, the MELQO team recommends caution in comparing findings across countries and highlights how further analysis will be needed on the increasing number of MODEL datasets from different countries that are currently being developed (Ibid.). This is consistent with findings from validation of the IDELA instrument from which many of MODEL's items were borrowed (Wolf et al., 2017). However, based on other studies of cross-context validity within the same country, findings should be more comparable for groups within the same country context (Raikes, 2019; Raikes, Tanzania; Wolf et al., 2017).

As of 2019, MODEL had been used in nearly 30 countries on three continents (ecdmeasure.org, 2019). Across those experiences, the MELQO team has established evidence of construct validity, inter-rater reliability in specific contexts, and internal reliability; requested further expert support in establishing cross-cultural validity based on heterogenous contextual experiences; and has not yet established predictive validity, or whether performance on MODEL tasks predicts children's future development into primary school (UNESCO, UNICEF, Brookings Institution, and World Bank, 2017). In addition, a 2021 analysis aimed at identifying a core set of assessment items for pre-primary age children describes how five of the literacy-related tasks in MODEL were psychometrically robust across at least 12 countries that had utilized the assessment tool (Pushpuratnam et al., 2021). The items included letter identification, listening comprehension, initial sound discrimination, letter sound identification, and name writing.

After the initial pilot with 200 children in Laos, the World Bank supported MOES to refine the instruments through a series of workshops with the MOES early childhood, training and curriculum, and monitoring and evaluation units as well as

representatives from the health sector and a selection of practicing ECD teachers (World Bank, 2016c, pp. 20-21). Subsequently, the test modules were rolled out through the large-scale assessment in a sample of 7,520 children ages 2-5 years old in the north of the country, as described above (Ibid.).

The direct assessment was adjusted for the LEARN Project in a few ways. First, one of the items in the MOES assessment (initial sound identification) was dropped and replaced with a new item (word segmentation) for reasons discussed below. Second, LEARN added a picture vocabulary test of receptive vocabulary at T1-T3, because this was an important skill emphasized in the programmatic interventions.[11]

The literacy-related domains, constructs, tasks, and items that were included in the LEARN assessment are captured in **Table 5** below; for a breakdown of all domains included in the full direct assessment, refer to http://ecdmeasure.org/.

[11] Because the LEARN sample was older on average than the MOES sample, the study also added semantic fluency, non-word decoding, and familiar word reading automaticity assessments at T3 to capture higher-level skills in the older children in the sample. However, these are not covered in this thesis because they exhibited strong floor effects, and because they cannot be subjected to gain score analysis since they were only collected at endline.

**Table 5: Literacy Domains, Constructs, Tasks, and Items in the Direct Assessment**

| Construct | Tasks | Items |
|---|---|---|
| Concepts about Print | Book Awareness | If you were going to read this book, can you show me how you would open it so you can read it? |
| | | Can you show me where I should start reading? |
| | Print Directionality | If I start to read here, on the first word, where do I continue reading? |
| Phonological Awareness | Initial Sound Discrimination | Which of the following words starts with the letter / sound 'N'? |
| | | Which of the following words starts with the letter / sound 'L'? |
| | | Which of the following words starts with the letter / sound 'P'? |
| | Word Segmentation | Please pronounce the last word in 'Ka Tai' |
| | | Please pronounce the last word of 'Hong Mor' |
| | | Please pronounce the last word of 'Mark Ban' |
| | | Please pronounce the last word of 'Khao Khoai' |
| | | Please pronounce the last word of 'Lai Seur' |
| Alphabet Knowledge | Letter Naming | What letter is this? [child is shown list of 20 consonants] |
| Language Comprehension | Listening Comprehension | Who stole the cat's hat? |
| | | What was the color of the hat? |
| | | Why was the cat chasing the mouse? |
| | | Where did the cat trap the mouse? |
| | | Why did the cat decide not to eat the mouse? |
| Emergent Writing | Name Writing | Can you write your name here in any way you know? |

| Construct | Tasks | Items |
|---|---|---|
| Receptive Vocabulary | Picture Vocabulary | Please point to the word '...' [child is shown 27 cards with 4 pictures on each card and asked to point to the specified target word] |
| Expressive vocabulary | Semantic Fluency | Name the things you see in this picture [child is shown a picture of a typical village scene] |
| Decoding | Nonsense word reading | Please read this word [child is shown list of 20 'decodable' words] |
| Word Recognition / Automaticity | Familiar word reading | Please read this word [child is shown list of 20 common sight words] |

### *4.6. Multi-Language Assessment Protocol*

Because the sample of children in the study came from diverse language backgrounds, a multi-language protocol was devised for administering the assessment items. Multi-language assessment is increasingly recognized as an appropriate approach in contexts where multiple home languages are present but where the language of instruction and assessment is an unfamiliar national or European language. The primary purposes of a multi-language assessment approach are three-fold: To provide a more comfortable and familiar testing environment for children; to truly assess their underlying knowledge rather than only their ability to express that knowledge in an L2; and to understand if young children's skills across languages are predictive of later achievement, for example, if their expressive language scores across multiple languages help to explain their reading outcomes in the early years of formal schooling (RTI International, 2019).

For the LEARN study, the multi-language protocol included utilizing data collectors who spoke the same language as the children to the extent possible and selecting the most appropriate language of assessment at different times. For instance, at baseline and midline, native Lao speakers were assessed in Lao for all tasks, while

non-Lao speakers were always provided with explanations of instructions in their home language, and were then assessed only in Lao for items where there was a clear wish by the government to assess their Lao skills (such as initial sound discrimination and word segmentation of Lao words); only in their home language where the assessment was attempting to uncover their underlying conceptual knowledge or executive function regardless of language (e.g., forward and backward digit spans – not covered in this thesis); and in both languages where it was deemed that assessing in both would reveal important comparative data about their abilities in their home language versus the language of instruction (e.g., listening comprehension). At endline, all children were assessed only in Lao with the expectation that their skill level in Lao would be sufficient to enable them to answer basic tasks by the end of grade 1.

**Table 6** below summarizes the assessment languages used for the literacy-related tasks in the assessment.

**Table 6: Language of Assessment Protocol for Literacy Tasks**

| Tasks | Native Lao Speakers Were Assessed in… | Non-Lao Speakers Were Assessed in… | |
|---|---|---|---|
| | Lao | Lao | Non-Lao Language |
| Familiarity with print | X | X | |
| Initial sound discrimination | X | X | |
| Word segmentation | X | X | |
| Letter name knowledge | X | X | |
| Listening comprehension | X | X | X |
| Beginning writing | X | X | |
| Receptive vocabulary | X | X | |
| Semantic fluency | X | X | |
| Nonsense word reading | X | X | |
| Familiar word reading | X | X | |

### *4.7. Approaches to Data Collection*

#### 4.7.1. Enumerator Selection and Training

The survey team was set up and trained with technical guidance from AIR and under my oversight and with field-level monitoring by the LEARN Monitoring and Evaluation Coordinator. The same core team of survey supervisors were present at all waves (and had also led the MOES ECD diagnostic assessment), while some of the same enumerators were present at different waves and some were not, due to challenges in maintaining long-term enumerator staff for data collection tasks that are only periodic. In addition, it was not possible to recruit a multi-lingual team of enumerators that spoke all languages for all villages due to the limited human resource pool with the required qualifications in Laos. As an alternative, the survey team leaders recruited local language interpreters in the target districts or villages during data

collection; this was often an unemployed teacher from a district center or a leader from the same village.

Before T1 data collection, AIR and LEARN staff conducted a two-day ToT for the survey managers. Subsequently, the AIR staff and survey managers conducted five days of enumerator training including one day of field practice in rural villages outside Vientiane before data collection began in the target districts. Only the top-performing candidates who demonstrated good technique and pairwise reliability during practice data collection were contracted for the survey work. The training processes and composition of the team for T2 and T3 was roughly the same as T1 (Indochina Research Ltd., 2017, 2018a, and 2018b). Each team that visited a school consisted of one team leader and three enumerators, while data collection overall was overseen by a survey manager and two quality control supervisors.

#### 4.7.2. Data Collection Procedures

Data were collected by the survey firm at each wave on tablets using CS-pro software with digitized versions of the data collection tools that had been pre-tested and were further revised after initial field testing of instruments. Using a tablet-based approach resulted in inherently cleaner data than a paper-based assessment, because features such as drop-down lists of possible results and skip logic prevented enumerators from entering incorrect values (Indochina Research, 2016).

Once data collection was completed at each wave, the same supervisors conducted data checks and exported the cleaned datasets separately for each instrument into SPSS and STATA. The external evaluator, AIR, then conducted further checking, cleaning and collation of the datasets into a merged file. At each wave, I inspected the dataset for any irregularities, which were then rectified by the survey firm or AIR, as required.

### *4.8. Data Cleaning and Preparation*

Upon receipt of the final datasets from the external research partner, I prepared them for analysis by ascertaining which background variables were most complete and correct (for example, where information on child sex was collected at multiple different times); and identifying and fixing missing data where possible (i.e., where a variable such as 'village' was missing data at T1 that could be obtained from information collected on the same variable during T2 or T3).

Considerable time was needed to double check and correct the final treatment status variable for all the villages and children in the dataset. This required cross-checking between participation variables collected for each child during each data collection wave, verifying those against project enrollment records for the summer pre-primary treatment, and determining the best course of action to take when there was a discrepancy between the data sources.

After cleaning the data in the original background variables, I computed composite background variables and indices that would be useful to analysis. This included parents' education levels and a socio-economic status (SES) index.

### *4.9. Ethical Considerations*

This study was conducted under rigorous ethical standards. It was approved through the UCL-Institute of Education (IOE) ethical review process and by the Institutional Review Board at AIR. It was administered under Plan International's guidelines on child protection, and the strict Government of Lao approval processes for research in schools and villages. In line with the United Kingdom's UK's *Ethical Guidelines for Education Research* (BERA, 2018), ethical issues that were addressed during the study included gatekeeper approval; children protection; voluntary informed consent and the right to withdraw; incentives; confidentiality; harm arising from participation in research; and my role as an 'insider' or practitioner researcher.

Without the approval of key gatekeepers in the MOES, research in Lao PDR is not allowed to proceed. With this in mind, the study team obtained prior approval in

writing from MOES to conduct the study, which was transmitted to provincial, district, village, and school government levels prior to data collection.

Because this study involved children, all data collectors were required to sign Plan International's Child Protection policy as a condition of their contracts, and any violations of the policy would have led to disciplinary action and possibly dismissal. Data collection team leaders were responsible for monitoring child protection issues during data collection, with oversight from AIR and Plan staff, and no violations of the policy were reported.

Voluntary, informed consent was sought from the parents/guardians of all participating children using a standard script at each wave of data collection. All study subjects had the right to refuse their child's participation in the first place or have the child withdraw at any point during the assessment.

Small incentives were given to children who participate in the study, including modest school-related items in keeping with local economic circumstances such as notebooks and pencils. These were considered proportionate to the circumstances and not large enough to warp participants' consent to participate or the answers they gave to assessment questions.

The study design did not allow for a completely anonymous data collection process because the same children were assessed at baseline and endline. Nevertheless, the research team ensured the confidentiality and privacy of information that was collected by assigning child codes for the purposes of data analysis, maintaining the name list separately from the codes, and only granting a handful of staff access to the original list of names. In addition, the study dataset used for analysis was fully anonymized, using only ID codes and no individually identifying names.

The nature of this research did not constitute an elevated risk of harm to participants, as it did not require participants to provide any sensitive, embarrassing or potentially legally jeopardizing information. The data collected on student outcomes was never

linked back to individual children or used to shame any child. Nevertheless, the research did run the risk of placing children in an unfamiliar assessment environment, particularly at baseline when they were unused to school-related interactions. This was mitigated to some extent through training of data collectors in techniques such as play-based introductions to the study.

The study considered potential harm that could arise related to randomized assignment of villages with the use of a no-intervention control group, which meant that some children would not benefit from the intervention during the life of the LEARN Project. It was deemed an ethical priority to establish effectiveness of the summer pre-primary model first and then consider scale-up, in line with MOES preferences to pilot new approaches before expanding them. As such, the study had a strong emphasis on testing impact and identifying recommendations for policymakers and practitioners to guide future scale-up to additional locations.

Finally, my role as a practitioner researcher – engaged as both project director and evaluator of the same project through my doctoral thesis – was a key ethical consideration in the study, although education research ethics guidelines typically do not offer explicit advice on how to address it (Zeni, 2001; BERA, 2018). Practitioner research is a type of applied inquiry, carried out by 'insiders,' that is concerned with understanding and strengthening professional practice (Gillman and Swain, 2006; Menter et al., 2011). Rigorous studies, such as RCTs, that involve practitioner-academic collaborations are typically conducted in pursuit of solutions to complex problems where the practitioners are able to tap into their direct, lived knowledge to identify and then, importantly, to act on the recommendations that are uncovered (Sawtell, 2018). At the same time, practitioner research designed to evaluate specific interventions may involve some form of self-evaluation by the implementers themselves and has thus been criticized for the potential bias and conflicts of interest that can arise (Robson, 2011; Gillman and Swain, 2006).

In the case of the LEARN study, I was simultaneously an 'insider' through my responsibility for implementation of the intervention and oversight of the external

evaluation partner and survey firm; and an 'outsider' both as a foreigner to the Lao linguistic and educational context and as a doctoral student. In the insider role, I led procurement of the evaluation and data collection services through AIR and Indochina Research Ltd.; guided development of the overall sampling approach in anticipation of common real-world challenges with RCTs; and contributed practical advice to tool refinement, field testing, and data collector training in line with the intervention design and the Lao context. I also helped troubleshoot real-world challenges that emerged during data collection, such as sample attrition between assessment waves and verification of treatment uptake.

At the same time, there were certain 'firewalls' between myself and the design and implementation of the study that helped promote a greater degree of objectivity. These firewalls included a competitive procurement process for the research and survey services that was governed by a multi-member procurement panel; actual data collection activities that I did not join so as not to influence the process; and initial preparation of the dataset that was completed by the external survey firm and research partner. My analysis of the data for this thesis was also guided by standards of practice in my professional field, a substantive review of the literature to identify the predictors of most significance in the Lao context, as well as oversight of the analytical process by my doctoral supervisors at the UCL, Institute of Education.

Although my role as a practitioner researcher could be viewed as an ethical shortcoming of the study, this kind of insider-outsider research in international education, conducted in collaboration between practitioners and external evaluators, has also been credited as "more sensitive to local context, while retaining systematic rigour and an important degree of detachment from the culture and worldview being studied" (McNess, Arthur, and Crossley, 2015, p. 300). My position as a doctoral researcher enabled me to lend a more critical eye to the study and its findings by playing the role of the 'subverter' in a rapidly globalizing education assessment field – an insider with intimate knowledge of the field attempting to reform and improve from within (Soudien, 2009). Thus, practice informed research and research informed practice in the quest for better outcomes and assessment approaches for

marginalized children, in a manner that would likely not have been possible for external evaluators alone.

## 5. Study Findings

### *5.1. Analytical Approaches*

Although data were collected on a range of pre-numeracy, executive function, and socio-emotional skills in the LEARN evaluation, this section focuses on literacy outcomes, which is the emphasis of the thesis.

First, a descriptive analysis of the assessment scores at baseline, midline, and endline is provided on the entire original sample of 673 children, in addition to reporting on the internal validity of the assessment tasks. This descriptive analysis places an emphasis on differences by ethnicity.

Next, the section presents the impact of the summer pre-primary program using analysis of gain scores. Cluster-robust multiple regression models control for covariates that were found to be imbalanced between treatment and control groups at baseline and to account for clustering of similar outcomes among children in the same schools. Implications of the findings are then discussed in the final chapter of the thesis, with a focus on answering the research question on the reliability and validity of the assessment approach in the Lao context and recommendations for future practice.

### *5.2. Administration of Subtasks and Computation of Subtask Scores*

This section describes how each task was administered and how scores were calculated on each task, as well as the internal reliability of each at baseline, midline and endline.

*Literacy Domain Score.* A literacy index was created for the purposes of analysis in this thesis using the same scaled means approach as the indices in the four-country validation study (MELQO, no date) and the World Bank ECD diagnostic assessment in Laos (World Bank, 2016c). Only the six tasks that were assessed at all three waves and are typically included in the MODEL literacy domain were included in the calculation: Concepts about print, initial sound identification, word segmentation, letter name knowledge, listening comprehension (in Lao), and name writing. The receptive vocabulary task was not included in the index, as this would render the results incomparable with other similar assessments used in Laos and elsewhere; and the listening comprehension assessment in ethnic languages was also not included, because it was administered only to non-Lao children and only at baseline and midline.

First, the scores on the six tasks were summed, and means were calculated on the summed scores. Then, to create the index, the means on each task were added together and the average was taken. This approach has the advantage of placing scores on all tasks onto a standardized scale of 0-1. Other approaches to creating the index were considered, including using z-scores to normalize the mean scores (Raikes, 2018), but scaled means was considered the most appropriate to communicating results in a real-world context because the resulting statistics are more intuitive to policymakers and educators, and more comparable to the other studies in Laos.

Internal reliability when combining all six of the items in each of the tasks included in the domain score, as measured by Cronbach's Alpha, was very good at baseline ($\alpha$=.87), midline ($\alpha$=.83) and endline ($\alpha$=.87).

*Concepts about Print.* Administered in Lao only for both Lao and non-Lao speaking children at baseline, mid-line and endline, the concepts about print task (also sometimes referred to as familiarity with print) was assessed through three items: Demonstrating how to correctly orient and open a story book, pointing to the start of the story, and indicating text directionality. The overall score on the item was

calculated by summing the three tasks and then taking the average to achieve a scale between 0 and 1.

Internal reliability on this task as measured by Cronbach's Alpha was poor at baseline (α=.44) adequate at midline (α=.61) and poor at endline (α =.56).

*Initial Sound Discrimination.* Administered only in Lao to all students at baseline, mid-line and endline, the initial sound discrimination task was a test of phonological awareness and consisted of three items. In the first item, children were told three words out loud and asked to identify which one out of the three started with the 'N' sound. This was repeated for the sounds 'L' and 'P.' Words selected for the test were simple nouns or verbs that would generally be familiar to young children. The overall average score was calculated by summing the scores on the individual tasks and taking the average to create a scaled score between 0 and 1.

Internal reliability on this task was good at baseline (α=.70) and adequate at midline (α=.61). At endline, internal reliability had improved to good (α=.75).

*Word Segmentation.* The MODEL child direct assessment tool originally adapted for Laos through the World Bank-supported Early Childhood Education (ECE) Project initially included the test of initial sound discrimination followed by a test of initial sound identification, in which the instructions stated that children could name the letter names or the letter sounds in a set of simple Lao words. However, based on experiences during baseline data collection for the World Bank assessment, the survey team leaders felt that the test of initial sound identification would be confusing for enumerators and children, because there is only one, highly standardized way of naming letters in the Lao alphabet, and because of the syllabic nature of the writing system.

As such, the test of initial sound identification was dropped in favor of a word segmentation task that the survey team leaders felt would be more appropriate for the phonological structure of Lao. In the test, children were asked to repeat the

second word in a set of two compound words after the enumerator said each word combination aloud: For example, the word ‘Tai’ in ‘Ka Tai’ (rabbit). Compound words were selected based on vocabulary that would be familiar to and simple enough for young children and that represented a range of different consonant sounds.

This task was administered in Lao only, for all children, at baseline, mid-line, and endline. The task consisted of five compound words, and the overall score was calculated by summing the five and dividing by the total number of items to create a score from 0 to 1.

Baseline internal reliability was very good (α=.89); and at midline (α=.94) and endline (α=.94) it was excellent.

*Letter Name Knowledge.* In the test of letter name knowledge, children were shown a sheet of 20 consonants[12] from the Lao alphabet, not in alphabetical order, and asked to state the name of each consonant (see **Figure 8** below). This test was administered in Lao only, for all children, at baseline, mid-line, and endline. A stop rule was applied after the first five incorrect responses. The overall score was created by summing the scores on each individual letter item and then taking the average to create a scaled score from 0 to 1.

Internal reliability on this task was very good baseline α=.85; good at midline α=.78; and very good at endline α=.87.

**Figure 8: Letter Name Knowledge Test**

[12] The Lao script does not have upper and lowercase letters, so unlike typical MODEL or EGRA assessments, no distinction could be drawn by letter case in constructing this task.

| | |
|---|---|
| ຂ | ຈ |
| ຜ | ພ |
| ກ | ບ |
| ມ | ຫ |
| ນ | ປ |
| ຢ | ຖ |
| ງ | ຜ |
| ລ | ສ |
| ຫ | ຄ |
| ຜ | ຊ |

*Listening Comprehension.* In the listening comprehension task, children were first read aloud a short story (English translation in **Figure 9** below). They were then asked to verbally answer five questions about the story that were asked of them aloud. Four out of the five comprehension questions required factual recall, while the final question required students to draw an inference.

**Figure 9: Listening Comprehension Passage in English**

**The Mouse and the Cat**

Once upon a time there was a fat cat. He always wore a red cap. Once when he was sleeping, a small mouse came silently and stole the hat. The cat woke up to see his hat gone, got very angry and started chasing the mouse. After a while, the mouse was trapped under a table and could not find any way to escape. So the mouse cried to the cat, "Please don't eat me cat. If you spare my life, I will return your hat." So, after getting back his hat the cat said, "Never touch my hat again" and he went back to sleep in a happy mood.

**Questions**

1. Who stole the cat's hat?
2. What was the color of the hat?
3. Why was the cat chasing the mouse?
4. Where did the cat trap the mouse?
5. Why did the cat decide not to eat the mouse?

At baseline and mid-line, this test was administered in Lao first to all children, and then in ethnic languages for children whose home language was not Lao. At endline, it was administered only in Lao to all children. The overall score on this task was computed by summing the scores on the five individual items and taking the average to create a scaled score from 0 to 1.

Internal reliability when administered in Lao was good at baseline (α=.79); very good at midline (α=.82); and very good at endline (α=.82). When administered in ethnic languages, internal reliability was very good at baseline (α=.81) and good at midline (α=.78).

*Name Writing.* The name writing assessment consisted of asking children to write their first names, with scoring on a scale of one to five as follows:

1. Does not write
2. Writes scribbles but no discernible letters
3. Writes letter-like marks
4. Writes letters but not his/her complete or correct name
5. Writes some letters in name
6. Writes name correctly

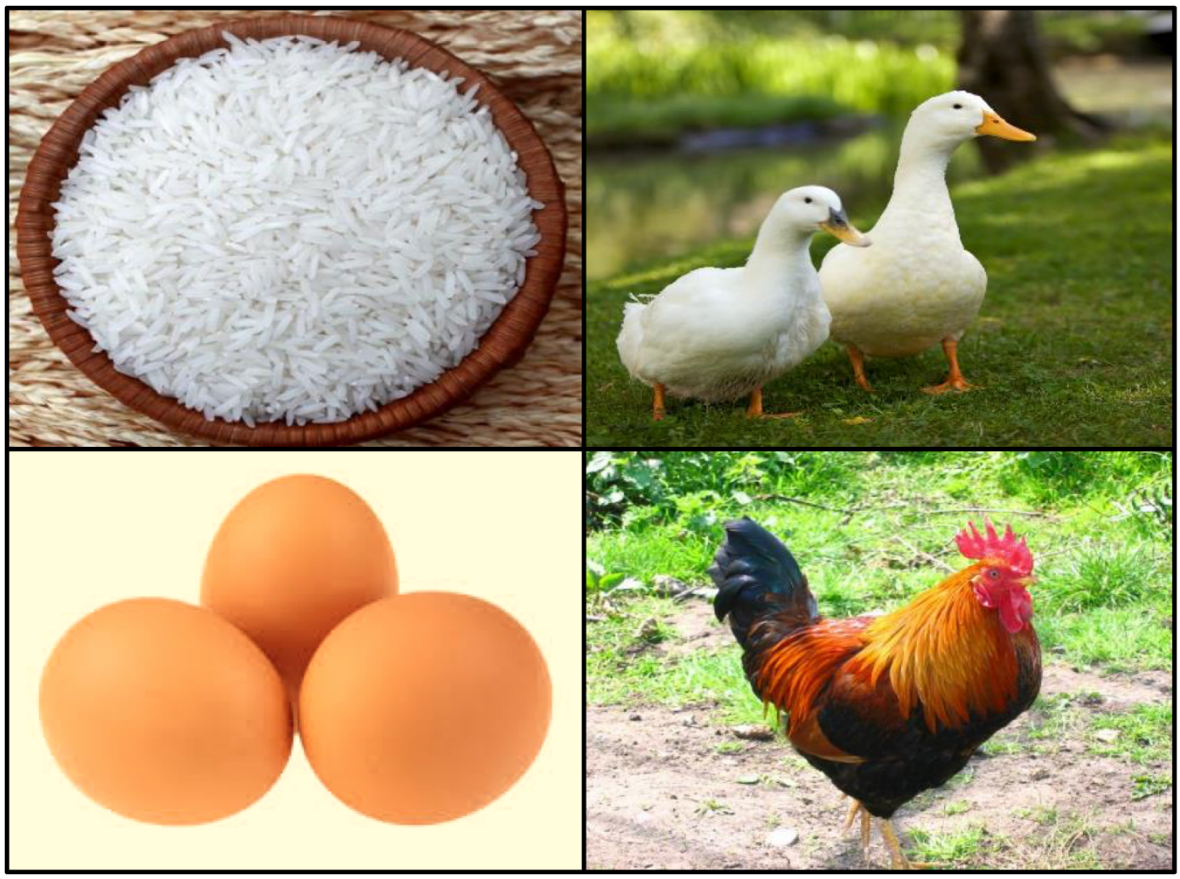
The mean was taken to create a score on a scale from 0-1. Internal reliability cannot be calculated for this task because it consists of only one item.

*Receptive Vocabulary.* In the receptive vocabulary task, there were 27 separate items (**Figure 10**). For each item, children were shown a set of four pictures on a card, told the name of the target word in Lao, and asked to point to the picture of the target word amongst the set of four pictures. Prior to developing this task, the research team reviewed existing assessment literature from Laos to identify if a receptive vocabulary test had been developed and validated in the Lao context previously. When it was determined that no suitable test existed, the team developed the task from scratch, identifying vocabulary items and distractors that would display a suitable range of difficulty and discrimination for the target age ranges in the rural context of the assessment.

The overall score on this task was calculated by taking the sum of the individual items and dividing by 27 to achieve a scaled score from 0 to 1. This test was administered in Lao only, for all children – in other words, children were told the name of the target word in Lao and not in their home language – at baseline, mid-line and endline.

Internal reliability was excellent at baseline ($\alpha$=.94) and midline ($\alpha$=.93) and very good at endline ($\alpha$=.89) on this task.

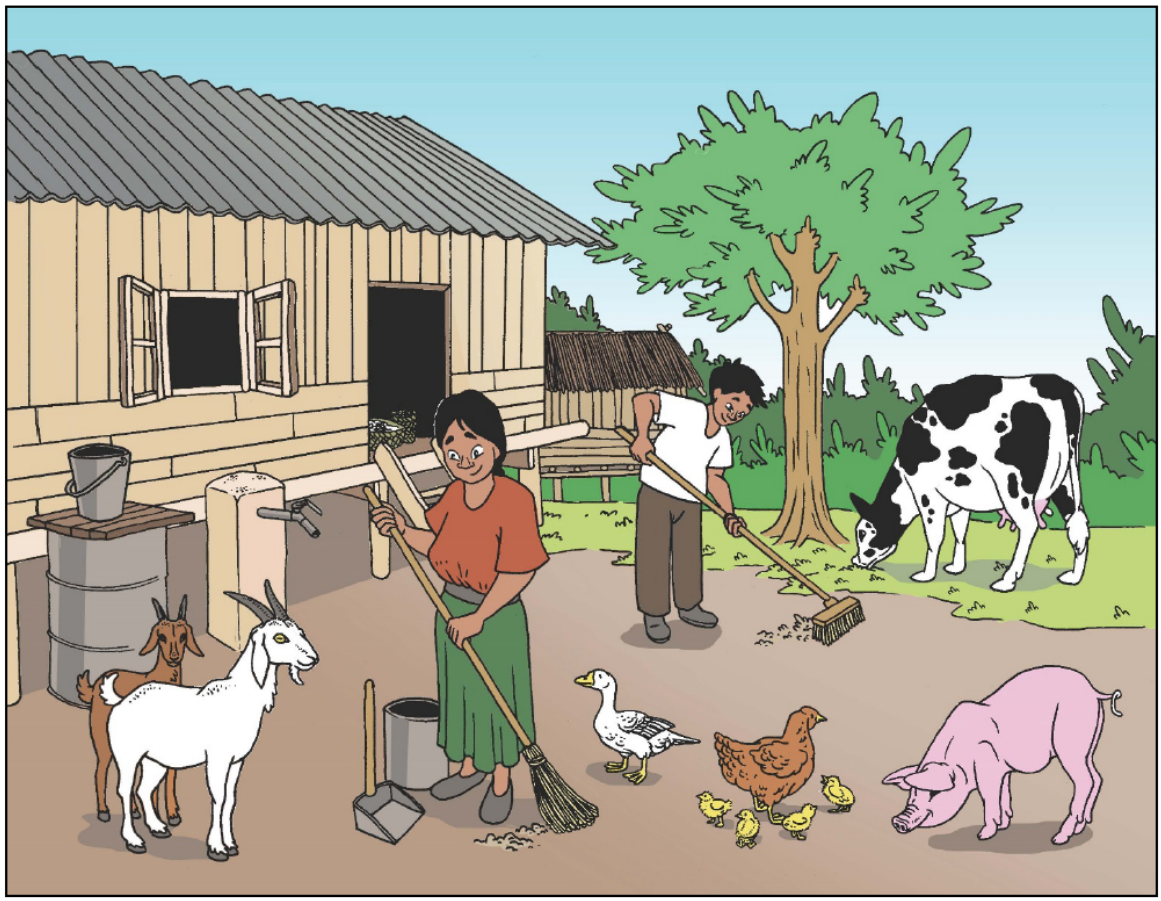
**Figure 10: Sample Receptive Vocabulary Item with Target Word 'Chicken'**



*Source: Plan International in Laos, 2016. Used with permission.*

*Semantic Fluency.* In the test of semantic fluency, which was administered only at endline, children were shown a color illustration of a typical Lao village scene that included a house, people, farm animals, trees, etc. (**Figure 11**). They were given 60 seconds to name any words that came to mind based on the picture. Enumerators were instructed to count the number of correct words related to the picture, and the number of words unrelated to the picture but that were nevertheless actual words. Children's raw score on this task was calculated by summing their score on words related to the picture plus words unrelated to the picture. As this item did not have a pre-established ceiling, each child's summed score was divided by the highest summed score on the task to create a scale from zero to one. Internal reliability could not be calculated because this task included only one item.

**Figure 11: Semantic Fluency Illustration**



*Source: American Institutes for Research and Catholic Relief Services, 2017. Used with permission.*

*Most Used Words.* In the test of most used words, children were shown the list below of 20 high frequency Lao vocabulary words from the grade 2 Lao language curriculum (**Figure 12**), which had been used in a previous assessment of second grade children in the LEARN Project, and asked to read the words aloud. For each word, children were given one point if they read it correctly and zero points if they read it incorrectly. To create a scaled score, children's score on all 20 words was totaled and then divided by 20.

This task was administered only at endline, and endline internal reliability was excellent at (α=.96), most likely because of the extensive zero scores.

**Figure 12: List of Most Used Words**

| # | Lao Word | English Equivalent | Transliteration |
|---|---|---|---|
| 1 | ອາ | Aunt | Aa |
| 2 | ດີ | Good | Dee |
| 3 | ເອ | [Varies] | Ae |
| 4 | ໄປ | Go | Pai |
| 5 | ມາ | Come | Maa |
| 6 | ຂ້ອຍ | I | Khoi |
| 7 | ດັ້ງ | Nose | Dang |
| 8 | ງາມ | Beautiful / Nice | Nyam |
| 9 | ໄກ່ | Chicken | Kai |
| 10 | ທຸ່ງ | Flag | Thoung |
| 11 | ຝົນຕົກ | Rain | Fontok |
| 12 | ອ້າຍ | Brother | Ai |
| 13 | ໝູ່ເພື່ອນ | Friend | Moopheuan |
| 14 | ເສື້ອ | Clothes | Seua |
| 15 | ເອື້ອຍ | Sister | Euay |
| 16 | ໂສ້ງ | Pants | Song |
| 17 | ລ້າງມື | Hand washing | Langmuu |
| 18 | ແມ່ | Mother | Mae |
| 19 | ຕັ່ງ | Chair | Tang |
| 20 | ປຶ້ມຫັດອ່ານ | Story book | Peumhat-anh |

*Decodable Words.* In the test of decodable words, children were shown a list of 20 nonsense words (**Figure 13**) that have no meaning but follow morphological

conventions of Lao, which had been used in a previous assessment of second grade children in the LEARN Project. Children were asked to read the words aloud. For each word, children were given one point if they read it correctly and zero points if they read it incorrectly. To create a scaled score between 0 and 1, children's score on all 20 words was totaled and the average was taken.

This task was administered only at endline, and endline internal reliability was excellent (α=.97), again most probably because of the extensive zero scores.

**Figure 13: List of Decodable Words**

| # | Decodable Word | Approximate Transliteration |
|---|---|---|
| 1 | ດິ | di |
| 2 | ຕຶ | tu |
| 3 | ຜຶ | fu |
| 4 | ຊ | se |
| 5 | ຖ | thu |
| 6 | ບີ | boe (with mid-tone marker) |
| 7 | ຢ | ye |
| 8 | ດດ | dae |
| 9 | ໄ | ai (with falling tone marker) |
| 10 | ຢາບ | yab |
| 11 | ໂ | fo (with falling tone marker) |
| 12 | ຟາມ | fam |
| 13 | ຫ | he |
| 14 | ງ່ | ngu (with mid tone marker) |
| 15 | ຜ໋ | fo (with falling tone marker) |
| 16 | ງດ | nged |
| 17 | ຮາຍ | hai |
| 18 | ຍ້ | nya (with falling tone marker) |
| 19 | ໂ | kho (with mid tone marker) |
| 20 | ຜ | fe |

### *5.3. Overall Scores at Baseline, Midline, and Endline*

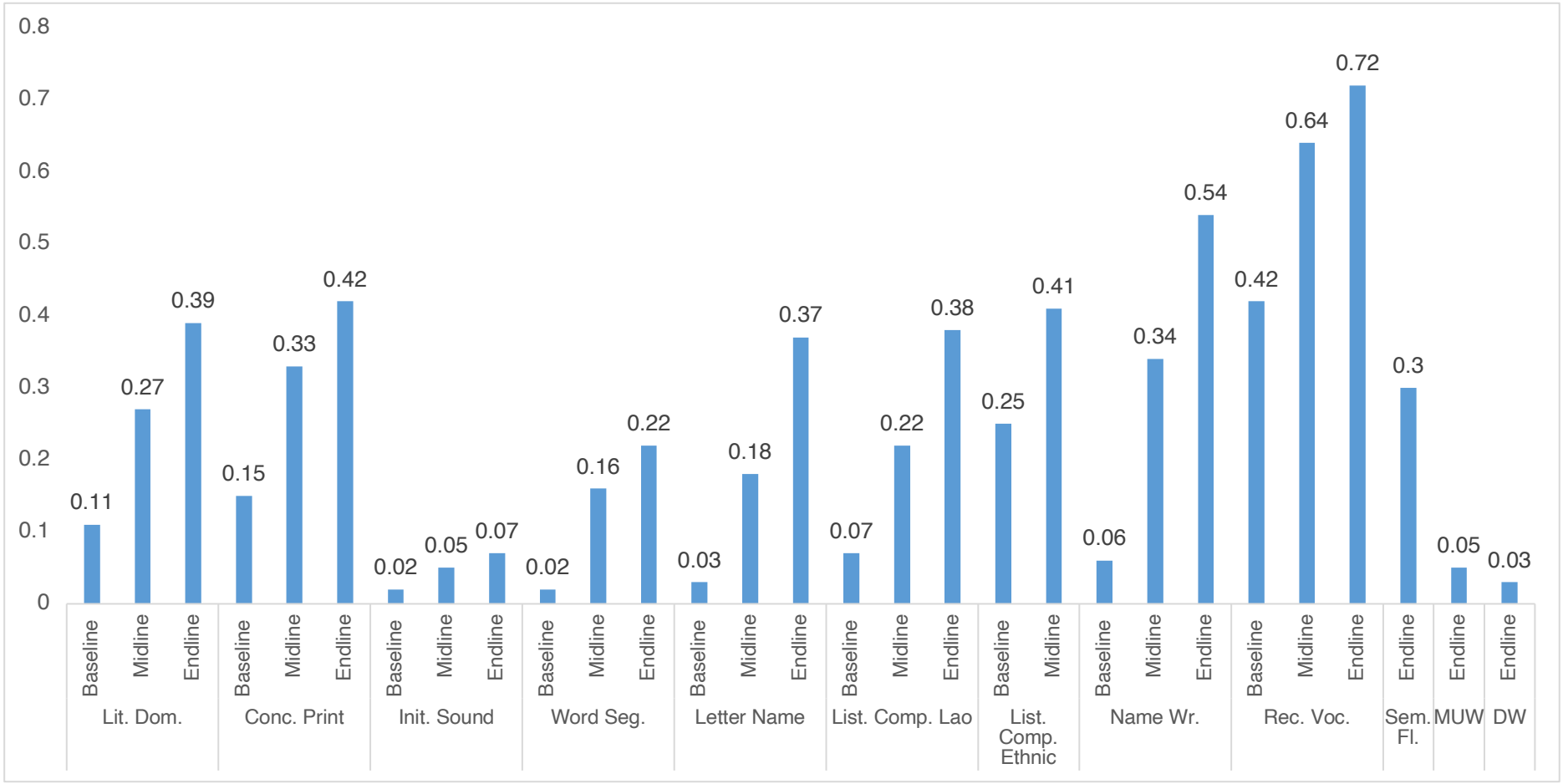
This section provides an analysis of average scores at baseline, midline, and endline on the literacy domain and the eight tasks described above across the original sample of 673 children. Utilizing the original sample for the purposes of descriptive analysis is useful because the larger sample affords greater statistical power in exploring differences by ethnic sub-groups. However, it is important to note that at midline and endline, some of the children in the sample had participated in the LEARN summer pre-primary model or a government pre-primary service. This may have altered the average scores in the sample at those two time points. Nevertheless, examining trajectories is instructive to the overall discussion of the research questions in this thesis, and therefore descriptive analysis is provided across all three waves and not only at baseline.

As described in **Figure 14** below and **Table 23** in **Annex 3**, average scores generally improved between baseline, midline, and endline across the tasks, as would be expected developmentally. Strong floor effects were prevalent at baseline on five tasks, including initial sound identification, word segmentation, letter name knowledge, listening comprehension in Lao, and name writing. The floor effects persisted across all three waves for initial sound identification and improved only modestly at midline and endline for word segmentation.

Children scored highest, on average, on the receptive vocabulary task, followed by name writing and concepts about print. However, as receptive vocabulary was multiple-choice, some of the higher performance may have been due to chance. Listening comprehension scores for ethnic children when assessed in their own languages were roughly as high at baseline as the listening comprehension scores across the sample in Lao at midline. The overall literacy domain score started relatively low – reflecting the widespread floor effects at baseline on most of the tasks it indexes – but had more than tripled by endline.

Out of the tasks added at endline, a reasonable range of scores was found only for semantic fluency, while most children scored zero on the most used words and decodable words tasks.

**Figure 14: Average Baseline, Midline, and Endline Standardized Scores on All Tasks, Original Sample[13]**




[13] Lit. dom.=literacy domain; conc. print=concepts about print; init. sound=initial sound; word seg.=word segmentation; letter name=letter name knowledge; list. comp. Lao=listening comprehension in Lao; list. comp. ethnic=listening comprehension in ethnic languages; name writ.=name writing; rec.voc.=receptive vocabulary; sem. fl.=semantic fluency; MUW=most used words; DW=decodable words.

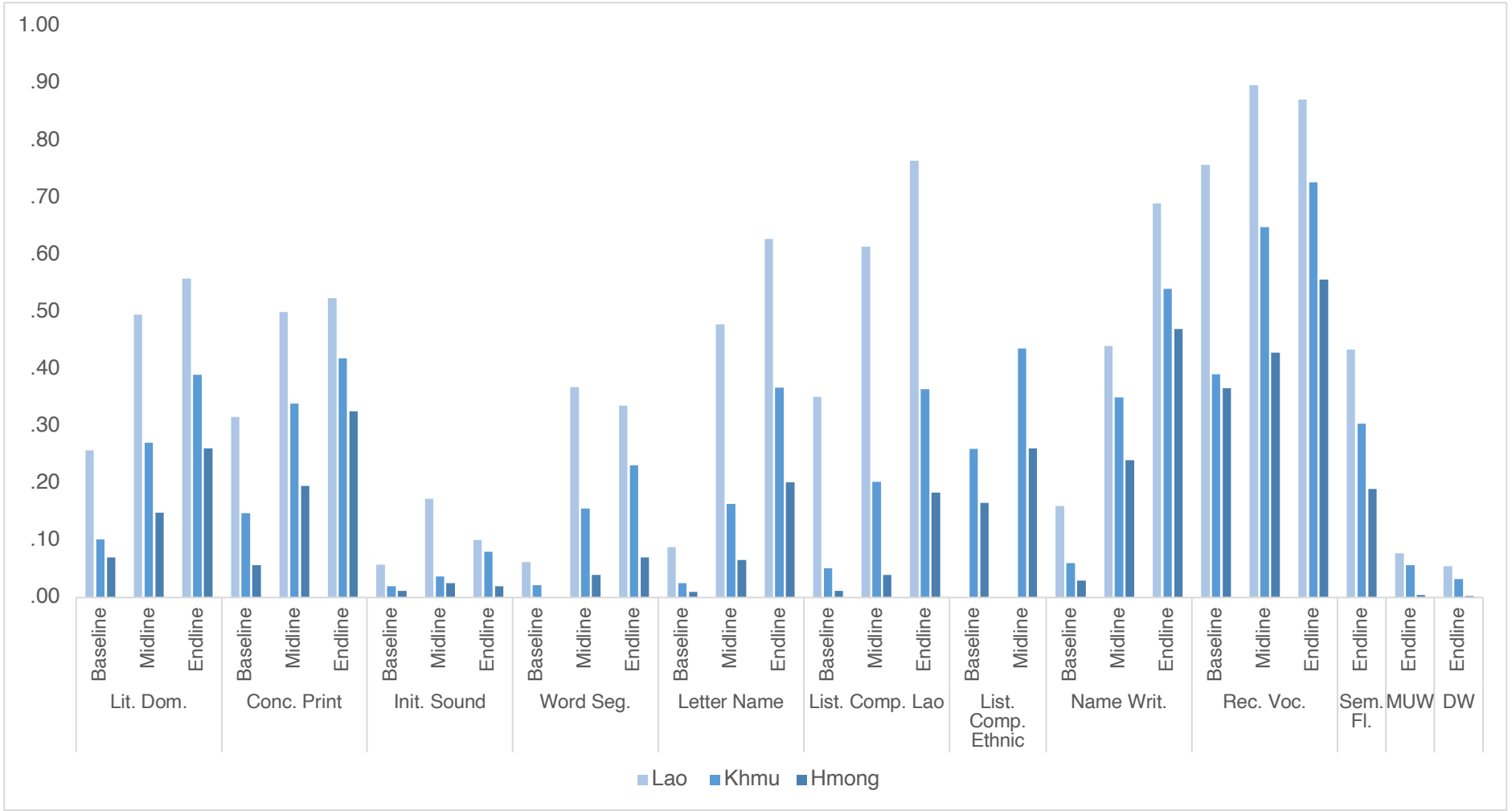
### 5.3.1. Overall Scores at Baseline, Midline, and Endline, by Ethnicity

Analysis of baseline, midline, and endline findings on the original sample of children by ethnicity – Lao, Khmu, and Hmong – reveals a consistent pattern. As indicated numerically in **Table 24** in **Annex 3** and graphically in **Figure 15** below, Lao children overwhelmingly scored highest across all tasks at all three waves, followed by Khmu children. Lao children's score improvements over time outstripped their Khmu and Hmong peers most notably in the letter name knowledge and listening comprehension in Lao tasks. Hmong children were at a clear disadvantage across the board, even scoring lower than Khmu children on listening comprehension in ethnic languages, although it was expected that being assessed in their own home language would have acted as an equalizer.

The score differences by ethnicity were highly significant using one-way Analysis of Variance (ANOVA) across all tasks and all three waves. The exception is initial sound identification, where all children were at the floor at baseline and experienced only limited improvements over time, regardless of ethnicity (although the differences by ethnicity were still moderately statistically significant). Similar to the overall descriptive analysis presented above, floor effects on the word segmentation task were still evident, although Lao children – and Khmu children to a lesser extent – did improve over time.

It is also important to note that none of the ethnic groups had achieved a high level of mastery of the skills assessed in the study by the end of grade 1 (at endline), except for receptive vocabulary, where Lao children had nearly reached the ceiling.

**Figure 15: Average Baseline, Midline, and Endline Standardized Scores on All Tasks, Original Sample, by Ethnicity[14]**




[14] Lit. dom.=literacy domain; conc. print=concepts about print; init. sound=initial sound; word seg.=word segmentation; letter name=letter name knowledge; list. comp. Lao=listening comprehension in Lao; list. comp. ethnic=listening comprehension in ethnic languages; name writ.=name writing; rec.voc.=receptive vocabulary; sem. fl.=semantic fluency; MUW=most used words; DW=decodable words.

## *5.4. The Effects of the Two Program Models on Children's Learning Outcomes*

After describing learning outcomes at T1-T3 in the section above, we now turn to an analysis of outcomes by treatment group. First, the section describes the analytic sample, explores sample attrition, and discusses baseline equivalence between treatment and control groups on key covariates. The analysis presented here follows guidelines related to sample attrition, baseline equivalence and analytical approaches for RCTs and quasi-experimental designs (QEDs) established by the U.S. Department of Education's Institute of Education Sciences (IES) (IES, 2020), which are commonly applied to similar education-related studies in developing countries. The analysis employs gain scores using multiple regression to control for covariates that lacked equivalence between the treatment and control groups at baseline and adjusts for the clustered nature of the data within villages.

### 5.4.1. The Analytic Sample

The analytic sample used for the analysis presented in this section includes only the 490 children (discussed in **Section 4.3** on sampling) who could be unequivocally verified as having participated in the summer pre-primary treatment or the control. As such, it is an analysis of compliers. This approach was deemed appropriate to the real-world nature of the study because Lao policymakers would find it most relevant to look at the effects of the summer pre-primary on children who actually participated in the intervention. Removing non-compliers from the analysis also helps to address the concern about the wide age range in the sample of assessed children, which did not match the narrower age range of children who would have been accepted to participate in the summer pre-primary in line with the official age of subsequent grade 1 enrollment.

In addition, since this analysis employs gain scores as the outcome measure, only the children who were present for data collection at all three waves are included, because attempting to calculate gains using one or more missing scores would result in misleading zero or negative scores. Accordingly, this is an analysis of compliers

and non-attriters. Finally, the analysis includes only Khmu and Hmong children, because, after first attempting to run analyses with all three ethnicities, it was found that the numbers of Lao children who remained in the analytical sample were too small to provide interpretable results.

### 5.4.2. Analysis of Attrition

Before delving into analysis of outcomes, this section first presents a detailed review of attrition from the sample of compliers, including overall attrition as well as differential attrition by treatment group and ethnicity. Attrition analysis is considered the critical first step in evaluating whether a study design is robust enough to rely on the outcome estimates that it produces, or if it suffers from unacceptable levels of potential attrition bias. This type of analysis also informs interpretation of findings in light of the characteristics of children who remained in the sample compared to those who did not.

For the purposes of this analysis, an attriter is defined as a child in the sample of children who took up their correct assignment, but who was assessed fewer than the full three assessment waves conducted in the study. Although they are later dropped from the analytic sample due to small numbers, Lao children are retained in the analysis of attrition because it is informative to compare rates for all three ethnicities.

The overall attrition rate was 12.1% (59 children out of the 490 total in the complier sample). Using standards established by the IES, this overall attrition rate falls within acceptable bounds for the threat of bias using conservative assumptions (i.e., below approximately 55% attrition) (IES, 2020). When looking at overall patterns of attrition according to the number of direct assessments, 11 children (1.2%) participated in only one direct assessment; 48 (9.8%) participated in two; and 431 (88%) participated in all three.

The acceptable standard of differential attrition rates between treatment and control groups, using conservative assumptions, is below roughly 6% (Ibid.). As shown in

**Table 7** below, the overall attrition rate for children who participated in the summer pre-primary treatment was 10.1% compared to 14.0% for children who participated in the control group, or a differential attrition rate of 3.9%. This difference was non-significant using chi square (Chi sq.) analysis.

**Table 7: Attrition by Treatment Group**

| | | Control | Treatment | Overall | Diff. Attrit. | Chi Sq. | df | *p* |
|---|---|---|---|---|---|---|---|---|
| **Attriter** | N | 34 | 25 | 59 | 3.9% | 1.732 | 1 | .188 |
| | % | 14.0% | 10.1% | 12.0% | | | | |
| **Non-** | N | 209 | 222 | 431 | | | | |
| **Attriter** | % | 86.0% | 89.9% | 88.0% | | | | |

It is also informative to compare baseline status on key covariates between attriters and non-attriters, as this helps to identify if the children who were not present for some part of the study varied in a systematic way from those who attended all three assessment waves. The analysis (presented in **Table 8**) revealed that both groups were statistically equivalent on SES, sex, and baseline literacy domain scores; but that attriters were significantly more likely to be older and to have mothers with lower education levels.

**Table 8: Baseline Covariate Equivalence for Attriters vs. Non-Attriters**

| | Attriter | | Non-Attriter | | *t* /Chi Sq. | df | *p* |
|---|---|---|---|---|---|---|---|
| | N | % or Mean | N | % or Mean | | | |
| **Age in Months** | **37** | **64.51** | **431** | **62.43** | **1.97** | **466** | **.050** |
| SES Scale (0-16) | 37 | 5.46 | 431 | 6.17 | -1.40 | 466 | .163 |
| % Female | 37 | 51.8% | 431 | 49.0% | 0.16 | 1 | .690 |
| **Mother's Education Scale (0-3)** | **36** | **0.53** | **413** | **1.00** | **-2.81** | **3** | **.005** |
| Baseline Literacy Domain Score | 31 | .10 | 426 | .11 | -0.59 | 455 | .558 |

When specifically comparing attrition rates by ethnicity (**Table 9**), there were also significant and large differences using chi square, with only 72.6% of Hmong children participating in all three waves compared to 90.8% of Khmu children and 88.9% of Lao children.

**Table 9: Attrition by Ethnicity**

| | | Lao | Khmu | Hmong | Total | Chi Sq. | df | *p* |
|---|---|---|---|---|---|---|---|---|
| **Attriter** | N | 5 | 34 | 20 | 59 | 19.2 | 2 | .000 |
| | % | 11.1% | 9.1% | 27.4% | 12.0% | | | |
| **Non-Attriter** | N | 40 | 338 | 53 | 431 | | | |
| | % | 88.9% | 90.9% | 72.6% | 88.0% | | | |

This indicates that analysis of treatment effects for Hmong children, in particular, should be interpreted carefully, with attention to the demographic characteristics and learning outcomes of the 27.4% of Hmong children who failed to participate during all three assessments.

Although the sample sizes are small and not all findings are significant, negative effect sizes on covariates and learning outcomes for attriters by ethnicity are consistent. For example, the Hmong attriters had a lower baseline literacy domain score than non-attriters, at .06 compared to .09 (*d*=-0.38). The Hmong attriters were also older (*d*=0.70), and their mothers less educated (*d*=-0.67), than the children who were present at all waves.

### 5.4.3. Analysis of Baseline Equivalence

The summer pre-primary treatment versus control RCT was compromised when actual assignment varied from original assignment in non-random ways, as described in the **Section 4.3.3** on adherence to sampling assignment. As such, the study is subject to IES stipulations around compromised RCTs or QEDs, namely that baseline equivalence must be established before proceeding with analysis of outcomes, and that lack of baseline equivalence must be corrected for using statistical adjustment such as ordinary least squares (OLS) regression with covariates if the effect size of the difference falls within advised limits (IES, 2020).

The final, analytical sample included only non-Lao, non-attriter children who complied with their original treatment status. This ultimately resulted in a sample of

202 children in the treatment and 171 children in the control group for a total of 391 who were included in regression analyses (**Table 10**).

**Table 10: Reduction in Sample from Original to Analytic Sample**

| N | Original Assignment | Compliers | Non-Attriters | Non-Lao Only |
|---|---|---|---|---|
| **Control** | 307 | 247 | 209 | 179 |
| **Treatment** | 366 | 243 | 222 | 212 |
| ***Total*** | *673* | *490* | *431* | *391* |

Using IES (2020) guidance on establishing baseline equivalence using a key baseline outcome indicator, a Cohen's *d* effect size for difference between the treatment and control group on the baseline literacy domain score variable were examined. An effect size of lower than $d$=.05 is considered inconsequential; between $d$=.05 and $d \leq$.25 is deemed to require statistical adjustment; and larger than $d$=.25 is considered to invalidate the study design because the groups are too unequal at baseline for statistical adjustment to be adequate (Ibid.). **Table 11** demonstrates the effect size in the present study, indicating that the difference on the literacy domain score at baseline for the summer pre-primary treatment versus control comparison was non-significant and had an effect size of .07. This falls within the range that requires statistical adjustment.

**Table 11: Literacy Domain Score Baseline Equivalence, Treatment vs. Control**

| | N[15] | Mean | SD | t | df | *p* | *d* |
|---|---|---|---|---|---|---|---|
| **Control** | 177 | .10 | .09 | 0.713 | 385 | .476 | 0.07 |
| **Treatment** | 210 | .09 | .09 | | | | |

After finalizing the sample as in the table above, gain scores were subjected to multi-level regression analysis to control for covariates that were found to lack equivalence

[15] Numbers do not sum to 391 because some children lacked data on the literacy domain score.

at baseline, to identify how those covariates had affected the significance and magnitude of the gains. Potential covariates, including age, sex, ethnicity, mother's education, and socio-economic status were first analyzed for mean differences between the summer pre-primary control and treatment groups using *t* or chi square statistics and for effect sizes using Cohen's *d.* These covariates were selected based on child-level demographic characteristics found to be predictive of child development in similar LMICs (Pisani, Borisova, and Dowd, 2018) and aligned with an analysis of the construct validity of the MODEL tool used in the LEARN evaluation in Laos (Gomez, Brown, and Spier, 2020). The latter investigated the associations between the measurement outcomes and child demographic characteristics such as child age, sex, socio-economic status, and levels of parent education. The authors found, for example, that socio-economic status was strongly predictive of performance on the vocabulary and numeracy tasks, while mothers' education predicted outcomes on the literacy tasks related to phonological awareness, letter name knowledge, and listening comprehension.

As indicated in **Table 12** below, the effect sizes for age, % Hmong, mother's education, and SES at baseline required statistical adjustment and needed to be controlled for in the regression analysis.

**Table 12: Baseline Equivalence on Key Covariates, Summer Pre-primary Comparison and Treatment Groups**

| | | N | Mean | Mean Diff. | Std. Dev. | F or Chi Sq. | *t* | df | *p* | *d* |
|---|---|---|---|---|---|---|---|---|---|---|
| Age in Months | Control | 179 | 62.16 | -0.66 | 6.575 | 2.38 | -1.04 | 389 | .300 | -0.11 |
| | Treat. | 212 | 62.82 | | 5.994 | | | | | |
| % Female | Control | 179 | 50.80% | 1.70% | | 0.12 | | 1 | .726 | 0.04 |
| | Treat. | 212 | 49.10% | | | | | | | |
| % Hmong | Control | 179 | 11.20% | 4.40% | | 1.60 | | 1 | .206 | 0.13 |
| | Treat. | 212 | 15.60% | | | | | | | |
| **Mother's Ed. (0-3)** | **Control** | **171** | **1.09** | **0.29** | **0.906** | **1.38** | **3.03** | **371** | **.003** | **0.32** |
| | **Treat.** | **202** | **0.80** | | **0.938** | | | | | |
| **SES (0-16)** | **Control** | **179** | **6.32** | **0.62** | **2.615** | **9.18** | **2.18** | **389** | **.030** | **0.22** |
| | **Treat.** | **212** | **5.70** | | **3.076** | | | | | |

This study takes advantage of the longitudinal, repeated measures design to present gain scores overall and by ethnicity, as well as by the summer pre-primary control and treatment groups, between baseline and midline, midline and endline, and baseline and endline. Gain scores – which are calculated by subtracting a pre-test score from a post-test score – are accepted by IES as an analytical approach for compromised RCTs and QEDs (IES, 2020). They are also commonly used in similar studies in the international development sector, including numerous evaluations of the impact of school readiness and early grade reading programs that employ longitudinal, repeated measures designs.[16]

Gain scores are useful in studies where the treatment and control groups demonstrate baseline imbalance on key dependent variables, because they allow analysis of "which groups made greater or less progress over time than others, regardless of whether one of the groups began with a higher score" (Chemonics and

[16] See, for example, RTI International, 2016b; RTI International, 2010; Chemonics and School-to-School International, 2016; Dowd et. al., 2013; and World Bank, 2018b.

School to School International, 2017, p. 28). In addition, these types of scores are well suited to measuring the intent of the summer pre-primary model, which aims to provide children with a school readiness boost immediately before grade 1 to close equity gaps for disadvantaged children. Gain scores also help to identify how much that boost persisted through grade 1.

To calculate gain scores, attriters were excluded, as is common practice in similar studies with analogous analytical approaches (see, for example, World Bank, 2018b; Levesque, Bardack, and Chigeda, 2020; and School-to-School International, 2017a). Gains on all literacy-related variables (the scaled scores between 0 and 1) were calculated between T1-T2, T2-T3 and T1-T3 by subtracting the first score from the second in each case. Next, means were computed for the treatment and control groups on each gain score and used for the visual representations of gains in the tables and figures below as well as calculation of treatment effect sizes.

### 5.4.4. Mean Gains and Effect Sizes

This section first lays out an analysis of differences in mean gains between the treatment and control groups, as well as the effect sizes of those differences using Cohen's *d*. Baseline scores are also provided for informational purposes, as they constitute the starting point for any subsequent gains or losses. Visualization of the mean gains by treatment status is provided in **Figure 16** below, while numerical values are included in **Table 25** in **Annex 3.**

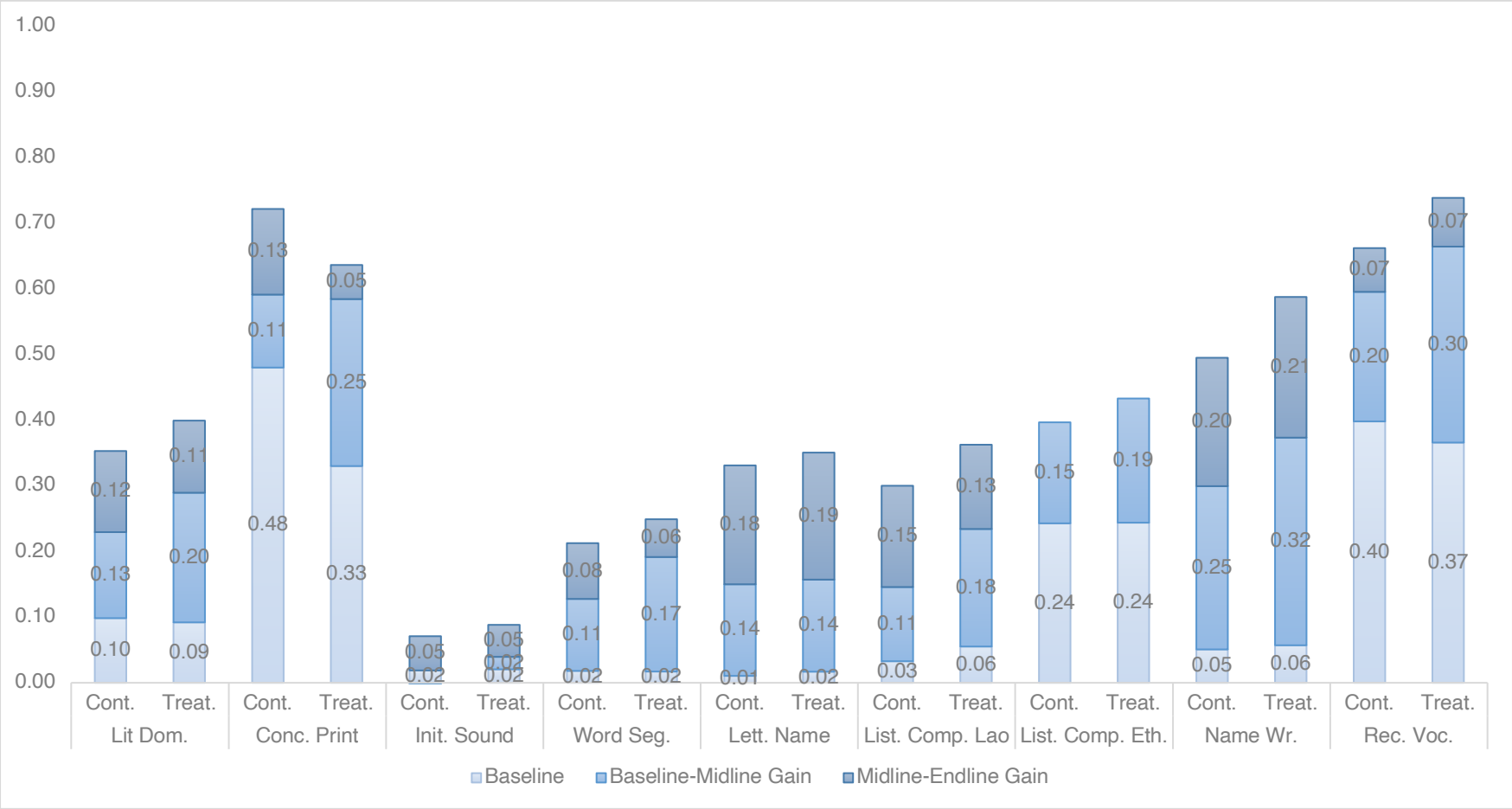
Most notable are the differences between treatment and control groups on literacy domain gains between baseline and midline and baseline and endline, with the treatment group making consistently greater gains; the concepts about print gains between baseline and midline; the listening comprehension (in Lao) gains between baseline and midline; and the receptive vocabulary gains between baseline and midline and midline and endline. These range from *d*=.24 to *d*=.42, all small effects using Cohen's (1988) taxonomy. There are also positive trends in the baseline-endline gains on concepts about print; the baseline-midline gains on word

segmentation; and the name writing gains between baseline and midline and midline and endline.

In general, these findings suggest that the intervention was most effective in giving children in the treatment group a school readiness boost between baseline and midline, but that there was some fadeout of effects between midline and endline. In several cases, that fadeout counteracted the initial boost and substantially reduced the overall baseline-endline effect sizes. In general, the findings also indicate limited impact of the intervention on the tasks associated with phonological awareness, letter knowledge, listening comprehension in ethnic languages, and name writing.

These findings will be further examined in **Section 5.4.6** through multi-level regression that controls for imbalanced baseline covariates and adjusts for school-level clustering.

**Figure 16: Mean Gain Scores, Control vs. Treatment[17]**




[17] Cont.=control; treat.=treatment; lit. dom.=literacy domain; conc. print=concepts about print; init. sound=initial sound; word seg.=word segmentation; lett. name=letter name knowledge; list. comp. Lao=listening comprehension in Lao; list. comp. ethnic=listening comprehension in ethnic languages; name writ.=name writing; rec.voc.=receptive vocabulary.

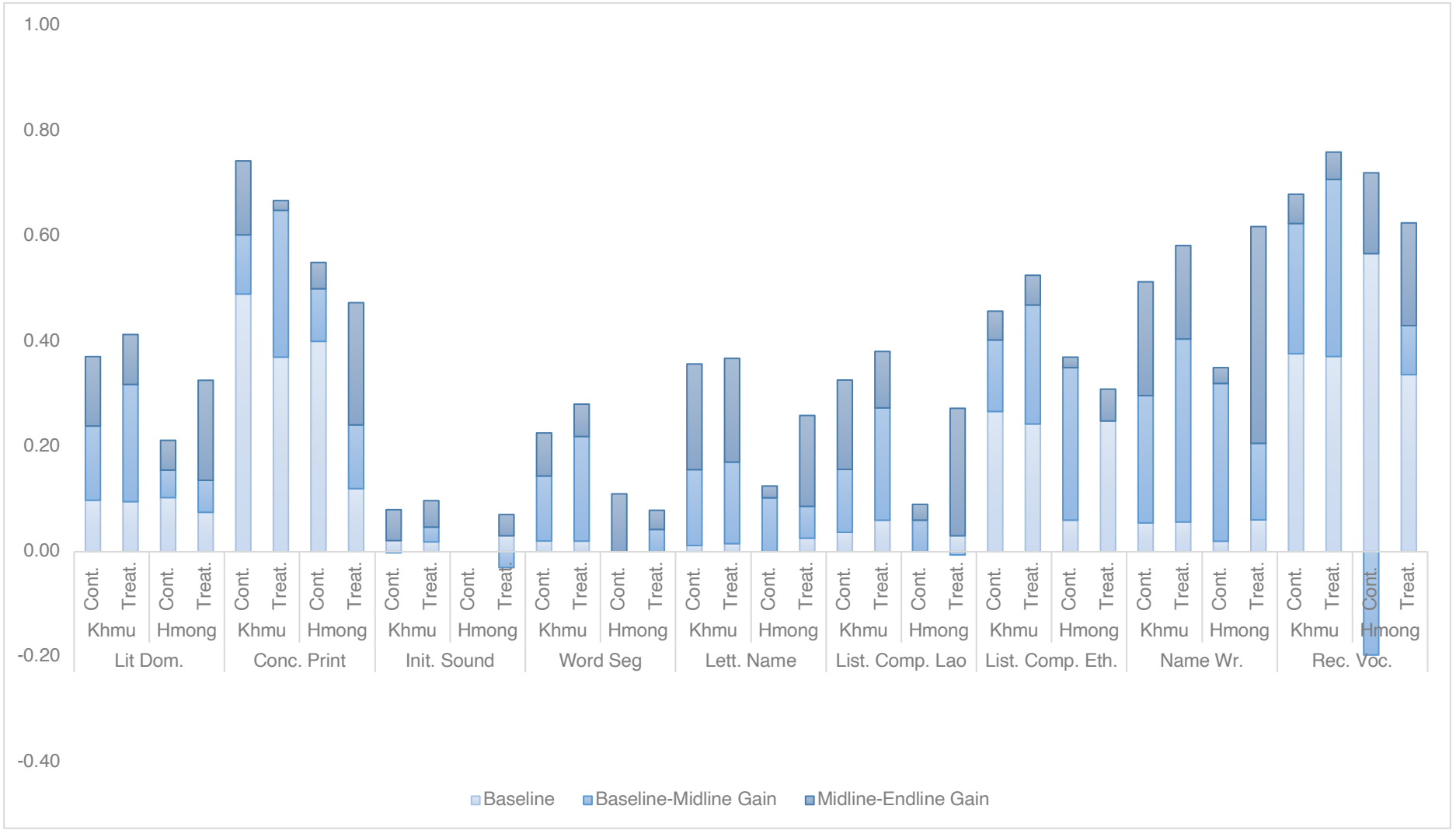
### 5.4.5. Mean Gains and Effect Sizes by Treatment and Ethnicity

An analysis of mean gains by treatment and ethnicity reveals additional nuances in the outcomes. This section compares effect sizes of the treatment-control differences *within* each ethnic group – Khmu control vs. Khmu treatment; and Hmong control vs. Hmong treatment – using mean gain scores and Cohen's *d*. Visualization of the mean gains by ethnicity is provided in **Figure 17** below and numerical values are in **Table 26** in **Annex 3.**

This analysis identifies substantial effects in areas such as the literacy domain, concepts about print, letter knowledge, listening comprehension (both in Lao and in ethnic languages), name writing, and receptive vocabulary tasks. These effects varied by ethnicity, but in general, the effect sizes were larger for Khmu children between baseline and midline but larger for Hmong children between midline and endline and baseline and endline. Using's Cohen's 1988 taxonomy, the baseline-midline effects for Khmu children would generally be considered small – for example, *d*=.45 for the literacy domain and *d*=.48 for concepts about print. In contrast, the effect sizes for Hmong children between midline and endline and baseline and endline are consistently medium to large. This includes *d*=.81 for the literacy domain gain between midline and endline and *d*=.80 between baseline and endline; *d*=.54 for concepts about print gains between midline and endline and *d*=.74 between baseline and endline; *d*=.59 for the midline-endline gain on letter knowledge; *d*=.72 for the midline-endline gain and *d*=.52 for the baseline-endline gain for listening comprehension; and the list goes on in **Table 26**. However, it is important to note that the sample sizes of Hmong children were small, so these effects should be interpreted cautiously.

As above, findings related to treatment effects by ethnicity will be further probed through multi-level regression analysis in **Section 5.4.6.**

**Figure 17: Mean Gain Scores, Control vs. Treatment, by Ethnicity**



### 5.4.6. Findings of Multilevel Regression

After exploring analysis of mean gains and effect sizes, the thesis now turns to multilevel regression to estimate treatment effects. All analysis reported in this section was conducted using the Stata 17 statistical package, with the *xtmixed* command for multilevel mixed-effects linear regression. A 2-level random intercept model was conducted where children (level 1) were nested within villages (level 2) to account for variation in the gain scores between villages, given that gain scores are correlated among children within villages. Cluster robust standard errors were calculated to provide unbiased estimation of the standard errors for model coefficients. The analysis was run separately to assess the dependent variables – gain scores – across 1) baseline and midline, 2) midline and endline and 3) baseline and endline. This analysis looked first at differences in gains between treatment and control groups. Second, the analysis utilized a treatment X ethnicity interaction term to explore the interaction between treatment status and Khmu or Hmong ethnicity with respect to learning gains.

The regression models were built by sequentially entering variables as follows:

- **Null Model:** Dependent variable (gain score), no independent variables
- **Model 1:** Null model + baseline covariates with effect sizes greater than d=.05[18]
- **Model 2:** Model 1 + a treatment dummy variable coded as summer pre-primary control group=0 and summer pre-primary treatment group=1
- **Model 3:** Model 2 + a treatment X ethnicity interaction term, with treated Hmong children=1 and all other children=0

Baseline covariates were unstandardized, resulting in unstandardized regression coefficients (*b)*. For ease of interpretation:

- A positive *b* for **treatment** indicates a positive overall treatment effect

[18] Refer to **Section 5.4.3** for further discussion on selection of these baseline covariates.

- A positive *b* for the **interaction** term indicates a positive effect of the treatment for Hmong children when compared to Khmu children
- A negative *b* for **treatment** indicates a negative overall treatment effect
- A negative *b* for the **interaction** term indicates a negative effect of the treatment for Hmong children when compared to Khmu children

Intra-class correlations are reported for the null models to understand the extent to which outcomes were correlated among children in the same villages, prior to any accounting for covariates or treatment status. Then, the regression outputs from Models 2 and 3 (not the Null Model or Model 1) are presented in tables because they represent the main thrust of the thesis' examination of outcomes by treatment group and by treatment and ethnicity. Statistically significant covariates from Model 2 are also described to explore the demographic factors that predict student performance.

#### 5.4.6.1. Literacy Domain

Intra-class correlations (ICCs) in the null models for the literacy domain gains were .297 between baseline and midline, .115 between midline and endline, and .264 between baseline and endline. As outlined in **Table 13** below, when examining baseline-midline literacy domain gains, there are significant estimated treatment effects overall in Model 2 (*b*=.075, standard error (SE)=.028, *p*=.008) and Model 3 (*b*=.096, SE=.032, *p*=.003). In other words, Model 3 predicts that children who participated in the treatment will have gained nearly .1 points more (out of 1) between baseline and midline than children who did not participate. There is no significant interaction between treatment and ethnicity in Model 3. Model 2 indicates significant associations between the gain score and the child's age, Hmong ethnicity, and mother's education level between baseline and midline.

No significant associations are found between treatment and gains, and no treatment-ethnicity interactions, between midline and endline or baseline and endline. There are, however, some significant associations between the gain score

and Hmong ethnicity and mother's education in Model 2 between baseline and endline.

## Table 13: Literacy Domain Score Regression Outputs[19]

**Baseline-Midline Gain**

| | Model 2 | | | | | Model 3 | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Fixed Effects** | ***b*** | **SE** | ***p*** | **[95% CI]** | | ***b*** | **SE** | ***p*** | **[95% CI]** | |
| Age | **.004** | **.001** | **.004** | **.001** | **.007** | **.004** | **.001** | **.003** | **.001** | **.006** |
| Hmong | **-.127** | **.045** | **.005** | **-.217** | **-.038** | -.030 | .078 | .701 | -.183 | .123 |
| Mother's Ed. | **.028** | **.011** | **.009** | **.007** | **.049** | **.029** | **.011** | **.006** | **.008** | **.051** |
| SES | .004 | .004 | .363 | -.004 | .012 | .004 | .004 | .350 | -.004 | .012 |
| Treatment | **.075** | **.028** | **.008** | **.019** | **.130** | **.096** | **.032** | **.003** | **.034** | **.159** |
| Treat. X Hmong | | | | | | -.159 | .090 | .078 | -.335 | .018 |
| Constant | -.153 | .084 | .069 | -.318 | .012 | -.157 | .080 | .050 | -.314 | .000 |
| **Random Effects** | | **Var.** | **SE** | **[95% CI]** | | | **Var.** | **SE** | **[95% CI]** | |
| Village-Level | | .006 | .002 | .004 | .010 | | .006 | .002 | .004 | .011 |
| Child-Level | | .023 | .002 | .018 | .028 | | .022 | .002 | .018 | .027 |
| | | **ICC** | **SE** | **[95% CI]** | | | **ICC** | **SE** | **[95% CI]** | |
| Village-Level ICC | | .214 | .051 | .131 | .330 | | .224 | .051 | .140 | .338 |

**Midline-Endline Gain**

| | Model 2 | | | | | Model 3 | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Fixed Effects** | ***b*** | **SE** | ***p*** | **[95% CI]** | | ***b*** | **SE** | ***p*** | **[95% CI]** | |
| Age | -.001 | .002 | .644 | -.005 | .003 | -.001 | .002 | .672 | -.004 | .003 |
| Hmong | .014 | .039 | .716 | -.062 | .091 | **-.078** | **.024** | **.001** | **-.126** | **-.030** |
| Mother's Ed. | .014 | .010 | .160 | -.005 | .033 | .012 | .010 | .201 | -.007 | .031 |
| SES | .000 | .003 | .910 | -.007 | .006 | -.001 | .004 | .780 | -.008 | .006 |
| Treatment | -.019 | .024 | .441 | -.067 | .029 | -.040 | .026 | .124 | -.091 | .011 |
| Treat. X Hmong | | | | | | **.162** | **.051** | **.001** | **.062** | **.261** |
| Constant | .170 | .132 | .198 | -.089 | .428 | .180 | .132 | .171 | -.078 | .438 |
| **Random Effects** | | **Var.** | **SE** | **[95% CI]** | | | **Var.** | **SE** | **[95% CI]** | |
| Village-Level | | .003 | .001 | .002 | .008 | | .003 | .001 | .001 | .007 |
| Child-Level | | .030 | .002 | .025 | .035 | | .030 | .002 | .025 | .035 |
| | | **ICC** | **SE** | **[95% CI]** | | | **ICC** | **SE** | **[95% CI]** | |
| Village-Level ICC | | .103 | .038 | .048 | .206 | | .079 | .038 | .030 | .191 |

**Baseline-Endline Gain**

| | Model 2 | | | | | Model 3 | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Fixed Effects** | ***b*** | **SE** | ***p*** | **[95% CI]** | | ***b*** | **SE** | ***p*** | **[95% CI]** | |
| Age | .003 | .002 | .125 | -.001 | .008 | .003 | .002 | .125 | -.001 | .008 |
| Hmong | **-.118** | **.047** | **.013** | **-.211** | **-.025** | -.119 | .064 | .065 | -.245 | .008 |
| Mother's Ed. | **.042** | **.010** | **.000** | **.023** | **.061** | **.042** | **.010** | **.000** | **.023** | **.061** |
| SES | .004 | .004 | .276 | -.004 | .013 | .004 | .004 | .279 | -.004 | .013 |
| Treatment | .057 | .033 | .088 | -.008 | .122 | .057 | .037 | .130 | -.017 | .130 |
| Treat. X Hmong | | | | | | .001 | .094 | .991 | -.183 | .185 |
| Constant | -.002 | .141 | .987 | -.278 | .273 | -.002 | .141 | .987 | -.279 | .274 |
| **Random Effects** | | **Var.** | **SE** | **[95% CI]** | | | **Var.** | **SE** | **[95% CI]** | |
| Village-Level | | .009 | .002 | .005 | .014 | | .009 | .002 | .005 | .015 |
| Child-Level | | .037 | .003 | .031 | .044 | | .037 | .003 | .031 | .044 |
| | | **ICC** | **SE** | **[95% CI]** | | | **ICC** | **SE** | **[95% CI]** | |
| Village-Level ICC | | .192 | .045 | .118 | .296 | | .192 | .047 | .117 | .300 |

[19] Mother's ed.=mother's education; SES=socio-economic index; treat. X Hmong=interaction term; ICC=intra-class correlation; *b*=unstandardized beta; SE=cluster robust standard error; CI=confidence interval; var.=variance.

#### 5.4.6.2. Concepts about Print

ICCs in the null models for the concepts about print gains were .120 between baseline and midline, .111 between midline and endline, and .121 between baseline and endline. **Table 14** below captures the regression findings for gains on the concepts about print task. When examining baseline-midline gains, Model 2 identifies significant associations between children's age and the baseline-midline gain. In addition, Model 2 identifies significant treatment effects (*b*=.146, SE=.043, *p*=.001) and Model 3 echoes this significant treatment effect (*b*=.169, SE=.046, *p*=.000). There is no significant interaction between treatment and ethnicity in Model 3.

For the midline-endline gain, Model 3 finds a significant negative treatment effect overall (*b*=-.117, SE=.050, *p*=.018) but a positive treatment effect for Hmong children (*b*=.289, SE=.133, *p*=.030). Once again, there is a significant association between children's age and the gain score in Model 2.

With regard to overall baseline-endline gains, no significant associations are predicted between concepts about print gains and demographic characteristics, treatment status, or the treatment-ethnicity interaction.

**Table 14: Concepts about Print Regression Outputs**

**Baseline-Midline Gain**

| | Model 2 | | | | | Model 3 | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Fixed Effects** | ***b*** | **SE** | ***p*** | **[95% CI]** | | ***b*** | **SE** | ***p*** | **[95% CI]** | |
| Age | **.006** | **.003** | **.036** | **.000** | **.012** | **.006** | **.003** | **.035** | **.000** | **.012** |
| Hmong | -.105 | .068 | .121 | -.238 | .028 | -.007 | .142 | .960 | -.286 | .272 |
| Mother's Ed. | .003 | .020 | .863 | -.036 | .043 | .006 | .020 | .786 | -.034 | .046 |
| SES | .000 | .007 | .986 | -.014 | .014 | .001 | .007 | .943 | -.013 | .014 |
| Treatment | **.146** | **.043** | **.001** | **.061** | **.230** | **.169** | **.046** | **.000** | **.079** | **.259** |
| Treat. X Hmong | | | | | | -.167 | .150 | .266 | -.461 | .127 |
| Constant | -.266 | .192 | .166 | -.643 | .110 | -.277 | .193 | .150 | -.654 | .100 |
| **Random Effects** | | **Var.** | **SE** | **[95% CI]** | | | **Var.** | **SE** | **[95% CI]** | |
| Village-Level | | .005 | .005 | .001 | .039 | | .005 | .005 | .001 | .039 |
| Child-Level | | .106 | .011 | .087 | .131 | | .106 | .011 | .086 | .130 |
| | | **ICC** | **SE** | **[95% CI]** | | | **ICC** | **SE** | **[95% CI]** | |
| Village-Level ICC | | .045 | .047 | .006 | .284 | | .043 | .046 | .005 | .287 |

**Midline-Endline Gain**

| | Model 2 | | | | | Model 3 | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Fixed Effects** | ***b*** | **SE** | ***p*** | **[95% CI]** | | ***b*** | **SE** | ***p*** | **[95% CI]** | |
| Age | **-.006** | **.003** | **.060** | **-.013** | **.000** | -.006 | .003 | .062 | -.013 | .000 |
| Hmong | .098 | .077 | .202 | -.053 | .250 | -.067 | .117 | .566 | -.295 | .162 |
| Mother's Ed. | .026 | .022 | .228 | -.017 | .069 | .024 | .022 | .287 | -.020 | .068 |
| SES | .008 | .007 | .249 | -.005 | .021 | .007 | .007 | .300 | -.006 | .020 |
| Treatment | -.079 | .046 | .088 | -.169 | .012 | **-.117** | **.050** | **.018** | **-.214** | **-.020** |
| Treat. X Hmong | | | | | | **.289** | **.133** | **.030** | **.028** | **.551** |
| Constant | 0.44 | 0.22 | 0.05 | 0 | 0.88 | .456 | .226 | .043 | .014 | .899 |
| **Random Effects** | | **Var.** | **SE** | **[95% CI]** | | | **Var.** | **SE** | **[95% CI]** | |
| Village-Level | | .009 | .007 | .002 | .044 | | .007 | .008 | .001 | .069 |
| Child-Level | | .116 | .010 | .098 | .136 | | .116 | .010 | .098 | .136 |
| | | **ICC** | **SE** | **[95% CI]** | | | **ICC** | **SE** | **[95% CI]** | |
| Village-Level ICC | | .074 | .056 | .016 | .285 | | .054 | .063 | .005 | .385 |

**Baseline-Endline Gain**

| | Model 2 | | | | | Model 3 | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Fixed Effects** | ***b*** | **SE** | ***p*** | **[95% CI]** | | ***b*** | **SE** | ***p*** | **[95% CI]** | |
| Age | .000 | .003 | .997 | -.007 | .007 | .000 | .003 | .984 | -.007 | .007 |
| Hmong | -.005 | .063 | .932 | -.130 | .119 | -.071 | .059 | .225 | -.187 | .044 |
| Mother's Ed. | .031 | .018 | .089 | -.005 | .066 | .030 | .018 | .105 | -.006 | .066 |
| SES | .006 | .007 | .360 | -.007 | .020 | .006 | .007 | .373 | -.007 | .020 |
| Treatment | .068 | .044 | .123 | -.018 | .154 | .052 | .051 | .309 | -.048 | .153 |
| Treat. X Hmong | | | | | | .114 | .109 | .295 | -.100 | .328 |
| Constant | 0.18 | 0.21 | 0.41 | -0.24 | 0.59 | .181 | .214 | .396 | -.237 | .600 |
| **Random Effects** | | **Var.** | **SE** | **[95% CI]** | | | **Var.** | **SE** | **[95% CI]** | |
| Village-Level | | .009 | .005 | .003 | .029 | | .008 | .005 | .002 | .030 |
| Child-Level | | .102 | .010 | .084 | .123 | | .102 | .010 | .085 | .123 |
| | | **ICC** | **SE** | **[95% CI]** | | | **ICC** | **SE** | **[95% CI]** | |
| Village-Level ICC | | .079 | .048 | .023 | .237 | | .072 | .049 | .018 | .245 |

#### 5.4.6.3. Initial Sound Identification, Word Segmentation, and Letter Name Knowledge

The initial sound identification, word segmentation, and letter name knowledge tasks exhibited similar findings and are presented together to avoid repetition.

ICCs in the null models for the initial sound identification gains were .011 between baseline and midline, .016 between midline and endline, and .054 between baseline and endline. For word segmentation, the ICCs were .298 between baseline and midline, .183 between midline and endline, and .117 between baseline and endline. For letter name knowledge, the ICCs were .278, .103, and .260, respectively.

As captured in **Table 15** and **Table 16** below, for the initial sound identification and word segmentation tasks, no significant associations are predicted between gain scores and child demographic characteristics, treatment status, or treatment-ethnicity interaction. This holds true for the baseline-midline, midline-endline, and baseline-endline gains alike. There are some significant associations between gains and Hmong ethnicity and mother's education between different data collection waves in some of the Model 2 regression outputs.

**Table 17** also captures a lack of significant predicted associations between treatment status and the treatment-ethnicity interaction for the letter name knowledge task but does identify some significant associations with characteristics such as age, Hmong ethnicity, and mother's education between different data collection waves in some of the Model 2 findings.

## Table 15: Initial Sound Identification Regression Outputs

**Baseline-Midline Gain**

| | Model 2 | | | | | Model 3 | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Fixed Effects** | ***b*** | **SE** | ***p*** | **[95% CI]** | | ***b*** | **SE** | ***p*** | **[95% CI]** | |
| Age | .000 | .001 | .844 | -.002 | .002 | .000 | .001 | .867 | -.002 | .002 |
| Hmong | -.027 | .022 | .212 | -.070 | .015 | .017 | .015 | .258 | -.013 | .047 |
| Mother's Ed. | **.029** | **.011** | **.007** | **.008** | **.050** | **.030** | **.011** | **.005** | **.009** | **.051** |
| SES | .000 | .003 | .867 | -.006 | .005 | .000 | .003 | .935 | -.006 | .005 |
| Treatment | .026 | .018 | .152 | -.009 | .061 | .036 | .021 | .080 | -.004 | .077 |
| Treat. X Hmong | | | | | | -.077 | .044 | .078 | -.162 | .008 |
| Constant | -0.04 | 0.08 | 0.62 | -0.19 | 0.1105 | -.043 | .075 | .563 | -.189 | .103 |
| **Random Effects** | | **Var.** | **SE** | **[95% CI]** | | | **Var.** | **SE** | **[95% CI]** | |
| Village-Level | | .000 | .001 | .000 | 42.006 | | .000 | .001 | .000 | .070 |
| Child-Level | | .026 | .007 | .016 | .042 | | .026 | .006 | .016 | .042 |
| | | **ICC** | **SE** | **[95% CI]** | | | **ICC** | **SE** | **[95% CI]** | |
| Village-Level ICC | | .005 | .034 | .000 | .999 | | .012 | .034 | .000 | .766 |

**Midline-Endline Gain**

| | Model 2 | | | | | Model 3 | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Fixed Effects** | ***b*** | **SE** | ***p*** | **[95% CI]** | | ***b*** | **SE** | ***p*** | **[95% CI]** | |
| Age | .002 | .002 | .371 | -.002 | .005 | .002 | .002 | .367 | -.002 | .005 |
| Hmong | -.018 | .021 | .384 | -.058 | .022 | **-.043** | **.021** | **.038** | **-.084** | **-.002** |
| Mother's Ed. | .021 | .013 | .107 | -.005 | .047 | .021 | .013 | .116 | -.005 | .047 |
| SES | .005 | .006 | .388 | -.006 | .016 | .005 | .006 | .417 | -.006 | .016 |
| Treatment | .002 | .024 | .940 | -.044 | .048 | -.004 | .027 | .887 | -.058 | .050 |
| Treat. X Hmong | | | | | | .041 | .044 | .349 | -.045 | .128 |
| Constant | -0.09 | 0.11 | 0.42 | -0.32 | 0.1324 | -.089 | .116 | .439 | -.316 | .137 |
| **Random Effects** | | **Var.** | **SE** | **[95% CI]** | | | **Var.** | **SE** | **[95% CI]** | |
| Village-Level | | .000 | .000 | .000 | .000 | | .000 | .000 | .000 | .000 |
| Child-Level | | .053 | .009 | .037 | .075 | | .053 | .009 | .037 | .075 |
| | | **ICC** | **SE** | **[95% CI]** | | | **ICC** | **SE** | **[95% CI]** | |
| Village-Level ICC | | .000 | .000 | .000 | .000 | | .000 | .000 | .000 | .000 |

**Baseline-Endline Gain**

| | Model 2 | | | | | Model 3 | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Fixed Effects** | ***b*** | **SE** | ***p*** | **[95% CI]** | | ***b*** | **SE** | ***p*** | **[95% CI]** | |
| Age | .002 | .002 | .282 | -.002 | .005 | .002 | .002 | .289 | -.002 | .005 |
| Hmong | **-.047** | **.023** | **.042** | **-.092** | **-.002** | -.023 | .020 | .243 | -.062 | .016 |
| Mother's Ed. | **.049** | **.015** | **.001** | **.019** | **.079** | **.050** | **.015** | **.001** | **.019** | **.080** |
| SES | .006 | .005 | .185 | -.003 | .016 | .007 | .005 | .179 | -.003 | .016 |
| Treatment | .033 | .025 | .186 | -.016 | .082 | .039 | .029 | .184 | -.018 | .095 |
| Treat. X Hmong | | | | | | -.039 | .047 | .406 | -.132 | .053 |
| Constant | -0.15 | 0.1 | 0.14 | -0.34 | 0.048 | -.149 | .098 | .130 | -.341 | .044 |
| **Random Effects** | | **Var.** | **SE** | **[95% CI]** | | | **Var.** | **SE** | **[95% CI]** | |
| Village-Level | | .000 | .000 | .000 | .000 | | .000 | .000 | .000 | .000 |
| Child-Level | | .057 | .010 | .041 | .080 | | .057 | .010 | .041 | .080 |
| | | **ICC** | **SE** | **[95% CI]** | | | **ICC** | **SE** | **[95% CI]** | |
| Village-Level ICC | | .000 | .000 | .000 | .000 | | .000 | .000 | .000 | .000 |

## Table 16: Word Segmentation Regression Outputs

**Baseline-Midline Gain**

| | Model 2 | | | | | Model 3 | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Fixed Effects** | ***b*** | **SE** | ***p*** | **[95% CI]** | | ***b*** | **SE** | ***p*** | **[95% CI]** | |
| Age | .002 | .002 | .292 | -.002 | .007 | .002 | .002 | .302 | -.002 | .007 |
| Hmong | **-.115** | **.047** | **.014** | **-.206** | **-.023** | **-.085** | **.039** | **.030** | **-.161** | **-.008** |
| Mother's Ed. | .024 | .023 | .288 | -.020 | .069 | .024 | .023 | .282 | -.020 | .069 |
| SES | .001 | .007 | .890 | -.012 | .014 | .001 | .007 | .882 | -.012 | .014 |
| Treatment | .074 | .059 | .213 | -.042 | .190 | .080 | .065 | .217 | -.047 | .208 |
| Treat. X Hmong | | | | | | -.047 | .077 | .538 | -.198 | .103 |
| Constant | -0.1 | 0.14 | 0.66 | -0.3 | 0.22 | -.065 | .142 | .649 | -.343 | .213 |
| **Random Effects** | | **Var.** | **SE** | **[95% CI]** | | | **Var.** | **SE** | **[95% CI]** | |
| Village-Level | | .031 | .012 | .014 | .067 | | .031 | .012 | .014 | .067 |
| Child-Level | | .072 | .014 | .049 | .105 | | .072 | .014 | .049 | .105 |
| | | **ICC** | **SE** | **[95% CI]** | | | **ICC** | **SE** | **[95% CI]** | |
| Village-Level ICC | | .300 | .095 | .150 | .510 | | .300 | .095 | .151 | .509 |

**Midline-Endline Gain**

| | Model 2 | | | | | Model 3 | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Fixed Effects** | ***b*** | **SE** | ***p*** | **[95% CI]** | | ***b*** | **SE** | ***p*** | **[95% CI]** | |
| Age | .002 | .003 | .582 | -.005 | .008 | .002 | .003 | .595 | -.005 | .008 |
| Hmong | -.002 | .072 | .983 | -.144 | .140 | .053 | .139 | .705 | -.220 | .326 |
| Mother's Ed. | .016 | .023 | .485 | -.029 | .061 | .017 | .023 | .463 | -.028 | .062 |
| SES | -.001 | .008 | .890 | -.016 | .014 | -.001 | .008 | .914 | -.016 | .015 |
| Treatment | -.035 | .068 | .607 | -.169 | .099 | -.023 | .074 | .758 | -.168 | .123 |
| Treat. X Hmong | | | | | | -.090 | .161 | .577 | -.405 | .225 |
| Constant | -0 | 0.22 | 0.93 | -0.4 | 0.41 | -.023 | .219 | .916 | -.453 | .407 |
| **Random Effects** | | **Var.** | **SE** | **[95% CI]** | | | **Var.** | **SE** | **[95% CI]** | |
| Village-Level | | .035 | .013 | .016 | .073 | | .035 | .013 | .016 | .074 |
| Child-Level | | .165 | .019 | .132 | .205 | | .164 | .019 | .132 | .205 |
| | | **ICC** | **SE** | **[95% CI]** | | | **ICC** | **SE** | **[95% CI]** | |
| Village-Level ICC | | .173 | .058 | .087 | .316 | | .175 | .059 | .087 | .320 |

**Baseline-Endline Gain**

| | Model 2 | | | | | Model 3 | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Fixed Effects** | ***b*** | **SE** | ***p*** | **[95% CI]** | | ***b*** | **SE** | ***p*** | **[95% CI]** | |
| Age | .005 | .003 | .143 | -.002 | .011 | .005 | .003 | .147 | -.002 | .011 |
| Hmong | -.117 | .063 | .064 | -.241 | .007 | -.049 | .127 | .699 | -.298 | .200 |
| Mother's Ed. | **.047** | **.020** | **.016** | **.009** | **.086** | **.048** | **.020** | **.013** | **.010** | **.087** |
| SES | .005 | .008 | .549 | -.011 | .020 | .005 | .008 | .521 | -.010 | .020 |
| Treatment | .046 | .051 | .365 | -.053 | .146 | .062 | .057 | .277 | -.050 | .174 |
| Treat. X Hmong | | | | | | -.116 | .141 | .410 | -.392 | .160 |
| Constant | -0.2 | 0.21 | 0.45 | -0.6 | 0.26 | -.168 | .212 | .429 | -.584 | .248 |
| **Random Effects** | | **Var.** | **SE** | **[95% CI]** | | | **Var.** | **SE** | **[95% CI]** | |
| Village-Level | | .013 | .006 | .005 | .032 | | .013 | .006 | .005 | .033 |
| Child-Level | | .136 | .017 | .106 | .174 | | .136 | .017 | .106 | .174 |
| | | **ICC** | **SE** | **[95% CI]** | | | **ICC** | **SE** | **[95% CI]** | |
| Village-Level ICC | | .085 | .043 | .030 | .217 | | .085 | .044 | .030 | .221 |

## Table 17: Letter Name Knowledge Regression Outputs

**Baseline-Midline Gain**

| | Model 2 | | | | | Model 3 | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Fixed Effects** | ***b*** | **SE** | ***p*** | **[95% CI]** | | ***b*** | **SE** | ***p*** | **[95% CI]** | |
| Age | **.004** | **.002** | **.014** | **.001** | **.007** | **.004** | **.002** | **.015** | **.001** | **.007** |
| Hmong | **-.084** | **.041** | **.039** | **-.164** | **-.004** | .009 | .056 | .870 | -.101 | .119 |
| Mother's Ed. | **.056** | **.016** | **.000** | **.025** | **.087** | **.057** | **.016** | **.000** | **.026** | **.088** |
| SES | .006 | .006 | .303 | -.006 | .018 | .006 | .006 | .290 | -.005 | .018 |
| Treatment | .014 | .044 | .755 | -.073 | .100 | .034 | .049 | .485 | -.062 | .130 |
| Treat. X Hmong | | | | | | **-.148** | **.073** | **.043** | **-.292** | **-.005** |
| Constant | -0.2 | 0.1 | 0.06 | -0.4 | 0 | -.195 | .097 | .045 | -.386 | -.005 |
| **Random Effects** | | **Var.** | **SE** | **[95% CI]** | | | **Var.** | **SE** | **[95% CI]** | |
| Village-Level | | 0.02 | 0.01 | 0.01 | 0.03 | | .018 | .006 | .010 | .033 |
| Child-Level | | .046 | .007 | .034 | .063 | | .046 | .007 | .034 | .063 |
| | | **ICC** | **SE** | **[95% CI]** | | | **ICC** | **SE** | **[95% CI]** | |
| Village-Level ICC | | .279 | .070 | .164 | .434 | | .280 | .072 | .162 | .439 |

**Midline-Endline Gain**

| | Model 2 | | | | | Model 3 | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Fixed Effects** | ***b*** | **SE** | ***p*** | **[95% CI]** | | ***b*** | **SE** | ***p*** | **[95% CI]** | |
| Age | -.001 | .003 | .610 | -.007 | .004 | -.001 | .003 | .646 | -.007 | .004 |
| Hmong | -.092 | .052 | .080 | -.194 | .011 | **-.168** | **.038** | **.000** | **-.243** | **-.093** |
| Mother's Ed. | .005 | .016 | .767 | -.027 | .036 | .004 | .016 | .807 | -.027 | .035 |
| SES | .002 | .005 | .741 | -.009 | .012 | .001 | .005 | .785 | -.009 | .012 |
| Treatment | .014 | .041 | .736 | -.066 | .094 | -.004 | .044 | .936 | -.089 | .082 |
| Treat. X Hmong | | | | | | .132 | .083 | .111 | -.030 | .293 |
| Constant | 0.26 | 0.2 | 0.19 | -0.1 | 0.65 | .267 | .199 | .179 | -.122 | .656 |
| **Random Effects** | | **Var.** | **SE** | **[95% CI]** | | | **Var.** | **SE** | **[95% CI]** | |
| Village-Level | | .010 | .005 | .004 | .025 | | .009 | .005 | .003 | .025 |
| Child-Level | | .077 | .007 | .064 | .093 | | .078 | .007 | .065 | .093 |
| | | **ICC** | **SE** | **[95% CI]** | | | **ICC** | **SE** | **[95% CI]** | |
| Village-Level ICC | | .112 | .049 | .045 | .249 | | .102 | .050 | .037 | .248 |

**Baseline-Endline Gain**

| | Model 2 | | | | | Model 3 | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Fixed Effects** | ***b*** | **SE** | ***p*** | **[95% CI]** | | ***b*** | **SE** | ***p*** | **[95% CI]** | |
| Age | .003 | .003 | .442 | -.004 | .010 | .003 | .003 | .448 | -.004 | .009 |
| Hmong | **-.181** | **.059** | **.002** | **-.298** | **-.065** | **-.161** | **.079** | **.043** | **-.316** | **-.005** |
| Mother's Ed. | **.060** | **.018** | **.001** | **.024** | **.096** | **.060** | **.018** | **.001** | **.024** | **.096** |
| SES | .007 | .007 | .314 | -.007 | .021 | .007 | .007 | .310 | -.007 | .021 |
| Treatment | .024 | .059 | .686 | -.092 | .140 | .029 | .067 | .669 | -.102 | .159 |
| Treat. X Hmong | | | | | | -.033 | .113 | .770 | -.255 | .189 |
| Constant | 0.08 | 0.23 | 0.73 | -0.4 | 0.53 | .078 | .232 | .735 | -.376 | .532 |
| **Random Effects** | | **Var.** | **SE** | **[95% CI]** | | | **Var.** | **SE** | **[95% CI]** | |
| Village-Level | | .031 | .008 | .019 | .051 | | .032 | .008 | .019 | .051 |
| Child-Level | | .096 | .007 | .083 | .112 | | .096 | .007 | .083 | .112 |
| | | **ICC** | **SE** | **[95% CI]** | | | **ICC** | **SE** | **[95% CI]** | |
| Village-Level ICC | | .246 | .053 | .157 | .363 | | .247 | .054 | .157 | .365 |

#### 5.4.6.4. Listening Comprehension in Lao

ICCs in the null models for the listening comprehension in Lao gains were .188 between baseline and midline, .074 between midline and endline, and .146 between baseline and endline. Model 3 in the multilevel regression analysis for listening comprehension in Lao (**Table 18**) predicts a significant association between treatment and gain scores between baseline and midline (*b*=.114, SE=.047, *p*=.016) and a significant treatment-ethnicity interaction favoring Khmu children (*b*=-.268, SE=.089, *p*=.003). Demographic characteristics are also significantly associated with gains between different waves in Model 2 for a few variables, including age, Hmong ethnicity, and mother's education. However, no significant associations are found between treatment and gains, and no treatment-ethnicity interactions, between midline and endline or baseline and endline.

## Table 18: Listening Comprehension in Lao Regression Outputs

**Baseline-Midline Gain**

| | Model 2 | | | | | Model 3 | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Fixed Effects** | ***b*** | **SE** | ***p*** | **[95% CI]** | | ***b*** | **SE** | ***p*** | **[95% CI]** | |
| Age | **.008** | **.002** | **.002** | **.003** | **.013** | **.007** | **.002** | **.002** | **.003** | **.012** |
| Hmong | **-.203** | **.063** | **.001** | **-.326** | **-.079** | -.042 | .057 | .464 | -.153 | .070 |
| Mother's Ed. | .018 | .016 | .263 | -.014 | .050 | .020 | .016 | .207 | -.011 | .052 |
| SES | .003 | .006 | .563 | -.008 | .015 | .004 | .006 | .520 | -.008 | .015 |
| Treatment | .077 | .042 | .066 | -.005 | .158 | **.114** | **.047** | **.016** | **.021** | **.206** |
| Treat. X Hmong | | | | | | **-.268** | **.089** | **.003** | **-.442** | **-.094** |
| Constant | -0.38 | 0.15 | 0.01 | -0.67 | -0.08 | -.385 | .147 | .009 | -.673 | -.098 |
| **Random Effects** | | **Var.** | **SE** | **[95% CI]** | | | **Var.** | **SE** | **[95% CI]** | |
| Village-Level | | .013 | .005 | .007 | .027 | | .014 | .005 | .007 | .027 |
| Child-Level | | .058 | .007 | .046 | .074 | | .057 | .007 | .045 | .072 |
| | | **ICC** | **SE** | **[95% CI]** | | | **ICC** | **SE** | **[95% CI]** | |
| Village-Level ICC | | .183 | .061 | .092 | .332 | | .191 | .062 | .097 | .342 |

**Midline-Endline Gain**

| | Model 2 | | | | | Model 3 | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Fixed Effects** | ***b*** | **SE** | ***p*** | **[95% CI]** | | ***b*** | **SE** | ***p*** | **[95% CI]** | |
| Age | -.001 | .003 | .851 | -.007 | .006 | .000 | .003 | .885 | -.006 | .006 |
| Hmong | -.017 | .074 | .818 | -.163 | .128 | **-.152** | **.074** | **.040** | **-.297** | **-.007** |
| Mother's Ed. | .026 | .016 | .100 | -.005 | .056 | .023 | .016 | .137 | -.007 | .054 |
| SES | -.006 | .006 | .313 | -.016 | .005 | -.007 | .005 | .233 | -.017 | .004 |
| Treatment | -.035 | .036 | .337 | -.107 | .036 | -.068 | .036 | .056 | -.137 | .002 |
| Treat. X Hmong | | | | | | **.242** | **.114** | **.033** | **.019** | **.465** |
| Constant | 0.21 | 0.21 | 0.32 | -0.2 | 0.61 | .223 | .208 | .283 | -.184 | .631 |
| **Random Effects** | | **Var.** | **SE** | **[95% CI]** | | | **Var.** | **SE** | **[95% CI]** | |
| Village-Level | | .006 | .003 | .002 | .018 | | .004 | .003 | .001 | .015 |
| Child-Level | | .083 | .006 | .071 | .096 | | .083 | .006 | .071 | .097 |
| | | **ICC** | **SE** | **[95% CI]** | | | **ICC** | **SE** | **[95% CI]** | |
| Village-Level ICC | | .066 | .035 | .023 | .178 | | .044 | .030 | .011 | .158 |

**Baseline-Endline Gain**

| | Model 2 | | | | | Model 3 | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Fixed Effects** | ***b*** | **SE** | ***p*** | **[95% CI]** | | ***b*** | **SE** | ***p*** | **[95% CI]** | |
| Age | **.007** | **.003** | **.023** | **.001** | **.013** | **.007** | **.003** | **.022** | **.001** | **.013** |
| Hmong | **-.214** | **.068** | **.002** | **-.347** | **-.081** | **-.204** | **.048** | **.000** | **-.299** | **-.110** |
| Mother's Ed. | **.045** | **.017** | **.007** | **.013** | **.077** | **.045** | **.017** | **.007** | **.012** | **.078** |
| SES | .001 | .006 | .925 | -.010 | .011 | .001 | .006 | .925 | -.010 | .012 |
| Treatment | .041 | .045 | .361 | -.047 | .130 | .044 | .048 | .361 | -.050 | .138 |
| Treat. X Hmong | | | | | | -.017 | .121 | .887 | -.254 | .220 |
| Constant | -0.18 | 0.2 | 0.38 | -0.58 | 0.22 | -.180 | .203 | .375 | -.577 | .218 |
| **Random Effects** | | **Var.** | **SE** | **[95% CI]** | | | **Var.** | **SE** | **[95% CI]** | |
| Village-Level | | .013 | .004 | .007 | .025 | | .013 | .005 | .007 | .026 |
| Child-Level | | .093 | .007 | .081 | .107 | | .093 | .007 | .080 | .107 |
| | | **ICC** | **SE** | **[95% CI]** | | | **ICC** | **SE** | **[95% CI]** | |
| Village-Level ICC | | .125 | .038 | .068 | .219 | | .126 | .040 | .067 | .226 |

#### 5.4.6.5. Listening Comprehension in Ethnic Languages

Regression analysis of listening comprehension in ethnic languages is only possible for baseline-midline gains because the task was not administered at endline. The ICC in the null model for the listening comprehension in ethnic languages gains was .189 between baseline and midline. **Table 19** below indicates a significant association between treatment and the gain score between baseline and midline in Model 3 (*b*=.120, SE=.061, *p*=.047) and a very large, significant treatment-interaction favoring Khmu children (*b*=-.529, SE=.211, *p*=.012). There are no significant associations between any of the child demographics and the gain score in the Model 2 regression findings.

**Table 19: Listening Comprehension in Ethnic Languages Regression Outputs**

| | Baseline-Midline Gain | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Model 2 | | | | | Model 3 | | | | |
| **Fixed Effects** | ***b*** | **SE** | ***p*** | **[95% CI]** | | ***b*** | **SE** | ***p*** | **[95% CI]** | |
| Age | .001 | .003 | .829 | -.005 | .007 | .000 | .003 | .961 | -.006 | .006 |
| Hmong | -.092 | .130 | .479 | -.347 | .163 | .231 | .176 | .189 | -.114 | .576 |
| Mother's Ed. | .026 | .022 | .231 | -.017 | .069 | .031 | .021 | .149 | -.011 | .072 |
| SES | -.009 | .008 | .289 | -.024 | .007 | -.007 | .008 | .367 | -.023 | .009 |
| Treatment | .046 | .063 | .461 | -.076 | .169 | **.120** | **.061** | **.047** | **.001** | **.239** |
| Treat. X Hmong | | | | | | **-.529** | **.211** | **.012** | **-.942** | **-.115** |
| Constant | .167 | .198 | .401 | -.222 | .556 | .144 | .190 | .448 | -.228 | .517 |
| **Random Effects** | | **Var.** | **SE** | **[95% CI]** | | | **Var.** | **SE** | **[95% CI]** | |
| Village-Level | | .031 | .010 | .017 | .057 | | .029 | .009 | .016 | .052 |
| Child-Level | | .116 | .011 | .096 | .139 | | .112 | .010 | .094 | .135 |
| | | **ICC** | **SE** | **[95% CI]** | | | **ICC** | **SE** | **[95% CI]** | |
| Village-Level ICC | | .211 | .054 | .124 | .336 | | .204 | .053 | .119 | .328 |

#### 5.4.6.6. Name Writing

ICCs in the null models for the name writing gains were .181 between baseline and midline, .054 between midline and endline, and .231 between baseline and endline. As shown in **Table 20**, there is a significant association between treatment and gain scores between baseline and midline in Model 3 (*b*=.117, SE=.059, *p*=.046) but no significant treatment-ethnicity interaction. There are no other significant associations between treatment and gain scores, or interactions between treatment and ethnicity, for the midline-endline or baseline-endline gain scores. Age and mother's education were predictive of gains at a few points in Model 2.

## Table 20: Name Writing Regression Outputs

**Baseline-Midline Gain**

| | Model 2 | | | | | Model 3 | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Fixed Effects** | ***b*** | **SE** | ***p*** | **[95% CI]** | | ***b*** | **SE** | ***p*** | **[95% CI]** | |
| Age | **.006** | **.003** | **.063** | **.000** | **.012** | .006 | .003 | .065 | .000 | .011 |
| Hmong | -.060 | .097 | .537 | -.249 | .130 | .166 | .191 | .384 | -.208 | .540 |
| Mother's Ed. | **.069** | **.024** | **.005** | **.021** | **.117** | **.073** | **.025** | **.003** | **.024** | **.121** |
| SES | .008 | .007 | .222 | -.005 | .021 | .009 | .007 | .181 | -.004 | .022 |
| Treatment | .065 | .055 | .237 | -.043 | .174 | **.117** | **.059** | **.046** | **.002** | **.232** |
| Treat. X Hmong | | | | | | -.376 | .199 | .058 | -.765 | .013 |
| Constant | -0.22 | 0.2 | 0.27 | -0.61 | 0.17 | -.235 | .192 | .221 | -.612 | .142 |
| **Random Effects** | | **Var.** | **SE** | **[95% CI]** | | | **Var.** | **SE** | **[95% CI]** | |
| Village-Level | | .020 | .006 | .011 | .036 | | .020 | .006 | .011 | .037 |
| Child-Level | | .106 | .010 | .088 | .128 | | .104 | .010 | .086 | .125 |
| | | **ICC** | **SE** | **[95% CI]** | | | **ICC** | **SE** | **[95% CI]** | |
| Village-Level ICC | | .160 | .044 | .091 | .266 | | .161 | .046 | .090 | .273 |

**Midline-Endline Gain**

| | Model 2 | | | | | Model 3 | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Fixed Effects** | ***b*** | **SE** | ***p*** | **[95% CI]** | | ***b*** | **SE** | ***p*** | **[95% CI]** | |
| Age | .001 | .004 | .881 | -.007 | .008 | .001 | .004 | .826 | -.007 | .008 |
| Hmong | .053 | .112 | .636 | -.166 | .272 | **-.171** | **.057** | **.003** | **-.283** | **-.058** |
| Mother's Ed. | .012 | .024 | .632 | -.036 | .059 | .009 | .024 | .691 | -.037 | .056 |
| SES | .003 | .009 | .762 | -.016 | .021 | .002 | .010 | .849 | -.017 | .021 |
| Treatment | .019 | .058 | .741 | -.095 | .133 | -.033 | .061 | .593 | -.153 | .087 |
| Treat. X Hmong | | | | | | **.403** | **.145** | **.005** | **.120** | **.687** |
| Constant | 0.12 | 0.26 | 0.66 | -0.4 | 0.64 | .135 | .267 | .612 | -.388 | .658 |
| **Random Effects** | | **Var.** | **SE** | **[95% CI]** | | | **Var.** | **SE** | **[95% CI]** | |
| Village-Level | | .014 | .008 | .005 | .041 | | .008 | .006 | .002 | .038 |
| Child-Level | | .198 | .016 | .169 | .231 | | .199 | .016 | .170 | .233 |
| | | **ICC** | **SE** | **[95% CI]** | | | **ICC** | **SE** | **[95% CI]** | |
| Village-Level ICC | | .065 | .035 | .022 | .177 | | .037 | .030 | .007 | .166 |

**Baseline-Endline Gain**

| | Model 2 | | | | | Model 3 | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Fixed Effects** | ***b*** | **SE** | ***p*** | **[95% CI]** | | ***b*** | **SE** | ***p*** | **[95% CI]** | |
| Age | .007 | .004 | .120 | -.002 | .015 | .006 | .004 | .119 | -.002 | .015 |
| Hmong | -.030 | .113 | .791 | -.252 | .192 | .005 | .207 | .979 | -.400 | .411 |
| Mother's Ed. | **.071** | **.020** | **.000** | **.032** | **.111** | **.072** | **.020** | **.000** | **.033** | **.110** |
| SES | .011 | .008 | .187 | -.005 | .027 | .011 | .008 | .190 | -.005 | .027 |
| Treatment | .077 | .069 | .266 | -.058 | .212 | .084 | .078 | .279 | -.068 | .237 |
| Treat. X Hmong | | | | | | -.058 | .242 | .810 | -.533 | .416 |
| Constant | -0.09 | 0.28 | 0.75 | -0.64 | 0.46 | -.092 | .280 | .743 | -.640 | .457 |
| **Random Effects** | | **Var.** | **SE** | **[95% CI]** | | | **Var.** | **SE** | **[95% CI]** | |
| Village-Level | | .038 | .008 | .024 | .058 | | .038 | .009 | .024 | .061 |
| Child-Level | | .155 | .012 | .133 | .179 | | .154 | .012 | .133 | .179 |
| | | **ICC** | **SE** | **[95% CI]** | | | **ICC** | **SE** | **[95% CI]** | |
| Village-Level ICC | | .196 | .045 | .122 | .298 | | .199 | .047 | .122 | .308 |

#### 5.4.6.7. Receptive Vocabulary

For the final task subjected to multi-level regression analysis was receptive vocabulary. ICCs in the null models for the receptive vocabulary gains were .333 between baseline and midline, .167 between midline and endline, and .259 between baseline and endline. As shown in **Table 21**, there was a significant association between treatment and baseline-midline gain score in Model 2 (*b*=.120, SE=.054, *p*=.027) and from baseline to endline in Model 2 (*b*=.118, SE=.045, *p*=.009). There was also a significant association between gain scores and Hmong ethnicity across the Model 2 regression outputs. However, there were no significant treatment-ethnicity interactions at any point, and no significant association between treatment and gain scores between midline and endline.

## Table 21: Receptive Vocabulary Regression Outputs

**Baseline-Midline Gain**

| | Model 2 | | | | | Model 3 | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Fixed Effects** | ***b*** | **SE** | ***p*** | **[95% CI]** | | ***b*** | **SE** | ***p*** | **[95% CI]** | |
| Age | .001 | .003 | .638 | -.004 | .006 | .001 | .003 | .612 | -.004 | .006 |
| Hmong | **-.275** | **.098** | **.005** | **-.467** | **-.082** | **-.340** | **.108** | **.002** | **-.550** | **-.129** |
| Mother's Ed. | .002 | .017 | .923 | -.032 | .036 | .001 | .017 | .964 | -.033 | .035 |
| SES | .009 | .009 | .315 | -.009 | .027 | .009 | .009 | .319 | -.009 | .027 |
| Treatment | **.120** | **.054** | **.027** | **.014** | **.226** | .106 | .058 | .068 | -.008 | .219 |
| Treat. X Hmong | | | | | | .104 | .174 | .549 | -.236 | .444 |
| Constant | .089 | .172 | .603 | -.248 | .427 | .092 | .173 | .594 | -.247 | .431 |
| **Random Effects** | | **Var.** | **SE** | **[95% CI]** | | | **Var.** | **SE** | **[95% CI]** | |
| Village-Level | | .024 | .006 | .014 | .040 | | .023 | .006 | .013 | .039 |
| Child-Level | | .088 | .008 | .074 | .104 | | .088 | .008 | .074 | .104 |
| | | **ICC** | **SE** | **[95% CI]** | | | **ICC** | **SE** | **[95% CI]** | |
| Village-Level ICC | | .213 | .048 | .134 | .321 | | .207 | .049 | .128 | .319 |

**Midline-Endline Gain**

| | Model 2 | | | | | Model 3 | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Fixed Effects** | ***b*** | **SE** | ***p*** | **[95% CI]** | | ***b*** | **SE** | ***p*** | **[95% CI]** | |
| Age | -.001 | .003 | .680 | -.007 | .005 | -.001 | .003 | .706 | -.007 | .005 |
| Hmong | **.108** | **.053** | **.041** | **.004** | **.211** | .047 | .082 | .566 | -.113 | .207 |
| Mother's Ed. | .000 | .014 | .973 | -.027 | .026 | -.001 | .014 | .923 | -.028 | .026 |
| SES | -.007 | .005 | .186 | -.017 | .003 | -.007 | .005 | .178 | -.018 | .003 |
| Treatment | -.004 | .032 | .899 | -.067 | .059 | -.018 | .033 | .583 | -.083 | .047 |
| Treat. X Hmong | | | | | | .102 | .105 | .334 | -.105 | .308 |
| Constant | 0.18 | 0.2 | 0.37 | -0.2 | 0.57 | .181 | .198 | .362 | -.208 | .570 |
| **Random Effects** | | **Var.** | **SE** | **[95% CI]** | | | **Var.** | **SE** | **[95% CI]** | |
| Village-Level | | .005 | .003 | .001 | .017 | | .005 | .003 | .002 | .018 |
| Child-Level | | .051 | .006 | .040 | .064 | | .050 | .006 | .039 | .064 |
| | | **ICC** | **SE** | **[95% CI]** | | | **ICC** | **SE** | **[95% CI]** | |
| Village-Level ICC | | .090 | .055 | .026 | .268 | | .096 | .056 | .029 | .275 |

**Baseline-Endline Gain**

| | Model 2 | | | | | Model 3 | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Fixed Effects** | ***b*** | **SE** | ***p*** | **[95% CI]** | | ***b*** | **SE** | ***p*** | **[95% CI]** | |
| Age | .001 | .003 | .851 | -.005 | .006 | .001 | .003 | .787 | -.005 | .006 |
| Hmong | **-.187** | **.075** | **.013** | **-.334** | **-.040** | **-.328** | **.042** | **.000** | **-.411** | **-.245** |
| Mother's Ed. | .004 | .012 | .721 | -.020 | .028 | .002 | .012 | .870 | -.022 | .026 |
| SES | .004 | .006 | .468 | -.007 | .016 | .004 | .006 | .508 | -.007 | .015 |
| Treatment | **.118** | **.045** | **.009** | **.030** | **.206** | .086 | .050 | .084 | -.012 | .184 |
| Treat. X Hmong | | | | | | **.233** | **.108** | **.030** | **.022** | **.444** |
| Constant | 0.22 | 0.17 | 0.21 | -0.1 | 0.56 | .229 | .175 | .193 | -.115 | .572 |
| **Random Effects** | | **Var.** | **SE** | **[95% CI]** | | | **Var.** | **SE** | **[95% CI]** | |
| Village-Level | | .015 | .005 | .008 | .029 | | .014 | .005 | .007 | .027 |
| Child-Level | | .068 | .008 | .054 | .087 | | .068 | .009 | .053 | .087 |
| | | **ICC** | **SE** | **[95% CI]** | | | **ICC** | **SE** | **[95% CI]** | |
| Village-Level ICC | | .179 | .057 | .093 | .317 | | .170 | .057 | .085 | .310 |

## 6. Discussion

This final section of the thesis delves into the implications of the study. First, a description of overall assessment findings in the sample is summarized. Next, effects of the summer pre-primary treatment are examined, with an attention to differences by ethnicity, to answer the main research questions. After that, the thesis shares recommendations for policymakers and practitioners to inform the design and implementation of similar interventions in the future. Finally, suggestions for researchers and evaluators in Laos and similar contexts are provided to improve future assessment practice, considering the unique nature of languages and writing systems in Southeast and South Asia.

### *6.1. Assessment Findings Overall*

In general, the overall assessment findings point to four main conclusions. First, most children in the study were at the floor on most tasks at baseline when their average age was approximately 5 years 3 months and before exposure to school readiness services. Children found the phonological awareness tasks extremely difficult, and this was the case even for Lao children who would have had the advantage of possessing more 'ear' for the sounds of the language. In addition, performance was very low on tasks related to emergent decoding (letter knowledge) and emergent writing (name writing). Lao children were somewhat of an exception: They could name around three letters on average and in some cases could scribble or write marks, indicating limited familiarity with the alphabet and pre-writing at baseline.

Despite poor outcomes on most tasks at baseline, the receptive vocabulary and listening comprehension tasks demonstrated that some of the children in the sample did possess nascent oral language skills, when assessed in Lao for receptive vocabulary and in their home language for comprehension. In the case of receptive vocabulary, this was limited to vocabulary words that would be highly familiar in a

rural village setting; and for listening comprehension it was restricted to factual recall with answers requiring limited spoken language. There was also evidence for some skills in concepts about print at baseline, particularly the ability to correctly orient and open a book, but not in items requiring understanding of print conventions.

Second, and as would be expected for a typical developmental trajectory, children realized steady improvements over time on nearly all tasks. At midline, when children were approximately 6 years and 3 months old on average, performance had increased markedly in concepts about print, emergent writing, and receptive vocabulary, while learners continued to struggle overall with the phonological awareness, letter knowledge, and listening comprehension in Lao tasks. At endline, when children were 6 years and 9 months old on average, the same positive trend continued on most tasks. The word segmentation, letter knowledge, and listening comprehension in Lao tasks had become easier, with around a quarter of children able to correctly segment compound words; and more than half of children able to recognize the first few letters of the alphabet and answer simple factual recall questions with one-word answers. The addition of the semantic fluency task at endline indicated that children were able to express some oral vocabulary relevant to a familiar village scene.

Despite these improvements over time, however, the overall average score topped 50% correct on only two tasks: Name writing and receptive vocabulary. The initial sound discrimination task remained particularly difficult from baseline through endline, except for a handful of Lao children. In addition, the widespread floor effects on the two higher-level skills added at endline – most used words and decodable words – indicated that only a very limited number of children had moved into the sight word / emergent decoding stage by the end of grade 1, and they were mostly Lao.

The third conclusion is that Lao children were at a distinct advantage across all assessment tasks, with Khmu children nearly catching up over time but Hmong children never able to close the gap with their Lao and Khmu peers. Lao children

started out at a higher level on all tasks and improved steadily at each subsequent wave. They performed particularly well on letter knowledge, listening comprehension in Lao, and name writing, and were close to the ceiling on receptive vocabulary by endline. In addition, they improved at a faster rate than their Khmu and Hmong peers on the phonological awareness tests even though their overall scores were not high. These findings align with the bottom-up and contrastive linguistic aspects of literacy development highlighted in the literature review, in which Lao children would bring substantial prior linguistic knowledge and Khmu children would benefit somewhat from the similarities between Lao and Khmu languages, while Hmong children would be at the greatest disadvantage both linguistically and geographically. The findings are also not surprising from a social constructivist perspective, which predicts that children's learning outcomes would vary based on background factors such as mothers' education levels, which were generally lowest among Hmong children.

The fourth and final conclusion, which can be extrapolated from the findings on listening comprehension – the one task that was administered both in Lao and in ethnic languages – is that the language of test administration may have made a substantial difference for non-Lao children. In line with the literature review on multi-language assessment, the findings on this task indicate that administering the full set of assessment items in ethnic languages for non-Lao children could have provided a much more accurate picture of their underlying knowledge and skills in their home languages rather than just their ability to express themselves in a second language. This conclusion is supported by psychometric testing conducted on the assessment items in the LEARN study, which identified that non-Lao children's limited knowledge of Lao language had substantial implications for their outcomes in the assessment (Gomez, Brown, and Spier, 2020).

However, it appears that adjusting the language of assessment would not fully address the problem, as there was still a fundamental disadvantage for non-Lao children even when assessed in a language they understood. The disadvantage was especially apparent for Hmong children, whose listening comprehension performance in their home language did not approach the same level as Lao

children's comprehension in Lao at baseline or midline. This finding points to the imperative for providing targeted learning support to level the playing field for non-Lao children with their Lao peers, over and above simply assessing them in their home languages.

### *6.2. Effects of Treatments Overall and by Ethnicity*

This section discusses answers to the two main research questions guiding the study related to treatment effects overall and by ethnicity and delves into the implications of demographic predictors and sample attrition for interpreting the findings.

***Research Question 1:*** *What is the effectiveness of the LEARN Project summer pre-primary model in improving children's Lao language school readiness at the start of grade 1 and sustaining their gains through the end of grade 1, in comparison to a control group?*

***Research Question 2:*** *How does the effectiveness of the model vary according to children's ethnicity?*

The multi-level regression analysis of gain scores in **Section 5.4.6**, coupled with Cohen's *d* effect sizes of the unadjusted differences in mean gains between the treatment and control groups (**Table 25** in **Annex 3**), reveal a pattern in which treatment was associated with larger overall gains between baseline and midline for six out of nine of the assessment scores: The literacy domain (*d*=0.36*)*, concepts about print (*d*=0.42), listening comprehension in Lao (*d*=0.24), listening comprehension in ethnic languages (*d*=.09), name writing (*d*=0.18), and receptive vocabulary (*d*=29). These are all classified as small effects using Cohen's taxonomy.

The modest gains in these six areas had typically faded between midline and endline. However, the only fadeout effect that was statistically significant in the multilevel analysis was found in concepts about print, with the treatment group

gaining less than the control group between midline and endline (*d*=-0.22). The lack of statistical significance in most of the fadeout effects in the regression analyses makes it challenging to draw definitive conclusions, although the trend is clear. The negative effect sizes of the differences in mean gains also lend some weight to the finding that the treatment and control groups had largely converged by endline across the assessment tasks.

Treatment effects also appeared to be diluted by this fadeout phenomenon when examining aggregate gains between baseline and endline, although trends for the treatment group were positive, and a significant association between treatment and gains was found in one case: Receptive vocabulary between baseline and endline (*d*=0.37). The effect sizes based on mean differences in gains between the treatment and control groups for other scores, such as the literacy domain (*d*=.25), concepts about print (*d*=0.20) and name writing (*d*=0.19), again suggest that some overall effects between baseline and endline were present although not as substantial as the initial baseline-midline gains.

Using multilevel regression, significant associations were not found between treatment status and gains on three tasks – initial sound identification, word segmentation and letter name knowledge – between any of the data collection waves. Initial sound identification continued to exhibit the extreme floor effects between all waves described elsewhere in the thesis.

Nevertheless, despite the lack of statistical significance in the associations between treatment and gains for these three tasks, trends were similar: Positive gains between baseline and midline, negative gains between midline and endline, and positive gains again between baseline and endline for the treatment group. These trend-related findings are echoed by some small but positive effect sizes when analyzing average gain scores without regression adjustments, such as for initial sound identification (baseline-midline *d*=.13; baseline-endline *d*=.09) and word segmentation (baseline-midline *d*=.19; baseline-endline *d*=.10).

Turning to treatment effects by ethnicity, the picture is mixed. There were no significant interactions between treatment and ethnicity for any of the gains between baseline-midline, midline-endline, or baseline-endline for six out of the nine scores: The literacy domain, initial sound identification, word segmentation, letter name knowledge, name writing, and receptive vocabulary. Significant interactions by ethnicity were only found between midline and endline for concepts about print, favoring Hmong children; and between baseline and midline for the two listening comprehension tasks, in both cases favoring Khmu children. The coefficients for the differences in gains by ethnicity were substantial, ranging from *b*=.289 for concepts about print between midline and endline; *b*=-.268 for listening comprehension in Lao between baseline and midline; and *b*=-.529 for listening comprehension in ethnic languages between baseline and midline.

Some evidence of positive treatment effects for Khmu children between baseline and midline is substantiated by the Cohen's *d* effect sizes for the unadjusted differences in mean gains between the treatment and control groups by ethnicity (**Table 26** in **Annex 3**). For example, this includes *d*=.45 for Khmu children on the literacy domain gain and *d*=.48 for concepts about print.

The opposite is suggested by midline-endline gains, where the effect sizes in the unadjusted gains by treatment and ethnicity provide some evidence of positive effects for Hmong children. For instance, the effect size of the midline-endline gain for treated Hmong children was *d*=.81 for the literacy domain gain; *d*=.54 for concepts about print; *d*=.59 for letter name knowledge; *d*=.72 for listening comprehension in Lao; and *d*=.79 for name writing. These are among the largest effect sizes in the study and suggest that Hmong children who participated in the treatment made a substantial positive turnaround between midline and endline.

The associations between treatment and ethnicity for the overall gains between baseline and endline are inconclusive because of the widespread lack of significant associations in the multilevel regressions. At the same time, the effect sizes in the unadjusted mean differences between treatment and control by ethnicity are

substantial and often favor Hmong children (for example, *d*=.80 for Hmong children compared to *d*=.21 for Khmu children on the baseline-endline literacy domain gain; *d*=.74 for Hmong children compared to *d*=.13 for Khmu children on the concepts about print gain; and *d*=1.00 for Hmong children compared to *d*=.31 for Khmu children on receptive vocabulary).

The effect sizes of the differences in gains by ethnicity suggest that the summer pre-primary intervention was more effective at giving Khmu children a school readiness boost at the start of grade 1 (between baseline and midline). Conversely, the intervention appeared to provide a larger boost to Hmong children by the end of grade 1 (between midline and endline) and that also contributed to larger effect sizes for Hmong children than Khmu children overall between baseline and midline. However, because the of the largely inconclusive findings treatment x ethnicity interactions in the regression analyses, the effect sizes by ethnicity can only be interpreted as suggestive.

The demographic covariates most commonly associated with learning gains in the regression analysis include age (positively), Hmong ethnicity (negatively) and mother's education (positively). These findings indicate that slightly older Khmu children with more educated mothers are predicted to perform best overall – a conclusion partially aligned with the 2020 psychometric analysis of the LEARN dataset by Gomez, Brown, and Spier, which found that mother's education levels positively predicted outcomes in key literacy tasks.

Taken together, the effect size findings and the findings related to demographic covariates suggest that Hmong children initially make slower progress than Khmu children but start to reap the benefits of the summer pre-primary as they move through grade 1. This phenomenon could perhaps be due to the linguistic, geographic, and systemic isolation of Hmong communities described in the literature review, which would have presented steeper readiness gaps for Hmong children to overcome at the start of primary school.

Intra-class correlations were generally highest on the baseline-midline gain scores, lowest for the midline-endline scores, and moderately high on the overall baseline-endline scores. The ICCs indicate that scores were most similar for children within the same schools between baseline-midline for the literacy domain (.297), word segmentation (.298), letter name knowledge (.278) and receptive vocabulary (.333) tasks. These findings are not unexpected, given the extensive influence of language on children's outcomes in the study, and the strong tendency for linguistic homogeneity in the target locations, where each village was either predominantly Khmu or predominantly Hmong with little mixing between the groups.

It is also important to review the two types of sample attrition experienced in the study – attrition due to known breaks in randomization or inability to verify treatment status; and due to failure to participate in one or more data collection waves – and the implications for interpretation of findings. Differential attrition was not a significant factor between treatment and control groups, and it was not related to key demographic covariates except for mother's education. Nevertheless, imperfect adherence to randomization means that the analysis could not meet the true 'gold standard' of RCT analysis and instead had to rely on quasi-experimental analytical approaches. This warrants cautious interpretation of any positive treatment effects.

The highly significant differences in attrition rates by ethnicity indicate that the Hmong children who remained in the analytic sample differed systematically from the Hmong children who dropped out, because the Hmong attriters were statistically older and had less-educated mothers, as well as lower baseline literacy skills. This fact complicates interpretation of findings related to the associations between treatment and ethnicity. Had those more disadvantaged Hmong children been included in the analytic sample, the baseline-midline advantages for treated Khmu children may have been even greater. Conversely, the apparent midline-endline and overall baseline-endline advantages for treated Hmong children may have disappeared.

These findings have several practical uses for policymakers and practitioners considering supporting similar interventions in the future. A key conclusion is that the summer pre-primary is associated with small but statistically significant gains between baseline and midline in the literacy domain, concepts about print, listening comprehension in Lao, listening comprehension in ethnic languages, name writing, and receptive vocabulary scores. Concepts about print, listening comprehension, and vocabulary were more explicitly taught in the course, suggesting that explicit instruction contributed to greater gains (see the discussion in **Section 6.3.4** on content validity below for further discussion on this point).

However, the positive effects in these tasks were counteracted by negative gains for the treatment group on the same tasks between midline and endline, which also generally overpowered any positive effects between baseline and endline. The exception is in expressive vocabulary, where a significant association between treatment and gains was found overall between baseline and endline. This indicates either that the increases in gains generated through the intervention were too modest to endure, or that something happened to both treatment and control children throughout grade 1 that caused the gains to converge.

In addition, no associations between treatment and gains between were found in the initial sound identification, word segmentation, or letter name knowledge tasks. This indicates a lack of impact of the intervention on other skills deemed to be important for school readiness and success in early primary by the global education community. It also suggests that future interventions should focus more on explicitly building those types of skills if the Lao government considers them to be integral to school readiness. In the case of the two tasks related to phonological awareness, the extent of zero scores indicates that the problem may have been in children's skills development – perhaps necessitating greater attention to building phonological skills in the summer pre-primary – but also in the limited of validity of the subtasks in the Lao language. The latter point will be further unpacked in **Section 6.3** in the discussion on cross-language validity.

The findings also underscore that the summer pre-primary was generally more effective at boosting school readiness before grade 1 for Khmu children. At the same time, the trends were more positive for Hmong children by endline at the end of grade 1. This lends credence to a central assumption of the summer pre-primary, that it could be beneficial in improving learning outcomes for non-Lao groups at various intervals in early primary school.

However, because Lao children were dropped from analysis due to small sample sizes, the thesis could not validate the hypothesis that the intervention would be at least as effective for non-Lao as for Lao children. In future programs, it may be informative to try to include larger samples of Lao children, perhaps focusing on the most disadvantaged Lao groups who would not otherwise benefit from school readiness service delivery, to stay true to the fundamental objective of the summer pre-primary model.

Overall, these findings lend cautious support to the Lao government's intentions to scale up the summer pre-primary model through the next education sector plan and the GPE-KIX innovation funding to boost school readiness at the start of primary school for the children who need it most but receive it least. At the same time, the findings raise some questions about the intervention's ability to increase children's gains across a wider range of literacy skills as they progress into and through first grade, although this finding is complicated in some ways by subtask validity issues that will be discussed further later.

In addition, the phenomenon of differential attrition by the most disadvantaged Hmong children indicates that future interventions may need to do more to enroll and retain the most marginalized Hmong children and ensure they are present for all assessment waves. More could also be done to understand why those Hmong children were difficult to reach – for example, if it was due to economic migration during the summer, being over- or under-age for the intervention, or illness or other demand-side issues – and identify specific ways to adjust the pre-primary model accordingly.

Finally, the consistent predictive value of mother's education levels across many of the assessment tasks underscores the need for broad-based education and development efforts reaching the most marginalized women, and perhaps to look at including targeted interventions to increase mothers' basic education levels alongside the summer pre-primary model. In fact, basic education interventions for mothers were considered in LEARN's initial design phases because of their potential – based on global evidence – that they would have positive knock-on effects for children's development and learning. However, they were not pursued due to the desire to focus on direct interventions with children.

### 6.2.1. Situating Treatment Effects in the Context of Similar Interventions

This section now discusses the estimated treatment effects of the LEARN evaluation in light of available reporting on outcomes from other school readiness interventions in Turkey, Tanzania, Lao PDR, Ethiopia, and Mozambique. This helps to situate the LEARN study within a body of similar research, to inform policymaker and practitioner decisions about relative effectiveness.

As described in **Section 3.2**, the summer pre-primary model in the LEARN Project was adapted from the original approach developed by AÇEV in Turkey. In 2011, the Turkish pilot of this model underwent a quasi-experimental assessment of short-term school readiness outcomes, with a pre-test administered shortly before course implementation and a post-test immediately afterwards (Bekman, Koç, and Taylan, 2011). The evaluation used an assessment battery linked to the curriculum content of the course, administered to a panel of 92 treatment children and 93 control children.

At endline, the treatment group had significantly higher mean scores than the control group in pre-numeracy, with a treatment average score of 32.2 compared to 15.9 for the control, out of 58 possible points; and in pre-literacy, with a treatment mean of

38.1 compared to a control mean of 21.9, out of 81 possible points.[20] Significant improvements were also found for semantic competence and listening comprehension, but not for receptive vocabulary (although the trend was positive). The authors conclude that the intervention led to the strongest gains in pre-numeracy because that was the most concerted area of focus in the course, while improvements in pre-literacy, vocabulary, and other areas were less extensive, and more specific future attention to language development would be warranted.

The nine-week AÇEV summer pre-primary model for Syrian refugee children and children in their host communities in Turkey also underwent a pre-post assessment of the activities using similar tools to the AÇEV assessment cited above. This assessment found that both Syrian and local children improved against the comparison group in pre-literacy and pre-numeracy skills, with the native Turkish children performing better than the Syrian children (Sezen, no date). For example, the average increase in pre-literacy scores for Syrian children between baseline and endline was 7.8 points against an increase of 3.2 for children in the control group; while for native children the increase was 9.6 points against an increase of .9 points for the control group.[21]

The 2016 Plan International pilot in Bokeo was also evaluated using a quasi-experimental (non-random, pre-post, panel design) assessment of 82 treatment and 40 control children (Bekman and Diri, 2016). A similar assessment battery to the original one in Turkey was used, focusing on pre-literacy and pre-numeracy skills, directly linked to course content. All instructions, test items, and answers to those test items were in children's home languages. The baseline was administered immediately before the course started and the endline immediately after the course ended. At endline, the treatment group's mean score was 49.7% compared to a

[20] Effect sizes were not reported and cannot be calculated based on information provided in the report.
[21] Effect sizes were not reported and cannot be calculated based on information provided in the report.

control mean of 29.9% on the literacy items (a difference of 19.8%) and 37.1% compared to 27.3% for the control on the numeracy items (a difference of 9.8%).[22]

The program in Tanzania underwent an assessment in 2017, with a comparison of outcomes at the start of grade 1 for children who had participated in the accelerated readiness course versus those who attended a formal pre-school or a no-intervention control group. A total of 1,191 children across 42 primary schools in 14 districts were assessed. The evaluation report does not describe the assessment tools utilized, domains tested, or effect sizes of the differences in outcomes. However, the study found that children who had attended the ASR program had substantially higher skill levels than the children in the control group, and moderately higher skills than children who attended formal pre-school for at least a year (EQUIP-Tanzania, no date).

A small-scale, non-randomized assessment in 2016 of the UNICEF ASR pilot in Ethiopia utilized a curriculum-based assessment tool with a total of 330 children in 28 schools, including 9 schools with summer ASR compared to 5 with 0 class (Spier, 2019). Students in the ASR group performed significantly better than students in the 0 Class group in Mathematics (10% higher, *d=0.44*) and in Literacy (11% higher, *d=0.48*) – indicating that the summer intervention may boost these domains of school readiness beyond the levels achieved by the 0 class alone.

The UNICEF and Save the Children ASR program in Mozambique was evaluated using an RCT design, with 30 schools assigned to the treatment group and 30 to the control (Bonilla et. al., 2018). The same panel of children were assessed with the IDELA instrument in three waves: At baseline in November 2017 before ASR interventions, at midline after ASR activities finished and prior to the start of grade 1, and at endline at the end of grade 1. The literacy domain included tasks related to oral vocabulary, print awareness, letter name knowledge, first letter sounds, emergent writing, and listening comprehension. The researchers did not describe

[22] Effect sizes were not reported and cannot be calculated based on information provided in the report.

language-related aspects of the evaluation in detail, such as the language used to deliver assessment instructions and tasks, or that children could use to produce responses, and whether those languages matched children's mother tongues.

Using intention-to-treat (ITT) analysis, the mid-term found a six percent increase (from 42% to 48%, or .36 standard deviation (SD)) in the total IDELA score for children in the treatment group compared to no increase (46% at both baseline and midline) for the control group. In addition, there was a 15 percent (.88 SD) total increase for the treatment group when considering only the children who participated in the intervention using local average treatment effect (LATE) analysis.

As the authors report, these "overall scores were driven by the impacts on specific sub-constructs: Emergent numeracy (ITT: 0.35 SD; LATE: 0.87 SD), emergent literacy (ITT: 0.30 SD; LATE: 0.75 SD), and motor skills (ITT: 0.38 SD; LATE: 0.93 SD)" (Ibid., p. 48). The authors note that, although there were significant improvements for the treatment group, the absolute gains were modest and overall levels of school readiness for both the treatment and control groups were still below expectations.

The 2019 endline of the Mozambique intervention (Bonilla et al., 2019) found highly significant positive effects of the treatment on the overall IDELA score as well as all its constituent domains except for socio-emotional development. Using ITT, the endline identified a 9-percentage point impact in the total IDELA score (from 41% to 56%, or 0.52 SD for the treatment group, compared to 45% to 50% for the control group). Using LATE analysis, there was a 17-point increase (0.93 SD). As in the midline, the overall scores were driven by the impacts in specific constructs: "emergent numeracy (ITT: 0.55 SD; LATE: 0.98 SD), emergent literacy (ITT: 0.39 SD; LATE: 0.70 SD), and motor skills (ITT: 0.44 SD; LATE: 0.78 SD)" (Ibid., pp. 35-36). The study emphasizes the striking point that – contrary to the fadeout effects often found in pre-school programs – the improvements in outcomes persisted and, in some cases, grew larger by endline.

The authors conclude by reporting that, despite the substantial effect sizes, children continued to struggle the most with emergent literacy, and the average literacy score for children in the treatment group never topped 50% correct. To situate this finding, they cite evidence from similar IDELA assessments around the world, which indicate average increases in emergent literacy scores from 45% correct for 5.5-year-olds at baseline to 55% correct for 6.5-year-olds when assessed later (Pisani, Borisova, and Dowd, 2018, in Bonilla et al., 2019).

The LEARN findings at baseline roughly align with the diagnostic baseline findings of the World Bank ECE project in Laos (World Bank, 2016c), though in some cases the average scores in the World Bank study were moderately higher. For example, the reported baseline average literacy domain score for five-year-olds in that assessment was approximately .25 out of 1, compared to .11 in the LEARN study.[23] Five-year-olds in the World Bank study who were able to answer any of the concepts about print questions scored just over .4 out of 1 overall on the task, as opposed to .15 in the LEARN study. 98% of children scored zero on the initial sound discrimination tasks and the overall average score for children who could answer any of the questions was .08 out of 1, which echoes the LEARN findings.

No comparisons can be drawn to the World Bank's findings on the word segmentation task because it was not administered in that assessment. 88% of children scored zero on the letter name knowledge task, and the children who could respond named just over 2.5 letters out of 20 at baseline. The overall average score when assessed in Lao on the listening comprehension task was .07 out of 1. When assessed in ethnic languages, out of the 5-year-olds who could respond, the average score was nearly .45 out of 1.

[23] The World Bank report provides graphs of baseline scores only and does not provide exact statistics; thus, in this section, World Bank findings should be taken as best estimates based on visual information in the graphs. In addition, the World Bank reports findings broken out by age, and for the purposes of this section, only the five-year-olds are discussed because they are closest in age to the LEARN baseline sample. The World Bank also reports by ethnicity, but as averages across all age groups and not just for five-year-olds, so those findings are not discussed in this section.

Although not yet formally published, the baseline-midline gains in the World Bank's cluster randomized control longitudinal study of ECE project impact with over 7,520 children in 376 villages were informally shared through education sector working groups in Laos (World Bank, 2018a). Using ITT analysis on five-year-olds at midline, the assessment found gains in the overall numeracy score of .05 out of 1 across the board for children in the control group, children in CCDGs and children in the MAT interventions – in other words, no direct evidence of impact on math skills. Gains in the overall literacy score were .05 out of 1 for the control group, .07 for the CCDGs, and .03 for the MAT interventions, indicating a slightly larger effect on literacy of the CCDGs.[24] It is important to note the caveat that startup of interventions in the project was delayed, and the midline was conducted after less than a full school year of implementation had taken place.

Direct comparisons between the outcomes from these activities and the LEARN summer pre-primary are not possible due to differences in contexts, intervention designs and objectives, and learning assessment approaches. Nevertheless, it is informative to discuss similarities and differences in a general sense.

The key take-away from review of the findings of the AÇEV-supported interventions in Turkey is that they only contributed to modest increases in literacy skills, similar to the LEARN outcomes. As the authors point out, this suggests that the literacy-focused content in the model could be strengthened and targeted. The Bokeo report underscores the same recommendation, noting that the summer pre-primary model adapted for Laos omitted the supplemental language module that is typically included in the Turkish model (Bekman and Diri, 2016) – a shortcoming that carried over into the summer pre-primary model in LEARN.

The assessment of the Bokeo pilot also reinforces the fact that the language used when delivering assessment instructions and administering test items (in this case,

[24] Statistical significance of the differences between control groups and the two treatment groups was not reported. Effect sizes were also not reported and cannot be calculated based on information provided in the presentation.

children's home languages), as well as the direct link between the course content and the assessment tasks, may have contributed to an assessment approach that was more valid for the context and the intervention, with fewer floor effects than in the LEARN evaluation.

The Tanzania evaluation identified positive outcomes from an accelerated intervention in comparison to a longer and more formal pre-school service, suggesting that a shorter, intensive model may be more effective than a longer approach. The Ethiopia evaluation of a very similar intervention to LEARN's summer pre-primary found a small effect size for the literacy outcomes using Cohen's 1998 taxonomy. The World Bank's midline in Lao PDR also points to similar small gains, of between .03 and .07 out of 1 on the literacy domain for the two treatment groups after less than a school year of intervention, although gains may have been larger if the study had reported outcomes for compliers in addition to using ITT analysis. In fact, the World Bank team overseeing the evaluation identified the need for strengthened monitoring and documentation of actual uptake as implementation rolled out as a core element of interpreting the impact evaluation findings based on who fully participated and who did not (Danchev, 2018).

In contrast, the larger reported effect sizes from Mozambique suggest stronger impacts of that accelerated summer intervention than the LEARN summer pre-primary was able to achieve using similar assessment tasks. It is not clear from the available reporting, however, how language issues may have affected the outcomes in Mozambique. At the same time, the Mozambique experience underlines how improvements in literacy skills in similar programs and populations around the world are often modest in absolute terms, and that six-year-old children typically score around half correct on these types of tests. In this sense, the LEARN assessment findings are lower than the global averages but not far off, with children scoring .39 out of 1 on average on the literacy domain at endline, when they were approximately six to seven years old.

In general, the review of outcomes from similar interventions identifies that the LEARN summer pre-primary impact was aligned with global experiences, although lower than the average findings from pre-primary programs using similar assessment tools. The findings highlight the importance of deliberately strengthening the literacy content of the course. The review also suggests, however, that it may be difficult to disentangle LEARN's modest outcomes from the challenges caused by assessing children in unfamiliar languages and highlights the importance of describing how language issues were addressed during assessment like the one in Mozambique to facilitate interpretation of results. Finally, the common challenge related to verifying uptake in the World Bank and LEARN assessments points to the need to build in careful verification mechanisms in future evaluations in Laos and analogous contexts, to mitigate against attrition bias and difficulties in applying RCT analytical approaches. However, if these real-world challenges are potentially too difficult to overcome when working in remote, marginalized communities in settings with incomplete education administrative data, future programs may want to consider alternative evaluation approaches that do not rely as extensively on sample randomization and verification.

### 6.2.2. Treatment Effects and the Pre-School Fade-out Phenomenon

Similar to the LEARN evaluation results, several studies have documented a 'fadeout' phenomenon in pre-school programs, in which scores for control groups eventually converge with treatment groups as children progress in primary school.

For instance, a landmark longitudinal RCT of the Head Start school readiness program in the United States followed children from ages 3-4 through the end of grade 3 after they had participated in one or two years of the program (Puma et al., 2012). The outcomes of interest included cognitive development, social-emotional development, health status, and parenting practices. The study found significant immediate impacts of the program on children's developmental outcomes, including literacy skills. However, during follow-up assessments in kindergarten, grade 1, and grade 3, the children in the control arms had largely caught up to the children in the

treatment arms. The authors point out that this convergence of outcomes has been a common finding among similar analyses of the longer-term effects of school readiness interventions in the United States.

Bailey et al. (2017) review the existing literature, also mainly from the US, on persistence of learning effects over time. They attempt to make sense of the fadeout phenomenon by describing three general models of interventions: Skill-building, foot-in-the-door, and sustaining. Their analysis postulates that, for the effects of skill-building programs to persist and avoid convergence with outcomes for non-treated children, interventions must focus on so-called 'trifecta' skills: Ones that are malleable, fundamental, *and* would not have developed on their own. As the authors note, "basic language and literacy skills are clearly fundamental and malleable but do not make the trifecta list because they develop from natural experiences under most counterfactual conditions or are specifically targeted in universally available early formal or informal learning environments" (Ibid., p. 20).

Foot-in-the-door interventions, the second model discussed, are not necessarily designed for persistence of outcomes. Instead, they are important because of their timing related to developmental windows of opportunity or vulnerability. They provide learners with enough knowledge and skills to prevent them from common risks at certain life stages, such as dropout or failure to enroll or progress in school. Fundamentally, they help children advance to a subsequent stage, but do not necessarily lead to effects that will continue to distinguish intervention participants from control groups over time.

The third model addressed by Bailey et al. is one of sustaining environments, in which "enriched post-intervention environments can be consciously planned and implemented, for example, by providing high-quality elementary school instruction that complements what has been taught before" (p. 25). The authors point out how the children from marginalized backgrounds who most need the boost in skills provided by pre-school interventions are also the most likely to enter lower-quality

primary school classrooms, thereby diluting the effects of the earlier support in the absence of deliberate sustaining efforts.

There is limited published literature interrogating the issue of fadeout in LMICs through longitudinal designs. Recent findings from an evaluation of the two-year Tayari pre-primary program in Kenya (Piper, 2018b; and Piper and Merseth, 2019) are perhaps most relevant to the LEARN study because the evaluation used a longitudinal assessment of core school readiness skills with a version of MODEL. Using ITT analysis, the study found significant effects at midline and endline – for example, an effect size of .31 SD in the literacy domain – which had disappeared at follow-up, roughly two years after the intervention ended.

The authors suggest that the effects may not have persisted because children who did not participate in Tayari nevertheless eventually learned the same skills in early primary school. They also question whether the evaluation measured skills that could reasonably be expected not to converge. This raises the potential need to re-think assessment approaches if the intention is to measure persistence of skills over time rather than (or in addition to) proximal benefits.

A longitudinal, cluster RCT of a program in Malawi tested the immediate and longer-term effects of a program of teacher training for community-based childcare centers combined with parenting education (Ozler et al., 2018). Similar to Tayari, the evaluation identified strong positive effects in the short term, but those had faded by the follow-up at 18 and 36 months post-baseline, when children were 4.5-6.5 and 6-8 years old, respectively. The key reason the authors postulated for this fadeout was that “the children transitioned into a low-quality primary school system, which had no coordination of curricula with treated preschools and could not provide an appropriate learning environment for the children to build sustainably on the gains they made earlier” (Ibid., pp. 464-465).

A core conclusion from this discussion on pre-school fadeout is that it is a common phenomenon in the early childhood development and school readiness sphere,

including LMIC contexts, and therefore it is not surprising that convergence of literacy outcomes occurred in LEARN. Beyond this, the Bailey et al. analysis is particularly germane in helping to better conceptualize the purpose of the summer pre-primary model and therefore review whether its assessment approach was fit for purpose. In general, a large part of the LEARN evaluation design implicitly assumed the summer pre-primary was a skill-building program that would generate malleable and fundamental skills – and ones which would not eventually emerge in grade 1 for all students. If the intention was mainly to evaluate whether the model contributed to longer-term impacts in early primary school, LEARN may have needed to assess different types of skills that met the trifecta definition.

In reality, the summer pre-primary had more of the characteristics of a foot-in-the-door intervention designed to present highly disadvantaged children with a window of opportunity to enter grade 1 on time and avoid early dropout and repetition. Although those types of outcomes are not discussed in this thesis because of the focus on literacy skills, AIR's external evaluation identified that the summer pre-primary did, in fact, increase on-time enrollment in grade 1, though it had no significant effects on grade 1 completion and only limited effects on on-time enrollment in grade 2 (American Institutes for Research, 2019). It may be useful for policymakers and practitioners to carefully consider how they communicate about the benefits of the model if it is, indeed, aimed mainly at providing enrollment-related benefits.

This way of thinking about the summer pre-primary as predominantly a foot-in-the-door model is supported by a 2020 analysis of ECD and primary school education management data from LMICs. For example, a recent think piece identified that persistent poor performance of primary education systems in the early grades is largely a consequence of inadequate ECD and school readiness service provision, as well as chronic high grade 1 intake, over-enrollment in grade 1, and inadequate progression to grade 2 (Crouch, Olefir, Saeki, and Savrimootoo, 2020). The grade 1 'bulge' results in age and ability heterogeneity, contributing to inefficiency in early primary as teachers struggle to assimilate a wide range of students. These findings underscore the importance of LEARN-style interventions as exactly the leg up

needed to improve efficiency-related outcomes for marginalized children within generally inefficient systems.

Reflecting on the Bailey et al. analysis as well as the findings from Malawi, the low quality of the learning environments and the curricular mismatch that met children after they finished the summer pre-primary and progressed into grade 1 may have been at play as well. The 2020 analysis by Crouch et al. seconds the suggestion that issues of pre-school fadeout may be partly resolved by greater cohesion between ECD and early primary curricula and pedagogical approaches. In fact, a final, critical policy implication stemming from the LEARN experience could be the need to simultaneously provide a school readiness boost to disadvantaged children while also creating a better curriculum link between pre-school and grade 1 and investing in higher quality early primary school in the most marginalized areas of Lao PDR. As described in **Section 3.3,** data from LEARN's monitoring system and qualitative research lend credence to the need for better coherence between ECD services and early primary school in the project's target areas, with absenteeism more than doubling during spot checks in grade 1 when compared to absentee rates during the summer course (Yang, 2019). In addition, teachers reported difficulties with maintaining child-friendly teaching and learning during grade 1 due to curriculum constraints (St. George and Khoonbarthao, 2018).

### *6.3. Implications for Evaluators and Researchers*

This section returns to the sub-research question guiding this thesis:

***Research Question 3:*** How do the unique features of the Lao language affect measurement of child-level outcomes and how can those features be more effectively addressed in global assessment approaches?

To answer this question, the section provides feedback on the reliability and validity of the MODEL assessment approach in the Lao language and similar linguistic contexts, and the implications for future evaluators and researchers. Where relevant,

the section also provides related recommendations to practitioners, including teachers, teacher educators, and curriculum designers working on school readiness in Laos. The section reflects on the literature review as part of this discussion, and particularly the elements related to the unique aspects of the Lao language, to better inform researchers, policymakers, and practitioners seeking to build on the successes of the summer pre-primary model and address its gaps.

### 6.3.1. Internal Reliability

Categorization of tasks in the assessment by their levels of internal reliability reveals that the majority were consistently 'good' or above across the three data collection waves, although in some cases the nearly universal zero scores complicate interpretation of the reliability alphas.

Concepts about print was an exception. This task exhibited poor internal reliability at both T1 and T3, which could suggest that the three items in this task were not consistently measuring the same underlying skill. It could also be providing evidence that, in the study context, young children are much more likely to be able to correctly orient a book but are slow to develop understanding of print conventions, and therefore combining those three results items in a problematic measure of the construct of print familiarity for that set of children.

In future assessments, evaluators could look more closely at the ways in which enumerators were scoring each individual item to see if the problem could lie there and if more training is required. They could also consider adding more items to this task to determine if that would increase internal reliability, such as items on book concepts (e.g., identifying the front cover/back cover/spine) or reading concepts (e.g., pointing to an individual letter or telling the difference between text and illustration). Adding items that might typically be used in alphabetic, non-*scriptio continua* languages – such as telling the difference between a letter and a word or sentence or identifying punctuation marks – would not be recommended, however, as these are not as relevant in the Lao writing system.

### 6.3.2. Face Validity and Cross-Context Validity

Although they ostensibly exhibited strong face validity – after all, it is the objective of most education systems, including the Lao system, to develop skills like letter knowledge and listening comprehension around the start of primary school – the deeper validity of the tasks in the assessment was problematic in a number of ways. The levels of variation and sensitivity of the tasks for the Lao children in the sample indicated a fair amount of contextual validity for that population. At the same time, the extent of zero scores for non-Lao children across assessment items at baseline, and in some cases continuing into the midline, calls into question the cross-context validity of the assessment for Khmu and Hmong children. In fact, the concepts about print task at endline and the listening comprehension in ethnic languages task at midline were the few where results approached a normal distribution across the sample, largely because zero scores and overall poor performance predominated for everyone except Lao children on most of the other tasks at T1 and T2.

This pitfall related to extensive floor effects is common in assessments of children at the so-called 'bottom of the pyramid,' where assessment items that may have worked in other contexts do not "travel well across cultural contexts" (Hornberger, 2018, p. 154). This introduces a type of bias in which "a particular group within an overall population distribution is prevented from being able to provide useful data, perhaps because they are below the floor or above the ceiling of the test" (Care, Robertson, and Ferido, 2018, p. 183). The better performance and wider variation on the receptive vocabulary and listening comprehension tasks at baseline indicate that future assessments with similar children could benefit from including more oral language tasks, taking the "overall language ability" approach espoused by Snow, Burns, and Griffin (1998, p. 111).

For example, the semantic fluency task could be included from the beginning rather than only at endline, and children could be given more than one scene to respond to. The receptive vocabulary assessment, which is not typically part of MODEL or

EGRA but was developed specifically for LEARN, could be used in future assessments after being subjected to more rigorous psychometric testing, including analysis of item difficulty, item discrimination, and the effectiveness of multiple-choice distractors. Evaluators could also use tasks from USAID's oral language assessment modules, including tests of expressive vocabulary such as naming common objects or retelling a story in the child's own words (Chiappetta, 2019), which in Laos would need to be administered in home languages at least for young children. Developing a wider variety of tasks to capture the range of skills that could be expected in the study population would enhance cross-context validity and provide more actionable evidence than merely revealing that most children could not answer most of the questions (Chiappetta, 2019; Dowd and Pisani, 2018).

### 6.3.3. Ecological Validity

The LEARN assessment may have also suffered from limited ecological validity related to unfamiliar and intimidating testing environments, especially at baseline before children had widespread exposure to formal schooling. Hmong children tend to live in remote, isolated communities that typically receive few outside visitors and have limited interaction with other language groups. They also tend to lack ECD and school readiness services that would have familiarized children with a learning and testing environment prior to taking part in the assessment. These ethnographic disconnects would have rendered the baseline assessment environment particularly alien to the normal ecology in which children would find themselves in their village (Maddox, 2018). Thus, in many ways, the assessment was measuring their reaction to an unfamiliar situation instead of producing a meaningful reflection of their true skills in the naturalistic settings of their home environments (Hornberger, 2018).

Future assessment practice for these and similar groups of children would benefit from building in more time for assessors to build rapport with children and their families – including warm-up games to introduce children to the testing environment and familiarize them with the testing tasks – and ongoing encouragement during test administration (Maddox, 2018; Vagh and Sharma, 2018). More attention would need

to be paid to selecting enumerators with skill and previous experience in interacting with young children and the ability to speak the same language; and more time spent training and practicing child-friendly assessment approaches would need to be built into evaluation timelines and budgets.

In fact, rather than applying superficial fixes in the form of more enumerator training or child-friendly games, critics of external, high-stakes summative evaluations have called for complementing them with classroom-based formative assessment conducted in a more similar environment to the real-world learning situations that children face (Wiliam and Thompson, 2007; Baird, Andrich, Hopfenbeck, and Stobart, 2016). Formative assessments have the advantage of providing ongoing feedback on performance to teachers and allowing them to adjust their instructional practice to support children before gaps widen irreparably (Kanjee, 2018).

In the context of early reading interventions, this approach has been adopted successfully in countries with some cultural and linguistic similarities to Laos, such as Nepal (RTI International, 2020). In that case, a modified, classroom-based early grade reading assessment was employed annually by teachers to gauge learning progress, with involvement of children's parents. This type of assessment was introduced to a limited extent in the LEARN intervention models but could have been used more extensively to provide a fuller picture of children's learning levels over time.

#### 6.3.4. Content Validity

Another area of concern in the LEARN evaluation was the content validity of the assessment tool: Whether what was assessed was what was taught in – or could realistically be affected by – the interventions. Content validity is always a concern when using international assessment approaches to measure the impact of a particular intervention, as "it is often difficult to express the impact of the project on a common global scale" (Crawfurd et. al., 2019). Programmatic evaluation was, in fact, not the primary goal of MODEL, which aimed instead to provide education systems

with broad diagnostics to aid in national and international monitoring of the status of early childhood development (Raikes et al., 2019). Indeed, the effort to develop a core set of psychometrically robust pre-primary assessment items based on experience to date with MODEL underscored that the core items will likely need to be supplemented with additional ones to capture aspects such as program impact and locally contextualized skills (Pushparatnam et al., 2021).

As discussed in **Section 6.2.2** on pre-school fadeout, content validity is an even more critical point to consider for accelerated, foot-in-the-door interventions. By nature, these provide short doses of interventions and may require larger sample sizes and/or instruments that more directly measure changes in the targeted skills than longer interventions where more time is available for children to develop early literacy competencies.

**Table 22** below lays out the overlap between what was directly taught in the summer pre-primary and the skills that were measured in the assessment, indicating that four out of the ten skills were explicitly included in the course. Out of those four, three exhibited positive associations between treatment status and gain scores based on the multilevel regression findings, while the fourth – expressive vocabulary – was not subjected to gain scores analysis. A similar mismatch was identified in the Bokeo pilot, which highlighted that literacy-related content was limited in the intervention design and that carried into the LEARN summer pre-primary approach as well (Bekman and Diri, 2016).

This mismatch between the intervention and the test battery calls into question the content validity of the assessment tool and raises the possibility that future assessments of similar interventions should be better tailored to their actual content to provide useful evidence to policymakers considering whether to adopt the models at larger scale. This could be done in a balanced way, retaining the broad skills that are assessed globally but adding assessment items that will be sensitive to the domains covered in the intervention without testing the exact content that is taught day to day (IES, 2020). In addition, attention could be paid to including ‘trifecta’ skills

in the assessment that will help to avoid convergence in assessment results over time.

**Table 22: Overlap between Interventions and Assessed Skills**

| Skill Area | Did the Course Teach This Skill? |
|---|---|
| Concepts about Print | Yes |
| Initial Sound Discrimination | No |
| Word Segmentation | No |
| Letter Knowledge | No |
| Name Writing | No |
| Listening Comprehension | Yes |
| Receptive Vocabulary | Yes |
| Expressive Vocabulary | Yes |
| Most Used Words | No |
| Decodable Words | No |

### 6.3.5. Predictive Validity

The literature on the MODEL and similar child assessment indicates that there is still work to be done to establish predictive validity in LMICs – that is, their usefulness, through longitudinal studies, in predicting children's performance on later and higher-level skills (Fernald, Prado, Kariger, and Raikes, 2017). There is also a largely unanswered question in the literature on multi-language assessment regarding the predictive value of children's skills when assessed in their home language with respect to later skills development in the language of instruction (Piper, Schroeder, and Trudell, 2016).

In addition, a small but growing number of studies have begun to demonstrate the importance of taking cross-language transfer of literacy skills over time into account in study design, implementation, and analysis in LMICs. In a 2015 meta-analysis, Chan and Sylva (2015) cautiously put forward a conceptual framework in which

emergent literacy for second language learners is a two-domain construct that includes both code-related and oral language skills, with tone included in the repertoire of code-related skills in tonal languages. They argue for a nuanced approach to future research with longitudinal studies that look at predictive pathways between code-related and oral language skills, as well as cross-language transfer of skills, as children progress on an L1 to L2 path through the early years and into primary school.

Programs in Africa have begun to probe these questions of cross-language skills transfer more rigorously. For example, an RCT in Kenya identified large improvements both in children's Kiswahili *and* in their English phonological awareness and letter-sound knowledge after they had received an eight-week training course in phonological awareness and letter knowledge in Kiswahili, indicating a substantial transmission of these skills between the two languages (effect sizes between .37 to .95) (Wawire and Kim, 2018). A 2016 study, also in Kenya, found significant cross-language transfer in word readings skills between children's home languages, Kiswahili, and English (Piper, Schroeder, and Trudell, 2016). A third Kenyan study found that children's phonological awareness and receptive vocabulary skills in their home languages were strongly predictive of their literacy skills in English and Kiswahili at later points in time (Jasinska, Wolf, Jukes, and Dubeck, 2019).

The LEARN study offers limited opportunities to probe these questions with regard to the literacy domain, because only a few higher-level skills were introduced at endline (most used words and decodable words), and they exhibited extreme floor effects. What is more, due to programmatic and budgetary constraints the elapsed time between T1 and T3 was only 1.5 years, leaving limited space for children to have developed higher-level skills. Unsurprisingly, examining the interactions between baseline and midline literacy domain scores (the domain score is an index of emergent literacy skills) and an endline decoding index (the average of the decodable and most used words tasks) highlights that there were limited correlations between the two (see **Figure 18** and **Figure 19** below). Lao children were the

exception, with a correlation of $r$=.44 ($p$ =.000), while Khmu and Hmong children had such extensive zero scores on the decoding skills that there were only very weak correlations.

**Figure 18: Scatterplot of the Relationship between Baseline Literacy Domain Score and Endline Decoding Index, All Children**

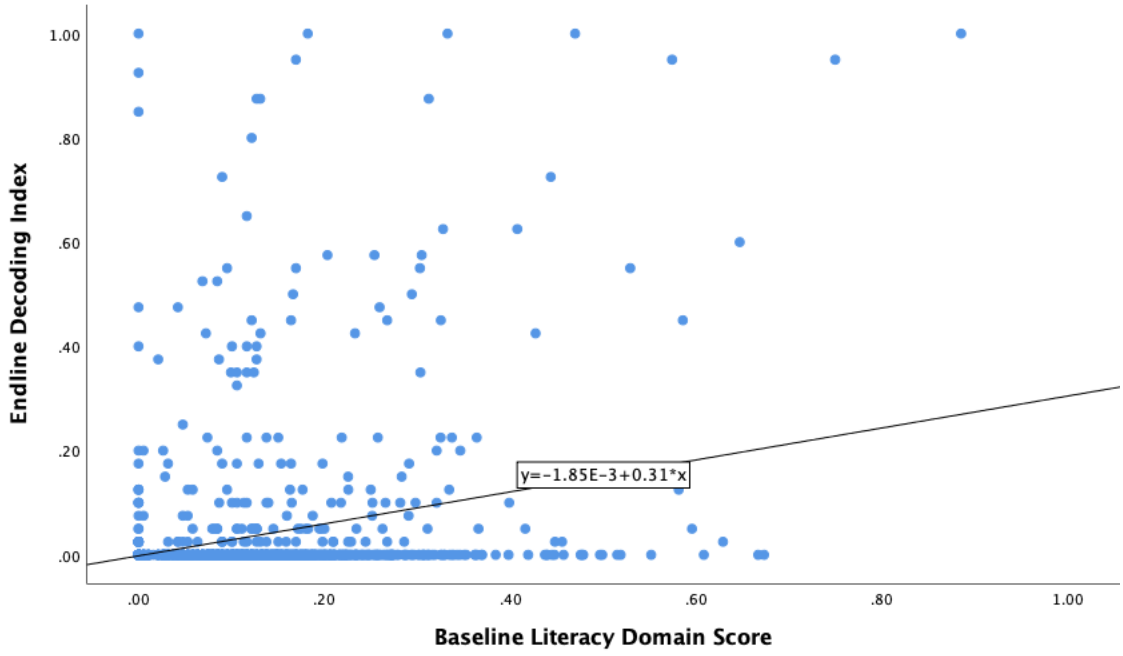
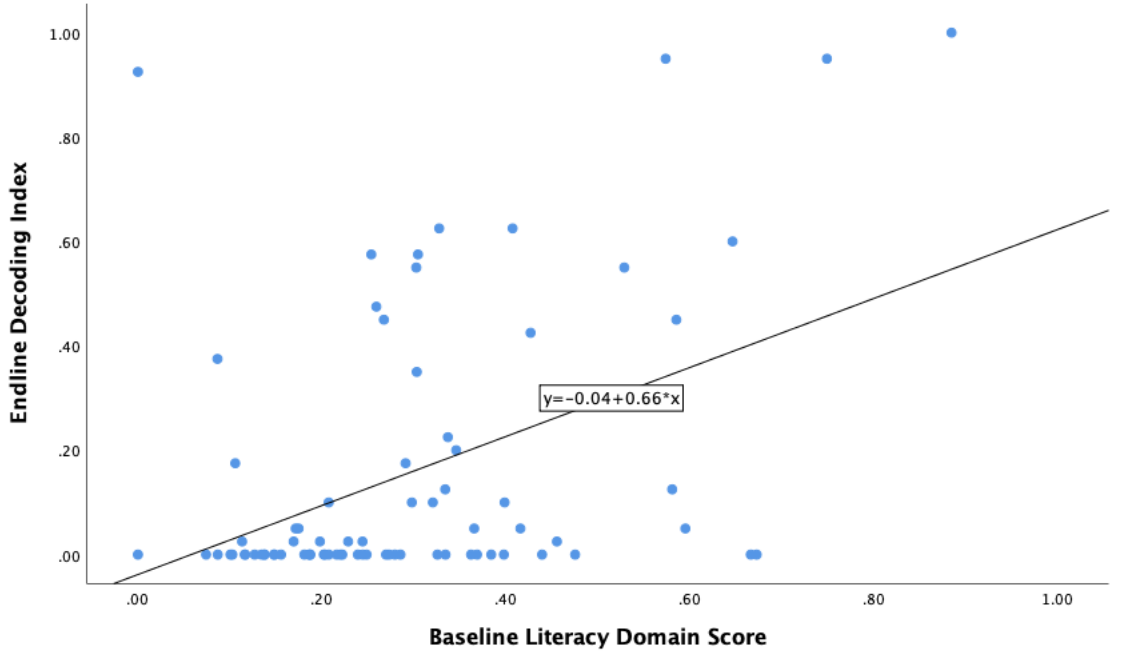



**Figure 19: Scatterplot of the Relationship between Baseline Literacy Domain Score and Endline Decoding Index, Lao Children**



Finally, only one literacy task was administered both in children's home language and in Lao – listening comprehension – limiting the ability to measure how well home language performance predicted performance in the language of instruction. Looking at correlations between scores on this task when ethnic children were assessed in their home language and in Lao identifies weak correlations between baseline-baseline, baseline-midline, and baseline-endline scores ($r=.28$ or lower). However, the correlations are slightly stronger when comparing midline performance in ethnic languages with midline performance in Lao ($r=.45$, $p=.000$). Future assessments should consider a longer time period to cover a wider developmental range of literacy skills and assessing all tasks in both home languages and the language of instruction to establish how children's skills are transferring between languages.

Indeed, language use in assessment is another major area of concern, and perhaps the most important one for this thesis. Proponents of MELQO caution that the MODEL assessment's psychometric properties may be influenced by the level of children's familiarity with the language used (Raikes, 2017). Despite attempts to conduct a multi-language assessment, the LEARN study still suffered from inevitable language-related constraints. Enumerators usually did not speak the same ethnic languages as the target children, and it was not politically feasible to write instructions in local languages like Hmong and Khmu so that the instructions could be delivered consistently.

Relying on locally recruited interpreters on the spot provided some level of support for children who struggled to understand or speak Lao, but this was also problematic because quality could vary, and it was not possible for enumerators to verify the accuracy of instructions or responses in languages they themselves did not understand. It was likely that this affected Hmong children more extensively than Khmu children, as it is more common to be able to find Khmu speakers who can interpret accurately in Lao than Hmong speakers who can do so.

A 2020 investigation (Gomez, Brown, and Spier) of the factor structure and measurement invariance of the assessment instrument used in the LEARN Project

evaluation sheds further light on how the instrument functioned in the multi-lingual Lao context. This analysis focused on identifying the factor structure of the measure and how well it held up in comparisons of outcomes among children from different ethnic and age groups. The authors found that the instrument reliably measured subtasks in the numeracy, executive function, and (receptive) vocabulary domains. At the same time, the authors concluded that literacy "is not reliably measured by items from MELQO, perhaps due to differences in children's proficiency with the Lao language. In fact, items from the Literacy domain,[25] which are the most dependent on children's skills in Lao language, had the lowest rate of correct responses compared with the other domains" (p. 14). They suggest that assessing non-Lao children in Lao to evaluate their literacy skills results in a non-reliable measure and noted the importance of assessing in children's first languages.

Although empirically substantiated by the factor and measurement invariance analysis, this recommendation to assess in first languages may be challenging to implement in practical terms, given that Lao is the official language of instruction and considering the Lao government's interest in measuring children's skills in Lao as well as political sensitivities around language of instruction. Instead, workarounds may be required to ensure the quality and validity of the assessments administered in Lao to non-Lao children. To address this challenge, in future assessments it would be advisable to adopt an approach to standardizing the use of language, such as video or audio recordings of instructions being given in local languages by native speakers, which are first back-translated to ensure accuracy. It would also be recommended to invest time and funds in working with survey firms to identify and thoroughly train local language recruiters ahead of time, if resources permit, so that at least instructions can be given effectively and uniformly in local languages.

In addition, further study on the interaction between the language spoken by the enumerator and the language spoken by the child – and how these combinations

[25] The literacy domain included initial sound discrimination, word segmentation (referred to as "phonological awareness" by the authors), letter name knowledge, and listening comprehension. Concepts about print and name writing were deleted because as "their factor loadings were lower than 0.4 or loaded high on more than one factor" (Gomez, Brown, and Spier, 2020, p. 11).

affect assessment results – may be illuminating. As highlighted in a 2020 compilation (Sha and Gabel) of research on the role of language in surveys, systematically examining these factors reveals how much language can substantially alter how researchers interact with study subjects. Language can also influence how well study subjects are able to engage with the study content and form responses, and whether enumerators code those responses as 'correct.'

Beyond the obvious improvements related to data enumeration that could be made, another area for future enhancement could be in adopting a conceptual scoring approach rather than limiting the assessment to mono-lingual scoring alone (RTI, 2018b; Bedore et al., 2005). "Conceptual scoring, which considers the number of concepts for which a child has a word in any language instead of counting language-specific lexical items," has been shown to provide more accurate assessments of bilingual children's total vocabulary and not just language-specific vocabulary (Gross, Buac, and Kaushanskaya, 2014, p. 575).

This can be operationalized by providing full or partial credit for correct answers that children provide in their home language in addition to any answers in Lao and must be preceded by an exercise to identify the range of possible labels for each vocabulary word in local vernaculars (Vagh and Sharma, 2018). This approach would also require enumerators who speak the same language as children and therefore can correctly score their answers in local languages *in situ;* or a methodology for audio recording answers and scoring them later.

### 6.3.6. Cross-Language Validity

In addition, the cross-language validity of the assessment in the Lao linguistic context must be comprehensively interrogated with an eye towards improving assessment practice in similar settings in the future. There is increasing recognition that assessment approaches first developed in high-income countries – and the assumptions that underpinned them – should be critiqued for their relevance across LMIC contexts. Although there has been some acknowledgement of the limitations of

cross-context and cross-language comparisons of results from assessments like EGRA (Bartlett, Dowd, and Jonason, 2015), there is nevertheless an implicit assumption, using typical early grade reading skills progression in the United States as a yardstick, that children around the world should be achieving a roughly similar level of oral reading fluency within the first one to two grades of primary school (USAID, 2019).

As the literature review on alpha-syllabic languages demonstrated, first or second grade may still be too early in languages with more complex orthographies and syllable structures as well as large consonant, vowel, and diacritic inventories, such as those found in South and Southeast Asia (Nag, 2007; Tiwari, 2011; RTI International, 2016a). Moreover, the literature on alpha-syllabaries underscores how critical it is to pay attention to the complexity of syllabograms in the language, and the typical instructional progression of teaching more and more complex syllable structures, when designing assessment items. For example, in the Indian alpha-syllabaries, some *akshara* are more complex and tend to appear later in reading instruction than syllable units that are simpler and higher-frequency, and it is not appropriate to include those complex syllable blocks in school readiness or very early reading assessments (Vagh and Sharma, 2018).

These points demand a re-think of globalized assessment regimes such as MODEL and EGRA in a context like Laos – as well as the design of curriculum and teaching and learning interventions – where tasks related to early decoding skills were clearly too difficult for the target population even by the end of first grade. Further analysis is required to determine when children, and particularly those in highly disadvantaged contexts, can reasonably be expected to develop enough skill in decoding in the Lao writing system to make subtasks like these meaningful. In addition, future assessments must include an analysis of phonological and syllabic complexity in Lao when selecting appropriate morphemes to include in tasks such as initial sound discrimination, word segmentation, most used words, and decodable words. For younger children, morphemes must be simple in orthography and phonology,

paralleling children's developmental progression and the sequence typically included in early instruction.

The literature on alpha-syllabic languages indicates that phonological and syllabic awareness must be explicitly taught to young children, and that these skills may also need to be included in standard assessment approaches to capture the skills that are predictive of reading progression in those writing systems, but that careful adaptations are required for each linguistic context. Examining existing assessment batteries for early reading, Nag et al.'s 2014 meta-analysis identified that well-established tools to measure young children's oral language skills – such as the ability to distinguish unique sounds in the language – are lacking. The authors recommended that assessments should include "psycholinguistic measures of skills that are relevant for the language of literacy instruction in developing countries" and not just a simple translation of tools from other contexts (p. 17).

In alpha-syllabic languages, this could include syllable knowledge (understanding of how syllables are blocked together to form words) and syllable-sound awareness (knowledge of common consonant-vowel combinations that are grouped syllabically). The authors importantly point out that the type of language and writing system also impacts the age and pace at which children can develop syllable, phoneme, and word recognition skills and thus when those skills should logically be taught and assessed. This may occur early in primary school in languages that have simple symbol-sound mappings and/or have small symbol inventories; but later in languages with complex symbol-sound connections and large symbol sets, or so-called orthographic depth, such as the alpha-syllabaries of South and Southeast Asia (Nag, 2017; Joshi, 2010) and Arabic (Abdelhadi, Ibrahim, and Eviatat, 2011).

Tasks that have been used to assess phonological and syllabic awareness in similar contexts to Laos include the tasks that were incorporated in the Nakamura, Joshi, and Ji study (2018) and Kasisopa et al. (2018), such as oral phoneme or syllable deletion. Other options include oral phoneme blending, syllable blending, identification of the beginning syllable (onset), and identification of the final syllable

(rhyme) (Nag, 2017). These tasks might also be meaningfully taught in the summer pre-primary to deepen children's skills in areas that the LEARN assessment identified as large gaps in the study population. For Hmong children, instruction would need to focus on developing an ear for the contrastive vowel lengths, final consonants and tone sounds that exist in Lao but not their native language.

Any such skills must be selected with careful attention to the morphology of Lao, however, where many words are mono-syllabic and therefore tasks such as syllable deletion would require careful construction, while a phoneme deletion task may only function well when enumerators are trained on how to instruct children and how to score correctly. In fact, the same caveat applies to the phonological awareness tests that functioned poorly in the LEARN study – particularly the initial sound discrimination task but also word segmentation, to a lesser extent – which indicate that further thinking and consultation with linguists is required to improve these tasks or identify suitable alternatives that are valid for the language and writing system and can be reliably assessed.

Orthographic knowledge is another critical aspect of the Lao writing system about which typical MODEL and EGRA assessments are largely silent. Apart from the letter test, there is no widespread approach to assessing children's knowledge of the individual symbols in the language and the spatial and sequential rules for combining them that apply to alpha-syllabaries with complex orthographic features.

Possibilities can be drawn from a few other EGRA assessments that have taken the unique orthographic conventions of the writing systems in Southeast Asia into account. A case in point is the EGRA in Cambodia, which tested children's knowledge not just of independent consonants and vowels but also of consonant subscripts and dependent vowels. This assessment found that grade 1 students had more difficulty identifying the more complex letter and diacritic forms than the simple forms (RTI International, 2018a), as would be predicted by the literature on skills progression in reading in alpha-syllabaries.

Similarly, EGRAs in Nepal have included *matra* identification tasks that assess children's knowledge of how dependent vowel markings combine with consonants to form syllables (RTI International, 2014). Like vowel diacritics in Lao, Nepali script includes a set of 10 dependent vowels (*matras)*, which do not appear independently but are ligatured to consonants. These consonant + dependent vowel combinations form syllables that are the building blocks of Nepali words and are explicitly taught in early primary school.

Keeping this linguistic feature in mind, a *matra* test was used in addition to a letter test in a 2014 EGRA assessment of students in Nepal. In this task, students had to "produce the sounds associated with each *matra* as quickly and accurately as they could within one minute, yielding a score of correct *matras* per minute (cmpm)" (RTI International, 2014, p. 9). As might be expected, across grades 2 and 3, students performed better on naming letters in isolation than on the relatively more orthographically complex reading of *matra*, where zero scores were prevalent.

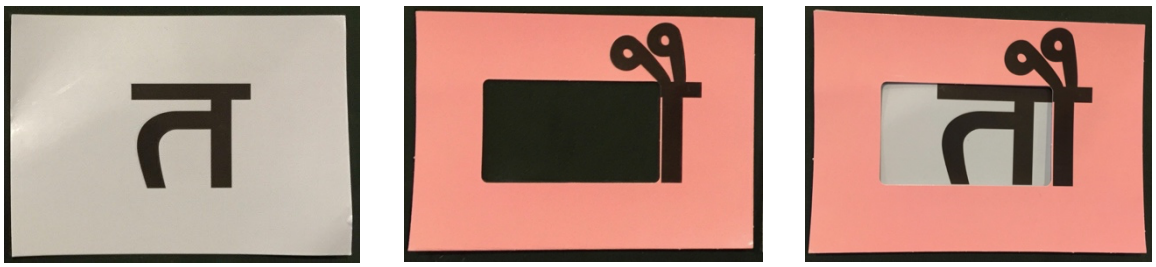
In addition, an orthographic choice task was used in a study on how well emergent literacy skills predicted word and text reading in Korean (Kim, 2015). In this task, children were presented with consonant-vowel syllable combinations and were asked to identify which were permitted according to the orthographic rules of the language and which violated those rules. For example, "in Korean, a vertical vowel letter is only permitted on the right of the preceding consonant, not on the left. Therefore, ㅏ오 is illegal, whereas 와 is legal" (Ibid., p. 466). The study found that orthographic awareness is a strong enough contributor to text-reading fluency that it should be explicitly taught and can be deliberately assessed, particularly for early grade learners. A similar assessment of language-specific symbol knowledge was utilized for Kannada and Telugu, Indian alpha-syllabic languages with large orthographic inventories similar to Lao (Nakamura, De Hoop, and Holla, 2019).

These examples from similar languages and writing systems in Cambodia, Nepal, Korea, and India point to potential tests of orthographic knowledge that could be used in Laos. For example, in Lao the vowel ໄ 'ai' always appears before the base

consonant, while the vowel า ‘aa’ always appears after the consonant. Other vowels can only appear as subscripts or superscripts around the consonant, such as the subscript vowel ‘uu’ in the word ครู (teacher) or the superscript vowel ‘ii’ in the word ปี (year).

A task could be constructed in which orthographically correct and incorrect vowel placements are shown and children are asked to indicate which ones are legal or illegal. Alternatively, children could be given a set of vowel and consonant manipulatives and asked to create legal combinations of consonants and vowels with the vowel letter or diacritic in the correct position relative to the consonant, like a pedagogical approach used to teach consonant + *matras* in early primary in Nepal. **Figure 20** below illustrates how this is done: The children are given consonant cards and ‘window’ cards with dependent vowel diacritics and asked to form legal syllables by combining the two.

**Figure 20: Illustration of Nepali Matras**



**Nepali consonant ‘t’** **Nepali *matra* ‘au’** **Nepali syllable ‘tau’**

*Source: The author, 2019*

These orthographic skills need not only be assessed but could also become a useful element of teaching and learning in early primary school. However, the experiences from the other countries and the literature review on the orthographic complexity of alpha-syllabaries provide a cautionary tale that any such orthographic choice items must be appropriately tailored to the developmental and linguistic status of participating children, or they will be too difficult. In addition, enumerators would

need careful training in how to score these items, and the tasks would need to be evaluated for reliability and validity before use at a wider scale.

Another unique feature of the Lao writing system, which is now found in only a handful of the world's languages, is its *scriptio continua* nature. Challenges related to *scriptio continua* text did not affect the LEARN assessment extensively, because it did not include any tests of reading connected text. However, for future assessments that include slightly older children, it may be possible to develop and validate assessments that help measure the extent to which they are able to recognize word boundaries rapidly and correctly in continuous text. Future assessments may also consider including tests of morphemic awareness, such as the ability to use lexical or contextual cues to determine the correct meaning in sentences where word boundaries could be placed differently and result in different combinations of meaning. An example of a word boundary recognition test could entail showing children who are readers a basic sentence in *scriptio continua* and ask them to correctly place word boundaries, such as the following (with its English translation):

ຂ້ມັກກິນປາ
ຂ້/ມັກ/ກິນ/ປາ

Iliketoeatfish
I/like/to eat/fish

The literature review also identified that including word boundaries when teaching reading in the early primary grades may be more effective in helping children distinguish between words than utilizing *scriptio continua* for younger students.

A virtually un-researched area of inquiry pertains to the valid assessment of tone in languages that utilize lexical tone to distinguish meaning. For this purpose, it is useful to employ a continuum of how assessments treat tonality, from 'tone-blind' to 'tone-aware.' The LEARN assessment – and other, similar assessments in the Lao context – was largely tone-blind. Enumerators were not explicitly trained or provided

uniform guidance on how to score children's responses in cases where they could not fully produce the correct tonality in Lao, used regional rather than Vientiane standard tones, or could not produce tones at all.

In addition, analysis of the tonal complexity of words used in the phonological awareness, listening comprehension in Lao, receptive vocabulary, most used words, and decodable words tests was not conducted. In fact, the decodable words assessment actively violated Lao tone rules, which are closely linked to how a word is spelled. Nor were efforts put in place to develop sub-tasks to assess children's tonality in Lao language or consider how the tones in their native languages might affect their ability to grasp tone in Lao. Furthermore, there was no assessment of children's ability to perceive and reproduce long and short vowel lengths, to which the correct application of tone in Lao is intricately tied.

To ensure that future assessments move from tone-blind to tone-aware, several steps could be taken. First, enumerators could be provided with a clear protocol on how to score children's production of tone, which is particularly important for L2 learners whose L1 accent and tonal repertoire will impact the way they pronounce tone in Lao. Thus, the protocol should include guidance on what counts as the 'correct' tone (with space for regional variations), and possibly include a system of partial credit, for example if children incorrectly apply or vocalize tone but pronounce the base syllable correctly according to its constituent phonemes.

The decodable words task presented particular challenges for children to correctly complete it and enumerators to appropriately score the production of tone, because there is no clear way to apply tone to nonsense words that do not follow Lao spelling rules. Nevertheless, decodable words tests are useful because they assess children's true decoding skills rather than their sight word knowledge. One alternative to dropping the decodable words test entirely in Laos would be to ensure that all decodable words in the task follow spelling and vowel length rules despite being nonsense words, and then assess whether children are able to apply the correct tone according to those rules.

A workaround that was applied in assessment of children's production of tones in a test of Cantonese syllables used a legal syllable with tone marked, and asked children to delete the onset sound while pronouncing the correct rime and tone for the syllable (Luk and Bialystock, 2008). Another test of decoding skills in Chinese tone languages – using legal combinations of characters with tones indicated – demonstrated the importance of including a large set of items to allow for more discrimination and paying careful attention to the tones, onsets, and rimes of each pseudo-word (Leong, Chen, and Tang, 2005). This was more straightforward in Cantonese because it is possible to indicate each tone in the language with a number; perhaps only nonsense words with the two explicitly marked Lao tones could be included as a solution, although this would limit the repertoire of words available for the task.

Second, an analysis of the tonal complexity of vocabulary words and listening comprehension passages should be part of the process for constructing tasks and items. The same attention to orthographic and tonal complexity should be paid when constructing listening comprehension passages. Reading comprehension passages are typically leveled using readability formulae that take into account a combination of word and sentence complexity, and this could also apply to listening comprehension. In the case of alpha-syllabic, tonal languages, this should be expanded to examine syllable, word, and sentence length (Room to Read, 2018) in addition to orthographic and tonal complexity.

For example, assessors could consult linguists and teachers to identify the tones that are typically easier for young second language learners to hear and reproduce – which may be different for Khmu and Hmong children – and include words and passages that incorporate those easier tones. Then, as children get older and instruction of tones becomes more advanced, more complex tone elements could be added. Curriculum and instruction in the summer pre-primary and into grade 1 could likewise adopt similar guidelines for selecting vocabulary that aligns with children's skills with tone at different developmental and educational stages, and particular

attention to helping Khmu children recognize tones in the first place and Hmong children to differentiate between the tones of Hmong and Lao.

Third, tasks that specifically assess children's tonemic awareness could be developed, again in consultation with linguists and educators. This could include tests similar to those identified in the literature review, in which children listen to minimal pairs of the same morpheme spoken with two different tones and are asked to distinguish them. Use of simple minimal pairs to help children distinguish between different tones could also be introduced as part of the teaching curriculum in the summer pre-primary and grade 1.

Most prior examples of the use of minimal pairs in learning assessments come from Chinese, such as a receptive tone awareness assessment used for Mandarin Chinese pre-readers that presented the same sets of morphemes with two different tones applied (McBride-Chang et al., 2008). A recorded voice then spoke one of the morpheme-tone combinations, and children were shown pictures of the items with the same morpheme but different tones and asked to point to the picture with the word that had been spoken aloud. To illustrate this, an example is 娘 /fu5/ (woman), and 父 /fu6/ (father). 48 items were included, with deliberate attention to the frequency of different tone types, such as mid level-mid level and low falling-low rising. The study identified that tone awareness was associated with Chinese character reading abilities, indicating that explicit assessment of tone can yield informative assessment results for the purposes of improving reading instruction in tonal languages.

A similar tone awareness task was developed for Mandarin Chinese, in which "children were given 10 pair of characters that had the same combinations of phonemes but composed of different tones. Then, they were asked to judge whether the two characters 'sound' the same or different. Three training trials were presented before the 10 experimental items" (Luk and Bialystok, 2008, p. 275). The items displayed only moderate internal reliability, however, with the authors postulating that this occurred because "tone is not an independent phonological construct. It binds

phonology and semantics and is deeply related to both of these linguistic structures," and also because tone sounds in Chinese overlap with each other (Ibid.). This points to the importance of thorough linguistic and psychometric evaluation of any minimal pairs task that might be added to future assessments in Laos.

Assessors may also want to consider standardizing the tasks by providing audio recordings of the minimal pairs to avoid variations in pronunciation by different enumerators. In fact, the literature review identified that non-native speakers may perform better on the task when shown both audio and video of the person speaking, as this provides visual cues regarding which tone is being expressed. In line with this finding, teachers could also be encouraged to combine both facial expressions and spoken utterances of tone to facilitate tone awareness skills for Khmu and Hmong children.

Children's chronemic awareness – which should also be a core element of instruction for young children learning Lao as a second language – could be assessed through a similar minimal pairs approach. In the case of assessments in Lao, this would focus on awareness of contrastive vowel length, which has a significant bearing on tone application and therefore meaning making in the language. For example, children could be asked to listen to and differentiate between minimal pairs of the same vowel sound in short and long form. Piloting would need to be done to identify the approach that is most valid for Lao phonological structure, such as listening to contrasting vowel sounds as individual phonemes on their own or within syllables.

Finally, the LEARN assessment did not incorporate measures of oral reading fluency (ORF) because the target population would have been at the floor on such an assessment. However, for the future it is recommended to implement longitudinal assessments in Laos that follow children through at least grade two and ideally into grade three. These assessments should expand on thinking already done around how best to assess fluency, accuracy, and comprehension and establish optimal

ORF benchmarks in tonal languages (Abadzi, 2011; Abadzi, 2013; Room to Read, 2018).

For example, ORF rates for children who read Lao with comprehension were around 60 words per minute for grade 2 students but fell to just over 20 words per minute for grade 3 children in an assessment conducted by Room to Read (Jukes, Cummiskey, and Jargano, 2018). This may have occurred because there were so many zero scores in grade 2 that the children who were included in the ORF calculations were only the high-performing, rapid readers; and possibly because as children become more competent readers of tonally complex text in grade 3 and beyond, they read more slowly but each word encodes more meaning (Ibid.). These and other points about assessment of oral reading fluency in tonal languages remain to be further empirically substantiated through future research.

### *6.4. Section Conclusions: Strengthening Future Assessments in the Lao Context*

This section concludes by summarizing the recommendations for strengthening similar future assessments in the Lao context based on experiences in the LEARN study. The study findings suggest that, although the assessment tool and evaluation approach used in LEARN were a good starting point, substantial modifications will be required to capture more useful data going forward.

First, although it is a specified goal of the education system to promote Lao language skills in early primary, the reality is that young children in disadvantaged, non-Lao communities are unlikely to have sufficient skills in Lao language before primary school and into first grade. Because the language of assessment was such a critical impediment to measuring children's true knowledge and skills, a multi-language protocol should be employed in which children are tested in their home languages through the 'overall language ability' approach (Snow, Burns, and Griffin, 1998) at least until they reach a minimum threshold of familiarity with Lao as the language of instruction. Tasks assessing Lao language skills could be added towards the end of grade 1 and into grades 2 and 3 to avoid the pitfall of extensive floor effects found

prior to grade 1 in the LEARN study. Ideally, children should also continue to be assessed in their home languages in addition to Lao on language-related tasks throughout the first two grades of primary school, to contribute to more effective measurement of their underlying conceptual knowledge and better understanding of cross-language transfer between home languages and the language of instruction. Targeted recruitment of Hmong- and Khmu-speaking enumerators, approaches to standardize language use in the assessment instructions, better training in scoring of non-Lao students' performance when they are assessed in Lao, and use of conceptual scoring approaches (Gross, Buac, and Kaushanskaya, 2014) could also help to increase the appropriateness of future evaluations.

On a second and related point, the assessment tool should be revised to ensure greater relevance to the population of children and their developmental stage. The literature review highlights that young, non-Lao learners are expected to be at the very early language-related stage of development at most, and not yet at the code-related stage by around the age of five or six, before formal school readiness instruction and without significant prior exposure to spoken Lao. As such, the assessment should include ample opportunities for children to demonstrate baseline status and progression over time in language-related skills before moving to code-related skills. The listening comprehension and receptive vocabulary tests that functioned relatively well in the LEARN evaluation could be supplemented by application of the semantic fluency test at all data collection waves and not just at endline; incorporation of an additional expressive vocabulary task from the IDELA tool that involves naming common market items and animals (Save the Children, 2017); and/or deployment of other items from USAID's oral language assessment modules, such as retelling a story in the child's own words (Chiappetta, 2019).

The initial sound identification test of phonological awareness functioned poorly at all waves and most likely should be dropped in future assessments or considered only after children have experienced explicit instruction in phoneme isolation and blending. The word segmentation task proved very difficult at baseline but became easier over time, in line with the literature review finding that the syllabic awareness

skills measured through word segmentation are relevant to acquisition of skilled reading in alpha-syllabic languages like Lao. In consultation with linguists, future assessments could consider adding further syllabic awareness assessments, such as syllable blending, syllable deletion, or onset/ryme tasks to identify common syllabograms, as discussed in the Ji (2018), Kasisopa et al. (2018), and Nag (2017) studies. These tasks would need to be designed with due attention to the complexity of the syllabograms and their appropriateness to the grade and developmental levels of the children (Vagh and Sharma, 2018), starting with oral skills and only moving into written syllabic tasks in grades 2-3.

Given the complexities of learning to read in alpha-syllabic, tonal languages highlighted in the literature review, non-Lao children would also not be expected to develop extensive code-related alphabetic, orthographic, and decoding skills until grades 2-3. Going forward, code-related tasks such as letter name knowledge and word-level decoding should not be introduced until after grade 1, while tasks to assess orthographic awareness could be introduced at around the same time to better measure this core skill for alpha-syllabic languages. This could include tests similar to the ones that were conducted in Korean for legal consonant-diacritic combinations (Kim, 2015) or in Kannada and Telugu for awareness of diacritic symbols (Nakamura, De Hoop, and Holla, 2019). Considerations related to the *scriptio continua* nature of Lao are likewise not likely to become relevant until non-Lao children have firmly entered the decoding stage in grades 2-3; at that time, tasks could be considered for discerning word boundaries or using lexical and contextual cues to distinguish meaning where word boundaries are ambiguous.

Future assessments should also be designed to ensure that they are more 'tone-aware' through careful attention to the tonal complexity of words in language-related tasks such as listening comprehension; issues of tonality in scoring of responses as correct or incorrect for non-Lao speakers; and assessment of tone-related skills such as the ability to distinguish different tones and contrastive vowel length in oral language, similar to what has been done in Mandarin Chinese (McBride-Chang et al., 2008; Luk and Bialystok, 2008). The decodable word reading task should be

dropped entirely unless Lao linguists can identify a way to construct the items so that tone rules are not violated and so that enumerators can correctly score the application of tone in 'nonsense' words. For reading comprehension tests of older children – recommended to start no earlier than grade 2 or 3 for non-Lao learners – future assessments could focus on establishing benchmarks of oral reading fluency that specifically consider the tonal complexity of the words being assessed and how factors such as tone type and morpheme length interact with oral reading fluency rates in Lao.

Finally, future interventions should duly consider issues of content validity, ecological validity, and use of formative assessment to make the testing approach more relevant to the summer pre-primary course objectives and non-Lao children's naturalistic environments, and to provide more actionable assessment data to teachers on an ongoing basis than what is typically provided by higher-stakes summative assessments such as the one used in LEARN. Promoters of the summer pre-primary should also consider framing and evaluating the model as primarily a 'foot-in-the-door' intervention that is designed to provide children with a boost of school readiness at the start of grade 1; or contemplate identifying what would constitute 'trifecta' skills in the Lao context, which are both malleable to instruction and would not be expected to develop on their own for all children through formal learning in primary school. Future evaluations should also contemplate increasing study sample size to the extent possible to mitigate the risks of sample attrition and dilution of statistical power; dedicate time and resources to careful verification of children's actual uptake of the treatment or control status; and identify ways to reduce differential attrition by the most vulnerable Hmong children from both the intervention and the assessment sample.

Taken together, these modifications to the LEARN evaluation approach should contribute to a more contextually and linguistically valid assessment that provides disadvantaged, non-Lao learners more opportunities to demonstrate their true skills and enables program implementers to better pinpoint the effects of the summer pre-primary model for the marginalized children it explicitly aims to support.

### *6.5. Study Limitations*

This section discusses study limitations before turning to overall conclusions in **Section 7.** This study suffered from four main limitations: Imperfect adherence to randomization; imperfect uptake of experimental treatments and difficulties in verifying uptake status; challenges common to longitudinal designs, including practice effects and sample attrition; and small sample size leading to shortcomings in disaggregation of findings for different ethnic groups.

Imperfect adherence to randomized assignments meant that the actual treatment groups could have varied in non-random ways, and made traditional RCT analytical approaches, such as Intention to Treat, problematic to apply. This compromised randomization also threatened the internal validity of the study and the ability to attribute outcomes to the interventions.

In addition, significant difficulties were faced in verifying actual treatment uptake for all groups in the study based on variables collected through the survey instruments. Although the LEARN team could use project enrollment records to confirm uptake of children in the summer pre-primary program, this was not possible in the control schools, resulting in incomplete or contradictory data for a number of children in the sample and a quandary about how to consider those children for analytic purposes.

Despite their methodological strengths, longitudinal designs suffer from two main potential limitations. The first is practice effects, in which participants are able to improve their responses in later waves because they already have experience answering the same questions in previous waves. This was mitigated to some extent by the fact that there were large floor effects at baseline, indicating that many children simply could not engage with the assessment questions and were therefore unlikely to remember the test items. Another mitigating factor, which is often built into longitudinal designs (Tolmie, Muijs, and McAteer, 2011), was the long duration of time between follow-up waves. In this case, there was one year between baseline and midline, and six months between midline and endline. In addition, the type of

performance measures used in the study are less prone to practice effects from repeated use, because the “complexity … of the variables being measured reduce carry-over and practice effects” (Ibid., p. 43).

Another weakness of longitudinal designs relates to sample attrition, in which the participants who drop out of a sample may systematically differ from those who remain in the study and therefore sample bias is introduced. This can be a particular risk in studies based in educational institutions such as schools, where the children who drop out are likely to come from home circumstances that are less stable or able to support children’s education (Ibid.).

Finally, the small size of the final analytic sample resulted in Lao children being removed and precluded testing the hypothesis that the intervention was at least as effective for non-Lao children as it was for native Lao speakers. The relatively small sample sizes also weakened the statistical power in sub-group comparisons of effects between Khmu and Hmong speakers.

These limitations were mitigated to the extent possible in reporting this thesis, as described in other sections, but must be borne in mind when interpreting findings.

## 7. Conclusions

This thesis has identified several areas where the LEARN Project summer pre-primary holds promise as a means of closing the school readiness gap for the most marginalized children in Lao PDR – results that are roughly in line with outcomes from similar programs around the world. The intervention was particularly effective in increasing children's gains in areas that were explicitly taught, including concepts about print, listening comprehension, and receptive vocabulary. Conversely, the model was less effective in areas that were not deliberately taught, such as letter name knowledge; or that were problematic from a linguistic perspective, such as phonological awareness.

There is some evidence that the summer pre-primary intervention benefitted Khmu children more than Hmong children between baseline and midline, while at the same time there are trends indicating that Hmong children in the treatment group may have excelled to some extent between midline and endline. Nevertheless, the higher rate of attrition from the sample of more marginalized Hmong children highlights the importance of targeted measures to ensure that such learners consistently participate. The consistent predictive nature of mother's education as a covariate also calls attention to the potential for interventions that raise women's educational levels to have positive knock-on effects for their children as well.

The thesis outlines how the summer pre-primary model might most effectively be marketed – and measured – as an intervention that helps disadvantaged children get their 'foot in the door' during a time of otherwise precarious transition from home to school. At the same time, to identify gains that persist into primary school, future assessments may need to measure skills that would not have developed otherwise. Moreover, future interventions may need to do more to strengthen the coherence between the school readiness and the grade 1 curricula as well as the quality of learning in first grade.

Lastly, the thesis raises specific suggestions for improving the validity of school readiness assessment approaches in the Lao context and for accelerated interventions. In particular, the thesis breaks new ground about MODEL's linguistic validity in assessing skills that are the defining features of Lao, such as syllabic structure and tonality. These recommendations could apply equally to improvements in the instructional content of the summer pre-primary and into grade 1 for disadvantaged non-Lao children. The LEARN experience also underscores the need to apply linguistically and contextually appropriate assessment approaches, such as hiring enumerators who speak the same language as children; finding ways to ensure that language is used uniformly in all assessment interactions; assessing young children in familiar languages; and applying conceptual scoring that captures children's underlying knowledge in different languages. The literature review highlights, however, how these choices are not simply assessment-related but also have political implications that must be addressed carefully. Finally, the limited match between what the assessment measured and what was taught in the summer pre-primary course highlights the need to supplement core items in the MODEL tool with items that are more directly relevant to measuring intervention impact.

Persistent poor performance by disadvantaged non-Lao children is a defining feature of the Lao education system from pre-primary age well into the first few grades of primary school. The thesis has put forward suggestions to ensure that assessment is better adapted to the reality these children face and therefore more likely to detect meaningful skill levels and impacts. However, a final but important conclusion is that it will not be enough to apply these recommendations to future assessment practice alone: They should also be considered and evaluated as curriculum and instructional improvements for disadvantaged, non-Lao populations. Baird et al. (2017) refer to this as systemic validity, in which tests are designed to be worth teaching to. This establishes a virtuous cycle in which well-designed assessments engender instructional strategies to develop the skills that those assessments are intended to measure. As the 2015 Incheon Declaration proclaimed, "no education target should be considered met unless met by all" (UNESCO, p. 7). By the same token, no

assessment paradigm should be considered fully valid, unless valid and useful for all.

## Annex 1: Glossary of Terms

- Non-Lao: In the Lao context and in this thesis, the term ‘non-Lao’ refers to any ethnic group other than the majority Lao-Tai ethnic group. In other contexts, this might typically be referred to as an ‘ethnic minority,’ but the latter term is not favored for in use in Laos.
- Home Language: This refers to the main language spoken by children at home. In other contexts, this might typically be referred to as ‘mother tongue,’ a term that is also not used in the Lao context.

## Annex 2: Acronyms

| | |
|---|---|
| AÇEV: | Anne Çocuk Eğitim Vakfı (Mother Child Education Foundation) |
| AIR: | American Institutes for Research |
| ANCOVA: | Analysis of Covariance |
| ANOVA: | Analysis of Variance |
| ASLO: | Assessment of Student Learning Outcomes |
| ASR: | Accelerated School Readiness |
| CCDG: | Community Child Development Groups |
| Chi Sq.: | Chi Square |
| COVID: | Coronavirus Disease 2019 |
| df: | Degrees of Freedom |
| DESB: | District Education and Sports Bureau |
| DIBELS: | Dynamic Indicators of Basic Early Literacy Skills |
| Diff.: | Differential |
| ECD: | Early Childhood Development |
| ECE: | Early Childhood Education |
| EGRA: | Early Grade Reading Assessment |
| EMIS: | Education Management Information System |
| ESDP: | Education Sector Development Plan |
| GER: | Gross Enrollment Ratio |
| GPE: | Global Partnership for Education |
| IDELA: | International Development and Early Learning Assessment |
| INGO: | International Non-governmental Organization |

| | |
|---|---|
| ID: | Identification |
| IES: | Institute of Education Sciences |
| IOE: | Institute of Education |
| ITT: | Intention to Treat |
| KIX: | Knowledge Innovation Exchange |
| L1: | First Language |
| L2: | Second Language |
| LATE: | Local Average Treatment Effect |
| LDC: | Least Developed Country |
| LEARN: | Lao Educational Access, Research and Networking Project |
| LMIC: | Lower Middle-Income Country |
| MAT: | Multi-age Teaching |
| MELQO: | Measuring Early Learning Quality and Outcomes |
| MODEL: | Measurement of Development and Early Learning |
| MOES: | Ministry of Education and Sports |
| MTR: | Mid-term Review |
| N: | Number |
| n.s.: | Non-significant |
| NLS: | New Literacy Studies |
| OLS: | Ordinary Least Squares |
| ORF: | Oral Reading Fluency |
| PDR: | People's Democratic Republic |
| PESS: | Provincial Education and Support Service |
| QED: | Quasi-Experimental Design |

| | |
|---|---|
| RAN: | Rapid Automized Naming |
| RCT: | Randomized Control Trial |
| RTI: | Research Triangle Institute International |
| RIES: | Research Institute for Education Sciences |
| RPA: | Romanized Popular Alphabet |
| SD: | Standard Deviation |
| SDG: | Sustainable Development Goal |
| SE: | Standard Error |
| SES: | Socio-economic Status |
| SVR: | Simple View of Reading |
| T1: | Time 1 |
| T2: | Time 2 |
| T3: | Time 3 |
| T4: | Time 4 |
| ToT: | Training of Trainers |
| UCL: | University College London |
| UK: | United Kingdom |
| UNESCO: | United Nations Educational, Scientific and Cultural Organization |
| UNICEF: | United Nations Children's Fund |
| USAID: | United States Agency for International Development |
| WPM: | Words Per Minute |

## Annex 3: Data Tables

### Table 23: Average Baseline, Midline, and Endline Standardized Scores on All Tasks, Original Sample Overall

| | N | Mean | SD |
|---|---|---|---|
| Baseline Literacy Domain | 626 | 0.11 | 0.11 |
| Midline Literacy Domain | 640 | 0.27 | 0.20 |
| Endline Literacy Domain | 632 | 0.39 | 0.24 |
| Baseline Concepts about Print | 631 | 0.15 | 0.23 |
| Midline Concepts about Print | 640 | 0.33 | 0.31 |
| Endline Concepts about Print | 632 | 0.42 | 0.29 |
| Baseline Initial Sound Identification | 630 | 0.02 | 0.12 |
| Midline Initial Sound Identification | 640 | 0.05 | 0.16 |
| Endline Initial Sound Identification | 632 | 0.07 | 0.21 |
| Baseline Word Segmentation | 629 | 0.02 | 0.12 |
| Midline Word Segmentation | 640 | 0.16 | 0.33 |
| Endline Word Segmentation | 632 | 0.22 | 0.37 |
| Baseline Letter Name Knowledge | 628 | 0.03 | 0.11 |
| Midline Letter Name Knowledge | 640 | 0.18 | 0.30 |
| Endline Letter Name Knowledge | 632 | 0.37 | 0.39 |
| Baseline Listening Comprehension in Lao | 627 | 0.07 | 0.19 |
| Midline Listening Comprehension in Lao | 640 | 0.22 | 0.31 |
| Endline Listening Comprehension in Lao | 632 | 0.38 | 0.35 |
| Baseline Listening Comprehension in Ethnic Languages | 552 | 0.25 | 0.31 |
| Midline Listening Comprehension in Ethnic Languages | 584 | 0.41 | 0.33 |
| Baseline Name Writing | 626 | 0.06 | 0.17 |
| Midline Name Writing | 640 | 0.34 | 0.38 |
| Endline Name Writing | 632 | 0.54 | 0.45 |
| Baseline Receptive Vocabulary | 626 | 0.42 | 0.30 |

| | N | Mean | SD |
|---|---|---|---|
| Midline Receptive Vocabulary | 640 | 0.64 | 0.27 |
| Endline Receptive Vocabulary | 632 | 0.72 | 0.21 |
| Semantic Fluency | 465 | 0.30 | 0.20 |
| Most Used Words | 465 | 0.05 | 0.16 |
| Decodable Words | 465 | 0.03 | 0.13 |

**Table 24: Average Baseline, Midline, and Endline Standardized Scores on All Tasks, Original Sample by Ethnicity**

| | | Lao | | | Khmu | | | Hmong | | | ANOVA | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | N | Mean | St. Dev. | N | Mean | St. Dev. | N | Mean | St. Dev. | *F* | DF | *p* |
| **Literacy Domain** | Baseline | 57 | 0.26 | 0.16 | 481 | 0.10 | 0.10 | 88 | 0.07 | 0.07 | 67.78 | 2, 625 | .000 |
| | Midline | 56 | 0.50 | 0.23 | 492 | 0.27 | 0.19 | 92 | 0.15 | 0.14 | 60.07 | 2, 639 | .000 |
| | Endline | 56 | 0.56 | 0.20 | 490 | 0.39 | 0.24 | 86 | 0.26 | 0.18 | 29.56 | 2, 631 | .000 |
| **Concepts about Print** | Baseline | 58 | 0.32 | 0.33 | 485 | 0.15 | 0.22 | 88 | 0.06 | 0.14 | 24.14 | 2, 630 | .000 |
| | Midline | 56 | 0.50 | 0.34 | 492 | 0.34 | 0.31 | 92 | 0.20 | 0.26 | 17.84 | 2, 639 | .000 |
| | Endline | 56 | 0.52 | 0.28 | 490 | 0.42 | 0.30 | 86 | 0.33 | 0.23 | 8.23 | 2, 631 | .000 |
| **Initial Sound Identification** | Baseline | 58 | 0.06 | 0.18 | 484 | 0.02 | 0.11 | 88 | 0.01 | 0.08 | 3.13 | 2, 629 | .044 |
| | Midline | 56 | 0.17 | 0.29 | 492 | 0.04 | 0.13 | 92 | 0.03 | 0.14 | 20.89 | 2, 639 | .000 |
| | Endline | 56 | 0.10 | 0.25 | 490 | 0.08 | 0.22 | 86 | 0.02 | 0.12 | 3.49 | 2, 631 | .031 |
| **Word Segmentation** | Baseline | 58 | 0.06 | 0.19 | 483 | 0.02 | 0.12 | 88 | 0.00 | 0.00 | 4.59 | 2, 628 | .011 |
| | Midline | 56 | 0.37 | 0.42 | 492 | 0.16 | 0.32 | 92 | 0.04 | 0.18 | 18.61 | 2, 639 | .000 |
| | Endline | 56 | 0.34 | 0.43 | 490 | 0.23 | 0.38 | 86 | 0.07 | 0.22 | 10.21 | 2, 631 | .000 |
| **Letter Name Knowledge** | Baseline | 58 | 0.09 | 0.21 | 482 | 0.03 | 0.10 | 88 | 0.01 | 0.07 | 9.61 | 2, 627 | .000 |
| | Midline | 56 | 0.48 | 0.43 | 492 | 0.16 | 0.28 | 92 | 0.07 | 0.17 | 39.90 | 2, 639 | .000 |
| | Endline | 56 | 0.63 | 0.41 | 490 | 0.37 | 0.38 | 86 | 0.20 | 0.29 | 21.95 | 2, 631 | .000 |

| | | Lao | | | Khmu | | | Hmong | | | ANOVA | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | N | Mean | St. Dev. | N | Mean | St. Dev. | N | Mean | St. Dev. | *F* | DF | *p* |
| **List. Comp. in Lao** | Baseline | 57 | 0.35 | 0.34 | 482 | 0.05 | 0.15 | 88 | 0.01 | 0.08 | 88.85 | 2, 626 | .000 |
| | Midline | 56 | 0.61 | 0.29 | 492 | 0.20 | 0.29 | 92 | 0.04 | 0.14 | 79.29 | 2, 639 | .000 |
| | Endline | 56 | 0.76 | 0.27 | 490 | 0.36 | 0.34 | 86 | 0.18 | 0.27 | 56.09 | 2, 631 | .000 |
| **List. Comp. in Ethnic Lang.** | Baseline | | | | 464 | 0.26 | 0.31 | 88 | 0.17 | 0.27 | 6.93 | 1, 551 | .009 |
| | Midline | | | | 492 | 0.44 | 0.33 | 92 | 0.26 | 0.30 | 22.70 | 1, 583 | .000 |
| **Name Writing** | Baseline | 57 | 0.16 | 0.26 | 481 | 0.06 | 0.16 | 88 | 0.03 | 0.11 | 12.76 | 2, 625 | .000 |
| | Midline | 56 | 0.44 | 0.39 | 492 | 0.35 | 0.38 | 92 | 0.24 | 0.35 | 5.15 | 2, 639 | .006 |
| | Endline | 56 | 0.69 | 0.40 | 490 | 0.54 | 0.44 | 86 | 0.47 | 0.47 | 4.11 | 2, 631 | .017 |
| **Receptive Vocabulary** | Baseline | 57 | 0.76 | 0.21 | 481 | 0.39 | 0.29 | 88 | 0.37 | 0.29 | 45.13 | 2, 625 | .000 |
| | Midline | 56 | 0.90 | 0.09 | 492 | 0.65 | 0.26 | 92 | 0.43 | 0.28 | 62.64 | 2, 639 | .000 |
| | Endline | 56 | 0.87 | 0.14 | 490 | 0.73 | 0.20 | 86 | 0.56 | 0.23 | 44.84 | 2, 631 | .000 |
| **Semantic Fluency** | Endline | 42 | 0.43 | 0.20 | 360 | 0.30 | 0.19 | 63 | 0.19 | 0.02 | 20.73 | 2, 464 | 0.000 |
| **Most Used Words** | Endline | 42 | 0.08 | 0.22 | 360 | 0.06 | 0.17 | 63 | 0.00 | 0.03 | 3.39 | 2, 464 | 0.035 |
| **Decodable Words** | Endline | 42 | 0.05 | 0.19 | 360 | 0.03 | 0.13 | 63 | 0.00 | 0.02 | 2.21 | 2, 464 | 0.111 |

**Table 25: Mean Gain Scores and Effect Sizes, Treatment vs. Control**

| | | N | Mean | Std. Dev. | Mean Diff. | *d* |
|---|---|---|---|---|---|---|
| Baseline Literacy Domain | Cont. | 177 | .10 | .09 | -0.01 | -0.07 |
| | Treat. | 210 | .09 | .09 | | |
| Literacy Domain Gain Baseline-Midline | Cont. | 177 | .13 | .17 | 0.07 | 0.36 |
| | Treat. | 210 | .20 | .19 | | |
| Literacy Domain Gain Midline-Endline | Cont. | 179 | .12 | .18 | -0.01 | -0.08 |
| | Treat. | 212 | .11 | .18 | | |
| Literacy Domain Gain Baseline-Endline | Cont. | 177 | .25 | .23 | 0.05 | 0.25 |
| | Treat. | 210 | .31 | .22 | | |
| Baseline Concepts About Print | Cont. | 177 | .48 | .67 | -0.15 | -0.25 |
| | Treat. | 211 | .33 | .54 | | |
| Concepts About Print Gain Baseline-Midline | Cont. | 177 | .11 | .35 | 0.14 | 0.42 |
| | Treat. | 211 | .25 | .34 | | |
| Concepts About Print Gain Midline-Endline | Cont. | 179 | .13 | .36 | -0.08 | -0.22 |
| | Treat. | 212 | .05 | .37 | | |
| Concepts About Print Gain Baseline-Endline | Cont. | 177 | .24 | .35 | 0.07 | 0.20 |
| | Treat. | 211 | .31 | .32 | | |
| Baseline Initial Sound Identification | Cont. | 177 | .02 | .12 | 0.00 | 0.02 |
| | Treat. | 211 | .02 | .11 | | |
| Initial Sound Identification Gain Baseline-Midline | Cont. | 177 | .00 | .15 | 0.02 | 0.13 |
| | Treat. | 211 | .02 | .18 | | |
| Initial Sound Identification Gain Midline-Endline | Cont. | 179 | .05 | .21 | 0.00 | -0.02 |
| | Treat. | 212 | .05 | .25 | | |
| Initial Sound Identification Gain Baseline-Endline | Cont. | 177 | .05 | .24 | 0.02 | 0.09 |
| | Treat. | 211 | .07 | .26 | | |
| Baseline Word Segmentation | Cont. | 177 | .02 | .12 | 0.00 | -0.01 |
| | Treat. | 211 | .02 | .11 | | |
| Word Segmentation Gain Baseline-Midline | Cont. | 177 | .11 | .30 | 0.06 | 0.19 |
| | Treat. | 211 | .17 | .37 | | |

| | | N | Mean | Std. Dev. | Mean Diff. | *d* |
|---|---|---|---|---|---|---|
| Word Segmentation Gain Midline-Endline | Cont. | 179 | .08 | .43 | -0.03 | -0.06 |
| | Treat. | 212 | .06 | .47 | | |
| Word Segmentation Gain Baseline-Endline | Cont. | 177 | .19 | .38 | 0.04 | 0.10 |
| | Treat. | 211 | .23 | .40 | | |
| Baseline Letter Name Knowledge | Cont. | 177 | .01 | .06 | 0.01 | 0.10 |
| | Treat. | 210 | .02 | .07 | | |
| Letter Name Knowledge Gain Baseline-Midline | Cont. | 177 | .14 | .26 | 0.00 | 0.00 |
| | Treat. | 210 | .14 | .26 | | |
| Letter Name Knowledge Gain Midline-Endline | Cont. | 179 | .18 | .29 | 0.01 | 0.04 |
| | Treat. | 212 | .19 | .30 | | |
| Letter Name Knowledge Gain Baseline-Endline | Cont. | 177 | .32 | .38 | 0.02 | 0.05 |
| | Treat. | 210 | .34 | .36 | | |
| Baseline Listening Comprehension in Lao | Cont. | 177 | .03 | .12 | 0.02 | 0.16 |
| | Treat. | 210 | .06 | .15 | | |
| Listening Comprehension in Lao Gain Baseline-Midline | Cont. | 177 | .11 | .25 | 0.07 | 0.24 |
| | Treat. | 210 | .18 | .30 | | |
| Listening Comprehension in Lao Gain Midline-Endline | Cont. | 179 | .15 | .29 | -0.03 | -0.09 |
| | Treat. | 212 | .13 | .30 | | |
| Listening Comprehension in Lao Gain Baseline-Endline | Cont. | 177 | .27 | .33 | 0.04 | 0.12 |
| | Treat. | 210 | .31 | .35 | | |
| Baseline Listening Comprehension Ethnic | Cont. | 173 | .24 | .31 | 0.00 | 0.00 |
| | Treat. | 201 | .24 | .31 | | |
| Listening Comprehension Ethnic Gain Baseline-Midline | Cont. | 173 | .15 | .38 | 0.04 | 0.09 |
| | Treat. | 201 | .19 | .39 | | |
| Baseline Name Writing | Cont. | 177 | .05 | .13 | 0.01 | 0.05 |
| | Treat. | 210 | .06 | .14 | | |
| Name Writing Gain Baseline-Midline | Cont. | 177 | .25 | .36 | 0.07 | 0.18 |
| | Treat. | 210 | .32 | .37 | | |
| Name Writing Gain Midline-Endline | Cont. | 179 | .20 | .46 | 0.02 | 0.04 |

| | | N | Mean | Std. Dev. | Mean Diff. | *d* |
|---|---|---|---|---|---|---|
| | Treat. | 212 | .21 | .46 | | |
| Name Writing Gain Baseline-Endline | Cont. | 177 | .45 | .44 | 0.09 | 0.19 |
| | Treat. | 210 | .53 | .46 | | |
| Baseline Receptive Vocabulary | Cont. | 177 | .40 | .27 | -0.03 | -0.11 |
| | Treat. | 210 | .37 | .30 | | |
| Receptive Vocabulary Gain Baseline-Midline | Cont. | 177 | .20 | .37 | 0.10 | 0.29 |
| | Treat. | 210 | .30 | .34 | | |
| Receptive Vocabulary Gain Midline-Endline | Cont. | 179 | .07 | .22 | 0.01 | 0.03 |
| | Treat. | 212 | .07 | .25 | | |
| Receptive Vocabulary Gain Baseline-Endline | Cont. | 177 | .26 | .30 | 0.11 | 0.37 |
| | Treat. | 210 | .37 | .30 | | |

**Table 26: Mean Gain Scores and Effect Sizes, Treatment vs. Control, by Ethnicity**

| | | | N | Mean | Std. Dev. | Mean Diff. | *d* |
|---|---|---|---|---|---|---|---|
| Baseline Literacy Domain | Cont. | Khmu | 157 | .10 | .09 | 0.00 | -0.03 |
| | Treat. | | 177 | .10 | .09 | | |
| | Cont. | Hmong | 20 | .10 | .05 | -0.03 | -0.38 |
| | Treat. | | 33 | .07 | .09 | | |
| Literacy Domain Gain Baseline-Midline | Cont. | Khmu | 157 | .14 | .17 | 0.08 | 0.45 |
| | Treat. | | 177 | .22 | .20 | | |
| | Cont. | Hmong | 20 | .05 | .16 | 0.01 | 0.07 |
| | Treat. | | 33 | .06 | .11 | | |
| Literacy Domain Gain Midline-Endline | Cont. | Khmu | 159 | .13 | .19 | -0.04 | -0.20 |
| | Treat. | | 179 | .09 | .18 | | |
| | Cont. | Hmong | 20 | .06 | .10 | 0.13 | 0.81 |
| | Treat. | | 33 | .19 | .19 | | |
| Literacy Domain Gain Baseline-Endline | Cont. | Khmu | 157 | .27 | .23 | 0.05 | 0.21 |
| | Treat. | | 177 | .32 | .22 | | |
| | Cont. | Hmong | 20 | .11 | .15 | 0.14 | 0.80 |
| | Treat. | | 33 | .25 | .19 | | |
| Baseline Concepts About Print | Cont. | Khmu | 157 | .49 | .68 | -0.12 | -0.19 |
| | Treat. | | 178 | .37 | .56 | | |
| | Cont. | Hmong | 20 | .40 | .60 | -0.28 | -0.62 |
| | Treat. | | 33 | .12 | .33 | | |
| Concepts About Print Gain Baseline-Midline | Cont. | Khmu | 157 | .11 | .35 | 0.17 | 0.48 |
| | Treat. | | 178 | .28 | .35 | | |
| | Cont. | Hmong | 20 | .10 | .39 | 0.02 | 0.07 |
| | Treat. | | 33 | .12 | .22 | | |
| Concepts About Print Gain Midline-Endline | Cont. | Khmu | 159 | .14 | .36 | -0.12 | -0.34 |
| | Treat. | | 179 | .02 | .37 | | |
| | Cont. | Hmong | 20 | .05 | .36 | 0.18 | 0.54 |

| | | | N | Mean | Std. Dev. | Mean Diff. | *d* |
|---|---|---|---|---|---|---|---|
| | Treat. | | 33 | .23 | .32 | | |
| Concepts About Print Gain Baseline-Endline | Cont. | Khmu | 157 | .25 | .36 | 0.05 | 0.13 |
| | Treat. | | 178 | .30 | .33 | | |
| | Cont. | Hmong | 20 | .15 | .28 | 0.20 | 0.74 |
| | Treat. | | 33 | .35 | .28 | | |
| Baseline Initial Sound | Cont. | Khmu | 157 | .02 | .13 | 0.00 | -0.02 |
| | Treat. | | 178 | .02 | .11 | | |
| | Cont. | Hmong | 20 | .00 | .00 | 0.03 | 0.30 |
| | Treat. | | 33 | .03 | .13 | | |
| Initial Sound Gain Baseline-Midline | Cont. | Khmu | 157 | .00 | .16 | 0.03 | 0.17 |
| | Treat. | | 178 | .03 | .19 | | |
| | Cont. | Hmong | 20 | .00 | .00 | -0.03 | -0.30 |
| | Treat. | | 33 | -.03 | .13 | | |
| Initial Sound Gain Midline-Endline | Cont. | Khmu | 159 | .06 | .22 | -0.01 | -0.04 |
| | Treat. | | 179 | .05 | .26 | | |
| | Cont. | Hmong | 20 | .00 | .00 | 0.04 | 0.28 |
| | Treat. | | 33 | .04 | .18 | | |
| Initial Sound Gain Baseline-Endline | Cont. | Khmu | 157 | .06 | .25 | 0.03 | 0.11 |
| | Treat. | | 178 | .08 | .26 | | |
| | Cont. | Hmong | 20 | .00 | .00 | 0.01 | 0.06 |
| | Treat. | | 33 | .01 | .23 | | |
| Baseline Word Segmentation | Cont. | Khmu | 157 | .02 | .12 | 0.00 | 0.00 |
| | Treat. | | 178 | .02 | .12 | | |
| | Cont. | Hmong | 20 | .00 | .00 | 0.00 | |
| | Treat. | | 33 | .00 | .00 | | |
| Word Segmentation Gain Baseline-Midline | Cont. | Khmu | 157 | .12 | .32 | 0.08 | 0.21 |
| | Treat. | | 178 | .20 | .39 | | |
| | Cont. | Hmong | 20 | .00 | .00 | 0.04 | 0.29 |
| | Treat. | | 33 | .04 | .19 | | |

| | | | N | Mean | Std. Dev. | Mean Diff. | *d* |
|---|---|---|---|---|---|---|---|
| Word Segmentation Gain Midline-Endline | Cont. | Khmu | 159 | .08 | .44 | -0.02 | -0.04 |
| | Treat. | | 179 | .06 | .50 | | |
| | Cont. | Hmong | 20 | .11 | .31 | -0.07 | -0.23 |
| | Treat. | | 33 | .04 | .32 | | |
| Word Segmentation Gain Baseline-Endline | Cont. | Khmu | 157 | .20 | .38 | 0.06 | 0.14 |
| | Treat. | | 178 | .26 | .42 | | |
| | Cont. | Hmong | 20 | .11 | .31 | -0.03 | -0.11 |
| | Treat. | | 33 | .08 | .25 | | |
| Baseline Letter Name Knowledge | Cont. | Khmu | 157 | .01 | .06 | 0.00 | 0.06 |
| | Treat. | | 177 | .02 | .06 | | |
| | Cont. | Hmong | 20 | .00 | .00 | 0.03 | 0.29 |
| | Treat. | | 33 | .03 | .11 | | |
| Letter Name Knowledge Gain Baseline-Midline | Cont. | Khmu | 157 | .14 | .26 | 0.01 | 0.04 |
| | Treat. | | 177 | .16 | .27 | | |
| | Cont. | Hmong | 20 | .10 | .23 | -0.04 | -0.21 |
| | Treat. | | 33 | .06 | .18 | | |
| Letter Name Knowledge Gain Midline-Endline | Cont. | Khmu | 159 | .20 | .30 | 0.00 | -0.01 |
| | Treat. | | 179 | .20 | .30 | | |
| | Cont. | Hmong | 20 | .02 | .10 | 0.15 | 0.59 |
| | Treat. | | 33 | .17 | .31 | | |
| Letter Name Knowledge Gain Baseline-Endline | Cont. | Khmu | 157 | .34 | .39 | 0.01 | 0.03 |
| | Treat. | | 177 | .35 | .37 | | |
| | Cont. | Hmong | 20 | .13 | .27 | 0.11 | 0.39 |
| | Treat. | | 33 | .23 | .29 | | |
| Baseline Listening Comprehension in Lao | Cont. | Khmu | 157 | .04 | .13 | 0.02 | 0.16 |
| | Treat. | | 177 | .06 | .15 | | |
| | Cont. | Hmong | 20 | .00 | .00 | 0.03 | 0.31 |
| | Treat. | | 33 | .03 | .12 | | |
| | Cont. | Khmu | 157 | .12 | .26 | 0.09 | 0.33 |

| | | | N | Mean | Std. Dev. | Mean Diff. | *d* |
|---|---|---|---|---|---|---|---|
| Listening Comprehension in Lao Gain Baseline-Midline | Treat. | | 177 | .21 | .32 | | |
| | Cont. | Hmong | 20 | .06 | .18 | -0.07 | -0.44 |
| | Treat. | | 33 | -.01 | .13 | | |
| Listening Comprehension in Lao Gain Midline-Endline | Cont. | Khmu | 159 | .17 | .30 | -0.06 | -0.21 |
| | Treat. | | 179 | .11 | .29 | | |
| | Cont. | Hmong | 20 | .03 | .25 | 0.21 | 0.72 |
| | Treat. | | 33 | .24 | .32 | | |
| Listening Comprehension in Lao Gain Baseline-Endline | Cont. | Khmu | 157 | .29 | .34 | 0.03 | 0.09 |
| | Treat. | | 177 | .32 | .35 | | |
| | Cont. | Hmong | 20 | .09 | .14 | 0.15 | 0.52 |
| | Treat. | | 33 | .24 | .34 | | |
| Baseline Listening Comprehension Ethnic | Cont. | Khmu | 153 | .27 | .32 | -0.02 | -0.08 |
| | Treat. | | 168 | .24 | .31 | | |
| | Cont. | Hmong | 20 | .06 | .16 | 0.19 | 0.70 |
| | Treat. | | 33 | .25 | .32 | | |
| Listening Comprehension Ethnic Gain Baseline-Midline | Cont. | Khmu | 153 | .14 | .37 | 0.09 | 0.24 |
| | Treat. | | 168 | .23 | .38 | | |
| | Cont. | Hmong | 20 | .29 | .39 | -0.29 | -0.73 |
| | Treat. | | 33 | .00 | .40 | | |
| Baseline Name Writing | Cont. | Khmu | 157 | .05 | .14 | 0.00 | 0.01 |
| | Treat. | | 177 | .06 | .14 | | |
| | Cont. | Hmong | 20 | .02 | .06 | 0.04 | 0.30 |
| | Treat. | | 33 | .06 | .16 | | |
| Name Writing Gain Baseline-Midline | Cont. | Khmu | 157 | .24 | .35 | 0.11 | 0.29 |
| | Treat. | | 177 | .35 | .38 | | |
| | Cont. | Hmong | 20 | .30 | .42 | -0.15 | -0.47 |
| | Treat. | | 33 | .15 | .26 | | |
| Name Writing Gain Midline-Endline | Cont. | Khmu | 159 | .22 | .47 | -0.04 | -0.09 |
| | Treat. | | 179 | .18 | .43 | | |

| | | | N | Mean | Std. Dev. | Mean Diff. | *d* |
|---|---|---|---|---|---|---|---|
| | Cont. | Hmong | 20 | .03 | .36 | 0.38 | 0.79 |
| | Treat. | | 33 | .41 | .55 | | |
| Name Writing Gain Baseline-Endline | Cont. | Khmu | 157 | .46 | .44 | 0.07 | 0.15 |
| | Treat. | | 177 | .53 | .46 | | |
| | Cont. | Hmong | 20 | .33 | .46 | 0.23 | 0.51 |
| | Treat. | | 33 | .56 | .44 | | |
| Baseline Receptive Vocabulary | Cont. | Khmu | 157 | .38 | .27 | -0.01 | -0.02 |
| | Treat. | | 177 | .37 | .30 | | |
| | Cont. | Hmong | 20 | .57 | .23 | -0.23 | -0.83 |
| | Treat. | | 33 | .34 | .30 | | |
| Receptive Vocabulary Gain Baseline-Midline | Cont. | Khmu | 157 | .25 | .34 | 0.09 | 0.27 |
| | Treat. | | 177 | .34 | .32 | | |
| | Cont. | Hmong | 20 | -.20 | .39 | 0.29 | 0.73 |
| | Treat. | | 33 | .09 | .40 | | |
| Receptive Vocabulary Gain Midline-Endline | Cont. | Khmu | 159 | .06 | .22 | 0.00 | -0.02 |
| | Treat. | | 179 | .05 | .22 | | |
| | Cont. | Hmong | 20 | .15 | .26 | 0.04 | -0.13 |
| | Treat. | | 33 | .20 | .35 | | |
| Receptive Vocabulary Gain Baseline-Endline | Cont. | Khmu | 157 | .30 | .28 | 0.09 | 0.31 |
| | Treat. | | 177 | .39 | .28 | | |
| | Cont. | Hmong | 20 | -.04 | .27 | 0.33 | 1.00 |
| | Treat. | | 33 | .29 | .36 | | |

**Annex 4: Direct Assessment Data Collection Tool**

**Instructions:**

This document will allow you to assess the development and early learning of young children (ages 2 through to less than 5 years of age). Read all questions to children exactly as they appear. You will see two forms of type:

- **Bold type in boxes indicates things you, the assessor, must say to the child out loud. Please read this type aloud to the child completely and exactly as it appears. This is important to ensure that the data will be collected in a standardized manner across all children.**
- *Italic type indicates instructions for you. Do not read these instructions aloud to the child.*

Throughout the assessment, offer neutral encouragement to the child. Say things like, *'You are working very hard - keep it up!'* Do not indicate to the child that they correctly or incorrectly answered the question, except where indicated in practice trials. Give encouragement in between questions, rather than in the middle of questions. Do not give hints to questions or make facial expressions while the child is completing tasks.

Survey Location

| | |
|---|---|
| **Province name**: | **Province code**: |
| **District name**: | **District code**: |
| **Village name**: | **Village code**: |
| **Name of children** | **Ethnic group:** |
| **Parent Questionnaire ID** | |
| **Date of interview** | |

**Start time**: **Time finish**: **Total**:

| #1 Quantitative comparison of two sets | | | | | | |
|---|---|---|---|---|---|---|
| *Materials: Picture with two groups of tomatoes*<br>*STOP RULES: None* | | | | | | |
| | Instructions | Correct Answer | Check one box only | | | |
| | | | Correct (1) | Incorrect (0) | Stop (99) | |
| 1a | *Show picture of tomatoes*<br>*Ask the child to:* **Put your finger on the picture with the most tomatoes.** | Points to the most tomatoes | 1 | 0 | 99 | |

| #2 Measurement Vocabulary | | | | | | |
|---|---|---|---|---|---|---|
| *Materials: Four sheets of objects in three sizes; 1 set each of dogs, elephant, sticks, and trees*<br>*STOP RULES: None* | | | | | | |
| | Instructions | Correct Answer | Check one box only | | | |
| | | | Correct (1) | Incorrect (0) | | Stop (99) |
| 2a | *Show picture of 3 dogs*<br><br>*Say:* **I'm going to show you some pictures and ask you some questions.**<br><br>*Say:* **Put your finger on the largest dog.** | Points to largest dog | 1 | 0 | | 99 |
| 2b | *Show picture of 3 elephants*<br><br>*Say:* **Put your finger on the smallest elephant.** | Points to smallest elephant | 1 | 0 | | 99 |
| 2c | *Show picture of 3 sticks.*<br><br>*Say:* **Put your finger on the longest stick.** | Points to longest stick | 1 | 0 | | 99 |
| 2d | *Show picture of 3 trees.*<br><br>*Say:* **Put your finger on the shortest tree.** | Points to shortest tree | 1 | 0 | | 99 |

| #3 Shape Naming | | | | | | |
|---|---|---|---|---|---|---|
| *Materials: Use the sheet with a triangle, square, rectangle and circle.*<br>*STOP RULES: None* | | | | | | |
| | Instructions: *Show picture of the shapes*<br><br>*Say:* **Now we will look at some shapes** | Correct Answer | Check one box only | | | |
| | | | Correct (1) | Incorrect (0) | | Stop (99) |
| 3a | *Point to the circle*<br><br>*Say:* **What is the name of this shape** | Child say's circle | 1 | 0 | | 99 |
| 3b | *Point to the rectangle*<br><br>*Say:* ***What is the name of this shape*** | Child says rectangle | 1 | 0 | | 99 |
| 3c | *Point to the triangle*<br><br>*Say:* ***What is the name of this shape*** | Child says triangle | 1 | 0 | | 99 |
| 3d | *Point to the square*<br><br>*Say:* ***What is the name of this shape*** | Child says square | 1 | 0 | | 99 |

| #4 Spatial Vocabulary | | | | | | |
|---|---|---|---|---|---|---|
| *Materials: Spatial Vocabulary Sheet*<br>*STOP RULES: None* | | | | | | |
| | Instructions | Correct Answer | Correct (1) | Incorrect (0) | | Stop (99) |
| 4a | *Place Spatial Vocabulary Sheet in front of child. Say:* **Now I am going to ask you some questions about these pictures.**<br><br>*Say:* **Point to the picture with the ball on top of the chair.** | Picture with balloon the chair | 1 | 0 | | 99 |
| 4b | *Say:* **Point to the picture with the ball underneath the chair.** | Picture with ball under the chair | 1 | 0 | | 99 |
| 4c | *Say:* **Point to the picture with the ball in front of the chair.** | Picture with ball in front of the chair | 1 | 0 | | 99 |
| 4d | *Say:* **Point to the picture with the ball beside the chair.** | Picture with ball next to the chair | 1 | 0 | | 99 |

| #5 Verbal Counting | | | | | | |
|---|---|---|---|---|---|---|
| *Materials: None*<br>*Other notes: Prompt as necessary* **What comes after <<last number stated>>**<br>*Self-correcting allowed*<br>*STOP RULES: When a child states a number incorrectly or reaches 30.* | | | | | | |
| | Instructions | Correct Answer | Indicate highest number counted. Check STOP RULE USED if you stopped the child. | | | |
| | | | Highest Number | | | Stop (99) |
| 5 | *Say:* **Now we are going to play some counting games. The first game is a counting out loud game. How high can you count? Start at one and tell me.** | Counts accurately | | | | 99 |

| #6 Number Identification | | | | | | |
|---|---|---|---|---|---|---|
| *Materials: Number Sheet and a blank sheet of paper to cover a column*<br>*Notes: Self-correcting is allowed.*<br>*If the child gets stuck for more than 5 seconds, point to the next number and say:* **Let's try this one.**<br>*STOP RULES: Five numbers consecutively incorrect* | | | | | | |
| 6 | Instructions<br>*Place the Numbers Sheet in front of the child. Using another sheet of paper, cover the right column.*<br><br>*Say:* **Here are some numbers. I will point to a number and I want you to tell me the number. It's OK if you don't know all of them.**<br><br>*Point to the first number in the left column and ask the child:* **What number is this?**<br><br>*Continue pointing at each number down the column.* | | Correct (1) | Incorrect (0) | | Stop (99) |
| 6a | | 4 | 1 | 0 | | 99 |
| 6b | | 2 | 1 | 0 | | 99 |
| 6c | | 1 | 1 | 0 | | 99 |
| 6d | | 5 | 1 | 0 | | 99 |
| 6e | | 3 | 1 | 0 | | 99 |
| 6f | | 8 | 1 | 0 | | 99 |
| 6g | | 10 | 1 | 0 | | 99 |
| 6h | | 6 | 1 | 0 | | 99 |
| 6i | | 9 | 1 | 0 | | 99 |
| 6j | | 7 | 1 | 0 | | 99 |
| | *When the child finishes the last number in the left column, cover the left column and point to the first number in the right column.*<br>*Say:* **What number is this?** *Continue pointing at each number down the column.* | | | | | |
| 6k | | 13 | 1 | 0 | | 99 |
| 6l | | 11 | 1 | 0 | | 99 |
| 6m | | 14 | 1 | 0 | | 99 |
| 6n | | 12 | 1 | 0 | | 99 |
| 6o | | 18 | 1 | 0 | | 99 |
| 6p | | 17 | 1 | 0 | | 99 |
| 6q | | 15 | 1 | 0 | | 99 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 6r | | 19 | 1 | 0 | | 99 |
| 6s | | 20 | 1 | 0 | | 99 |
| 6t | | 16 | 1 | 0 | | 99 |
| | *Check if stop rule was used.* | | | | | |

| #7 Producing A Set | | | | | | |
|---|---|---|---|---|---|---|
| *Materials: 20 macaroni that can be used for counting*<br>*STOP RULE: If child cannot give you 3 items <u>and</u> cannot give you 6 items, move on to next assessment task (#4). If child misses only one of the first two items, proceed with item 7c.* | | | | | | |
| | Instructions | Correct Answer | Correct (1) | Incorrect (0) | | Stop (99) |
| 7a | *Arrange 20 macaroni randomly in front of the child.*<br>*Say:* **Now we'll play a game with stones. Please give me three macaroni.** | Hands or pushes over 3macaroni | 1 | 0 | | 99 |
| 7b | *Rearrange the 20 macaroni randomly again in front of the child.*<br>*Say:* **Now, please give me six macaroni.** | Hands or pushes over 6 macaroni | 1 | 0 | | 99 |
| 7c | *Rearrange the 20 macaroni randomly again in front of the child.*<br>*Say:* **Now, please give me fourteen macaroni.** | Hands or pushes over 14macaroni | 1 | 0 | | 99 |
| | *Check if stop rule was used.* | | | | | |

| #8 Addition with two sets | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| *Materials: 10 macaroni pieces that can be used for counting*<br>*Notes: If child counts ("one, two, three, four, five") ask again,* **How many will you have altogether?** If child shows quantity on fingers, ask **Can you say it?**<br>*STOP RULES: None* | | | | | | | | |
| | Instructions | Correct Answer | Correct (3) says " five" | Correct (2) but can' t say the total without | Correct (1) but shows five | Incorrect (0) | | Stop (99) |
| 8a | *Place ten macaroni pieces to the side of the child on the table/floor.*<br>*Say:* ***I am going to ask you a question about numbers. Here are some macaroni to help you. You can use them if you want to, but you don't have to. Listen very carefully to the question.***<br>*Say:* **If you have three macaroni…**<br>*Say:* **And I give you two more macaroni, how many macaroni will you have altogether?** | Says "Five"<br>Check one box | 3 | 2 | 1 | 0 | | 99 |
| 8b | Strategy (indicate how child solved the problem based on your observation) | Check one box | | Uses stones to solve problem (1) | Uses fingers to solve problem (2) | No apparent counting (just says five) (3) | | Unknown strategy (4) |
| | | | | 1 | 2 | 3 | | 4 |

#9 Spatial visualization

*Materials:*

- *A picture of the dog (what the puzzle would look like when put together)*
- *Laminated paper jigsaw puzzle with five pieces (i.e. cut along the black lines)*

*STOP RULES: If the child can't do any pieces within 2 minutes move on. However, if they have started to get it right let them go until they can't get another piece right in 2 minutes.*

| | Instructions | Correct Answer | Check one box only | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | No matching pieces (0) | 2 matching pieces (2) | 3 matching pieces (3) | 4 matching pieces (4) | 5 matching pieces (5) | | Stop (99) |
| 9 | *Show the picture of the puzzle to child and say:* **We are going to have some fun with this puzzle. This is a picture of a dog and we are going to try to make this picture with these pieces.**<br><br>*Show the child the puzzle pieces in a random order (but the right way up) and say:* **Can you try to join the pieces together to make this picture?** | The jigsaw puzzle pieces must be matching in right ways (in order to receive 4 or 5 scores all or 5 pieces must match each other. For instance, 2 pictures with 2 pieces not two separate ones would score 2 | 0 | 2 | 3 | 4 | 5 | | 99 |

| #10 Concepts about Print | | | | |
|---|---|---|---|---|
| *<u>Materials</u>**:** Age-appropriate book*<br>*Hold the book vertically by the outside edge with the opening towards you and the spine towards the child. Hand the book to the child in this position.* | | | | |
| 10 | **We are going to look at a book and I need your help. If you were going to read this book, can you show me how you would open it so you can read it?**<br>*Turn to the first page of the story. Say:*<br>**Can you show me where I should start reading?**<br>*Open on the next page and point to the first word on that page. Say:*<br>**If I start to read here, on the first word, where do I continue reading? Show me with your finger.** | Child opens book right side up and page open | 0=incorrect<br>1=correct<br>99 = Stop | |
| | | Childs points to some part of text on the first page | 0=incorrect<br>1=correct<br>99 = Stop | |
| | | Child points to text following identified text | 0=incorrect<br>1=correct<br>99 = Stop | |

#11 Initial Sound Discrimination

*Materials: None*

*Other Notes: Additional Prompts: Repeat the list of words ONCE per question if needed*

*STOP RULES: None*

*PRACTICE TRIAL:*

**Let's play another game. This is a listening game***.*

**I like Noy.**

**My name is "Noy". The word "Noy" pronounced with alphabet N.**

**For the next words I will pronounce, Can you please tell me what words pronounced with alphabet N?**

"< Non.", "< Saly >.", "<Pook>."

If child says, "Non," say, "That's right. Noy starts with the sound No, so she will like the word non.

If child gives an incorrect response or no response, say. "Noy would like the word "non" because "non" starts with the sound "no," just like her name.

| | Instructions | Correct Answer | Correct (1) | Incorrect (0) | | Stop (99) |
|---|---|---|---|---|---|---|
| 11a | *Say:* **Are you ready to play some more? Noy likes words that pronounced with N . Which of the following words pronounced with alphabet N? <<List of three words, one pronounced with N>>.** "< Nam>.", "< Khao >.", "< Pa >." | Matches the pronunciation with the word #1___Nam_____ | 1 | 0 | | 99 |
| 11b | *Say:* **Now I have another friend, Lee. Lee likes words that pronounced with alphabet letter L. Which of the following words pronounced with L? <<List of three words, one pronounced with L>>.** | Matches the pronunciation with the word #2____Len_____ | 1 | 0 | | 99 |

| | | | | | | |
|---|---|---|---|---|---|---|
| | ***"< Xang >.","< Len >.","< Meo >."*** | | | | | |
| 11c | *Say:* **Now I have another friend, Pet. Pet likes words that pronounced with alphabet letter P. Which of the following words pronounced with P? <<List of three words, one pronounced with same P>>.**<br>***"< Heun >", "< Koiy >.", "< Por >."*** | Matches the pronunciation with the word<br>#3___Por_____ | 1 | 0 | | 99 |

#12 Initial Sound Identification

*Materials: None*

*Other Notes: Additional Prompts: Repeat the list of words ONCE per question if needed*

*STOP RULES: If the child does not respond after 5 seconds, mark as "No response" and say the next prompt.*

PRACTICE TRIAL:

**This is another listening game:** I want you to pay attention to the two words I will say next and I will repeat it once. Then I'd like you to tell me how to pronounce the last word I mentioned. For example:
# the first two words is "Somsak"…. "Som-Sak"…. The word you should pronounce would be the last word, that is "Sak".
# now let's try this again, the next two words is "Kinkhao" … "Kin-Khao"…. The last word to pronounce is "Khao".

**Interviewer**: tell the kids to say "khao"

Then ask the kid if they understand what they need to do. If you see that the kid can do it well, start the actual quiz/game. If you see that the kids are still confused or don't know what to do yet, please explain to them the above 2 examples once again before start the game.

| | Instructions | Correct Answer | Correct (1) | Incorrect (0) | | Stop (99) |
|---|---|---|---|---|---|---|
| 12a | ***Interviewer pronounce****: please pronounce the last word of "Ka Tai"* | Pronounced #1_____Tai____ | 1 | 0 | | 99 |
| 12b | ***Interviewer pronounce*** *please pronounce the last word of "Hong Mor"* | Pronounced #2_____Mor___ | 1 | 0 | | 99 |
| 12c | ***Interviewer pronounce:*** *please pronounce the last word of "Mark Ban"* | Pronounced #3___Ban____ | 1 | 0 | | 99 |
| 12d | ***Interviewer pronounce:*** *please pronounce the last word of "Khao Khoai"* | Pronounced #4____Khoai____ | 1 | 0 | | 99 |
| 12e | ***Interviewer pronounce:*** *please pronounce the last word of "Lai Seur"* | Pronounced #5____Seur_____ | 1 | 0 | | 99 |

| #13 Letter Name Knowledge | | | | | | |
|---|---|---|---|---|---|---|
| *Materials: Letters Sheet and a blank sheet of paper to cover a column*<br><br>*Other Notes: Self-correcting is allowed.*<br>*If the child gets stuck for more than 5 seconds, point to the next letter and say:* ***Let's try this one.***<br><br>*STOP RULES: Five numbers consecutively incorrect* | | | | | | |
| | Instructions | | Correct (1) | Incorrect (0) | | Stop (99) |
| 13 | *Say:* **We will play an alphabet letter game now.**<br><br>*Place the Letters Sheet in front of the child. Using another sheet of paper, cover the right column.*<br><br>*Say:* **Here are some letters. I will point to the letters and I want you to tell me the name of the letter.**<br><br>*Point to the first letter in the left column and ask the child:* **What letter is this?**<br><br>*Continue pointing at each letter down the column.* | | | | | |
| 13a | | B | 1 | 0 | | 99 |
| 13b | | S | 1 | 0 | | 99 |
| 13c | | A | 1 | 0 | | 99 |
| 13d | | T | 1 | 0 | | 99 |
| 13e | | M | 1 | 0 | | 99 |
| 13f | | U | 1 | 0 | | 99 |
| 13g | | D | 1 | 0 | | 99 |
| 13h | | V | 1 | 0 | | 99 |
| 13i | | X | 1 | 0 | | 99 |
| 13j | | Q | 1 | 0 | | 99 |
| | *When the child finishes the last letter in the left column, cover the left column and point to the first letter in the right column.*<br><br>*Say:* **What letter is this?** *Continue pointing at each letter down the column.* | | | | | |
| 13k | | E | 1 | 0 | | 99 |
| 13l | | R | 1 | 0 | | 99 |
| 13m | | N | 1 | 0 | | 99 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 13n | | L | 1 | 0 | | 99 |
| 13o | | O | 1 | 0 | | 99 |
| 13p | | K | 1 | 0 | | 99 |
| 13q | | P | 1 | 0 | | 99 |
| 13r | | F | 1 | 0 | | 99 |
| 13s | | C | 1 | 0 | | 99 |
| 13t | | G | 1 | 0 | | 99 |
| | *Check if stop rule was used.* | | | | | |

#14 Listening Comprehension Story

*Materials: None*

*Other Notes: Additional Prompts: Each question may be repeated ONCE if needed.*

*STOP RULES: None*

STORY:

*Say:* **Now I am going to tell you an interesting story. After I have told you the story, I will ask you some questions. Listen carefully, okay?**

*Say:* **This story is called The Mouse and the Cat**
**Once upon a time there was a fat cat. He always wore a red hat. Once when he was sleeping, a small mouse came silently and stole the hat. The cat woke up to see his hat gone, got very angry and started chasing the mouse. After a while, the mouse was trapped under a table and could not find any way to escape. So the mouse cried to the cat, "Please don't eat me cat. If you spare my life, I will return your hat." So, after getting back his hat the cat said, "Never touch my hat again" and he went back to sleep in a happy mood.**

**Now I am going to ask you some questions about the story.**

*Ask each question slowly and clearly.*

| | Instructions | Correct Answer | Correct (1) | Incorrect (0) | | Stop (99) |
|---|---|---|---|---|---|---|
| 14a | *Say:* **Who stole the cat's hat?** | The mouse | 1 | 0 | | 99 |
| 14b | *Say:* **What was the color of the hat?** | Red | 1 | 0 | | 99 |
| 14c | *Say:* **Why was the cat chasing the mouse?** | Because the mouse took/stole its hat. | 1 | 0 | | 99 |
| 14d | *Say:* **Where did the cat trap the mouse?** | Under the table | 1 | 0 | | 99 |
| 14e | *Say:* **Why did the cat decide not to eat the mouse?** | Because the mouse gave back the hat | 1 | 0 | | 99 |

| #15 Name Writing | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| *Materials: One blank piece of paper (use following page), pencil or pen.*<br>*STOP RULES: Limit this section to 2 minutes from when the child begins writing. If the child does not write for a minute after your instructions, stop and say,*<br>***We're going to move on to our next game now.*** | | | | | | | | | |
| | Instructions | Nothing (0) | Scribbles, no discernable symbols (1) | Symbol-like marks (2) | Non-name letters (3) | Has letters in name, but name is not correct – letters are out of order; other mistakes are | Letters in name in correct order and orientation (5) | | Stop (99) |
| 15 | *Place the blank piece of paper and the pencil or pen in front of the child.*<br><br>*Say: Now we are going to* **write. Can you write your name here in any way you know? Don't worry if you can't do it well, just try your best.** | 0 | 1 | 2 | 3 | 4 | 5 | | 99 |
| | *Check if stop rule was used at one minute* | | | | | | | | |
| | *Check if stop rule was used at two minutes.* | | | | | | | | |

#16 Head, Toes, Knees, Shoulders Task

*Materials: None*

*STOP RULES:*

*If the child does not respond correctly after 3 practice sessions of the HEAD/TOES section, stop this task and move on to the next task.*

*If child does not get any of HEAD/TOES assessment correct, stop this task and move onto next task.*

TEACHING TRIAL (HEAD/TOES):

**Now we're going to play a game. Listen exactly to what I say: Touch your head.**

*Assessor physically touches his/her head with two hands. Wait until the child puts two hands on his/her own head.*

**Good! Now touch your toes.**

*Assessor physically touches his/her toes with two hands. Wait until the child puts two hands on his/her own toes.*

*Repeat the two commands up to two times with motions until the child imitates you correctly.*

PRACTICE TRIAL (HEAD/TOES):

**Now we're going to be a little silly and you do the OPPOSITE of what I say. When I say touch your HEAD, you touch your TOES. When I say touch your TOES, you touch your HEAD. So you do something OPPOSITE from what I say.**

**"Touch your head"? (***assessor DOES NOT touch head or toes***)**

**"Touch your toes"? (***assessor DOES NOT touch toes or head)*

*Say:* **That's right** *when the child responds correctly. If the child responds incorrectly, provide additional explanations up to 3 times before beginning the test portion:*
***Remember, when I say touch your head, you touch your toes.***
***Remember, when I say touch your toes, your touch your head.***

| | Instructions | Correct Answer | Correct (2) | Self-corrects (1) | Incorrect (0) | | Stop (99) |
|---|---|---|---|---|---|---|---|
| 16a | *Say:* **Now we will keep playing this game, listen** | Touches toes | 2 | 1 | 0 | | 99 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | **carefully and do the OPPOSITE of what I say.**<br><br>*DO NOT touch your head/toes and DO NOT provide feedback or extra explanations*<br><br>*Say:* **Touch your head** | | | | | | |
| 16b | *Say:* **Touch your toes** | Touches head | 2 | 1 | 0 | | 99 |
| 16c | *Say:* **Touch your toes** | Touches head | 2 | 1 | 0 | | 99 |
| 16d | *Say:* **Touch your head** | Touches toes | 2 | 1 | 0 | | 99 |
| 16e | *Say:* **Touch your toes** | Touches head | 2 | 1 | 0 | | 99 |
| | *Check if stop rule was used before assessment* | | | | | | |
| | *Check if stop rule was used after assessment* | | | | | | |

#17 Forward Digit Span

*Materials: None*

*Other Notes: If the child makes an error, supply the correct answer on the practice items only.*

*STOP RULES: None*

TEACHING AND PRACTICE TRIAL:

**In this game, I am going to say a list of numbers. After you hear the numbers, I want you to repeat them after me in the same order.**

**If I say 7...8, You say 7…8**

**Now you try a couple. Please listen carefully.**

*Pause for one second in between each number in the sequence. For example, « 4 » [pause] « 2 ».*

*Say:* ***4…2*** *Wait for child to respond. If the child makes an error, supply the correct answer. If the child answers correctly say,* **That's right.**

*Say:* ***6…1…3*** *Wait for child to respond. If the child makes an error, supply the correct answer. If the child answers correctly say,* **That's right.**

| | Instructions | Correct Answer | Correct (1) | Incorrect (0) | | Stop (99) |
|---|---|---|---|---|---|---|
| 17a | ***Okay, now let's do some more. Just listen carefully and do your best.***<br><br>*Pause for one second in between each number in the sequence*<br>*Say:* ***1…6*** | 1…6 | 1 | 0 | | 99 |
| | *Write in response:* | | | | | |
| 17b | *Say:* **5…2…9** | 5…2…9 | 1 | 0 | | 99 |
| | *Write in response:* | | | | | |
| 17c | *Say:* **8…3…1…4** | 8…3…1…4 | 1 | 0 | | 99 |
| | *Write in response:* | | | | | |
| 17d | *Say:* **1…2…4…7…3** | 1…2…4…7…3 | 1 | 0 | | 99 |
| | *Write in response:* | | | | | |

#18 Backward Digit Span

*Materials: None*

*Other Notes: If the child makes an error, supply the correct answer on the practice items only.*

*STOP RULES:* End this assessment if the child errs on both trials of the same digit set (but if child gets one of a set correct, continue assessment).

TEACHING AND PRACTICE TRIAL

*Say:* **Now we are going to play another number game. I'm going to give you a list of numbers and you are going to say them backwards. If I say 1, 2, You say 2, 1.**

*Pause for one second in between each number in the sequence. For example, « 1 » [pause] « 2 ».*

*Say:* **Now you try it/ The numbers are 1, 2.**

*If the child responds correctly, say:* **That's right!**

*If the child does not respond correctly, say:* **The numbers are: 1, 2. When I say them backwards, they are 2, 1.**

**Let's try another one: 4, 8, 3.**

*If the child responds correctly, say:* **That's right!**

*If the child does not respond correctly, say:* **The numbers are: 4, 8, 3. When I say them backwards, they are 3, 8, 4.**

| | Instructions | Correct Answer | Correct (1) | Incorrect (0) | | Stop (99) |
|---|---|---|---|---|---|---|
| 18a | *Say:* ***That was fun! Let's do more: Whatever I say, you should say it backwards.*** *Pause for one second in between each number in the sequence.*<br><br>*Say*: ***4…1*** | 1…4 | 1 | 0 | | 99 |
| | *Write in response:* | | | | | |
| 18b | *Say:* **6…2** | 2…6 | 1 | 0 | | 99 |
| | *Write in response:* | | | | | |
| 18c | *Say:* **3…5…6** | 6…5…3 | 1 | 0 | | 99 |

| | | | | | | |
|---|---|---|---|---|---|---|
| | *Write in response:* | | | | | |
| 18d | *Say:* **2…9…8** | 8…9…2 | 1 | 0 | | 99 |
| | *Write in response:* | | | | | |
| 18e | *Say:* ***4…9…2…7*** | 7…2…9…4 | 1 | 0 | | 99 |
| | *Write in response:* | | | | | |
| 18f | *Say:* ***1…6…4…5*** | 5…4…6…1 | 1 | 0 | | 99 |
| | *Write in response:* | | | | | |
| 18g | *Say:* ***8…3…6…1…5*** | 5…1…6…3…8 | 1 | 0 | | 99 |
| | *Write in response:* | | | | | |
| | *Check if stop rule was used* | | | | | |

| | #19 - Vocabulary Assessment | | | | |
|---|---|---|---|---|---|
| *Materials: Picture*<br><br>*STOP RULES: None* | | | | | |
| **Instructions**<br>*[Say in child's first language] Now I'm going to show you some pictures. I am going to say a word in Lao, and then I want you to point to the picture that shows the word I am saying. It's okay if you don't know all of the answers, just give it your best try.* | | Correct Answer | Correct = 1 | Incorrect = 0 | **Stop (99)** |
| 19a | Show picture "PV1" and say with children *"Show me [target word]."* | Eye | 1 | 0 | **99** |
| 19b | Show picture "PV2" and say with children *"Show me [target word]."* | Chalk | 1 | 0 | **99** |
| 19c | Show picture "PV3" and say with children *"Show me [target word]."* | Fish | 1 | 0 | **99** |
| 19d | Show picture "PV4" and say with children *"Show me [target word]."* | Balloon | 1 | 0 | **99** |
| 19e | Show picture "PV5" and say with children *"Show me [target word]."* | Bird | 1 | 0 | **99** |
| 19f | Show picture "PV6" and say with children *"Show me [target word]."* | Trees | 1 | 0 | **99** |
| 19g | Show picture "PV7" and say with children *"Show me [target word]."* | Truck | 1 | 0 | **99** |
| 19h | Show picture "PV8" and say with children *"Show me [target word]."* | Butterfly | 1 | 0 | **99** |
| 19i | Show picture "PV9" and say with children *"Show me [target word]."* | Chicken | 1 | 0 | **99** |
| 19j | Show picture "PV10" and say with children *"Show me [target word]."* | Basket | 1 | 0 | **99** |
| 19k | Show picture "PV11" and say with children *"Show me [target word]."* | Blackboard | 1 | 0 | **99** |
| 19l | Show picture "PV12" and say with children *"Show me [target word]."* | Chair | 1 | 0 | **99** |
| 19m | Show picture "PV13" and say with children *"Show me [target word]."* | Pickup Truck | 1 | 0 | **99** |
| 19n | Show picture "PV14" and say with children *"Show me [target word]."* | Water Bug | 1 | 0 | **99** |
| 19o | Show picture "PV15" and say with children *"Show me [target word]."* | Rice Paddy | 1 | 0 | **99** |
| 19p | Show picture "PV16" and say with children *"Show me [target word]."* | Goat | 1 | 0 | **99** |
| 19q | Show picture "PV17" and say with children *"Show me [target word]."* | Boat race | 1 | 0 | **99** |

| | | | | | |
|---|---|---|---|---|---|
| 19r | Show picture “PV18” and say with children *“Show me [target word].”* | Crocodile | 1 | 0 | **99** |
| 19s | Show picture “PV19” and say with children *“Show me [target word].”* | Bridge | 1 | 0 | **99** |
| 19t | Show picture “PV20” and say with children *“Show me [target word].”* | Monkey | 1 | 0 | **99** |
| 19u | Show picture “PV21” and say with children *“Show me [target word].”* | Fish Net | 1 | 0 | **99** |
| 19v | Show picture “PV22” and say with children *“Show me [target word].”* | Papaya | 1 | 0 | **99** |
| 19w | Show picture “PV23” and say with children *“Show me [target word].”* | Bamboo Shoot | 1 | 0 | **99** |
| 19x | Show picture “PV24” and say with children *“Show me [target word].”* | Comb | 1 | 0 | **99** |
| 19y | Show picture “PV25” and say with children *“Show me [target word].”* | Sing a Song | 1 | 0 | **99** |
| 19z | Show picture “PV26” and say with children *“Show me [target word].”* | Hat | 1 | 0 | **99** |
| 19aa | Show picture “PV27” and say with children *“Show me [target word].”* | Fence | 1 | 0 | **99** |
| ***Those are all of the pictures I have. You did a great job! Thank You*** | | | | | |

#20 - Semantic Fluency

**Instructions**

This is a semantic fluency test, which can be used to measure a facet of oral language proficiency for children in the pre-primary and early grades in Laos. The test can be conducted in any language required by the assessor. It should be conducted in one language at a time. Please provide some practice sessions for the child to understand the rules of the test and so that the data is reflective of a "free association" task, as opposed to a vocabulary task in which they simply name the items in the picture.

1. Give this test to one student at a time. Start with the scene covered up or face down while you explain the directions.

2. Say the following: *I am now going to show you a scene with lots of different images in it. When I say "begin," I want you to say out loud as many words as you can think of from the scene or any words that come to mind when looking at the scene. The words do not actually have to be in the scene. Try not to say full sentences or stories, but just try to list words as they come to your mind. You will have 60 seconds to say as many words as you can and you should say these words in Lao. Do you understand what you are supposed to do?*

3. Turn the sheet over so the child can see the image and say "*begin*" while at the same time starting your 60-second timer.

4. Once the test has begun, use a scrap of paper to make a slash for every word the child says in the correct language that corresponds to the scene. When the 60 seconds are up, tell the child to stop.

5. Administer this test to all students in Lao.

| **Introduction:** Enumerator count the words that child said and record by separate words inside picture and words outside picture | | **Words** |
|---|---|---|
| A | Number of words inside picture | |
| B | Number of words outside picture | |
| Total (Automatically counted by program) | | |

#21 - Most-Used Words

ອຸປະກອນ: ລາຍການຄຳສັບ
ເອົາບັນຊີລາຍການ “ຄຳສັບຕ່າງໆ ທີ່ຄິດຄົ້ນ” ທີ່ຫຸ້ມດ້ວຍຊອງໃສ ໃຫ້ກັບເດັກເບິ່ງ.
*ຂ້ອຍຢາກໃຫ້ເຈົ້າອ່ານລາຍການຄຳສັບໃຫ້ຂ້ອຍຟັງຕື່ມອີກ. ຄຳສັບເຫົ່ານີ້ ບໍ່ມີໃນປຶ້ມແບບຮຽນ, ມັນບໍ່ແມ່ນຄຳສັບທີ່ມີຕົວຈິງ ເພາະເປັນຄຳສັບທີ່ພວກເຮົາແຕ່ຂຶ້ນເອງ. ແຕ່ວ່າ ກໍ່ສາມາດອ່ານໄດ້. ເຈົ້າຈົ່ງຊີ້ໃສ່ ແລະ ອ່ານແຕ່ລະຄຳ ໂດຍເລີ່ມຕົ້ນຈາກນີ້* ***(ນັກສຳພາດຊີ້ໃສ່ຄຳທຳອິດ)*** *ແລະ ຍ້າຍໄປເທື່ອລະຄຳສັບແບບນີ້* ***(ນັກສຳພາດຊີ້ຄຳຕໍ່ໄປ ໂດຍເລີ່ມແຕ່ຊ້າຍຫາຂວາ).*** *ນ້ອງເຂົ້າໃຈບໍ? ເອົາລະ ເລີ່ມຕົ້ນໄດ້ແລ້ວ.*
ນັກສຳພາດ: ໝາຍທຸກຄຳສັບທີ່ເດັກອ່ານຜິດ. **<u>ຈົ່ງຈື່ວ່າ</u>**
ການອອກສຽງຄຳຕ່າງໆຕາມສຳນຽງທ້ອງຖິ່ນ ແມ່ນຍອມຮັບໄດ້,
ຖ້າວ່າເດັກອ່ານບໍ່ໄປຕາມລຳດັບ ເວົ້າກັບເດັກວ່າ “ນ້ອງອ່ານຄຳນີ້ໄດ້ບໍ່
(ນັກສຳພາດຊີ້ໃສ່ຄຳສັບທີ່ເດັກຂ້າມ) ແລ້ວບອກເດັກອ່ານຄຳຕໍ່ໄປຕາມລະດັບ
ແຕ່ຊ້າຍຫາຂວາ”. ນັກສຳພາດຕ້ອງແນ່ໃຈວ່າ,
ທຸກຄຳສັບທີ່ເດັກອ່ານຜິດແມ່ນຖືກໝາຍໄວ້ແລ້ວ ກ່ອນທີ່ຈະໄປເຮັດກິດຈະກຳກັບຂໍ້ຕໍ່ໄປ.

**<u>ສຳລັບເດັກທີ່ອ່ອນ:</u>**
ຖ້າວ່າເດັກລັງເລໃຈໃນຄຳສັບໃດຄຳສັບໜຶ່ງ ເປັນເວລາ 5 ວິນາທີ, ຈົ່ງຖາມເດັກວ່າ,
ມີຄຳໃດໃນບັນຊີລາຍການນີ້ທີ່ນ້ອງຮູ້ແດ່? ບອກຂ້ອຍແດ່ ຫຼື ອ່ານຄຳທີ່ນ້ອງຮູ້ເບິ່ງດຸ.
***ນັກສຳພາດ***: ຈົ່ງເວົ້າຊ້ຳໆອີກ ເພື່ອກະຕຸກຊຸກຍູ້ໃຫ້ເດັກສຶບຕໍ່ອ່ານ.
ຖ້າເດັກຍັງລັງເລໃຈຢູ່ເປັນເວລາ 10 ວິນາທີ ໃຫ້ເວົ້າກັບເດັກວ່າ “ຂອບໃຈຫຼາຍໆ
ເກັ່ງຫຼາຍທີ່ນ້ອງພະຍາຍາມເຮັດ ແລະ ເຮັດໄດ້ດີຫຼາຍແລ້ວ”. ຫຼັງຈາກນັ້ນ
ໃຫ້ຖືວ່າທຸກຄຳສັບທີ່ເດັກບໍ່ໄດ້ອ່ານ ຫຼື ບໍ່ພະຍາຍາມທີ່ຈະອ່ານ ແມ່ນຜິດ
ແລ້ວໝາຍຈຳນວນຄຳສັບທີ່ຜິດ ກ່ອນຈະຍ້າຍໄປຂໍ້ຕໍ່ໄປ.

| **<u>ຄຳແນະນຳ</u>**<br>*ປ່ອຍໃຫ້ເດັກອ່ານ ແລະ ຊີ້ໃສ່ຄຳສັບດ້ວຍຕົວເອງພາຍຫຼັງທີ່ແນະນຳກິດຈະກຳແລ້ວ* | | ຄຳຕອບທີ່ຖືກຕ້ອງ | ໃຫ້ຄະແນນ | ໃຫ້ຄະແນນ | ຂໍ້ມູນຫາຍ |
|---|---|---|---|---|---|
| 21a | *ນັກສຳພາດຊີ້ໃສ່ຄຳສັບ ແລ້ວບອກໃຫ້ເດັກອ່ານ* | ອາ | 1 | 0 | 99 |
| 21b | *ປ່ອຍໃຫ້ເດັກຊີ້ໃສ່ຄຳສັບ (ແນ່ໃຈວ່າແມ່ນແຕ່ຊ້າຍຫາຂວາ) ແລະ ອ່ານເອງ* | ດີ | 1 | 0 | 99 |
| 21c | *ປ່ອຍໃຫ້ເດັກຊີ້ໃສ່ຄຳສັບ (ແນ່ໃຈວ່າແມ່ນແຕ່ຊ້າຍຫາຂວາ) ແລະ ອ່ານເອງ* | ເອ | 1 | 0 | 99 |
| 21d | *ປ່ອຍໃຫ້ເດັກຊີ້ໃສ່ຄຳສັບ (ແນ່ໃຈວ່າແມ່ນແຕ່ຊ້າຍຫາຂວາ) ແລະ ອ່ານເອງ* | ໄປ | 1 | 0 | 99 |
| 21e | *ປ່ອຍໃຫ້ເດັກຊີ້ໃສ່ຄຳສັບ (ແນ່ໃຈວ່າແມ່ນແຕ່ຊ້າຍຫາຂວາ) ແລະ ອ່ານເອງ* | ມາ | 1 | 0 | 99 |
| 21f | *ປ່ອຍໃຫ້ເດັກຊີ້ໃສ່ຄຳສັບ (ແນ່ໃຈວ່າແມ່ນແຕ່ຊ້າຍຫາຂວາ) ແລະ ອ່ານເອງ* | ຂ້ອຍ | 1 | 0 | 99 |

| | | | | | |
|---|---|---|---|---|---|
| 21g | *ປ່ອຍໃຫ້ເດັກຊີ້ໃສ່ຄຳສັບ (ແນ່ໃຈວ່າແມ່ນແຕ່ຊ້າຍຫາຂວາ) ແລະ ອ່ານເອງ* | ດັງ | 1 | 0 | 99 |
| 21h | *ປ່ອຍໃຫ້ເດັກຊີ້ໃສ່ຄຳສັບ (ແນ່ໃຈວ່າແມ່ນແຕ່ຊ້າຍຫາຂວາ) ແລະ ອ່ານເອງ* | ງາມ | 1 | 0 | 99 |
| 21i | *ປ່ອຍໃຫ້ເດັກຊີ້ໃສ່ຄຳສັບ (ແນ່ໃຈວ່າແມ່ນແຕ່ຊ້າຍຫາຂວາ) ແລະ ອ່ານເອງ* | ໄກ່ | 1 | 0 | 99 |
| 21j | *ປ່ອຍໃຫ້ເດັກຊີ້ໃສ່ຄຳສັບ (ແນ່ໃຈວ່າແມ່ນແຕ່ຊ້າຍຫາຂວາ) ແລະ ອ່ານເອງ* | ທຸງ | 1 | 0 | 99 |
| 21k | *ປ່ອຍໃຫ້ເດັກຊີ້ໃສ່ຄຳສັບ (ແນ່ໃຈວ່າແມ່ນແຕ່ຊ້າຍຫາຂວາ) ແລະ ອ່ານເອງ* | ຝົນຕົກ | 1 | 0 | 99 |
| 21l | *ປ່ອຍໃຫ້ເດັກຊີ້ໃສ່ຄຳສັບ (ແນ່ໃຈວ່າແມ່ນແຕ່ຊ້າຍຫາຂວາ) ແລະ ອ່ານເອງ* | ອ້າຍ | 1 | 0 | 99 |
| 21m | *ປ່ອຍໃຫ້ເດັກຊີ້ໃສ່ຄຳສັບ (ແນ່ໃຈວ່າແມ່ນແຕ່ຊ້າຍຫາຂວາ) ແລະ ອ່ານເອງ* | ໝູ່ເພື່ອນ | 1 | 0 | 99 |
| 21n | *ປ່ອຍໃຫ້ເດັກຊີ້ໃສ່ຄຳສັບ (ແນ່ໃຈວ່າແມ່ນແຕ່ຊ້າຍຫາຂວາ) ແລະ ອ່ານເອງ* | ເສື້ອ | 1 | 0 | 99 |
| 21o | *ປ່ອຍໃຫ້ເດັກຊີ້ໃສ່ຄຳສັບ (ແນ່ໃຈວ່າແມ່ນແຕ່ຊ້າຍຫາຂວາ) ແລະ ອ່ານເອງ* | ເອື້ອຍ | 1 | 0 | 99 |
| 21p | *ປ່ອຍໃຫ້ເດັກຊີ້ໃສ່ຄຳສັບ (ແນ່ໃຈວ່າແມ່ນແຕ່ຊ້າຍຫາຂວາ) ແລະ ອ່ານເອງ* | ໂສ້ງ | 1 | 0 | 99 |
| 21q | *ປ່ອຍໃຫ້ເດັກຊີ້ໃສ່ຄຳສັບ (ແນ່ໃຈວ່າແມ່ນແຕ່ຊ້າຍຫາຂວາ) ແລະ ອ່ານເອງ* | ລ້າງມື | 1 | 0 | 99 |
| 21r | *ປ່ອຍໃຫ້ເດັກຊີ້ໃສ່ຄຳສັບ (ແນ່ໃຈວ່າແມ່ນແຕ່ຊ້າຍຫາຂວາ) ແລະ ອ່ານເອງ* | ແມ່ | 1 | 0 | 99 |
| 21s | *ປ່ອຍໃຫ້ເດັກຊີ້ໃສ່ຄຳສັບ (ແນ່ໃຈວ່າແມ່ນແຕ່ຊ້າຍຫາຂວາ) ແລະ ອ່ານເອງ* | ຕັ່ງ | 1 | 0 | 99 |
| 21t | *ປ່ອຍໃຫ້ເດັກຊີ້ໃສ່ຄຳສັບ (ແນ່ໃຈວ່າແມ່ນແຕ່ຊ້າຍຫາຂວາ) ແລະ ອ່ານເອງ* | ປື້ມຫັດອ່ານ | 1 | 0 | 99 |

#22 - Decodable Words

ອຸປະກອນ: ລາຍການຄຳສັບ
ເອົາບັນຊີລາຍການ “ຄຳສັບຕ່າງໆ ທີ່ຄິດຄົ້ນ” ທີ່ຫຸ້ມດ້ວຍຊອງໃສ ໃຫ້ກັບເດັກເບິ່ງ.
*ຈົ່ງເວົ້າວ່າ: ຂ້ອຍຢາກໃຫ້ເຈົ້າອ່ານລາຍການຄຳສັບໃຫ້ຂ້ອຍຟັງຕື່ມອີກ. ອີກຄັ້ງ, ຄຳສັບເຫຼົ່ານີ້ ບໍ່ມີໃນປຶ້ມແບບຮຽນ, ມັນບໍ່ແມ່ນຄຳສັບທີ່ມີຕົວຈິງ ເພາະເປັນຄຳສັບທີ່ພວກເຮົາແຕ່ຂຶ້ນເອງ. ແຕ່ວ່າ ກໍສາມາດອ່ານໄດ້. ເຈົ້າຈົ່ງຊີ້ໃສ່ ແລະ ອ່ານແຕ່ລະຄຳ ໂດຍເລີ່ມຕົ້ນຈາກນີ້* ***(ນັກສຳພາດຊີ້ໃສ່ຄຳທຳອິດ)*** *ແລະ ຍ້າຍໄປເທື່ອລະຄຳສັບແບບນີ້* ***(ນັກສຳພາດຊີ້ຄຳຕໍ່ໄປ ໂດຍເລີ່ມແຕ່ຊ້າຍຫາຂວາ).*** *ນ້ອງເຂົ້າໃຈບໍ? ເອົາລະ ເລີ່ມຕົ້ນໄດ້ແລ້ວ.*
***<u>ນັກສຳພາດ</u>***: ໝາຍທຸກຄຳສັບທີ່ເດັກອ່ານຜິດ. **<u>ຈົ່ງຈື່ວ່າ</u>**
ການອອກສຽງຄຳຕ່າງໆຕາມສຳນຽງທ້ອງຖິ່ນ ແມ່ນຍອມຮັບໄດ້.
ຖ້າວ່າເດັກອ່ານບໍ່ໄປຕາມລຳດັບ ເວົ້າກັບເດັກວ່າ “ນ້ອງອ່ານຄຳນີ້ໄດ້ບໍ່
***(ນັກສຳພາດຊີ້ໃສ່ຄຳສັບທີ່ເດັກຂ້າມ)*** ແລ້ວບອກເດັກອ່ານຄຳຕໍ່ໄປຕາມລະດັບ
ແຕ່ຊ້າຍຫາຂວາ”. ນັກສຳພາດຕ້ອງແນ່ໃຈວ່າ,
ທຸກຄຳສັບທີ່ເດັກອ່ານຜິດແມ່ນຖືກໝາຍໄວ້ແລ້ວ ກ່ອນທີ່ຈະໄປເຮັດກິດຈະກຳກັບຂໍ້ຕໍ່ໄປ.

**<u>ສຳລັບເດັກທີ່ອ່ອນ:</u>**
ຖ້າວ່າເດັກລັງເລໃຈໃນຄຳສັບໃດຄຳສັບໜຶ່ງ ເປັນເວລາ 5 ວິນາທີ, ຈົ່ງຖາມເດັກວ່າ,
ມີຄຳໃດໃນບັນຊີລາຍການນີ້ທີ່ນ້ອງຮູ້ແດ່? ບອກຂ້ອຍແດ່ ຫຼື ອ່ານຄຳທີ່ນ້ອງຮູ້ເບິ່ງດຸ.
***<u>ນັກສຳພາດ</u>***: ຈົ່ງເວົ້າຊ້ຳໆອີກ ເພື່ອກະຕຸກຊຸກຍູ້ໃຫ້ເດັກສຶບຕໍ່ອ່ານ.
ຖ້າເດັກຍັງລັງເລໃຈຢູ່ເປັນເວລາ 10 ວິນາທີ ໃຫ້ເວົ້າກັບເດັກວ່າ “ຂອບໃຈຫຼາຍໆ
ເກັ່ງຫຼາຍເທີ້ນ້ອງພະຍາຍາມອ່ານ ແລະ ເຮັດໄດ້ດີຫຼາຍແລ້ວ”.
ຫຼັງຈາກນັ້ນໃຫ້ໝາຍທຸກຄຳສັບທີ່ເດັກບໍ່ໄດ້ອ່ານ ຫຼື ບໍ່ພະຍາຍາມທີ່ຈະອ່ານ **ວ່າແມ່ນຜິດ**
ກ່ອນຈະຍ້າຍໄປຂໍ້ຕໍ່ໄປ.

| **<u>ຄຳແນະນຳ</u>**<br>*ປ່ອຍໃຫ້ເດັກອ່ານ ແລະ*<br>*ຊີ້ໃສ່ຄຳສັບດ້ວຍຕົວເອງພາຍຫຼັງທີ່ແນະນຳກິດຈະກຳແລ້ວ* | | ຄຳຕອບທີ່ຖືກຕ້ອງ | ຖືກ | ບໍ່ຖືກ | ຝ່ານບໍ່ |
|---|---|---|---|---|---|
| 21a | *ນັກສຳພາດຊີ້ໃສ່ຄຳສັບ ແລ້ວບອກໃຫ້ເດັກອ່ານ* | ດິ | 1 | 0 | 99 |
| 21b | *ປ່ອຍໃຫ້ເດັກຊີ້ໃສ່ຄຳສັບ*<br>*(ແນ່ໃຈວ່າແມ່ນແຕ່ຊ້າຍຫາຂວາ) ແລະ ອ່ານເອງ* | ຕີ | 1 | 0 | 99 |
| 21c | *ປ່ອຍໃຫ້ເດັກຊີ້ໃສ່ຄຳສັບ*<br>*(ແນ່ໃຈວ່າແມ່ນແຕ່ຊ້າຍຫາຂວາ) ແລະ ອ່ານເອງ* | ຝີ | 1 | 0 | 99 |
| 21d | *ປ່ອຍໃຫ້ເດັກຊີ້ໃສ່ຄຳສັບ*<br>*(ແນ່ໃຈວ່າແມ່ນແຕ່ຊ້າຍຫາຂວາ) ແລະ ອ່ານເອງ* | ເສ | 1 | 0 | 99 |
| 21e | *ປ່ອຍໃຫ້ເດັກຊີ້ໃສ່ຄຳສັບ*<br>*(ແນ່ໃຈວ່າແມ່ນແຕ່ຊ້າຍຫາຂວາ) ແລະ ອ່ານເອງ* | ຖຸ | 1 | 0 | 99 |
| 21f | *ປ່ອຍໃຫ້ເດັກຊີ້ໃສ່ຄຳສັບ*<br>*(ແນ່ໃຈວ່າແມ່ນແຕ່ຊ້າຍຫາຂວາ) ແລະ ອ່ານເອງ* | ເບີ | 1 | 0 | 99 |
| 21g | *ປ່ອຍໃຫ້ເດັກຊີ້ໃສ່ຄຳສັບ*<br>*(ແນ່ໃຈວ່າແມ່ນແຕ່ຊ້າຍຫາຂວາ) ແລະ ອ່ານເອງ* | ເຢະ | 1 | 0 | 99 |
| 21h | *ປ່ອຍໃຫ້ເດັກຊີ້ໃສ່ຄຳສັບ*<br>*(ແນ່ໃຈວ່າແມ່ນແຕ່ຊ້າຍຫາຂວາ) ແລະ ອ່ານເອງ* | ແດ | 1 | 0 | 99 |
| 21i | *ປ່ອຍໃຫ້ເດັກຊີ້ໃສ່ຄຳສັບ*<br>*(ແນ່ໃຈວ່າແມ່ນແຕ່ຊ້າຍຫາຂວາ) ແລະ ອ່ານເອງ* | ໄອ້ | 1 | 0 | 99 |

| | | | | | |
|---|---|---|---|---|---|
| 21j | *ບ່ອຍໃຫ້ເດັກຊີ້ໃສ່ຄຳສັບ (ແນ່ໃຈວ່າແມ່ນແຕ່ຊ້າຍຫາຂວາ) ແລະ ອ່ານເອງ* | ຢາບ | 1 | 0 | 99 |
| 21k | *ບ່ອຍໃຫ້ເດັກຊີ້ໃສ່ຄຳສັບ (ແນ່ໃຈວ່າແມ່ນແຕ່ຊ້າຍຫາຂວາ) ແລະ ອ່ານເອງ* | ໂຜ້ | 1 | 0 | 99 |
| 21l | *ບ່ອຍໃຫ້ເດັກຊີ້ໃສ່ຄຳສັບ (ແນ່ໃຈວ່າແມ່ນແຕ່ຊ້າຍຫາຂວາ) ແລະ ອ່ານເອງ* | ຝາມ | 1 | 0 | 99 |
| 21m | *ບ່ອຍໃຫ້ເດັກຊີ້ໃສ່ຄຳສັບ (ແນ່ໃຈວ່າແມ່ນແຕ່ຊ້າຍຫາຂວາ) ແລະ ອ່ານເອງ* | ເຫ | 1 | 0 | 99 |
| 21n | *ບ່ອຍໃຫ້ເດັກຊີ້ໃສ່ຄຳສັບ (ແນ່ໃຈວ່າແມ່ນແຕ່ຊ້າຍຫາຂວາ) ແລະ ອ່ານເອງ* | ງ່ | 1 | 0 | 99 |
| 21o | *ບ່ອຍໃຫ້ເດັກຊີ້ໃສ່ຄຳສັບ (ແນ່ໃຈວ່າແມ່ນແຕ່ຊ້າຍຫາຂວາ) ແລະ ອ່ານເອງ* | ຜ | 1 | 0 | 99 |
| 21p | *ບ່ອຍໃຫ້ເດັກຊີ້ໃສ່ຄຳສັບ (ແນ່ໃຈວ່າແມ່ນແຕ່ຊ້າຍຫາຂວາ) ແລະ ອ່ານເອງ* | ຕາ | 1 | 0 | 99 |
| 21q | *ບ່ອຍໃຫ້ເດັກຊີ້ໃສ່ຄຳສັບ (ແນ່ໃຈວ່າແມ່ນແຕ່ຊ້າຍຫາຂວາ) ແລະ ອ່ານເອງ* | ຮາຍ | 1 | 0 | 99 |
| 21r | *ບ່ອຍໃຫ້ເດັກຊີ້ໃສ່ຄຳສັບ (ແນ່ໃຈວ່າແມ່ນແຕ່ຊ້າຍຫາຂວາ) ແລະ ອ່ານເອງ* | ຍ້າ | 1 | 0 | 99 |
| 21s | *ບ່ອຍໃຫ້ເດັກຊີ້ໃສ່ຄຳສັບ (ແນ່ໃຈວ່າແມ່ນແຕ່ຊ້າຍຫາຂວາ) ແລະ ອ່ານເອງ* | ໂຂ່ | 1 | 0 | 99 |
| 21t | *ບ່ອຍໃຫ້ເດັກຊີ້ໃສ່ຄຳສັບ (ແນ່ໃຈວ່າແມ່ນແຕ່ຊ້າຍຫາຂວາ) ແລະ ອ່ານເອງ* | ເຜ | 1 | 0 | 99 |
| ***ທ້າຫມົດນີ້ແມ່ນຮູບພາບທີ່ຂ້ອຍມີ, ເຈົ້າຕອບໄດ້ດີຫລາຍ! ຂອບໃຈ*** | | | | | |

## Annex 5: References


Abadzi, H. (2013). “Literacy for All in 100 Days? A Research-based Strategy for Fast Progress in Low-income Countries.” Washington, DC: Global Partnership for Education.

Abadzi, H. (2012). “Developing Cross-Language Metrics for Reading Fluency Measurement: Some Issues and Options.” Washington, DC: Global Partnership for Education.

Abadzi, H. (2011). “Reading Fluency Measurements in EFA FTI Partner Countries: Outcomes and Improvement Prospects.” Washington, DC: EFA-FTI Secretariat.

Abdelhadi, S., Ibrahim, R., and Eviatar, Z. (2011). “Perceptual Load in the Reading of Arabic: Effects of Orthographic Visual Complexity on Detection.” *Writing Systems Research, 3*(2), pp. 117-127.

Adams, M. J. (1990). *Beginning to Read: Thinking and Learning about Print.* Cambridge, Massachusetts: MIT Press.

Asfaha, Y.M., Kurvers, J., and Kroon, S. (2009). “Grain Size in Script and Teaching: Literacy Acquisition in Ge’ez and Latin.” *Applied Psycholinguistics, 30*(4), pp. 709–724.

Alexander, P.A. and Fox, E. (2004). “A Historical Perspective on Reading Research and Practice.” In Ruddell, R.B. and Unrau, N.J. (Eds.). *Theoretical Models and Processes of Reading (5th Ed.)*. Newark, Delaware: International Reading Association, pp. 33-68.

American Institutes for Research (AIR) (2016). “Evaluation of the Accelerated School Readiness Intervention in Ethiopia: Cohort 1 School Readiness Outcome.” Unpublished report. Washington, DC: AIR.

American Institutes for Research (AIR) (2017). “Monitoring, Evaluation, and Research Activities for the Lao Educational Access, Research, and

Networking (LEARN) Project: Baseline Evaluation Report." Unpublished report. Washington, DC: AIR.

American Institutes for Research (AIR) (2019). "Monitoring, Evaluation, and Research Activities for the Lao Educational Access, Research, and Networking (LEARN) Project: Final Evaluation Report." Unpublished report. Washington, DC: AIR.

American Institutes for Research (AIR) (2018). "Monitoring, Evaluation, and Research Activities for the Lao Educational Access, Research, and Networking (LEARN) Project: Midline Evaluation Report." Unpublished report. Washington, DC: AIR.

Ansell, N., Froer, P., Huijsmans, R., and Rivers, I. (2018). "Education Systems, Aspiration and Learning in Remote Rural Settings: Draft Study Report." London: Brunel University.

Bailey, D., Duncan, G.J., Odgers, C.L., and Yu, W. (2017). "Persistence and Fadeout in the Impacts of Child and Adolescent Interventions." *Journal of Research on Educational Effectiveness, 10*(1), pp. 7-39.

Baird, J., Andrich, D., Hopfenbeck T., and Stobart, G. (2017). "Assessment and Learning: Fields Apart." *Assessment in Education: Principles, Policy & Practice, 4*(3), pp. 317-350.

Barma, N. and Oksen, S. (2014). "The Institution-Building Context in Lao PDR." In Barma, N., Huybens, E., and Viñuela, L. (Eds.) *Institutions Taking Root: Building State Capacity in Challenging Contexts.* Washington, DC: World Bank, pp. 111-118.

Bartlett, L., Dowd, A.J., and Jonason, C. (2015). "Problematizing Early Grade Reading: Should the Post-2015 Agenda Treasure What is Measured?" *International Journal of Educational Development, 40,* pp. 308-314.

Basic Education Quality and Access in Lao PDR (BEQUAL) (2018). "ASLO Grade 3 Headline Results and Implications." Vientiane, Lao PDR: BEQUAL.

Basic Education Quality and Access in Lao PDR (BEQUAL) (2018). “Implementation Plan Phase 1 (October 2018 – August 2020).” Vientiane, Lao PDR: BEQUAL.

Bauer, L. and Schanzenbach, D. (2016). “The Long-Term Impact of the Head Start Program.” Washington, DC: Brookings Institution.

Bedore, L.M., Pena, E.D., Garcia, M., and Cortez, C. (July 2005). “Conceptual versus Monolingual Scoring: When Does It Make a Difference?” *Language, Speech, and Hearing Services in Schools, 36*(3), pp. 188-200.

Bekman, S. and Diri., A. (2016). “The Effects of a Summer School Program on Lao Children.” Unpublished report. Vientiane, Lao PDR: Plan International.

Bekman, S., Koç, A.A., and Taylan, E.E. (2011). “Effectiveness of an Intervention Program for Six-year-olds: A Summer-school Model.” *European Early Childhood Education Research Journal, 19*(4), pp. 409-431.

Bekman, S., Koç, A.A., and Taylan, E.E. (2004). “A Model of Early Intervention in South-East Turkey: A Pilot Implementation of the ‘Summer Preschool.’” Istanbul: Boğaziçi University and Mother-Child Education Foundation (AÇEV).

Benveniste, L., Marshall, J., and Santibañez, L. (2007). “Teaching in Lao PDR.” Washington, DC: World Bank.

Black, M., Walker, S., Fernald, L., Andersen, C., DiGirolamo, A., Lu, C., McCoy, D., Fink, G., Shawar, Y., Shiffman, J., Devercelli, A., Wodon, Q., Vargas-Baron, E., and Grantham-McGregor, S. “Early Childhood Development Coming of Age: Science through the Life Course.” *Lancet 2017*(389), pp. 77–90.

Blythe, H.I., Liang, F., Zang, C., Wang, Z., Yan, G., Bai, X., and Liversedge, S.P. (2012). “Inserting Spaces into Chinese Text Helps Readers to Learn New Words: An Eye Movement Study.” *Journal of Memory and Language, 67,* pp. 241–254.

Bonilla, J., Spier, E., Carson, K., Ring, H., and Sirma, P. (2018). “Evaluation of the UNICEF Mozambique Accelerated School Readiness Pilot Programme: Midline Report.” Washington, DC: American Institutes for Research.

Bonilla, J., Spier, E., Carson, K., Ring, H., Belyakova, Y., Brodziak, I., and Adelman-Sil, E. (2019). “Evaluation of the UNICEF Mozambique Accelerated School Readiness Pilot Programme: Final Report.” Washington, DC: American Institutes for Research.

Brinkman, S. (2015). Meeting with Jodie Fonseca, 6 October.

British Educational Research Association (BERA). (2018). “Ethical Guidelines for Education Research.” London: BERA.

Britto, P. (2012). “School Readiness: A Conceptual Framework.” New York: UNICEF.

Burnham, D., Kasisopa, B., Reid, A., Luksaneeyanawin, S., Lacerda, F., Attina, V., Rattanasone, N.X., Schwarz, I., and Webster, D. (2015). “Universality and Language-specific Experience in the Perception of Lexical Tone and Pitch.” *Applied Psycholinguistics, 36,* pp. 1459–1491.

Burnham, D., Francis, E., Webster, D., Luksaneeyanawin, S., Attapaiboon, C., Lacerda, F., and Keller, P.K. (1996). “Perception of Lexical Tone across Languages: Evidence for a Linguistic Mode of Processing.” 4th International Conference on Second Language Processing, Philadelphia, USA.

Burnham, D., Lau, S., Tam, H., and Schoknecht, C. (2001). “Visual Discrimination of Cantonese Tone by Tonal but Non-Cantonese Speakers, and by Non-Tonal Language Speakers.” 4th International Conference on Auditory-Visual Speech Processing, Aalborg, Denmark.

Burnham, D. and Jones, C. (2002). “Categorical Perception of Lexical Tone by Tonal and Non-Tonal Language Speakers.” 9th Australian International Conference on Speech Science and Technology, Melbourne, Australia.

Care, E., Robertson, P., and Ferido, M. (2018). “Learning Progression Models that Support Learning at the Bottom of the Pyramid.” In Wagner, D.A, Wolf, S., and Boruch, R.F. (Eds.) *Learning at the Bottom of the Pyramid: Science, Measurement, and Policy in Low-Income Countries.* Paris: UNESCO, pp. 181-192.

Chall, J.S. (1996). *Stages of Reading Development (2nd Ed.).* Fort Worth, USA: Harcourt-Brace.

Chamberlain, J. (2017). “Contrastive Analysis of Lao with Selected Mon-Khmer, Sino-Tibetan and Hmong-Mien Languages.” Vientiane, Lao PDR: BEQUAL Project.

Chamberlain, J. (no date). “Enantiodromia in National Language Policy and Language Planning: The Case of the Lao PDR.” International Conference on Sino-Tibetan Languages and Linguistics. No location identified.

Chan, L.L.S and Sylva, K. (2015). “Exploring Emergent Literacy Development in a Second Language: A Selective Literature Review and Conceptual Framework for Research.” *Journal of Early Childhood Literacy, 15*(1), pp. 3–36.

Chen, X. and Pasquarella, A. (2017). “Learning to Read Chinese.” In Verhoeven, L. and Perfetti, C. (Eds.). *Learning to Read across Languages and Writing Systems.* Cambridge, UK: Cambridge University Press.

Chounlamany, K. and Kounphilaphanh, B. (2011). *New Methods of Teaching? Reforming Education in Lao PDR.* PhD Thesis. Umea, Sweden: Department of Education, Umea University.

Cincotta-Segi, A. (2014). “Language/ing in Education: Policy Discourse, Classroom Talk and Ethnic Identities in the Lao PDR.” In Sercombe, P. and Tupas, R. (Eds.). *Language, Education and Nation-building. Palgrave Studies in Minority Languages and Communities.* London: Palgrave Macmillan, pp. 106-130.

Cincotta-Segi, A. (2011a) “Signalling L2 Centrality, Maintaining L1 Dominance: Teacher Language Choice in an Ethnic Minority Primary Classroom in the Lao PDR.” *Language and Education, 25*(1), 19-31.

Cincotta-Segi, A. (2011) “Talking in, Talking around, and Talking about the L2: Three Literacy Teaching Responses to L2 Medium of Instruction in the Lao PDR.” *Compare: A Journal of Comparative and International Education, 41*(2), 195-209.

Cincotta-Segi, A. (2011b). "'The Big Ones Swallow the Small Ones.' Or Do They? Language-In-Education Policy and Ethnic Minority Education in the Lao PDR." *Journal of Multilingual and Multicultural Development,* 32(1), 1-15.

Cohen, J. (1998). *Statistical Power Analysis for the Behavioral Sciences (2nd Ed.).* Mahwah, New Jersey: Lawrence Erlbaum Associates.

Crawfurd, L., Hares, S., Le Nestour, A., and Rossiter, J. (2019). *Does Education Need a QALY and Is LAYS It?* [Online]. Retrieved from https://www.cgdev.org/blog/does-education-need-qaly-and-lays-it. Accessed 22 November 2019.

Crouch, L. (2019). "Making the Pyramid Less Pyramidal?" In Wagner, D.A, Wolf, S., and Boruch, R.F. (Eds.) *Learning at the Bottom of the Pyramid: Science, Measurement, and Policy in Low-Income Countries.* Paris: UNESCO, pp. 59-71.

Crouch, L., Olefir, A., Saeki, H., and Savrimootoo, T. (2020). "Déjà Vu All Over Again? Recent Evidence on Early Childhood and Early Grade Repetition in Developing Countries." *Journal of Multilingual and Multicultural Development,* 32(1), 1-15. *Prospects* (2020).

Crouch, L., Rolleston, C., and Gustafsson, M. (2020). "Eliminating Global Learning Poverty: The Importance of Equalities and Equity*." International Journal of Educational Development,* in press.

Cunnington, P. (2011). "Resettlement in Laos." Vientiane, Lao PDR: European Union.

Danchev, P. (2018). Personal communication to Jodie Fonseca, 5 July.

DaSilva, C. (2019). *Moving Beyond Mother Tongue: A Pro-multilingual Approach to Learning* [Online]. Retrieved from: https://www.globalpartnership.org/blog/moving-beyond-mother-tongue-pro-multilingual-approach-learning. Accessed 1 November 2019.

Dombey, H. (2009). "The Simple View of Reading." Brighton, UK: University of Brighton.

Dooley, K. (2015). “Literature Review: Teaching Tone in Lao Language Literacy to Students in Early Grades.” Vientiane, Lao PDR: Room to Read.

Dowd, A.J. and Pisani, L. (2018). “Early Learning and Development at the Bottom of the Pyramid.” In Wagner, D.A, Wolf, S., and Boruch, R.F. (Eds.) *Learning at the Bottom of the Pyramid: Science, Measurement, and Policy in Low-Income Countries.* Paris: UNESCO, pp. 159-173.

Dowd, A.J., Friedlander, E., Guajardo, J., Mann, N., and Pisani, L. (2013). “Literacy Boost Cross Country Analysis Results.” Washington, DC: Save the Children.

ECD Measure (no date). *Core Direct Assessment Item Documentation* [Online]. Retrieved from http://ecdmeasure.org/new-map/. Accessed 24 May 2019.

ECD Measure (2019). *MELQO Around the World* [Online]. Retrieved from http://ecdmeasure.org/new-map/. Accessed 24 May 2019.

Education Partnerships Group (2019). “Profile of Pre-primary Provision in Abidjan and Bouaké: A Review and Policy Implications for Establishing a Pilot.” London: Education Partnerships Group.

Educational Standards and Quality Assurance Center (ESQAC), and Emblen, V. (2011). “Research Report: Teaching Practices and Learning Outcomes, Grade 3, Two Remote Districts in Bokeo Province.” Vientiane, Lao PDR: ESQAC and Plan International.

Ehri, L.C. (2005). “Learning to Read Words: Theory, Findings, and Issues.” *Scientific Studies of Reading, 9*(2), pp. 167–188.

Enfield, N.J. (2010) “Language and Culture in Laos: An Agenda for Research.” *The Journal of Lao Studies, 1*(1), pp. 48-54.

Enfield, N.J. (2007). “A Grammar of Lao.” Berlin: Mouton de Gruyter Press.

Enfield, N. J. (2009). “‘Case Relations’ in Lao, a Radically Isolating Language.” In Malčukov, A.L., and A. Spencer, A. (Eds.). *The Oxford Handbook of Case*. Oxford: Oxford University Press.

Enfield, N.J. (2008). “Lao Linguistics in the 20th Century and Since.” In Y. Goudineau, Y., and Lorrillard, M. (Eds.), *Recherches Nouvelles Sur le Laos*. Paris: École Française d'Extrême-Orient, pp. 435-452.

EQUIP-Tanzania (no date). “Improving Equitable Access to Quality Pre-primary Education in Tanzania: The School Readiness Programme.” Dar-es-Salaam, Tanzania: EQUIP-Tanzania.

Esposito, C.M and Kham, S.D. (2012). “Contrastive Breathiness across Consonants and Vowels: A Comparative Study of Gujarati and White Hmong.” *Journal of the International Phonetic Association, 42*(2), pp.123-143.

Ethnologue (2019). *Languages of the World: Laos* [Online]. Retrieved from https://www.ethnologue.com/country/LA. Accessed 8 May 2019.

Fernald, L.C.H., Prado, E., Kariger, P., and Raikes, A. (2017). “A Toolkit for Measuring Early Childhood Development in Low- and Middle-Income Countries.” Washington, DC: World Bank.

Flewitt, R. and Ang, L. (2020). *Research Methods for Early Childhood Education.* London: Bloomsbury Academic.

Florit, E. and Cain, K. (2011). “The Simple View of Reading: Is It Valid for Different Types of Alphabetic Orthographies?” *Educational Psychology Review, 23,* pp. 553-576.

Fonseca, J. (2015). “Literature Review for the Design of the Lao Educational Access, Research and Networking Project.” Unpublished report. Vientiane, Lao PDR: Plan International.

Fonseca, J. (2019). “Promoting Evidence-based, Accelerated School Readiness at Scale in Lao PDR: Experiences from Plan International’s LEARN Project.” Comparative and International Development Society Annual Conference, San Francisco, 16 April 2019.

Frost, R. and Katz, L. (1989). “Orthographic Depth and the Interaction of Visual and Auditory Processing in Word Recognition.” *Memory & Cognition, 17*(3), 302-310.

Gee, J. P. (1998). “The New Literacy Studies: From ‘Socially Situated’ to the Work of the Social.” Madison, Wisconsin: University of Wisconsin.

Gertler, P. et. al. (2011). *Impact Evaluation in Practice.* Washington, DC: World Bank.

Gertler, P., Heckhman, J., Pinto, R., Zanolini, A., Vermeersch, C., Walker, S., Chang, S., and Grantham-McGregor, S. (2014). “Labor Market Returns to an Early Childhood Stimulation Intervention in Jamaica.” *Science, 344(*6187), pp. 998-1001.

Gillman, M. and Swain, J. (2006). *The SAGE Dictionary of Social Research Methods.* London: Sage.

Global Partnership for Education (2021). “Adapting, Testing and Scaling a Proven Summer Pre-primary Education Model in Cambodia, Lao PDR and Tanzania.” Washington, DC: Global Partnership for Education.

Gomez, J.A., Brown, J.L, and Spier, E. (2020): “Factor Structure, Measurement Invariance, and Construct Validity of MELQO Measure of Child Development and Early Learning in Laos.” *International Journal of School & Educational Psychology* (2020), 1-77.

Gove, A. and Wettenberg, A. (2011). “The Early Grade Reading Assessment: An Introduction.” In Gove, A. and Wettenberg, A. (Eds.) *The Early Grade Reading Assessment: Applications and Interventions to Improve Basic Literacy.* Research Triangle Park, NC: RTI International, pp. 1-38.

Government of Lao PDR (2000). “Decree 001.” Vientiane: Government of Lao PDR.

Graham, B.E. and van Ginkel, A.J. (2014) “Assessing Early Grade reading: The Value and Limits of ‘Words per Minute.’” *Language, Culture and Curriculum, 27*(3), pp. 244-259.

Gross, M., Buac, M., and Kaushanskaya, M. (2014). “Conceptual Scoring of Receptive and Expressive Vocabulary Measures in Simultaneous and Sequential Bilingual Children.” *American Journal of Speech-Language Pathology 574*(23), pp. 574–586.

Handke, J. (2017). *Linguistic Micro-Lectures: Finding Minimal Pairs.* Marburg, Germany: Marburg University.

Hao, Y. (2012). “Second Language Acquisition of Mandarin Chinese Tones by Tonal and Non-tonal Language Speakers.” *Journal of Phonetics, 40,* pp. 269–279.

Hart, C. (2018). *Doing a Literature Review: Releasing the Research Imagination (Second Edition).* London: Sage.

Helburn, S. (1995). “Cost, Quality and Child Outcomes in Child Care Centers.” Denver: University of Colorado Department of Economics.

High, P. (2008). “School Readiness: Technical Report.” *Pediatrics 121*(4), pp. 1,008-1,015.

HighScope (2021). *Perry Preschool Project* [Online]. Available from https://highscope.org/perry-preschool-project/. Accessed 8 August 2021.

Ho, C.S.-K. and Bryant, P. (1997) “Learning to Read Chinese Beyond the Logographic Phase.” *Reading Research Quarterly, 32,* pp. 276–289.

Hornberger, N.H. (2018). “Ethnography and Assessment at the Bottom of the Pyramid: Commentary.” In Wagner, D.A, Wolf, S., and Boruch, R.F. (Eds.) *Learning at the Bottom of the Pyramid: Science, Measurement, and Policy in Low-Income Countries.* Paris: UNESCO, pp. 154-155.

IDELA Network (2019). *Countries of Implementation* [Online]. Retrieved from https://idela-network.org/about/countries-where-idela-is-in-use/. Accessed 24 May 2019.

Indochina Research Ltd. (2016). “Completion Report: Data Collection for the LEARN Project Impact Evaluation – Baseline Survey 2016.” Unpublished report. Vientiane: Indochina Research Ltd.

Indochina Research Ltd. (2017). “Completion Report: Data Collection for the LEARN Project Impact Evaluation – Midline Survey 2017.” Unpublished report. Vientiane: Indochina Research Ltd.

Indochina Research Ltd. (2018a). “Completion Report: Data Collection for the LEARN Project Impact Evaluation – Endline Survey 2018.” Unpublished report. Vientiane: Indochina Research Ltd.

Indochina Research Ltd. (2018b). “Completion Report: Data Collection for the LEARN Project Impact Evaluation – Follow-up Survey 2018.” Unpublished report. Vientiane: Indochina Research Ltd.

Institute of Education Sciences (IES) (2020). “What Works Clearinghouse: Standards Handbook Version 4.1.” Washington, DC: United States Department of Education.

Jasinska, K.K, Wolf, S., Jukes, M.C.H., and Dubeck, M.M. (2019). “Literacy Acquisition in Multilingual Educational Contexts: Evidence from Coastal Kenya.” *Developmental Science, 22*(5), pp. 1-17.

Joshi, R.M. (2010). “Role of Orthography in Literacy Acquisition and Literacy Problems Among Monolinguals and Bilinguals.” In Aram, D. and Korat, O. (Eds.). *Literacy Development and Enhancement Across Orthographies and Cultures.* Boston: Springer*,* pp.167-176.

Joshi, R.M., Tao, S., Aaron, P.G., and Quiroz, B. (2012). “Cognitive Component of Componential Model of Reading Applied to Different Orthographies.” *Journal of Learning Disabilities, 45*(5), p. 480–486.

Joshi, R.M. and Aaron (2000). “The Component Model of Reading: Simple View of Reading Made a Little More Complex.” *Reading Psychology, 21,* pp. 85–97.

Jukes, M., Cummiskey, C., and Gargano, M.N. (2018). “Guidance Note: Setting Data-Driven Oral Reading Fluency Benchmarks.” Research Triangle Park, NC: RTI International.

Kagan, S., Moore, E., and Bredekamp, S. (eds.) (1995). “Reconsidering Children's Early Development and Learning: Toward Common Views and Vocabulary.” Washington, DC: National Education Goals Panel.

Kamerman, S. (2006). "A Global History of Early Childhood Education and Care.” Paris: UNESCO.

Kanjee, A. (2018). “Use of International Assessments to Support Learning at the Bottom of the Pyramid.” In Wagner, D.A, Wolf, S. and Boruch, R.F. (Eds.). *Learning at the Bottom of the Pyramid: Science, Measurement, and Policy in Low-Income Countries.* Paris: UNESCO, pp. 77-92.

Kasisopa, B., Reilly, R.G., Luksaneeyanawin, S., and Burnham, D. “Child Readers’ Eye Movements in Reading Thai.” *Vision Research, 123*(2016), pp. 8–19.

Kasisopa B., El-Khoury Antonios, L., Jongman, A., Sereno, J.A., and Burnham, D. (2018). “Training Children to Perceive Non-native Lexical Tones: Tone Language Background, Bilingualism, and Auditory-Visual Information.” *Frontiers in Psychology, 9*(1508), pp. 1-14.

Kim, Y.K. (2015). “Developmental, Component-Based Model of Reading Fluency: An Investigation of Predictors of Word-Reading Fluency, Text-Reading Fluency, and Reading Comprehension.” *Reading Research Quarterly, 50*(4), op. 459–481.

Kosonen, K. (2005) “Vernaculars in Literacy and Basic Education in Cambodia, Laos and Thailand.” *Current Issues in Language Planning, 6*(2), 122-142.

Kosonen, K. (2017). “Language of Instruction in Southeast Asia.” Paris: UNESCO.

Kurlantzick, J. (2017). *A Great Place to Have a War: America in Laos and the Birth of a Military CIA.* New York: Simon & Schuster.

Leong, C.K., Cheng, P.W., and Tan, L.H. (2005). “The Role of Sensitivity to Rhymes, Phonemes and Tones in Reading English and Chinese Pseudowords.” *Reading and Writing 18*, pp. 1–26.

Lervag, A., Hulme, C., and Melby-Lervag, M. (2018). “Unpicking the Developmental Relationship Between Oral Language Skills and Reading Comprehension: It’s Simple, But Complex.” *Child Development, 89*(5), pp. 1821–1838.

Levesque, K., Bardack, S., and Chigeda, A., (2020). “Tablet-based Learning for Foundational Literacy and Math: An 8-month RCT in Malawi.” San Francisco: Imagine Worldwide.

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.

Linan-Thompson, S. (2014). “Screening of School Readiness Skills: A Review of the Literature.” Washington, DC: USAID.

Liow, S.J.R. and Lee, L.C. (2004). “Metalinguistic Awareness and Semi-syllabic Scripts: Children’s Spelling Errors in Malay.” *Reading and Writing: An Interdisciplinary Journal, 17*, pp. 7–26.

Liu, Y., Wang, M., Perfetti, C.A., Brubaker, B. and Wu, S., and MacWhinney, B. (2011). “Learning a Tonal Language by Attending to the Tone: An In Vivo Experiment.” *Language Learning 61*(4), pp. 1119–114.

Lor, M. and Gao, G. (2020). “Hmong and Chinese Qualitative Research Interview Questions: Assumptions and Implications of Applying the Survey Back Translation Method.” In Sha, M. and Gabel, T. (Eds.) *The Essential Role of Language in Survey Research.* Research Triangle Park: RTI Press.

Luk, G., and Bialystok, E. (2008). “Common and Distinct Cognitive Bases for Reading in English–Cantonese Bilinguals.” *Applied Psycholinguistics 29*(02) pp. 269-289.

Maddieson, I. (2013). “Tone.” In Dryer, M.S. and Haspelmath, M. (Eds.) *The World Atlas of Language Structures Online.* Leipzig: Max Planck Institute for Evolutionary Anthropology. Retrieved from http://wals.info/chapter/13. Accessed 23 June 2019.

Maddox, B. (2018). “Ethnography and Assessment at the Bottom of the Pyramid.” In Wagner, D.A, Wolf, S., and Boruch, R.F. (Eds.) *Learning at the Bottom of the Pyramid: Science, Measurement, and Policy in Low-Income Countries.* Paris: UNESCO.

Magnuson, K.A., Ruhm, C., and Waldfogel, J. (2007). “The Persistence of Preschool Effects: Do Subsequent Classroom Experiences Matter?” *Early Childhood Research Quarterly 22*(2007), pp. 18–38.

Mang, E. (2001). “A Cross-language Comparison of Preschool Children’s Vocal Fundamental Frequency in Speech and Song Production.” *Research Studies in Music Education, 16*, p. 1-14.

Martinez, S., Naudeau, S., and Pereira, V. (2012). “The Promise of Preschool in Africa: A Randomized Impact Evaluation of Early Childhood Development in Rural Mozambique.” Washington, DC: World Bank and Save the Children.

MacKinnon, A. and Thepphasoulithone, P. (2014). “Educational Reform in Laos: A Case Study.” *International Journal of Educational Studies, 1*(1), pp. 19-34.

Matthews, P.H. (2014). *The Concise Oxford Dictionary of Linguistics (3 ed.).* Oxford, UK: Oxford University Press.

McBride-Chang, C., Tong, X., Shu, H., Wong, A.M.Y, Leung, K., and Tardif, T. (2008) “Syllable, Phoneme, and Tone: Psycholinguistic Units in Early Chinese and English Word Recognition.” *Scientific Studies of Reading, 12*(2), pp. 171-194.

McNess, E., Arthur, L., and Crossley, M. (2015). “Ethnographic ‘Dazzle’ and the Construction of the ‘Other’: Revisiting Dimensions of Insider and Outsider Research for International and Comparative Education.” *Compare, 45(*2), pp. 295-316.

Menter, I., Elliot, D., Hulme, M., Lewin, J., and Loweden, K. (2011). *A guide to practitioner research in education.* London: Sage.

Messerli P., Heinimann, A., Epprecht, M., Phonesaly, S., Thiraka, C., Minot, N. (2008). “Socio-Economic Atlas of the Lao PDR: An Analysis Based on the 2005 Population and Housing Census.” Vientiane, Lao PDR: Swiss National Center of Competence in Research North-South.

Ministry of Education and Sports (2009). “2009-2015 Education Sector Development Framework (ESDF).” Vientiane, Lao PDR: Ministry of Education and Sports.

Ministry of Education and Sports (2015a). “Draft Revised National Education Law.” Unpublished. Vientiane, Lao PDR: MOES.

Ministry of Education and Sports (2015b). “Education and Sector Development Plan 2016-20.” Vientiane, Lao PDR: MOES.

Ministry of Education and Sports (2018a). “ESDP 2016-2020: Mid-Term Review Report.” Vientiane, Lao PDR: MOES.

Ministry of Education and Sports (2018b). “ESDP Mid-Term Review Annex: Assessing Disparities in Access and Outcomes.” Vientiane, Lao PDR: MOES.

Ministry of Education and Sports, Department of External Relations (2017, January 31). Remarks at “The Introduction Workshop on Research Activities in the LEARN Project.” Vientiane, Lao PDR: MOES.

Ministry of Education and Sports, Department of Personnel and Organization (2018). “Report on the Survey of the Number of Teachers and Administration Offices across the Nation.” Vientiane, Lao PDR: MOES.

Ministry of Planning and Investment of Lao PDR (2015). “Fourth Population and Housing Census.” Vientiane, Lao PDR: Ministry of Planning and Investment.

Ministry of Planning and Investment of Lao PDR (2017). “Graduation from Least Developed Country Status, Lao PDR 2017: The 5th National Human Development Report.” Vientiane, Lao PDR: UNDP.

Mishra, R. and Stainthorp, R. (2007) “The Relationship between Phonological Awareness and Word Reading Accuracy in Oriya and English: A Study of Oriya-speaking Fifth-graders.” *Journal of Research in Reading, 30*(1), pp. 23–37.

Mixdorff, H., Charnvivit, P., and Burnham, D. (2005). “Auditory-visual Perception of Syllabic Tones in Thai.” 2005 Conference on Auditory-Visual Speech Processing, British Columbia, Canada.

Mixdorff, H., Hu, Y., and Burnham, D. (2005). “Visual Cues in Mandarin Tone Perception.” 9th European Conference on Speech Communication and Technology, Lisbon, Portugal.

Moats, L. (1998). “Teaching Decoding.” *American Educator, Spring/Summer 1998,* pp 1-13.

Mortensen, D. (2017). *Hmong-Mien Languages* [Online]. Retrieved from https://oxfordre.com/linguistics/view/10.1093/acrefore/9780199384655.001.0001/acrefore-9780199384655-e-341. Accessed 13 May 2019.

Moscoviz, Laura (2019). *Early Learning: A Focus on Accelerated School Readiness Programmes* [Online]. London: Education Partnerships Group. Available from https://epg.org.uk/early-learning-a-focus-on-accelerated-school-readiness-programmes/. Accessed 5 August 2021.

Mukari, S.Z.S., Yu, X., Ishak, W.S., and Mazlan, R. (2015) “Effects of Tonal Language Background on Tests of Temporal Sequencing in Children.” *International Journal of Audiology, 54(*12), pp. 951-957.

Munoz-Chereau, B., Ang, L., Dockrell, J., Outhwaite, L., and Heffernan, C. (2021). “Measuring Early Child Development across Low and Middle-income Countries: A Systematic Review.” *Journal of Early Childhood Research,* online edition.

Muroga, M., Zaw, H., Mizunoya, S., Lin, H., Brossard, M., and Reuge, N. (2020). “COVID-19: A Reason to Double Down on Investments in Pre-primary Education.” Innocenti, Italy: UNICEF Office of Research.

Myers, J. and Pinnock, H. (2017). “Guide to the Accelerated Education Principles.” Paris: UNESCO.

Nag, S. (2017). “Assessing Literacy in Developing Countries. Evidence Brief: What Is Available and What is Required?” London: Department for International Development.

Nag, S. (2017). “Learning to Read Kannada and Other Languages of South Asia.” In Verhoeven, L. and Perfetti, C. (Eds.). *Learning to Read across Languages and Writing Systems.* Cambridge: Cambridge University Press, pp. 82-103.

Nag, S. (2007). “Early Reading in Kannada: The Pace of Acquisition of Orthographic Knowledge and Phonemic Awareness.” *Journal of Research in Reading, 30*(1), pp. 7-22.

Nag, S., Caravolas, M., and Snowling, M. (2011). “Beyond Alphabetic Processes: Literacy and Its Acquisition in the Alphasyllabic Languages.” *Reading and Writing, 24*, pp. 615–622.

Nag, S. and Perfetti, C. (2014) “Reading and Writing: Insights from the Alphasyllabaries of South and Southeast Asia.” *Writing Systems Research, 6*(1), pp. 1-9.

Nag, S. and Snowling, M. (2012) “Reading in an Alphasyllabary: Implications for a Language Universal Theory of Learning to Read.” *Scientific Studies of Reading, 16*(5), pp. 404-423.

Nag, S., Treiman, R., and Snowling, M. (2010) “Learning to Spell in an Alphasyllabary: The Case of Kannada.” *Writing Systems Research, 2*(1), pp. 41-52.

Nag, S., Chiat, S., Torgerson, C., and Snowling, M. (2014). “Literacy, Foundation Learning and Assessment in Developing Countries: Final Report.” London: Department for International Development.

Nakamura, P.R., De Hoop, T., and Holla, C.U. (2019). “Language and the Learning Crisis: Evidence of Transfer Threshold Mechanisms in Multilingual Reading in South India.” *Journal of Development Studies,* 55*(11)*, pp. 2287–2305.

Nakamura, P.R., Koda, K., and Joshi, R. (2014). “Biliteracy acquisition in Kannada and English: A Developmental Study.” *Writing Systems Research, 6*(1), pp. 132-147.

Nakamura, P.R., Joshi, M., and Ji, X.R. (2018). “Investigating the Asymmetrical Roles of Syllabic and Phonemic Awareness in Akshara Processing.” *Journal of Learning Disabilities 51(*5), pp. 499-506.

Nakamura, P. and de Hoop, T. (2014). “Facilitating Reading Acquisition in Multilingual Environments in India (FRAME-India): Final Report.” Washington, DC: American Institutes for Research.

National Assembly of Lao PDR (2003). “Constitution of the Lao People’s Democratic Republic.” Vientiane: Government of Lao PDR.

National Assembly of Lao PDR (2007). “Law on Education (as Amended).” Vientiane, Lao PDR: Government of Lao PDR.

National Reading Panel (2000). “Teaching Children to Read: An Evidence-Based Assessment of the Scientific Research Literature on Reading and Its Implications for Reading Instruction.” Washington, DC: National Institutes of Health.

Nguyen, A.T., C.L. Ingram, C.L.J., and Pensalfini, J.R. (2008). “Prosodic Transfer in Vietnamese Acquisition of English Contrastive Stress Patterns.” *Journal of Phonetics, 36(2008),* pp. 158–190.

Office of Child Development (1970). “Project Head Start 1968: The Development of a Program.” Washington, DC: United States Department of Health, Education and Welfare.

Office of Head Start (2021). *Head Start History* [Online]. Washington, DC: United States Administration for Children and Families. Available from https://www.acf.hhs.gov/ohs/about/history-head-start. Accessed 5 August 2021.

Ovesen, J. “All Lao? Minorities in the Lao People’s Democratic Republic.” In Duncan, C.R. (Ed.) (2008). *Civilizing the Margins: Southeast Asian Government Policies for the Development of Minorities.* Singapore: National University of Singapore Press, pp. 214-240.

Ozler, B., Fernald, L.C.H, Kariger, P., McConnell, C., Neuman, M., and Fraga, E. (2018). “Combining Pre-school Teacher Training with Parenting Education: A Cluster-Randomized Controlled Trial.” *Journal of Development Economics, 133*(2018), pp. 448–467.

Perfetti, C.A. and Lesgold, A.M. (1977). “Coding and Comprehension in Skilled Reading and Implications for Reading Instruction.” In Resnick, L.B. and Weaver, P. (Eds.) *Theory and Practice of Early Reading.* Hillsdale, NJ: Erlbaum, pp. 57-84.

Perfetti, C. and Verhoeven, L. “Epilogue: Universals and Particulars in Learning to Read across Seventeen Orthographies.” In Verhoeven, L. and Perfetti, C. (Eds.) (2017). *Learning to Read across Languages and Writing Systems.* Cambridge: Cambridge University Press. pp 437-454.

Pholsena, V. (2005). “A Liberal Model of Minority Rights for an Illiberal Multiethnic State? The Case of the Lao PDR.” In Kymlicka, W. and He, B. (Eds.). *Multiculturalism in Asia.* Oxford, UK: Oxford University Press, pp. 80-119.

Pholsena, V. (2010). “Nation/Representation: Ethnic Classification and Mapping Nationhood in Contemporary Laos.” *Asian Ethnicity, 3*(2), 175-197.

Piper, B. (2018a). “What is ‘Learning’ in the Case of Marginalized Populations in Low-income Countries? Commentary.” In Wagner, D.A, Wolf, S., and Boruch, R.F. (Eds.) *Learning at the Bottom of the Pyramid: Science, Measurement, and Policy in Low-Income Countries.* Paris: UNESCO, pp. 24-26.

Piper, B. (2018b). “Tayari’s Longitudinal Endline Results: Program Impact and Implementation Decisions.” 62nd Annual Conference of the Comparative and International Education Society, Mexico City, Mexico, March 25-29, 2018.

Piper, B. and van Ginkel, A. (2017) “Reading the Script: How the Scripts and Writing Systems of Ethiopian Languages Relate to Letter and Word Identification.” *Writing Systems Research,* 9(1), pp. 36-59.

Piper, B. and King, K.M. (2019). “Longitudinal Impacts of the Medium-scale Tayari Pre-primary Intervention in Kenya: Resisting fadeout effects?” 63rd Annual Conference of the Comparative and International Education Society, San Francisco, California, USA, April 14-18, 2019.

Piper, B., Schroeder, L., and Trudell, B. (2016). “Oral Reading Fluency and Comprehension in Kenya: Reading Acquisition in a Multilingual Environment.” *Journal of Research in Reading, 39(*2), pp. 133–152.

Pisani, L., Borisova, I., and Dowd, A.J. (2018). “Developing and Validating the International Development and Early Learning Assessment (IDELA).” *International Journal of Educational Research*, 91, 1–15.

Pisani, L., Borisova, I., and Dowd, A.J. (2015). “International Development and Early Learning Assessment Technical Working Paper.” Washington, DC: Save the Children.

Plan International (2016). “Capacity Statement: Early Childhood Care and Development in Emergencies.” Woking, UK: Plan International.

Puma, M. Bell, S., Cook, R., Heid, C., Broene, P., Jenkins, F., Mashburn, A., and Downer, J. (2012). “Third Grade Follow-up to the Head Start Impact Study: Final Report.” Washington, DC: U.S. Department of Health and Human Services.

Pushparatnam, A., Bazaldua, D., Holla, A., Azevedo, J., Clarke, M., and Devercelli, A. (2021). “Measuring Early Childhood Development Among 4–6-Year-Olds: The Identification of Psychometrically Robust Items Across Diverse Contexts.” *Frontiers in Public Health 2021*(9), pp.

Qin, Z. and Mok, P.P. (2011). “Perception of Cantonese Tones by Mandarin, English and French Speakers.” International Conference of Phonetic Sciences XVII, Hong Kong, China.

Raikes, A. “Measuring Child Development and Learning.” *European Journal of Education, 52(*2017), pp. 511–522.

Raikes, A. (no date). “Results from Four MELQO Pilot Countries: Evidence of Within-Country Validity and Cross-Country Comparison of Child Development Data.” Lincoln, Nebraska: University of Nebraska and ECD Measure.

Raikes, A. (2019). Discussion with Jodie Fonseca, 16 April.

Raikes, A. and Koziol, N. (2018). “Guidance for MELQO Data Analyses at Country and Global Levels.” Lincoln, Nebraska: University of Nebraska and ECD Measure.

Raikes, A. and Sayre, R. (2019). *The Power of Local Engagement to Measure Early Childhood Education* [Online]. https://www.globalpartnership.org/blog/power-local-engagement-measure-early-childhood-education. Accessed 25 March 2021.

Raikes, A., Sayre, R. Davis, D., Anderson, K., Hyson, M., Seminario, E., and Burton, A. (2019). “The Measuring Early Learning Quality and Outcomes Initiative: Purpose, Process and Results.” *Early Years, 39*(4), pp. 360-375.

Rayner, K. and Reichle, E.D. (Eds.). (2010). “Models of the Reading Process.” *Wiley Interdisciplinary Reviews: Cognitive Science, 1*(6), 787-799.

Reynolds, A. and Temple, J. (1998). “Extended Early Childhood Intervention and School Achievement: Age Thirteen Findings from the Chicago Longitudinal Study.” *Child Development, 69*(1), pp. 231-246.

Rhyner, P.M., Haebig, E.K., and West, K.M. “Understanding Frameworks for the Emergent Literacy Stage.” In Rhyner, P.M. (Ed.) (2009). *Emergent Literacy and Language Development: Promoting Learning in Early Childhood*. New York: Guilford Press, pp. 5-35.

Roberts, David, Walter, S.L., and Snider, K. (2016). “Neither Deep nor Shallow: A Classroom Experiment Testing the Orthographic Depth of Tone Marking in Kabiye (Togo).” *Language and Speech, 9*(1), pp. 113–138.

Robson, C. (2011). *Real-World Research (Third Ed.).* Sussex: Wiley & Sons.

Room to Read. (2018). “Setting Data-Driven Oral Reading Fluency Benchmarks Guidance Note.” Washington, DC: Room to Read.

Rowsell, J. and Pahl, K. (2015). “Introduction.” In Rowsell, J. and Pahl, K. (Eds.). *The Routledge Handbook of Literacy Studies.* London: Routledge, pp. 1-16.

RTI International (2017). “All Children Reading – Asia EGRA Benchmarks and Standards Research Report.” Washington, DC: United States Agency for International Development.

RTI International (2018a). “All Children Reading – Cambodia Student Performance in Early Literacy: Baseline Report.” Washington, DC: United States Agency for International Development.

RTI International (2015a). “Assessment of Early Grade Reading in the Education Sector in Cambodia.” Washington, DC: United States Agency for International Development.

RTI International (2016a). “Early Grade Reading Assessment (EGRA) Toolkit, Second Edition.” Washington, DC: United States Agency for International Development.

RTI International (2010). “EGRA Plus: Liberia Program Evaluation Report.” Washington, DC: United States Agency for International Development.

RTI International (2018b). “Multi-Language Assessment (MLA) for Young Children: An Instrument in Development.” Presentation at the Innovations in Early Childhood Development Assessment workshop, 10 October 2018.

RTI International (2014). “Nepal Early Grade Reading Assessment (EGRA) Study.” Washington, DC: United States Agency for International Development.

RTI International (2016b). “The Primary Math and Reading (PRIMR) Initiative Endline Impact Evaluation – Revised Edition.” Washington, DC: United States Agency for International Development.

RTI International (2020). “USAID’s Early Grade Reading Program (EGRP) in Nepal: Final Progress Report.” Washington, DC: United States Agency for International Development.

Rumelhart, D.E. (1985). “Toward an Interactive Model of Reading.” In Ruddell, R. and Singer, H. (Eds). *Theoretical Models and Processes of Reading (3rd ed.).* Newark, Delaware: International Reading Association, pp. 722-750.

Rumelhart, D.E. (1980). “Schemata: The Building Blocks of Cognition.” In Spiro, R.J., Bruce, B.C., and Brewer, W.F. (Eds.). *Theoretical Issues in Reading Comprehension: Perspectives from Cognitive Psychology, Linguistics, Artificial Intelligence, and Education.* Hillsdale, NJ: Lawrence Erlbaum, pp. 33-58.

Rumelhart, D.E. and Ortony, A. (1977). “The Representation of Knowledge in Memory.” In Anderson, R. C., Spiro, J., and Montague, W. E. (Eds.). *Schooling and the Acquisition of Knowledge. London*: Routledge, pp. 99-136.

Sammons, P. (2010a). “Does Pre-school Make a Difference? Identifying the Impact of Preschool on Children’s Cognitive and Social Behavioural Development at Different Ages.” In Sylva, K., Malhuish, E., Sammons, P., Siraj-Blatchford, I., and Taggart, B. (Eds.). *Early Childhood Matters: Evidence from the Effective Pre-School and Primary Education Project.* London: Routledge.

Sammons, P. (2010b). “Do the Benefits of Pre-school Last? Investigating Pupil Outcomes to the End of Key Stage 2 (Aged 11).” In Sylva, K., Malhuish, E., Sammons, P., Siraj-Blatchford, I., and Taggart, B. (Eds.). *Early Childhood Matters: Evidence from the Effective Pre-School and Primary Education Project.* London: Routledge.

Save the Children (2016). “ELM – Emergent Literacy and Math Toolkit.” Washington, DC: Save the Children.

Save the Children (2017). “International Development and Early Learning Assessment (IDELA).” Washington, DC: Save the Children.

Save the Children (2019). “Request for Proposal (RFP) for the Baseline, Midline, Endline Evaluation of the USAID Learn to Read Activity.” Vientiane, Lao PDR: Save the Children.

Sawtell, M.B. (2018). *An Exploration of Practitioner-researcher Collaboration on Randomised Controlled Trials of Complex Interventions.* Ph.D. thesis, University College London.

Schlemmer, G. (2017). “Ethnic Belonging in Laos: A Politico-Historical Perspective.” In Bouté, V. and Pholsena, V. (Ed.). *Changing Lives: New Perspectives on Society, Politics and Culture in Laos.* Singapore: National University of Singapore Press, pp. 251-285.

School-to-School International (2017a). “GraphoGame™ Teacher Training Service: Evaluation Report.” Washington, DC: USAID.

School-to-School International (2017b). “Qysas (Stories): An Arabic Leveled Digital Library for Every Classroom.” Washington, DC: USAID.

School-to-School International (2017c). “Supporting Technology-Based Innovations to Improve Early Grade Reading Outcomes for Students who have Low Vision or are Blind.” Washington, DC: USAID.

Schweinhart, L. (2013). “Long-term Follow-up of a Preschool Experiment.” *Journal of Experimental Criminology 2013*(9), pp. 389–409.

Scott-Little, C., Kagan, S., and Frelow, V. (2006). “Conceptualization of Readiness and the Content of Early Learning Standards: The Intersection of Policy and Research?” *Early Childhood Research Quarterly 21*(2006), pp. 153–173.

Serikari, S. (2019). *Can School Readiness Programs Work? Learnings from Karnataka and Gujarat* [Online]. Available from https://www.centralsquarefoundation.org/school-readiness-programs-karnataka-gujarat/. Accessed 8 August 2021.

Sezen, S. (no date). “The Effects of Early Childhood Education Summer Schools on Syrian and Local Children: Assessment Report Summary.” Istanbul: AÇEV.

Sha, M. and Gabel, T. (Eds). (2020). “The Essential Role of Language in Survey Research.” North Carolina, USA: Research Triangle International.

Shafiq, M., Devercelli, A., and Valerio, A. (2018). “Are There Long-term Benefits from Early Childhood Education in Low- and Middle-income Countries? *Education Policy Analysis Archives, 26*(122), pp. 1-46.

Share, D.L. and Daniels, P.T. (2016) “Aksharas, Alphasyllabaries, Abugidas, Alphabets, and Orthographic Depth: Reflections on Rimzhim, Katz, and Fowler.” *Writing Systems Research, 8*(1), pp. 17-31.

SIL (2019). *SIL Glossary of Linguistic Terms* [Online]. Available from https://glossary.sil.org/term/tone. Accessed 26 May 2019.

Smith, T.M., Cobb, P., Farran, D.C., Cordray, D.S., and Munter, C. (2013). “Evaluating Math Recovery: Assessing the Causal Impact of a Diagnostic Tutoring Program on Student Achievement.” *American Educational Research Journal,* 50(2), pp. 397-428.

Snow, K. (2006) “Measuring School Readiness: Conceptual and Practical Considerations.” *Early Education and Development,* 17(1), pp. 7-41.

Snow, C., Burns, S., and Griffin, P. (1998). “Preventing Reading Difficulties in Young Children.” Washington, DC: National Academies of Sciences.

Spier, E. (2019). Meeting with Jodie Fonseca, 17 April.

St. George, E. (2015). “Overview of Strategic Issues in Basic Education: Findings of Literature Review of Donor Programs in Lao PDR.” Unpublished report. Vientiane, Lao PDR: Plan International.

St. George, E. and Koonbarthao, S. (2018). “Research Study into Teacher Motivation, Parent Engagement, Scalability, and Sustainability of the LEARN Project.” Unpublished report. Vientiane, Lao PDR: Plan International.

Stanovich, K.E. “Toward an Interactive-Compensatory Model of Individual Differences in the Development of Reading Fluency.” *Reading Research Quarterly, 16*(1), pp. 32–71.

Stevens, C., Keller, P., and Tyler, M. (2004). “Language Tonality and Its Effects on the Perception of Contour in Short Spoken and Musical Items.” 8th International Conference on Music Perception & Cognition, Evanston, Illinois, USA, August 3-7, 2004.

Street, B. (1994). (1994) “The New Literacy Studies: Implications for Education and Pedagogy.” *Changing English, 1*(1), pp. 113-126.

Stuart-Fox, M. (2005). “Politics and Reform in the Lao People’s Democratic Republic.” Perth, Australia: Asia Research Center.

Tiwari, S. (2011) “A Preliminary Investigation of Akshara Knowledge in the Malayalam Alphasyllabary: Extension of Nag’s (2007) Study.” *Writing Systems Research, 3*(2), pp. 145-151.

Teale, W.H. and Sulzby, E. (Eds.). (1986). *Emergent Literacy: Writing and Reading.* New Jersey: Ablex Publishing.

Tolmie, A., Muijs, D., and McAteer, E. (2011). *Quantitative Methods in Educational and Social Research using SPSS.* Berkshire, UK: Open University Press.

Tunmer, W. and Hoover, W. (2017). “The Cognitive Foundations of Reading and Its Acquisition: A Conceptual Framework for Reading Intervention.” Christchurch, New Zealand: University of Canterbury.

UNDP (2019a). *About Lao PDR* [Online]. Retrieved from http://www.la.undp.org/content/lao_pdr/en/home/countryinfo.html. Accessed 8 May 2019.

UNDP (2019b). *Lao People’s Democratic Republic: Human Development Indicators* [Online]. Retrieved from http://hdr.undp.org/en/countries/profiles/LAO. Accessed 8 May 2019.

UNDP (2018a). “Human Development Indices and Indicators, 2018 Statistical Update.” New York: UNDP.

UNDP (2018b). “Human Development Indices and Indicators: 2018 Statistical Update Country Explanatory Note, Lao People’s Democratic Republic.” New York: UNDP.

UNESCO (2015). “Education 2030 Incheon Declaration: Towards Inclusive and Equitable Quality Education and Lifelong Learning for All.” Paris: UNESCO.

UNESCO, UNICEF, Brookings Institution and the World Bank (2017). *Overview: Measuring Early Learning Quality and Outcomes* [Online]. Retrieved from: https://www.brookings.edu/research/overview-measuring-early-learning-quality-and-outcomes/. Accessed 9 May 2019.

UNHCR (2016). "The Case for Accelerated Education." Geneva: UNHCR.

UNICEF (2015). "Situational Analysis: Student Learning Outcomes in Primary Education in Lao PDR." Vientiane, Lao PDR: UNICEF.

UNICEF (2017). "UNICEF Annual Report 2016: Mozambique." Maputo: UNICEF.

UNICEF (2018). "UNICEF Annual Report 2017: Mozambique." Maputo: UNICEF.

UNICEF (2019a). "Innovation Case Study: Accelerated School Readiness Programme." New York: UNICEF.

UNICEF (2019b). "A World Ready to Learn." New York: UNICEF.

United Nations (2021). "The Sustainable Development Goals Report 2021." New York: United Nations.

United Nations Lao PDR (2018). *Lao PDR's Eligibility for Graduation from LDC Status Confirmed.* Retrieved from: http://www.la.one.un.org/media-center/news-and-features/375-lao-pdr-s-eligibility-for-graduation-from-ldc-status-confirmed. Accessed 6 May 2018.

USAID (2019). *Early Grade Reading Barometer: Actionable Data for a More Literate World* [Online]. Retrieved from https://earlygradereadingbarometer.org/pages/about_egra. Accessed 6 October 2019.

Vagh, S.B. and Sharma, D. (2018). "Localized Assessments Can Enhance our Understanding of Learning at the Bottom of the Pyramid: Insights from India." In Wagner, D.A, Wolf, S., and Boruch, R.F. (Eds.) *Learning at the Bottom of the Pyramid: Science, Measurement, and Policy in Low-Income Countries.* Paris: UNESCO, pp. 121-136.

Vasanta, D. (2004). “Processing Phonological Information in a Semi-syllabic Script: Developmental Data from Telugu.” *Reading and Writing: An Interdisciplinary Journal, 17,* pp. 59–78.

Wagner, D.A, Wolf, S., and Boruch, R.F. (Eds.) (2018). *Learning at the Bottom of the Pyramid: Science, Measurement, and Policy in Low-Income Countries.* Paris: UNESCO.

Wang, Y. and Kuhl, P.A. (2003). “Evaluating the ‘Critical Period’ Hypothesis: Perceptual Learning of Mandarin Tones in American Adults and American Children at 6, 10 and 14 Years of Age.” 15th International Congress of Phonetic Sciences, Barcelona, Spain.

Wang, Y., Spence, M.M., Jongman, A., and Sereno, J.A. (1999). “Training American Listeners to Perceive Mandarin Tones.” *Journal of the Acoustic Society of America, 106*(6), pp. 3649- 3658.

Wayland, R. and Guion, S.G. (2003). “Perceptual Discrimination of Thai Tones by Naive and Experienced Learners of Thai.” *Applied Psycholinguistics, 24,* pp. 113–129.

Wayland, R.P. and Guion, S.G. (2004). “Training English and Chinese Listeners to Perceive Thai Tones: A Preliminary Report.” *Language Learning 54*(4), pp. 681–712.

Wayland, R.P. and Li, B. (2008). “Effects of Two Training Procedures in Cross-language Perception of Tones.” *Wayland Journal of Phonetics 36*(2008), pp. 250–267.

Wawire, B.A. and Kim, Y.G. (2018). “Cross-Language Transfer of Phonological Awareness and Letter Knowledge: Causal Evidence and Nature of Transfer.” *Scientific Studies of Reading 22(*6), pp. 443–461.

West, J., Denton, K., and Germino-Hausken, E. (2000). “America’s Kindergarteners.” Washington, DC: National Center for Education Statistics.

Wiliam, D. and Thompson, M. (2007). “Integrating Assessment with Learning: What Will It Take to Make it Work?’ In D. Dwyer (Ed.), *The Future of Assessment: Shaping Teaching and Learning.* New York: Lawrence Erlbaum Associates,

Willems, D., Defrancq, B., Colleman, T., and Noël, D. (Eds.) (2003). *Contrastive Analysis in Language: Identifying Linguistic Units of Comparison.* Hampshire, UK: Palgrave Macmillan.

Willms, J.D. (2018). “Educational Prosperity: An Assessment Strategy for Supporting Student Learning in Low-income Countries.” In Wagner, D.A, Wolf, S., and Boruch, R.F. (Eds.) *Learning at the Bottom of the Pyramid: Science, Measurement, and Policy in Low-Income Countries.* Paris: UNESCO, pp. 97-113.

Wijayathilake, A.D.K., Parrila, R., Inoue, T., and Nag, S. “Cognitive Predictors of Word Reading in Sinhala.” In Nag, S. and Perfetti, C. Eds. (2014). *Reading and Writing: Insights from the Alphasyllabaries of South and Southeast Asia. Writing Systems Research, 6*(1), pp. 1-9.

Winskel, H., Radach, R., and Luksaneeyanawin, S. (2009). “Eye Movements when Reading Spaced and Unspaced Thai and English: A Comparison of Thai–English Bilinguals and English Monolinguals.” *Journal of Memory and Language, 61,* pp. 339–351.

Wolf, S., Halpin, P., Yoshikawa, H., Dowd, A. J., Pisani, L., and Borisova, I. (2017). “Measuring School Readiness Globally: Assessing the Construct Validity and Measurement Invariance of the International Development and Early learning Assessment (IDELA) in Ethiopia.” *Early Childhood Research Quarterly, 41,* pp. 21–36.

World Bank (2019). *Data: Lao PDR* [Online]. Washington, DC: World Bank. Retrieved from: https://data.worldbank.org/country/lao-pdr and https://databank.worldbank.org Accessed 11 February 2019.

World Bank (2021). *Evaluations – Early Childhood Development* [Online]. Washington, DC: World Bank. Available from

https://www.worldbank.org/en/programs/sief-trust-fund/brief/evaluations-early-childhood-development. Accessed 8 August 2021.

World Bank (2018a). “Lao PDR Early Childhood Education Project.” Presentation at the LEARN Project Steering Committee Meeting, Vientiane, 5 July 2018.

World Bank (2015). “Drivers of Poverty in Lao PDR.” Washington, DC: World Bank.

World Bank (2016a). “Lao PDR School Autonomy and Accountability SABER Country Report.” Washington, DC: World Bank.

World Bank (2016b). “Reducing Early Grade Drop Out and Low Learning Achievement in Lao PDR: Root-Causes and Possible Interventions.” Washington, DC: World Bank.

World Bank (2018b). “TEGRA Learning Gains Report from Midline to Endline.” Washington, DC: World Bank.

World Bank (2016c). “The Status of Early Childhood Health and Development in Northern Lao PDR: Baseline Results from the Early Childhood Education Project.” Washington, DC: World Bank.

World Bank, Inter-American Development Bank, and UNICEF (2018). “G20 Development Working Group: Investing in Early Childhood Development.” Washington, DC: World Bank.

World Health Organization, UNICEF, and World Bank (2018). “Nurturing Care for Early Childhood Development: A Framework for Helping Children Survive and Thrive to Transform Health and Human Potential.” Geneva: World Health Organization.

Yang, S. (2018). “LEARN Project Monitoring Data Report, June 2016 - November 2018.” Unpublished report. Vientiane, Lao PDR: Plan International.

Yeung, W.H., Chen, K.H., and Werker, J.F. (2013). “When Does Native Language Input Affect Phonetic Perception? The Precocious Case of Lexical Tone.” *Journal of Memory and Language, 68,* pp. 123–139.

Yip, M. (2002) *Tone.* Cambridge, UK: Cambridge University Press.

Zeni, J. (2001). *Ethical Issues in Practitioner Research.* New York: Teachers College Press.

Zsiga, E. and Rattima, N. (2007). “Tone Features, Tone Perception, and Peak Alignment in Thai.” *Language and Speech, 50*(3), pp. 343-383.