

March 2022

## Design Paper 12

# Language of Instruction in Schools in LMICs: A Systematic Review

**Pooja Nakamura, Zelealem Leyew, Adria Molotsky,  
Varsha Ranjit and Kevin Kamto**

## About CEDIL

The Centre of Excellence for Development Impact and Learning (CEDIL) is an academic consortium supported by the UK government through UK Aid. The mission of the Centre is to test innovative methodologies in evaluation and evidence synthesis and promote evidence-informed development. CEDIL-supported projects fall into three programmes of work: evaluating complex interventions, enhancing evidence transferability, and increasing evidence use.

## CEDIL design paper

CEDIL design papers present the protocols of studies supported by the CEDIL programme. They describe the objectives, the methodology and the implementation plan of impact evaluations, evidence syntheses, and methodological papers supported by CEDIL.

## About this design paper

This design paper was submitted to CEDIL by the “Language Transitioning Research Synthesis” Project S.209 team.

Suggested citation: Nakamura, P., Leyew, Z., Molotsky, A., Ranjit, V. and Kamto, K. (2022) ‘Language of Instruction in Schools in LMICs: A Systematic Review’, CEDIL Design Paper 12. Centre of Excellence for Development Impact and Learning (CEDIL), London and Oxford.

This project was funded by the Centre of Excellence for Development Impact and Learning (CEDIL), supported by UK aid from the UK Government. The views expressed in this research project paper do not necessarily reflect the UK Government's official policies or CEDIL.

All content is the sole responsibility of the authors. CEDIL has not commissioned a professional copy-edit of this paper. Any errors and omissions are the sole responsibility of the authors.





# Language of Instruction in Schools in LMICs: A Systematic Review

**Pooja Nakamura,<sup>1</sup> Zelealem Leyew,<sup>2</sup> Adria Molotsky,<sup>1</sup> Varsha Ranjit,<sup>1</sup> and Kevin Kamto<sup>1</sup>**

Submitted to the Coordinating Group of:

<input type="checkbox"/>	Crime and Justice
<input type="checkbox"/>	Education
<input type="checkbox"/>	Disability
<input checked="" type="checkbox"/>	International Development
<input type="checkbox"/>	Nutrition
<input type="checkbox"/>	Social Welfare
<input type="checkbox"/>	Other:

Plans to co-register:

<input checked="" type="checkbox"/>	No		
<input type="checkbox"/>	Yes	<input type="checkbox"/> Cochrane	<input type="checkbox"/> Other
<input type="checkbox"/>	Maybe		

Date Submitted: 30 June 2020

Date Revision Submitted: 21 October 2020

Approval Date:

Publication Date:

<sup>1</sup> American Institutes for Research (AIR).

<sup>2</sup> Addis Ababa University.

---

## BACKGROUND

---

### *The Problem*

Even though more children are in school than ever before, more than 250 million of them are not learning basic literacy and numeracy skills (World Bank, 2018; 2019). Several factors contribute to this state of learning poverty, ranging from the macro (national education and teacher education systems) to the micro (cognitive learning mechanisms). However, one factor central to learning, but often under researched, misunderstood, or overlooked, is the role of language in education. The language the child is taught in is closely linked to successfully acquiring academic and socioemotional skills. Yet, language of instruction (LOI) programmatic and policy choices are made arbitrarily in many countries and are also constantly shifting based on community demand, political realities, and donor policies.

Children across low- and middle-income countries (LMICs) are learning in multilingual contexts. This has wide-ranging social, economic, political, and educational consequences for learning. Reading research makes it clear that children will learn to read only language(s) they understand (NICHD Early Child Care Research Network, 2005; Ouellette, 2006; Hoover & Tunmer, 2020). Even if children learn to decode (sound out) words—with some degree of fluency—reading comprehension will remain unattainable without sufficient oral language comprehension.

A few recent systematic reviews highlight the importance of instruction in the mother tongue (or a language the child speaks and understands well) for quality learning outcomes in LMIC's (Evans & Acosta, 2020; Nag, Vagh, Dulay, & Snowling, 2019). The benefits to mother tongue-based multilingual education programs are multifaceted, including higher likelihood of girls and marginalized communities staying in school (Benson, 2005), increasing educational equity and maintenance of cultural and linguistic diversity (Ball, 2010), allowing parents and communities to participate in the learning process (Nag et al., 2019) as well as long-term cost benefits (Heugh, 2004; 2012). There are also clear cognitive benefits to learning to read in a known or familiar language, as the skills from the first language transfer and facilitate learning to read in a new language (Koda, 2008; Chung, Chen, & Geva, 2019). Furthermore, strong bilingual education models have significant positive effects on non-linguistic functions (Bialystok, 2005) and executive function skills (Bialystok, 2018) that lay a strong foundation for later socioemotional skills, as well as on academic achievement (Collier & Thomas, 2017).

At the same time, there is an ever-increasing demand from communities for education in the national or international postcolonial language (Coleman, 2011). The primary reason for this demand is the link between the postcolonial language and socioeconomic mobility (Azam, Chin, & Prakash, 2010). Other factors that complicate LOI choices include linguistically heterogeneous classrooms, in which there are multiple mother tongues in one school or area (Nakamura, Carson, Davis, Rai, & Todd, 2017; Reddy, 2011); as well as the fact that some mother tongue languages have no scripts, lack teaching and learning materials, have limited trained teachers, or lack political or community will to be implemented as languages in education (Piper, Zuilkowski, & Ong'ele, 2016; Trudell & Piper, 2013).

This leads to a situation in which decision-makers must reconcile *both* the well-documented benefits of mother-tongue instruction along with the quest for socioeconomic mobility

through a postcolonial or international (later acquired) language at earlier grades. Therefore, this systematic review will focus on Language of Instruction (LOI) choices in education programs and policies on student literacy outcomes in multilingual contexts in LMICs. In particular, we ask the question of whether or not mother tongue (or familiar language) instruction impacts reading outcomes, as well as aim to investigate the unanswered question of when to introduce or transition to additional languages of instruction to foster quality bilingual or multilingual reading outcomes.

## **Theory of Change for LOI Policies and Programs on Literacy Outcomes**

### *Theory Underlying Bilingual and Multilingual Literacy Acquisition*

In order to base our theory of change in theory (Brown, 2020), we developed a learning science<sup>3</sup> framework of the *cognitive mechanisms* that underpin literacy learning in bilingual and multilingual learning contexts. This provides a theoretical basis for how we expect LOI policy and program interventions to impact literacy outcomes in LMIC's.

The Cognitive Foundations of Reading and its Acquisition (CFRA) is a model that lays out the cognitive components required for successful reading in *monolingual learners* – and links those components to curriculum effectiveness and reading teacher knowledge and teaching effectiveness (Hoover & Tunmer, 2020). The Peter Effect in teaching reading is rooted in the principle that it is not possible to give what one doesn't have (Applegate & Applegate, 2004; Binks-Cantrell, Washburn, Joshi, & Hougén, 2012). Studies show that reading teacher effectiveness is significantly related to their own reading enthusiasm (Applegate & Applegate, 2004) as well as their own knowledge of the cognitive foundations of reading (Binks-Cantrell et al., 2012).

Learning science theories from various disciplines such as psychology and linguistics reveal that the underlying mechanisms of acquisition of reading skills in *bilingual and multilingual learners* is different from learning to read in monolingual learners in significant and predictable ways. First, learning to read in a second or later acquired language (referred to henceforth as L2/x) is significantly impacted by transfer of reading skills from a first language (L1)<sup>4</sup>. Second, L2/x learning is also significantly impacted by L2/x oral language skills, which are highly variable in L2/x learners compared to monolingual learners. This notion that L2/x reading skills are reliant on a combination of L1 reading skills and L2/x oral language skills is encapsulated in the linguistic interdependence hypothesis, the underlying proficiency hypothesis (Cummins, 1979; 1981) and the transfer facilitation model (TFM) of second language reading (Koda, 2005; 2007; 2008).

Chung, Chen and Geva (2019) provide an updated interactive framework for crosslinguistic transfer in L2/x reading, in which they posit that the relationship between L1 and L2/x reading skills is influenced by cognitive, linguistic, and metalinguistic factors such as language specific

---

<sup>3</sup> Learning science is the field of study that aims to “better understand the cognitive and social processes that result in the most effective learning, and to use this knowledge to redesign classrooms and other learning environments so that people can learn more deeply and more effectively (Sawyer, 2006, p. xi).

<sup>4</sup> We refer to L1 as the first language(s) of the child or any language(s) the child uses and understands with high levels of familiarity and proficiency, oftentimes called “mother tongue” (MT). We refer to L2/x as the second or later acquired language(s) of the child, i.e. any language(s) they are learning and have emerging levels of familiarity and proficiency with.

versus language neutral constructs, L1-L2/x distance, L1-L2/x proficiency and complexity. They extend the model to postulate that transfer is also impacted by socio-cultural factors such as age of beginning acquisition of the L2/x, immigration experience, educational settings, and extent of exposure to the L1 and L2/x.

Indeed, empirical evidence is accumulating for each of these factors. In a meta-analysis of the cognitive and linguistic sub-skills in transfer, Melby-Lervag and Lervag (2011) find that phonological awareness and decoding skills present significant correlations across L1-L2/x; but that these relationships are less (or not) present in oral language comprehension and reading comprehension subskills. Reflecting this need to start with a foundation in the L1 for successful outcomes in the L2/x, Collier and Thomas (2017) present the results from a 32-year longitudinal research study on bilingual education in the United States. The results reveal that it takes an average of six years of high-quality instruction in both the mother tongue (MT)/L1 as well as English/L2/x, with at least 50% of the instruction being conducted in MT/L1 for English learners to perform comparably with their monolingual peers on academic outcomes.

There are increasingly diverse research methods being employed to answer a variety of question related to biliteracy development and LOI transitioning policy and practice questions. For example, studies have recently begun using threshold methodologies to examine “how much” of a particular skill or knowledge is needed to benefit from transfer in biliteracy or bilingualism. In Northern England, De Cat, Gusnanto, and Serratrice (2018) utilize Cox proportional hazard regression models to identify a threshold of “bilingual experience” for early executive functioning skill benefits. In North America in a two-way Spanish-English bilingual immersion program, Feinauer, Hall-Kenyon, and Everson (2017) employ discontinuous change-point regression models to show that the relationship between L2 oral language skills and L2 reading development is not linear. Finally, in two LMIC contexts – India and Ethiopia – Nakamura and colleagues used structural break regression analyses to test whether there is an empirically determinable point of “sufficiency” in the L1 literacy skills to transfer and foster success in literacy skills in the L2/x (Nakamura, de Hoop, & Holla, 2018; Nakamura et al., 2019). The results showed that there was a nonlinearity in the relationship between first and second language reading scores in six languages pairs across these two countries, implying that there may be a point at which children are cognitively and linguistically ready for literacy instruction in L2/x as L1 reading skills reach a point of sufficient maturity for transfer to take place.

Beyond the cognitive underpinnings of biliteracy acquisition, certain aspects of the home environment are also identified as significant predictors of literacy and biliteracy acquisition. According to the Home Literacy Environment (HLE) model, informal language and literacy practices at home, i.e. those not directly related to engaging with print at home are predictive of concept of print and emergent literacy skills; whereas, formal language and literacy practices at home, i.e. those that are explicitly meant to teach children language and literacy skills are predictive of early decoding development (Senechal & LeFevre, 2002; 2014). Cross-country reviews also highlight that parental attitudes towards reading, number of books at home (indirect HLE factors) and literacy-linked activities at home (direct HLE factors) have a significant impact on reading outcomes (Park, 2008). Given the vast mismatches between home and school language in LMICs (Nag et al., 2019), as well as the generally lower rates of adult literacy (and thus parental literacy) in LMICs (Abadzi, 2003), the evidence underscores the additional risk of home literacy and language environments that do not have the resources necessary to support reading development in the language of the school – or any language

(Nag et al., 2019). However, studies also suggest that there is context-specificity in the relative importance of various dimensions of the home language and literacy environment on specific reading component outcomes (Park, 2008; Nag et al., 2019; Friedlander, 2020)

Classroom and teacher factors such as attendance (of both teachers and students), the lack of a safe learning space (Spier et al., 2019), nutritional inputs (Plaut et al., 2017), and availability of print (through digital media or not) are necessary factors for learning – however, they are not sufficient (Snilstveit et al., 2016). Pedagogical inputs (such as structured learning progressions, skill-based learning, teaching and learning at the “right” level) and teacher professional development are emerging as the most effective ingredients for translating access and safe learning spaces into quality learning outcomes (Evans & Popova, 2015; Evans & Acosta, 2020). In fact, Evans and Acosta (2020) underscore mother tongue instruction programs as one of the most effective elements of pedagogical interventions.

Individual differences, such as age and socioeconomic status also are known to moderate the relationship between teaching inputs and language and literacy outcomes. Although it is clear that language learning abilities decline as individuals get older (Flege, Yeni-Komshian, & Liu, 1999), there does not seem to be any conclusive evidence that there is a biologically based point at which a child’s (or individual’s) ability to learn a new language diminishes at a significantly higher rate than prior to such a point (Bialystok & Hakuta, 1999; Bialystok & Miller, 1999; Birdsong & Molis, 2001; Hakuta, Bialystok, & Wiley, 2003). Neurobiological studies reveal that age of acquisition does not alter the underlying brain structure of bilinguals (Frenck-Mestre et al., 2005; Friederici, Steinhauer, & Pfeifer, 2002). However, more intuitively, differential aspects of language learning (such as phonological and grammatic processing) are more susceptible to age-associated declines than others, such as semantic processing (Abutalebi, Cappa, & Perani, 2001; Hernandez, Martinez, & Kohnert, 2000; Weber-Fox & Neville, 1999, 2001). De Cat et al.’s (2009) study showed that SES was a significant correlate of bilingualism’s positive impacts; but found that the threshold effect of bilingual proficiency held even after SES was controlled for. In the United States and Canada, several evaluations of bilingual education program impacts on learning outcomes shows that although SES is a significant correlate of educational attainments, when controlled, bilingual children in bilingual programs outperform bilingual children in monolingual programs (Bialystok, 2018). Researchers in the West stress the critical difference between “bilingual education” (positive, additive connotation) and “education of bilingual children” (negative, subtractive connotation) (Bialystok, 2018). This distinction manifests itself in policies that either embrace bilingualism and multilingualism as a force for global integration versus those that use local language as steppingstones towards a different national or international language, which in turn contributes to motivation to learn and community and family involvement in the education system.

Taken together, these studies help us move towards the development of a middle range theory on multilingual education and biliteracy acquisition. However, it is still unclear how to construct an effective LOI policy, beyond noting that teaching a child in a language they are familiar with is critical for learning. There is little understanding of the mechanism of transfer of skills from one language to another, the “right” timing or skill level at which a child is most likely to benefit from learning a new language, or how to foster quality bilingual/multilingual outcomes after the initial year(s) of mother tongue instruction. As such, the policy question remains at what grade or point to transition students from one medium of instruction to another, and how to develop LOI decisions in LMIC’s that are likely to be most impactful in

improving literacy and biliteracy scores. It is also unknown whether such an LOI model is likely to be the same across different contexts. It is thus critical to consolidate evidence through a rigorous evidence synthesis.

### *Intervention logic framework*

Based on the problem statement above and the theoretical framework of bilingual/biliteracy learning, it is clear that L1 and L2 skills are significantly related in complex and constantly interacting ways that are important for the development of an intervention logic framework. As such, there are many ways that language transition interventions that aim to improve literacy skills may be effective.

In figure 1, we present our logic framework. We begin with the key assumption that the child has access to a learning program. Access can occur in the form of school infrastructure with teachers who do not use any technology, or a blended learning environment within a school or community building where teachers or teaching assistants may use some technology to enhance learning (e.g. the eSchool 360 model implemented by the Impact Network in Zambia), or an online/digital learning program that is – or could be – facilitated by a remote teacher or guide (e.g. Mindspark software), or an entire online/digital learning program that is self-guided or guided by a virtual built-in guide (e.g. the Google Bolo app). Another assumption underlying our logic framework is that teachers are willing and able to learn and change how they teach in line with new curricula, teaching materials, and pedagogies tailored to bilingual or multilingual students and varying language types. The introduction of revised LOI choices cannot be effective or adequately applied within the classroom without a revision in the teaching and learning materials to reflect this change. Further, since not all teachers may be fully fluent in the LOI choices or know how to teach those language(s), some may be required to obtain additional trainings to improve teaching knowledge or be re-assigned to schools where they can teach language(s) they are fluent in and trained to teach reading in.

LOI transition intervention activities can be manifested in many ways. For the purposes of this research, we operationalize LOI transition programs and policies as those that have any one or any combination of the following components:

- 1) An education program or policy that is implemented in a mother tongue (MT) or local language, which will then lead to a transition to (complete change) or addition of (adding as a subject or dual language instruction) teaching of a new language, which may or may not occur during the course of the program. These are important to examine as the skills taught and learned during the course of the program will have significant implications for “readiness” of transfer to the new later acquired language.
- 2) An education program or policy that is implemented in a language that is not the child’s “own” language (i.e. a language the child has enough proficiency to learn in). These programs are important to investigate as they constitute a child “transitioning” out of their own language into an education system in a new language right from the start of education.
- 3) An education program or policy that is implemented in which students transition from one LOI to another or add a new language as part of the medium of instruction during



the course of the program or policy.

In any of these scenarios students' learning acquisition process can be impacted by the language used for teaching – and as such, can have significant implications for effectiveness of learning to read. This is regardless of whether the LOI transition is the key component of the educational program.<sup>5</sup>

Furthermore, the program or policy intervention is most likely to succeed in improving learning outcomes if it has standards, curriculum, and trained teachers (the latter for classroom-based instruction, as opposed to technology-based instruction) that focus on the cognitive foundation models (CFRA) (Hoover & Tunmer, 2020) and/or the interactive models of reading transfer (Chung, Chen, & Geva, 2019). Although several programs may not explicitly have these theoretical frameworks named in their models, this is based on the theoretical premise that a curriculum or a teacher cannot give what they do not have (Applegate & Applegate, 2004; Binks-Cantrell et al., 2012).

These programmatic components (or activities) may improve the quality of teacher (or technology) knowledge and practices, increase child's motivation to learn (as this will maintain teaching at the "right level", Banerjee et al., 2016; Pritchett & Beatty, 2015), and also increase parental and community involvement in the child's education (Benson, 2005). Finally, all these will improve the effectiveness of the LOI decision, leading to impacts on the child's L1 literacy skills (reading or decoding based skills as well as oral language skills) and the child's L2/x literacy skills.

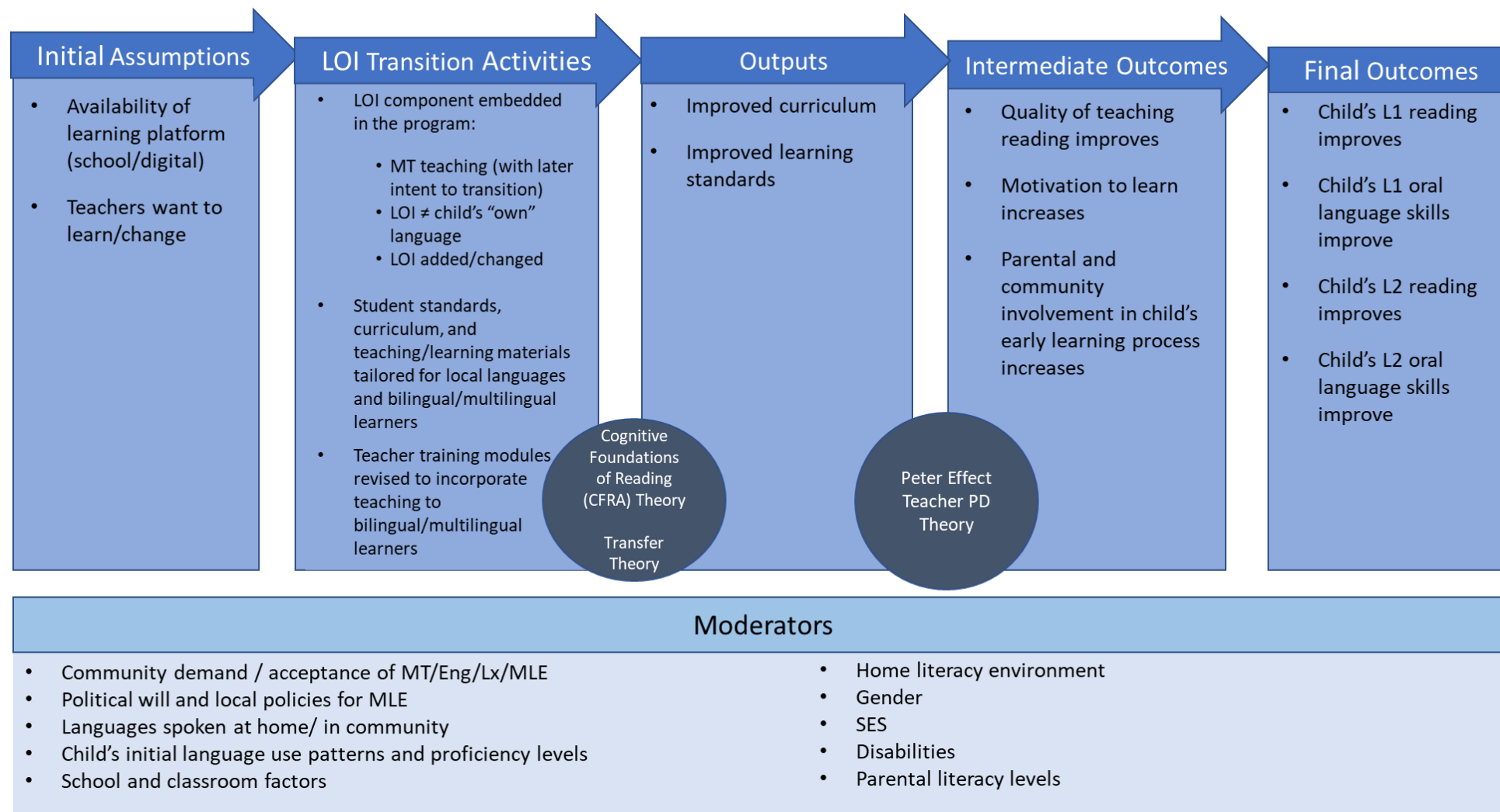
We also examine the role of several possible factors that are likely to moderate the likelihood that the intervention will improve literacy skills, including community demand for the L1 vs. the L2/x, local and national policies supporting the implementation of the program or policy, socioeconomic status, parental literacy/schooling level, language(s) spoken at home, home literacy environment (exposure to print), child's initial language use and proficiency level in language(s) of the school, gender, and disability.

Given that LOI choices touch several aspects of the education system, we aim to synthesize the linkages between the inputs in order to trace how components of LOI programming may impact different sections of the system. For instance, inputs in curricular choices in terms of timing and sequencing of skills in each language would impact standards and curriculum development decisions; whereas teacher training inputs would impact professional development modules, assignment of teachers to schools based on languages they speak versus languages students speak, and urban-rural teacher availability. All practice and policy recommendations will be interpreted within the theory of change, in order to further develop a middle-range theory for LOI decision making in LMICs that is reflective of both the micro psycholinguistic and learning science ingredients in improving learning outcomes as well as the macro sociolinguistic, socioeconomic, and political environment within which LOI policy and practice decisions are being made.

---

<sup>5</sup> Where possible, we will isolate the impact of LOI from impacts of other programming features such as learning materials, teacher professional development, etc. However, the likelihood of finding studies within our search for which we are able to do so is low as LOI policies are rarely, if ever, randomly assigned or assigned in isolation.

**Figure 1. Logic Framework for LOI Policies and Interventions on Literacy Outcomes**



---

## WHY IT IS IMPORTANT TO DO THE REVIEW

---

This systematic review will aim to help decision makers—ministries of education, teacher training institutes, community leaders, interested donors, and implementing and research organizations—understand and effectively use existing evidence related to multilingual education. Most reading programs being funded by large international donors such as USAID and UNICEF include programming in the mother tongue for at least the first few years of primary school (for example, the USAID-funded Creative-implemented Vamos Ler! project in Mozambique or READ II program in Ethiopia, or the USAID-funded RTI-implemented PRIMR program in Kenya). Yet, countries continue to shift policies at the national level (for example, Rwanda switched from a Kinyarwanda LOI policy to an English only policy in 2019 reversing a 2015 MT only policy reform, Edwards, 2019), and regional governments continue to change the years of transition of medium of instruction (e.g. India’s regional push towards regional language education despite community demands for sooner English medium of education, Amaravati, 2019; Gejji, 2019). Indeed, recent international education reports continue to emphasize that “learning poverty” cannot be solved without a better evidence base on how to tackle the role of language of instruction in multilingual LMIC educational settings (World Bank, 2018; 2019).

We will also examine evidence gaps that may hinder efforts to implement successful LOI policies in multilingual LMICs. We will focus particularly on the ministry of education in Ethiopia to guide the implementation of the country’s new three language policy that is part of the education roadmap reform being developed for rollout in the near future. Our conversations with key stakeholders in the Ethiopian education system, including the Ministry of Education team that is tasked with developing the roadmap development suggest there is an urgent need to gather and understand the evidence on how to implement this multilingual LOI policy effectively.

The generalizability of the findings across language types is informed by the framework that all writing systems of the world are divided into four main types (Nag & Perfetti, 2014; Perfetti, 2003): alphabetic, syllabic, alphasyllabic, and morphosyllabic. Although we will examine various local contextual factors that may hinder or facilitate the impact of the program, including linguistic complexities on the dimensions of orthographic depth, orthographic breadth, graphic complexity in both/all languages, we will be able to consolidate the findings for a broader middle-range theory for each of the four writing system types.

This study builds on recent systematic reviews that have shown that mother tongue instruction is critical for learning quality (Nag et al., 2019; Evans & Acosta, 2020); but is unique in that it will be the first to systematically review the evidence on how and when to add – or transition from – one language of instruction to another. Furthermore, the study utilizes a combination of methods, including critical discourse analyses to map policy and practice documents to the evidence generated from the systematic review. Finally, by exploring various psycholinguistic underpinnings of reading and examining a variety of sociolinguistic contexts of learning and drawing from our multi-disciplinary theoretical framework, it helps build middle-range theory on the mechanisms that may explain why certain LOI policies are likely to be more effective than others.

---

## OBJECTIVES

---

To address the evidence gap in making effective LOI decisions, we propose a systematic review<sup>6</sup> of the role of LOI choices in education programs and policies on literacy outcomes in multilingual educational contexts in LMICs. Grounded in the multidisciplinary theory of change described above of what factors link LOI choices and literacy outcomes, we will gather, organize, and synthesize the evidence on the specific role of the three LOI choices described in the ToC (teaching in MT with later transition, teaching in a non-MT language, or teaching in two or more languages at one time) and its impact on literacy and biliteracy outcomes. We will focus our systematic review and meta-analysis only on quantitative and qualitative intervention studies from LMIC's as these have the highest relevance for decision making in multilingual LMIC contexts. We will also only include languages that are relevant and commonly spoken in LMIC's. For example, we will likely include studies that examine Arabic to English transfer, but not Arabic to Swedish transfer. Against this backdrop, we pose the following research questions:

### ***Primary Research Questions:***

1. What are the short- and long-term impacts of **initial LOI** choices on literacy and biliteracy outcomes, and how do they differ across various LMIC contexts?
2. What are the short- and long-term impacts of **LOI transition** on literacy and biliteracy outcomes, and how do they differ across various LMIC contexts?

### ***Secondary Research Question:***

3. What are the overall messages of key donor and stakeholder agencies around LOI choices and learning outcomes? How do those messages link to the existing evidence?

### ***Additional Mapping Questions:***

4. What is the quality of the available evidence on the role of LOI choices on literacy outcomes?
5. What are the evidence gaps about the role of LOI choices in bilingual and multilingual educational contexts in LMICs?

---

## METHODOLOGY

---

In this section we provide detail on the methods that will be employed to answer **our primary research question**.

### ***Criteria for including and excluding studies***

#### *Types of study designs*

The primary research question on the effectiveness of interventions will be addressed using

---

<sup>6</sup> We follow the definition of systematic reviews and evidence syntheses from the Campbell Corporation: “a systematic review is an academic research paper . . . that uses a method called evidence synthesis to look for answers to a pre-defined question.”

quantitative experimental or quasi-experimental as well as qualitative studies that include a programmatic or policy intervention.

Specifically, we will include the following study designs for **quantitative studies**: (1) experimental designs using random assignment to the intervention and (2) quasi-experimental designs with nonrandom assignment (such as regression discontinuity designs, “natural experiments,” and studies in which participants self-selected into the program). Quasi-experimental studies must (1) collect longitudinal data (baseline and end line) or cross-sectional data (end line) from treatment and comparison groups and (2) use propensity score or another type of matching, difference-in-differences estimation, instrumental variables regression, multivariate cross-sectional regression analysis, or other forms of multivariate analysis (such as the Heckman selection model or multivariate ordinary least squares regression analysis). We will include studies with data collected at the individual level to ensure that the study focuses on child-level learning outcomes.

We will include each of the multivariate quasi-experimental methods to maximize the external validity of the systematic review. However, several of the quasi-experimental studies we propose to include may include only OLS regression analysis and, therefore, may not be able to provide unbiased impact estimates. In such cases, we exclude these studies from our meta-analysis. To mitigate concerns about internal validity of some of the included studies, we will conduct a risk of bias assessment and stratify our meta-analysis by identification strategy, where feasible, as in Brody et al. (2015). This stratified meta-analysis will enable us to assess the internal validity of the included studies with a high risk of bias by comparing the impact estimates in those studies with the impact estimates in studies with a low risk of bias (Chinen et al., 2017).

For **qualitative studies**, we will include any intervention studies that utilize the following illustrative methods: 1) case studies, 2) focus group discussions; 3) key informant interviews; and 4) observations of classrooms or community language use. At the abstract screening stage, we intend to include all qualitative studies that examine an intervention, regardless of methodology. Depending on the number of studies that are returned from the abstract screening stage, we will either select only those qualitative studies that are linked to a quantitatively examined intervention, in order to explain why or how that particular intervention may or may not be effective; or select all qualitative intervention studies for full-text review, which will provide a fuller picture of how and why multilingual or mother tongue instruction programs may or may not be effective in LMIC's in general.

#### *Types of participants and settings for both quantitative and qualitative studies:*

We will include studies that focus on interventions that include primary and secondary school aged children in low- and middle-income countries, as defined by the World Bank.<sup>7</sup> We will include studies about the effects of LOI choices regardless of the educational status or skill level of children at the time of the intervention. Only studies conducted between the years of 1995 and 2020 and published in English or Amharic will be considered.

---

<sup>7</sup> We are using the 2014 LMICs definition, which includes Argentina, Hungary, Seychelles, and Venezuela. These countries, however, were categorised as high-income countries in the July 2015 update.

In the case of qualitative studies, we will include intervention studies that have a focus on school-aged children in low- and middle-income countries, as defined by the World Bank.

### *Types of interventions*

The interventions included in this review will be LOI choices made by educational policies and programs that *directly* aim to increase children's literacy in bilingual or multilingual LMIC education contexts. These interventions include programs with one or more of the following components:

- Full early learning programs for mother tongue education or bilingual and multilingual children
- Official language of instruction policy (laws or de facto policy) changes
- Teacher training for mother tongue programs or bilingual or multilingual education programs<sup>8</sup>
- Standards development for mother tongue or bilingual or multilingual education programs
- Technology-based interventions for mother tongue or bilingual or multilingual education programs
- Mother tongue or bilingual teaching and learning materials
- Mother tongue or bilingual books
- Mother tongue or bilingual book clubs, libraries, community reading spaces, mobile book vans etc.
- Assessments used as part of mother tongue or bilingual programming
- Mother tongue or bilingual or multilingual radio or media programming

### *Types of comparison conditions*

Eligible comparison conditions will include no intervention, pipeline, or "business as usual." Where the typical comparison condition of 'no intervention' or 'business as usual' is not selected, eligible comparison conditions will include students before a LOI policy change within a region, students within regions with different or no LOI transition policies within a country, and students in schools who are not in a MT or regional language program. In the case of qualitative studies, a comparison will not be necessary.

### *Types of outputs and outcome measures*

We will include studies that focus on intermediate and/or final outcomes. Below we present definitions for each of the output and outcome variables for the primary research question:

#### **Outputs:**

- 1. Curriculum and Standards:** We define the curriculum as a document or framework that spells out a wide range of content, objectives, methodologies,

---

<sup>8</sup> Bilingual or multilingual programs are defined as any that have more than one language of instruction, have additional language subject classes, or the students are bilingual or multilingual, or students are in mixed-language/linguistically heterogeneous classrooms

resources, assessments, and organizational information regarding what a child is expected to learn. A curriculum should be closely aligned with content standards (what students should know and be able to do at a particular grade) and performance standards (a scale to indicate how, and what percentage of students are performing on the content standards). These standards in turn are of high quality when they are linked to the cognitive foundations of reading acquisition (Hoover & Tunmer, 2020).

### **Intermediate Outcomes:**

- 1. Reading teacher quality:** Based on the Peter Effect studies that a reading teacher cannot pass on what they themselves do not have, we define two main dimensions in reading teacher quality: a) the enthusiasm to teach reading as indicated by the habits and practices of the teacher (Applegate & Applegate, 2004) as well as b) the knowledge of the sub-constructs required for learning to reading (Binks-Cantrell et al., 2012).
- 2. Student motivation:** Defined as a child's motivation to attend school as well as to want to learn to read or engage with language, print, or stories. This will be measured in student's reading behaviours at home or in school and in the community, and attitudes towards reading.
- 3. Parental and community involvement:** Studies have found that there is a significant association between student's being taught in home and community languages and the parental involvement in the education system (Benson, 2007). We define this outcome as the frequency and quality of interactions between the parents and/or community members and teachers, as well as the amount of time spent by teachers involved with student's learning at home (helping with homework, learning *from* the student, supporting the student with their learning, asking questions about school, etc.)

### **Final Outcomes:**

We provide operational definitions for our final outcomes from a range of research on reading development that looks at reading across language and orthographic types as well as across L1/L2 learning status (Verhoeven & Perfetti, 2017; Koda & Zehler, 2008; UNESCO et al., 2019). Each of these skills will be considered for both L1 and L2/x:

- 1. Emergent literacy sub-skills:** Emergent literacy skills or "concept of print (or print awareness)", refer to the ability to understand the conventions and functions of print in one's own language(s), including the ability to distinguish pictures from letters or words, identify the beginning and ending of stories, etc.
- 2. Oral language skills:** These skills include a range of receptive and expressive oral language abilities, such as the ability to follow simple spoken instructions, identify the meaning of spoken word, explain the meaning of a word in their own words, retell a short story, understand explicit and implicit comprehension questions from a short listening passage, etc.
- 3. Metalinguistic awareness:** Metalinguistic awareness is the awareness of the functionally useful units of speech and print and the relationship between the two.

There are several sub-components of metalinguistic awareness, but we will focus on **phonological awareness**, or the awareness of speech units in any language and **morphological awareness**, or the awareness of morphological units in any language.

4. **Sound-symbol correspondences:** Oftentimes called letter naming or letter knowledge, sound-symbol correspondence skills refer to the ability to see a single printed letter, akshara, or character and be able to sound out the symbol.
5. **Decoding:** The ability to see a printed word or cluster of symbols and sound out the word or cluster of symbols. There are several paths that learners take to acquire this skill, but our primary concern will be on whether or not students are able to reach the entire phonological representation of the printed word, regardless of which path they take to achieve the skill.
6. **Oral reading fluency:** The ability to sound out a short passage or story with accuracy, speed, and prosody.
7. **Reading comprehension:** The ability to comprehend both explicit and implicit information presented in single or multiple phrases or sentences of text.

If data are not available for each of these subskills separately, based on the CFRA (Hoover & Tunmer, 2020), we will create composite scores for the emergent literacy and oral language measures (#1-3 above), and for the decoding scores (#4-6 above), and for the reading comprehension scores (#7). Understanding that there are usually only 4 questions on many EGRA scores for reading comprehension, we will consider either removing these items or merging them with the decoding scores for reliability, if necessary.

We will include only literacy outcomes even if the study looks at scores on other subjects.

Outcome measures will not be considered to filter qualitative studies, which will serve to address the secondary research question.

### ***Search Strategy***

We developed a search strategy in consultation with an information specialist. Our search strategy will enable us to identify relevant published and unpublished literature by focusing on relevant academic and institutional databases, citation tracking, and snowballing of references. We identified the following literature searches.

#### ***Electronic Sources***

Comprehensive database searches will include the following paid-access and free-access electronic databases:

#### ***Paid-Access Databases***

1. ERIC (Education Resources Information Center)
2. Education Source
3. EdWorkingPapers



4. EdArXiv
5. Registry of Efficacy and Effectiveness Studies (REES)
6. PsycINFO
7. JSTOR Arts & Sciences I-X Collections
8. JSTOR Business III Collections
9. SAGE Publications
10. ScienceDirect
11. Springer Science+Business Media
12. Taylor & Francis
13. Wiley
14. WorldCat

#### *Open-Access Databases*

15. Campbell Collaboration
16. Cochrane Library
17. 3ie Impact Evaluation Repository
18. Directory of Open Access Journals (DOAJ)
19. Directory of Open Access Books (DOAB)
20. Development Experience Clearinghouse (DEC)
21. Institute of Development Studies (eldis)
22. Education International

#### *Grey Literature*

Grey literature searches will include a review of institutional websites as well as generic searches such as via Google and Google Scholar. In addition to the following institutional websites, key papers will be examined in search of other relevant institutional sources.

#### *Institutional Websites or Research Funders*

1. The UK Department for International Development (DFID) (<https://www.gov.uk/government/organisations/departments-for-international-development>)
2. The US Agency for International Development (USAID) (<https://www.usaid.gov/>)
3. The Joint Libraries of the World Bank and International Monetary Fund (JOLIS) (<https://library.worldbankimflib.org/>)
4. J-PAL (<http://www.povertyactionlab.org>)
5. UNESCO (<https://en.unesco.org/>)
6. UNICEF (<https://www.unicef.org/>)
7. UNICEF Office of Research (<https://www.unicef-irc.org/>)
8. UNHCR (<https://www.unhcr.org/>)
9. Population Council (<https://www.popcouncil.org/>)
10. World Vision (<https://www.worldvision.org/>)
11. Save the Children (<https://www.savethechildren.org>)
12. Plan International (<https://plan-international.org/>)
13. Organization of American States (OAS) (<http://www.oas.org/en/>)
14. Global Partnership for Education (GPE) (<https://www.globalpartnership.org/>)

Forward and backward snowballing of the references of key papers will provide additional

studies for review that may not be found in database searches. Citation searches will be conducted in Google Scholar, Scopus, and Web of Science. This set of key papers will include:

### Key Papers

15. Anchor key papers identified by the authors. These papers are listed below.
16. Key papers identified by external funder (CEDIL).
17. Studies that pass the inclusion criteria.
18. Additional key papers identified from institutional websites.

Table 1 presents the list of key papers identified by authors and reviewers to be used in citation searches.

**Table 1. Key Papers**

Study
Piper, B., Zuilkowski, S., Ong'ele, S. (2016). Implementing mother tongue instruction in the real world: Results from a medium-scale randomized controlled trial in Kenya. <i>Comparative Education Review</i> , 60, 776-807.
Brunette, T., Piper, B., Jordan, R., King, S., & Nabacwa, R. (2019). The impact of mother tongue reading instruction in twelve Ugandan languages and the role of language complexity, socioeconomic factors, and program implementation. <i>Comparative Education Review</i> , 63, 591-612.
Banerji, R. Berry, J., & Shotland, M. (2017). The impact of maternal literacy and participation programs: Evidence from a randomized evaluation in India. <i>American Economic Journal: Applied Economic</i> , 9(4), 303-337.
Chicoine, L. (2019). Schooling with learning: The effect of free primary education and mother tongue instruction reforms in Ethiopia. <i>Economics of Education Review</i> , 69, 94-107.
Laitin, D. D., Ramachandran, R., Walter, S. (2019). The legacy of colonial language policies and their impact on student learning: Evidence from an experimental program in Cameroon. <i>Economic Development and Cultural Change</i> , 68 (1), 239-272.
Shin, J., Sailors, M., McClung, N., Hoffman, J. V., Pearson, D., Kaambankadzanja, D., Mwale, L. (2019). Access to local books: the effects of Read Malawi from a children's rights perspective.

### Search Strings

We developed search strings in collaboration with an information specialist. Each search string is adapted to fit the syntax of the database host in order to utilize Boolean operators (AND/OR), wildcards, truncation, and other database search features. The search strings are designed to return studies that include at least one keyword in the following four themes:

1. Participants: preschool, elementary school, pre-primary, kindergarten, primary school, early childhood
2. Literacy: reading, literacy, mother tongue, language of instruction, medium of instruction, cross-language transfer, language transition, reading transfer, multilingual education, bilingual education
3. Setting: low-income, middle-income, third world, developing, underdeveloped, LMIC, global south, Africa, Asia, LAC, Southeast Asia

In order to capture both quantitative and qualitative literature relevant to the primary research questions, we will not include search strings for study design, comparison condition, and outcome measures. Using these criteria in the search strategy would exclude relevant qualitative studies in addition to quantitative and mixed-methods studies that omit this information from the title and abstract. We will assess the inclusion criteria for study design, comparison condition, and outcomes during the screening of the studies.

### ***Screening***

Screening will take place in two phases: first on the basis of titles and abstracts and then on the basis of full texts.

#### ***Screening Phase 1***

After our initial search is completed, we will conduct a manual abstract review process. Each abstract will be reviewed independently by two trained reviewers. We will conduct the following steps in the abstract screening phase (based on Polanin et al. 2018).

1. Creation of the abstract screening tool. The tool has been created taking into consideration the following issues:
  - a. Questions are objective and “single-barrelled”
  - b. Questions include yes/no/don’t know structures, with dropdown options as needed
  - c. Detailed key for dropdown options (such as what is meant by “literacy outcome” or “school-age”)
  - d. Questions are ordered hierarchically to ensure that studies that are not meeting any one inclusion criteria are immediately removed from further screening
2. Review of the tool by the lead of our quantitative analysis
3. Train reviewers on the use of the tool
4. Conduct pilot tests of using the screener with the review team all screening the same 15-20 abstracts until consensus is reached on the tool.
5. Periodic check-ins to enhance intellectual buy-in as well as to ensure any disagreements are resolved

#### ***Screening Phase 2***

In the second phase, we will review the full text of all studies that pass Phase 1 screening. Multiple reviewers will independently identify and confirm the following information for each study:

#### ***Quantitative studies***

- Target population
- Intervention type
- Comparison group
- Quantitative methodology
- Data source
- Outcome measures

### *Qualitative studies*

- Target population
- Intervention type
- Qualitative methodology
- Gender norms
- Barriers to intervention effectiveness
- Facilitators of intervention effectiveness

Studies pass Phase 2 screening if the information pulled from each study passes the inclusion criteria listed in the previous section.

---

## **METHODOLOGY TO SYNTHESIZE THE LITERATURE**

---

### ***Quantitative Studies***

#### *Data Extraction*

Two team members with expertise in quantitative research will work independently to extract information from each quantitative study included in the review. Both team members will use a data extraction form and fill the data from the extraction form in a table. We will resolve disagreements through discussion.

#### *Risk of Bias Assessment*

We will determine the rigor of the quantitative studies using an adaptation of a set of criteria, to assess risk of bias in experimental and quasi-experimental studies (Hombrados & Waddington, 2012). While the risk of bias assessment is very labor intensive, the number of quantitative studies we expect to require this assessment is low. We will assess the risk of the following biases:

1. Selection bias and confounding, based on quality of identification strategy to determine causal effects and assessment of equivalence across the beneficiaries and nonbeneficiaries.
2. Performance bias, based on the extent of spillovers to students in comparison groups and contamination of the control or comparison group.
3. Outcome and analysis reporting biases, including:
  - a. The use of potentially endogenous control variables
  - b. Failure to report nonsignificant results
  - c. Other unusual methods of analysis
4. Other biases, including:
  - a. Motivation and courtesy biases (Hawthorn effect and John Henry effect)
  - b. Coherence of results
  - c. Retrospective baseline data collection
  - d. Differential attrition bias
  - e. Other biases, such as strong researcher involvement in the implementation of the intervention.

#### *Measures of Treatment Effects*

In accordance with, Chinen et al., 2017, we will extract information from each quantitative study to estimate the standardized effect sizes (for continuous variables) or odds ratios (for

binary variables) across studies. We will also calculate standard errors and confidence intervals where feasible.

We will report effect sizes as Hedge's  $g$  and will adjust effect sizes reported as Cohen's  $d$  to Hedge's  $g$ . We will use Hedges'  $g$  effect sizes (sample-size-corrected standardized mean differences (SMDs)) for continuous outcome variables, which measure the effect size in units of standard deviation of the outcome variable, and for binary outcomes, we will calculate odds ratios (ORs).

The SMD using Cohen's  $d$  is calculated by dividing the mean difference with the pooled standard deviation by applying the formula in Equation 1:

$$(1) \text{ SMD} = \frac{Y_t - Y_c}{S_p}$$

where  $Y_t$  refers to the outcome for the treatment group,  $Y_c$  refers to the outcome for the comparison group, and  $S_p$  refers to the pooled standard deviation.

The pooled standard deviation  $S_p$  will be calculated by applying either Equations 2 or 3:

$$(2) S_p = \frac{\sqrt{((SD_y)^2 * (nt + nc - 2)) - (\frac{\beta^2 * (nt * nc)}{nt + nc})}}{nt + nc}$$

$$(3) S_p = \frac{\sqrt{(nt - 1) * st^2 + (nc - 1) * sc^2}}{nt + nc - 2}$$

where  $SD_y$  refers to the standard deviation for the point estimate from the regression,  $n_t$  refers to the sample size for the treatment group,  $n_c$  refers to the sample size for the comparison group, and  $\beta$  refers to the point estimate. We will use Equation 2 for regression studies with a continuous dependent variable and Equation 3 when the study provides information about the standard deviation for the treatment group and the comparison group.

To transform Cohen's  $d$  into Hedge's  $g$ , we will use the small sample correction for the SMD by applying on formula 4:

$$(4) \text{ SMD}_{\text{corrected}} = \text{SMD}_{\text{uncorrected}} * \left(1 - \frac{3}{4 * (nt + nc - 2) - 1}\right)$$

Lastly, Equation 5 will estimate the standard error of the SMD:

$$(5) \text{ SE} = \sqrt{\frac{nt + nc}{nc * nt} + \frac{\text{SMD}^2}{2 * (nc + nt)}}$$

For studies using linear probability models, we will assume linearity in the estimation of standardized effects as in Brody et al. (2015). For example, if we observe a mean baseline value for the comparison group of 0.097 and an effect size of 5.1 percentage points, then we will assume that the follow-up value for the treatment group would be  $0.097 + 0.051 = 0.148$ , and we will assume that the follow-up value for the comparison group will be 0.097. To correct the standard errors for studies where the outcome variable is clustered at a level above the individual or household, we will use adjusted standard errors by applying corrections to the

standard errors and confidence intervals using the variance inflation factor (Higgins & Green, 2011):

$$(7) \ SE_{corrected} = SE_{uncorrected} \times \sqrt{1 + (m - 1) \times ICC}$$

where  $m$  is the number of observations per cluster, and ICC is the intracluster correlation coefficient. We will estimate the ICC for each of the relevant outcome measures for our sample of included quantitative studies and for which we are able to access the data on the outcome measures.

When we are unable to retrieve the missing data, we will impute effect sizes and associated standard errors based on the t or F statistic or p-values. We will use David Wilson's practical meta-analysis effect-size calculator to conduct such imputations. Where sample sizes for the treatment and the comparison group are not reported, we will assume equal sample sizes across the groups.

### *Methods for Handling Dependent Effect Sizes*

Where studies report more than one effect size on the basis of different statistical methods, we will follow the procedure as laid out in Chinen et al., 2017 and will select the effect size with the lowest risk of bias. Where studies report more than one effect size based on the same individuals, we will employ the robust variance estimation techniques to adjust for effect size dependency (Hedges, Tipton, & Johnson, 2010). When studies present multiple impact estimates for different variables measuring the same construct, we will use a sample-size weighted average to measure a "synthetic effect size." In cases where more than one study uses the same data set (e.g., national level EGRA data) to measure a literacy outcome, we will use the effect size from the study with the lowest risk of bias. If the risk of bias is the same, we will estimate an average effect size through inverse-weighted random effect meta-analysis. In cases where one study measures the same outcome at different points in time, we will extract the effect size by relying on the outcome measure that was measured closest to the time period of the measurement in other studies included in the same meta-analysis. In cases where studies include more than one treatment arm, we will include the effect size from the treatment arm with the lowest risk of bias. If the risk of bias is the same, we will use the effect size from the treatment arm that is most similar to the other programs included in the meta-analysis.

### **Meta-Analysis**

We will pool the results of the quantitative studies that focus on the same outcome variables and same intervention types using meta-analysis. In other words, we will conduct multiple different meta-analyses based on intervention type and outcome variable (described above which reading outcomes will be pooled if necessary). We will examine the heterogeneity of the effect sizes for each outcome across studies and use meta-regression to model the variation in effect size and will use forest plot visualization (Borenstein et al., 2009). We will use Stata to conduct the meta-analysis.

For the meta-analysis, we will include only studies with an emphasis on LOI choice that use one of the following designs: (1) experimental designs using random assignment to the intervention and (2) quasi-experimental designs with nonrandom assignment (such as

regression discontinuity designs, “natural experiments,” and studies in which participants self-select into the program).

Where possible, we will perform sensitivity analysis for potential moderators:

- Risk of bias status for each risk of bias category;
- Study design (randomized controlled trials versus quasi-experimental studies);
- Gender
- SES
- Parental literacy levels
- Alignment of language spoken at home with LOI
- Geography

We will use random-effects meta-analysis because the average effect of LOI choice is likely to differ across contexts due to differences in program design and contextual characteristics. We will supplement our random-effects meta-analysis with network meta-analysis to enable indirect comparisons of two treatments that have a common comparator.

We will also use stratified meta-analysis according to contextual and methodological moderator variables to investigate factors explaining heterogeneity. We will use two contextual moderating variables: (1) type of orthography and (2) grade.

### *Missing Data*

In cases where it is not feasible to estimate the effect size because of missing data, we will contact the researchers to request the missing information in order to calculate the effect sizes. If authors do not respond or do not provide sufficient information to calculate the effect size, we will not include the study in the meta-analysis. Even so, the study and its findings will still be discussed within our narrative write-up.

### ***Treatment of Qualitative Research***

Every study that is selected for full-text review, will undergo a full quality appraisal.

### *Quality Appraisal*

We will assess the quality of the qualitative studies using the nine-item Critical Appraisal Skills Programme Qualitative Research Checklist (Critical Appraisal Skills Programme, 2013), judging the adequacy of stated aims, the data collection methods, the analysis, the ethical considerations, and the conclusions drawn. For each item, one trained researcher on our team will independently fill out the appraisal to determine whether the study had adequately met the item and gave “yes,” “no,” or “can’t tell” response. Afterwards, they will then come together to discuss their responses to each item until they reach consensus. We will rate studies that score 0–2 “no” or “can’t tell” responses as low risk of bias, studies that score 3–5 “no” or “can’t tell” responses as medium risk of bias, and studies that score 6–9 “no” or “can’t tell” responses as high risk of bias.

**Table 2. Quality Appraisal Criteria**

Criteria	Coding
<i>Screening Question:</i> Is there a clear statement of study aims of the research?	Yes / Can't tell / No
<i>Screening Question:</i> Is a qualitative methodology appropriate?	Yes / Can't tell / No
Is it worth continuing?	Yes / Can't tell / No
Was the research design appropriate to address the aims of the research?	Yes / Can't tell / No
Was the recruitment strategy appropriate to address the aims of the research?	Yes / Can't tell / No
Were the data collected in a way that addressed the research question?	Yes / Can't tell / No
Has the relationship between researcher and participants been adequately considered?	Yes / Can't tell / No
Have ethical issues been taken into consideration?	Yes / Can't tell / No
Was the data analysis sufficiently rigorous?	Yes / Can't tell / No
Is there a clear statement of findings?	Yes / Can't tell / No
Is the research valuable?	Yes / Can't tell / No

After full-text review, we will conduct a thematic synthesis of the qualitative study findings. Each study's main findings will be coded to encapsulate the content of each findings (e.g. "the teacher joins Portuguese and the local language to help the student understand", "we use [the local language] only to pull the student from where he is and understand the subject"). These statements are then categorized into higher order themes (such as "use of local language in post-colonial language classes"). We will then extract implications for better understanding why or how multilingual education choices work in various contexts.

### *Critical Discourse Analysis*

In order to answer our secondary research question focused on the overall policy messages being conveyed by the key donor and stakeholder agencies around LOI choices, we will conduct a qualitative critical discourse analysis (CDA), through Systemic Functional Linguistics (SFL) analysis (including linguistic and textual analysis) and ideological analysis (Martin & Rose, 2007; van Dijk, 2006) on 2-3 key donor and MOE documents on LOI policies or strategies.

By using CDA, we will analyse the discourse in these LOI documents, and the LOI-related discourse that has been included and excluded by these donor institutions and understand the dynamics between development assistance network, donor institutions and, if feasible, nation-states. In addition, we will conduct CDA based on discursive psychology to understand the positions and power relations between these donor institutions and nation-states. The discursive nature helps explore the identity of stakeholders, their positions and their narration



within a social context (Hajer, 1995; Hewitt, 2009). Using the discursive tradition will help reveal how donors justify certain models of mother-tongue based education and how donors persuade nation-states on LOI policies. Together, the SFL and Ideological analyses approaches will help us identify how stakeholders present and/or consume a “shared set of ideas” about language and education; for example, how is evidence discussed and applied, how are groups conducting mother tongue education versus post-colonial medium of instruction discussed and described, etc.

SFL provides a framework through which we will analyse linguistic features used: how often do they use particular words? What affect do those words carry? Do the linguistic features carry explicit or implicit power structures in terms of LOI choices and consequences? What type of evidence is used to justify certain models of LOI? How is the concept of ‘mother tongue’ described by the donor institution versus the MoE and on what narrative ideology is this concept based on? Given that labelling is one of the very first steps in realizing how language is treated in any language policy or planning decisions (Kaplan & Baldauf, 1997), and terminological variance abounds in LOI policies worldwide, this will also shed considerable light on the links between how donors, other decision makers, and the layman consume and use information and evidence on language decisions in education.

These approaches will allow us to select, analyse, and interpret the essential messages embedded in donor policies and discourse around language issues in education policies. The primary theoretical premise of CDA is that language choices are shaped by, and shape, society and that language policy is influenced largely by historical, social, political and ideological environments in nation-states (Fairclough, 2009). As such, this approach will help various stakeholders – MOEs, teacher training institutes, implementing organizations - understand and to objectively evaluate donors’ monetary prioritizations; as well as for donors to review, and if necessary, revise or adapt their existing and future education policies to incorporate a more evidence-based view of language issues.

Once the CDA has been conducted, we will disseminate the findings through webinars, blog posts, social media, and individual communications to key stakeholders, including MOE personnel and key members from donor organizations and implementing partner organizations. The results will be shared in such a way that they are tailored to each stakeholder through policy briefs, brochures, and easy to access materials. The dissemination will be closely followed by co-interpretation and planning meetings or workshops. In these meetings, the stakeholders will be encouraged to closely analyse existing documents and discourse around the role of language in education – including assessing whether the discourse itself is absent – and then work together to determining changes that could likely steer the course of LOI policy in a direction more in line with the evidence, as well as highlight gaps that are still hindering decision makers’ abilities to move forward with more effective LOI policies and practices.

---

## **QUALITY ASSURANCE**

---

AIR has a rigorous system for quality assurance (QA), which ensures that we deliver high-quality products. Dr. Thomas de Hoop serves as our QA reviewer on this project; he has over 13 years of experience designing and managing large-scale systematic reviews in education in LMICs. He will be responsible for providing support to the team and for ensuring the quality

of the research materials produced. Specifically, Dr. de Hoop will support the project team by providing inputs on the design of evidence synthesis' protocol, refining work and analysis plans, reviewing analysis and results, making suggestions about interpretation of further analysis, and reviewing final research products including blogs, policy briefs, and the final synthesis report. Dr. de Hoop will sign off on all drafts and final drafts after quality assurance of each deliverable.

### **COVID-19 Risk Mitigation**

AIR has extensive experience facing and overcoming challenges associated with managing and conducting research. While the desk-based nature of the evidence synthesis greatly reduces the risks to the project, especially considering the current COVID-19 pandemic, we are aware that a few potential risks remain. For instance, there is the possibility that one or more of our project team members may be directly or indirectly affected by COVID-19, subsequently, reducing their ability to work on the project. However, we developed the project team such that every position has backup support from another staff member, so the project is unlikely to be delayed due to COVID-19 affecting any one team member. If we experience the unfortunate circumstance of multiple team members being affected by COVID-19 at the same time, this will require us to pull from our internal staffing networks to employ additional staff to support the project team with the research activities. Again, we do not foresee this resulting in any delays to the progress of the project.

---

## REFERENCES

---

- Abadzi, H. (2003). *Improving adult literacy outcomes: Lessons from cognitive research for developing countries*. Operations Evaluation Department. The World Bank: Washington, D. C.
- Abutalebi, J., Cappa, S. E., & Perani, D. (2001). The bilingual brain as revealed by functional neuroimaging. *Bilingualism: Language and Cognition*, 4, 179–190.
- Amaravati, I. (2019, September 11). English to be medium of instruction in Andhra government schools. *India TV News*. Retrieved from <https://www.indiatvnews.com/news/india/english-medium-of-instruction-andhra-government-schools-548970>
- Applegate, A. J., & Applegate, M. D. (2004). The Peter Effect: Reading habits and attitudes of teacher candidates. *The reading Teacher*, 57, 554-563.
- Azam, M., Chin, A., & Prakash, N. (2010). *The returns of English language skills in India* (IZA Discussion Paper 4802). Retrieved from <http://ftp.iza.org/dp4802.pdf>
- Ball, J. (2010). *Educational quality for children from diverse language backgrounds: mother tongue-based bilingual or multilingual education in the early years*. Presentation to UNESCO International Symposium: Translation and Cultural Mediation, Paris.
- Banerjee, A., Banerji, R., Berry, J., Duflo, E., Kannan, H., Mukherji, S., Shotland, M., Walton, M. (2016). Mainstreaming an effective intervention: Evidence from randomized evaluations of “teaching at the right level” in India. *NBER Working Paper Series No. 22746*.
- Banerji, R. Berry, J., & Shotland, M. (2017). The impact of maternal literacy and participation programs: Evidence from a randomized evaluation in India. *American Economic Journal: Applied Economic*, 9(4), 303-337.
- Benson, C. & Hakuta K., B (2005). *Girls, educational equity and mother tongue-based teaching*. Bangkok, Thailand: UNESCO Asia and Pacific Regional Bureau for Education.
- Bialystok, E., & Miller, B. (1999). The problem of age in second-language acquisition: Influences from language, structure, and task. *Bilingualism: Language and Cognition*, 2(2), 127–145. <https://doi.org/10.1017/S1366728999000231>
- Bialystok, E. (1991). Bilingual education for young children: Review of the effects and consequences. <https://doi.org/10.1080/13670050.2016.1203859>
- Benson, P. (2007). Autonomy in language teaching and learning. *Language Teaching*, 40(1), 21–40. <https://doi.org/10.1017/S0261444806003958>
- Bialystok, E. (2005). Consequences of bilingualism for cognitive development. In J. Kroll & A. M. B. de Groot (Eds.), *Handbook of Bilingualism: Psycholinguistic Approaches* (pp. 417–432). NY, New York: Oxford University Press.
- Bialystok, E. (2018). Bilingual education for young children: Review of the effects and consequences. *International Journal of Bilingual Education and Bilingualism*.
- Binks-Cantrell, E., Washburn, E. K., Joshi, M. R., & Hougen, M. (2012). Peter Effect in the preparation of reading teachers. *Scientific Studies of Reading*, 16 (6), 526-536.
- Birdsong, D., & Molis, M., (2001). *On the evidence for maturational constraints in second-language acquisition*. *Journal of Memory and Language*, 44, 235–249.

- Borenstein, M., Hedges, L.V., Higgins, J.P.T., and Rothstein, H. (2009). *Introduction to Meta-Analysis*. Chichester: Wiley.
- Brody, C., de Hoop, T., Vojtkova, M., Warnock, R., Dunbar, M., Murthy, P., et al. (2015). *Economic self-help group programs for improving women's empowerment: A systematic review*. Oslo, Norway: The Campbell Collaboration.
- Brown, A. N. (2020). *Using social science theories to design and evaluate development programs*. CEDIL Programme Lecture. Retrieved from: <https://cedilprogramme.org/wp-content/uploads/2020/05/Social-science-theory-for-design-and-evaluation.pdf>
- Brunette, T., Piper, B., Jordan, R., King, S., & Nabacwa, R. (2019). The impact of mother tongue reading instruction in twelve Ugandan languages and the role of language complexity, socioeconomic factors, and program implementation. *Comparative Education Review*, 63, 591-612.
- Chicoine, L. (2019). Schooling with learning: The effect of free primary education and mother tongue instruction reforms in Ethiopia. *Economics of Education Review*, 69, 94-107.
- Chinen, M., de Hoop, T., Alcázar, L., Balarin, M., & Sennett, J. (2017). Vocational and business training to improve women's labour market outcomes in low- and middle-income countries. Campbell Collaboration.
- Chung, S. C., Chen, X., & Geva, E. (2019). Deconstructing and reconstructing cross-language transfer. *Journal of Neurolinguistics*, 50, 149-161.
- Coleman, H. (2011). Developing countries and the English language. In H. Coleman (ed), *Dreams and realities: Developing countries and the English language* (pp. 1-14). London: The British Council.
- Collier, V. P., & Thomas, W. P. (2017). Validating the power of bilingual schooling: Thirty-two years of large-scale longitudinal research. *Annual Review of Applied Linguistics*, 37, 203-217.
- Critical Appraisal Skills Programme (CASP) (2013). Collaboration for Qualitative Methodologies. Available from <http://www.casp-uk.net/wp-content/uploads/2011/11/CASP-Qualitative-Research-Checklist-31.05.13.pdf>.
- Cummins, J. (1979). Linguistic interdependence and the educational development of bilingual children. *Review of Educational Research*, 49(2), 222-251.
- Cummins, J. (1981). The role of primary language development in promoting educational success for language minority students. In *Schooling and language minority students: A theoretical framework*. Sacramento, CA: California State Department of Education, 3-49. doi:10.13140/2.1.1334.9449
- De Cat, C., Gusnanto, A., Serratrice, L. (2018). Identifying a threshold for the executive function advantage in bilingual children. *Studies in Second Language Acquisition*, 40 (1), 119-151.
- Edwards, S. (2019). Rwanda's education U-turn sparks debate on mother tongue instruction. Devex. Retrieved from: <https://www.devex.com/news/rwanda-s-education-u-turn-sparks-debate-on-mother-tongue-instruction-96129>
- Evans, D. K., & Acosta, M. A. (2020). Education in Africa: What are we learning . *Center for Global Development Working Paper* .
- Evans, D., & Popova, A. (2015). What Really Works to Improve Learning in Developing Countries? An Analysis of Divergent Findings in Systematic Reviews (SSRN Scholarly Paper ID 2571082). Social Science Research Network. <https://papers.ssrn.com/abstract=2571082>

- Fairclough, N. (2009). *Language and globalization*. New York: Routledge.
- Feinauer, E, Hall-Kenyon, K, Everson, K. (2017). Rethinking the linguistic threshold hypothesis: Modelling the linguistic threshold among young Spanish–English Bilinguals. *Bilingualism: Language and Cognition*, 20, 886–902.
- Flege, J. E., Yeni-Komshian, G. H., & Liu, S. (1999). Age Constraints on Second-Language Acquisition. *Journal of Memory and Language*, 41(1), 78–104.  
<https://doi.org/10.1006/jmla.1999.2638>
- Frenck-Mestre, C. (2016). Eye-movement recording as a tool for studying syntactic processing in a second language: A review of methodologies and experimental findings: *Second Language Research*. <https://doi.org/10.1191/0267658305sr257oa>
- Friederici, A. D., Steinhauer, K., & Pfeifer, E. (2002). Brain signatures of artificial language processing: Evidence challenging the critical period hypothesis. *Proceedings of the National Academy of Sciences*, 99(1), 529–534.  
<https://doi.org/10.1073/pnas.012611199>
- Friedlander, E. W. (2020). The home literacy environment in rural Rwanda and its relationship to early grade reading. *Scientific Studies of Reading*, 24, 123-140.
- Gejji, A. (2019, September 7). Karnataka: Scrap English medium in govt schools, writers urge govt. *Times of India*. Retrieved from  
[https://timesofindia.indiatimes.com/city/bengaluru/karnataka-scrap-english-medium-in-govt-schools-writers-urge-govt/articleshow/71028126.cms?utm\\_source=contentofinterest&utm\\_medium=text&utm\\_campaign=cppst](https://timesofindia.indiatimes.com/city/bengaluru/karnataka-scrap-english-medium-in-govt-schools-writers-urge-govt/articleshow/71028126.cms?utm_source=contentofinterest&utm_medium=text&utm_campaign=cppst)
- Hajer, M. (1995). *The politics of environmental discourse: Ecological modernization and the policy process*. Clarendon Press: Oxford.
- Hakuta, K., Bialystok, K., & Wiley, E. (2003). Critical evidence: A test of the critical period hypothesis for second language acquisition. *Psychological Science*, 14, 31–38.
- Hedges, L. V., Tipton, E., & Johnson, M. C. (2010). Robust variance estimation in meta-regression with dependent effect size estimates. *Research Synthesis Methods*, 1(1), 39–65. <https://doi.org/10.1002/jrsm.5>
- Hernandez, A. E., Martinez, A., Kohnert, K. (2000). In search of the language switch: An fMRI study of picture naming in Spanish-English bilinguals. *Brain and Language*, 73(3), 421–431.
- Heugh, K. (2004). *Is Multilingual Education Really More Expensive? In: Making Multilingual Education a Reality for All*, ed. by Joachim F. Pfafe, pp 212-233.
- Heugh, K. (2012). Cost implications of the provision of mother-tongue and strong bilingual models of education in Africa. In A. Ouane & C. Glanz (Eds.), *Optimising learning, education and publishing in Africa: The language factor*. (pp. 255–289). Hamburg, Germany: UNESCO.
- Hewitt, S. (2009). *Discourse analysis and public policy research*. Center for Rural Economy Discussion Paper Series No. 24. Center for Rural Economy, New Castle. Newcastle, UK.
- Higgins, J. P. T. & Green, S., eds. (2011). *Cochrane Handbook for Systematic Reviews of Interventions*. Version 5.1.0 (updated March 2011). The Cochrane Collaboration Available at <http://www.cochrane-handbook.org/>.
- Hombrados, J. G., and Waddington, H., (2012). *Internal Validity in Social Experiments and Quasi-Experiments: An Assessment Tool for Reviewers*. Mimeo: 3ie.
- Hoover, W. A., & Tunmer, W. E. (2020). *The Cognitive Foundations of Reading and Its Acquisition: A Framework with Applications Connecting Teaching and Learning*. Springer.



- Introduction to Meta-Analysis | Wiley. (n.d.). Wiley.Com. Retrieved October 16, 2020, from <https://www.wiley.com/en-us/Introduction+to+Meta+Analysis-p-9780470057247>
- Kaplan, R. B. & Baldauf, R. B., Jr. (1997). *Language Planning: From practice to theory*. Clevedon, UK: Multilingual Matters.
- Koda, K. (2005). *Insights into second language reading*. Cambridge, UK: Cambridge University Press.
- Koda, K. (2007). Cross-linguistic constraints on second-language reading development. In K. Koda (Ed), *Reading and language learning* (pp. 1–44). Malden, MA: Blackwell Publishing.
- Koda, K. (2008). Impacts of prior literacy experience on second language learning to read. In K. Koda & A. M. Zehler (Eds.), *Learning to read across languages: Cross-linguistic relationships in first- and second language literacy development* (pp. 68–96). New York, NY: Routledge.
- Koda, K., & Zehler, A. (Eds.). (2008). *Learning to read across languages: Crosslinguistic relationships in first- and second-language literacy development*. New York, NY: Routledge.
- Laitin, D. D., Ramachandran, R., Walter, S. (2019). The legacy of colonial language policies and their impact on student learning: Evidence from an experimental program in Cameroon. *Economic Development and Cultural Change*, 68 (1), 239-272.
- Martin, J. R., & Rose, D. (2007). *Working with discourse: Meaning beyond the clause*, 2<sup>nd</sup> Edition. New York: Continuum.
- Melby-Lervag, M. & Lervag, A. (2011). Cross-linguistic transfer of oral language, decoding, phonological awareness and reading comprehension: a meta-analysis of the correlational evidence. *Journal of Research in Reading*, 34(1): 114-135.
- Nag, S., & Perfetti, C. A. (2014). Reading and writing: Insights from alphasyllabaries of South and Southeast Asia. *Writing Systems Research*, 6, 1-9.
- Nag, S., Vagh, S. B., Dulay, K. M., & Snowling, M. J. (2019). Home literacy, school language, and children literacy attainments: A systematic review of evidence from low- and middle-income countries. *The Review of Education*, 7, 91-150.
- Nakamura, P., Carson, K., Davis, D., Rai, N. & Todd, A. (2017). Language Mapping Study in Mozambique: Final Report. Prepared for USAID.
- Nakamura, P.R., de Hoop, T., & Holla, C.U. (2018). Language and the learning crisis: Evidence of transfer threshold mechanisms in multilingual reading in South India, *The Journal of Development Studies*, 55(11): 2287-2305.
- Nakamura, P. R., Bonilla, J., Mekonnen, B., Tefera, D., Gebrekidan, Z, & Turner, K. (2019). READ M&E: Language Transition Study Report. American Institutes for Research, Washington, D. C. Retrieved from: <https://www.air.org/sites/default/files/READ-ME-Lang-Transition-Rpt-508.pdf>
- NICHD Early Child Care Research Network. (2005). Pathways to reading: The role of oral language in the transition to reading. *Developmental Psychology*, 41, 428-442.
- Ouellette, G. P. (2006). What's meaning got to do with it: The role of vocabulary in word reading and reading comprehension. *Journal of Educational Psychology*, 98 (3), 554-566.
- Park, H. (2008). Home literacy environments and children's reading performance: A comparative study of 25 countries. *Education Research Evaluation*, 14, 489-505.

- Perfetti, C. A. (2003). The universal grammar of reading. *Scientific Studies of Reading*, 7, 3-24.
- Piper, B., Zuilkowski, S., Ong'ele, S. (2016). Implementing mother tongue instruction in the real world: Results from a medium-scale randomized controlled trial in Kenya. *Comparative Education Review*, 60, 776-807.
- Plaut, D., Thomas, M., Hill, T., Worthington, J., Fernandes, M., & Burnett, N. (2017). Getting to Education Outcomes: Reviewing Evidence from Health and Education Interventions. In D. A. P. Bundy, N. de Silva, S. Horton, D. T. Jamison, & G. C. Patton (Eds.), *Child and Adolescent Health and Development* (3rd ed.). The International Bank for Reconstruction and Development / The World Bank. <http://www.ncbi.nlm.nih.gov/books/NBK525274/>
- Pritchett, L., & Beatty, A. (2015). Slow down, you're going too fast: Matching curricula to student skills. *International Journal of Educational Development*, 40, 276-288.
- Reddy, P. (2011). *Biliteracy development in middle-school students from the slums of India* (Unpublished doctoral dissertation). Carnegie Mellon University, Pittsburgh, PA.
- Sawyer, R. K. (2006). Preface. In R. K. Sawyer (ed.), *The Cambridge handbook of learning Sciences* (pp. xi– 1). Cambridge, UK: Cambridge University Press.
- Sénéchal, M., & Le Fevre, J. (2002). Parental involvement in the development of children's reading skill: A five year longitudinal study. *Child Development*, 73, 445-460.
- Sénéchal, M., & Le Fevre, J. (2014). Continuity and change in the home literacy environment as predictors of growth in vocabulary and reading. *Child Development*, 85 (4), 1552–1568.
- Shin, J., Sailors, M., McClung, N., Hoffman, J. V., Pearson, D., Kaambankadzanja, D., Mwale, L. (2019). Access to local books: the effects of Read Malawi from a children's rights perspective
- Snilstveit, B., Stevenson, J., Menon, R., Phillips, D., Gallagher, E., Geleen, M., Maxwell Stamp, Jobse, H., Schmidt, T., Jimenez, E., (2016). The impact of education programmes on learning and school participation in low- and middle-income countries (2016th ed.). International Initiative for Impact Evaluation (3ie). <https://doi.org/10.23846/SRS007>
- Spier, E., Leenknicht, F., Carson, K., Bichay, K., & Faria, A.-M. (2019). Tipping the scales: Overcoming obstacles to support school readiness for all in low- and middle-income countries. *Early Years*, 39(3), 229–242. <https://doi.org/10.1080/09575146.2019.1576031>
- UNESCO UIS, USAID, USAID, DFID, ACER, and Bill and Melinda Gates Foundation et al., (2019). Global Proficiency Framework for Reading and Mathematics: Grades 2 to 6. Retrieved from: <https://www.edu-links.org/resources/global-proficiency-framework-reading-and-mathematics>
- Trudell, B., & Piper, B. (2013). Whatever the law says: language policy implementation and early-grade literacy achievement in Kenya. *Current Issues in Language Planning*, 15: 4-21.
- Hoover, W. A., & Tunmer, W. E. (2020). *The cognitive foundations of reading and its acquisition*. Literacy Studies, 20. Springer Nature: Switzerland.
- Van Dijk, T. A. (2006). Ideology and discourse analysis. *Journal of Political Ideologies*, 11 (2), 115-140.

- Verhoeven, L., & Perfetti, C. (Eds.). (2017). Learning to read across languages and writing systems. Cambridge, UK: Cambridge University Press.
- Weber-Fox, C., & Neville, H. J. (2001). Sensitive periods differentiate processing of open- and closed-class words: An ERP study of bilinguals. *Journal of Speech, Language, and Hearing Research*, 44, 1338–1353.
- World Bank. (2018). *The World development report: Learning to realize education's promise*. Washington, DC © World Bank. Retrieved from: <https://openknowledge.worldbank.org/bitstream/handle/10986/28340/9781464810961.pdf>
- World Bank. (2019). *Ending learning poverty: What will it take?*. World Bank, Washington, D.C. © World Bank. Retrieved from: <https://openknowledge.worldbank.org/handle/10986/32553>

---

## REVIEW AUTHORS

---

### Lead review author:

The lead author is the person who develops and coordinates the review team, discusses and assigns roles for individual members of the review team, liaises with the editorial base, and takes responsibility for the ongoing updates of the review.

<b>Name:</b>	Pooja Nakamura
<b>Title:</b>	Principal Investigator
<b>Affiliation:</b>	American Institutes for Research
<b>Address:</b>	1000 Thomas Jefferson Street, NW
<b>City, State, Province or County:</b>	Washington, DC
<b>Postal Code:</b>	20007
<b>Country:</b>	USA
<b>Phone:</b>	(202) 403-6716
<b>Email:</b>	<a href="mailto:pnakamura@air.org">pnakamura@air.org</a>

### Co-author(s):

<b>Name:</b>	Zealelem Leyew
<b>Title:</b>	Principal Investigator
<b>Affiliation:</b>	Addis Ababa University
<b>Address:</b>	Department of Linguistics Addis Ababa University P.O. Box 150125
<b>City, State, Province or County:</b>	Addis Ababa
<b>Postal Code:</b>	N/A
<b>Country:</b>	Ethiopia



Phone:	+254-011-660-1235
Email:	<a href="mailto:Zleyew4@gmail.com">Zleyew4@gmail.com</a>

<b>Name:</b>	Adria Molotsky
Title:	Economic Researcher
Affiliation:	American Institutes for Research
Address:	1000 Thomas Jefferson Street, NW
City, State, Province or County:	Washington, DC
Postal Code:	20007
Country:	USA
Phone:	(202) 403-5371
Email:	<a href="mailto:amolotsky@air.org">amolotsky@air.org</a>

<b>Name:</b>	Varsha Ranjit
Title:	Research Associate
Affiliation:	American Institutes for Research
Address:	1000 Thomas Jefferson Street, NW
City, State, Province or County:	Washington, DC
Postal Code:	20007
Country:	USA
Phone:	(202) 403-5546
Email:	<a href="mailto:vranjit@air.org">vranjit@air.org</a>

<b>Name:</b>	Kevin Kamto
Title:	Research Associate
Affiliation:	American Institutes for Research
Address:	1000 Thomas Jefferson Street, NW
City, State, Province or County:	Washington, DC
Postal Code:	20007
Country:	USA
Phone:	(202) 403-5023
Email:	<a href="mailto:kkamto@air.org">kkamto@air.org</a>

---

## ROLES AND RESPONSIBILITIES

---

- Content: Pooja Nakamura, Zelealem Leyew
  - Systematic review methods: Pooja Nakamura, Adria Molotsky, Varsha Ranjit, Kevin Kamto
  - Statistical analysis: Adria Molotsky, Kevin Kamto
  - Information retrieval: Pooja Nakamura, Zelealem Leyew, Adria Molotsky, Varsha Ranjit, Kevin Kamto
  - Quality review: Thomas de Hoop
- 

## APPENDIX 1: EXAMPLE SEARCH STRATEGY

---

Query
"Read*" OR "Literac*" OR "Mother Tongue*" OR "First Language*" OR "Second Language*" OR "third language*" OR "language of instruction" OR "medium of instruction*" OR "cross-language transfer" OR "cross-linguistic transfer" OR "language transition*" OR "L1-L2 transfer" OR "reading transfer" OR "cross-linguistic" OR "English language learners" OR "ELL*" OR "MTB-MLE" OR "multilingual education" OR "dual language*" OR "ESL" OR "English as a second language" OR "English as an additional language" OR "EAL" OR "Language Policy" OR "Bilingual Education" OR "postcolonial language" OR "colonial language" OR "English medium" OR "minority language*" OR "majority language*" OR "home language*" OR "language of the playground*" OR "community language*" OR "language of wider communication*" OR "language minority" OR "limited English proficiency" OR "Mother tongue based multilingual education" OR "dual language immersion" OR "translanguag*" OR "Language Transition"
"Preschool" OR "pre-school" OR "elementary school" OR "pre-primary" OR "kindergarten" OR "pre-k" OR "prek" OR "primary grade*" OR "primary school*" OR "early childhood"
"Low and middle income countries" OR "LMIC*" OR "develop* country" OR "global south" OR "Africa" OR "Asia" OR "LAC" OR "Southeast Asia" OR "Latin America and the Caribbean" OR "Latin America"

---

## APPENDIX 2: PILOT PHASE 1 SCREENING

---

If any of the preliminary exclusion criteria can be checked, the study is excluded and screening is stopped. For each of the criteria, do not exclude a study if there is not enough information in the title and abstract to exclude with certainty.

1. ☐ **Exclude** if date of publication is before 1995.
2. ☐ **Exclude** if language is not English or Amharic.
3. ☐ **Exclude** if setting is a high-income country.
4. ☐ **Exclude** if there is no reference to literacy, reading, or LOI.

Studies that remain are then assessed for their relevance to the primary or secondary research questions. Studies may be relevant to more than one of the research questions, for example, if they have a mixed-methods design or explore the secondary research question.

If a study is quantitative primary evidence, use the exclusion criteria that immediately follow. After the study is either included or excluded based on these criteria, proceed to screen the study on the basis of the second research question. If a study passes any of the three sets of inclusion criteria, it will be included in Phase 2 for full text screening.

5. ☐ Quantitative primary study (research question 1)

- 5a. ☐ **Exclude** if the study does not focus on evaluating an intervention.
- 5b. ☐ **Exclude** if the study does not use longitudinal or cross-sectional data collected at the individual level.
- 5c. ☐ **Exclude** if the study design is not experimental or if the study is quasi-experimental without propensity score or another type of matching, difference-in-difference estimation, instrumental variable regression, multivariate cross-sectional regression analysis, or other forms of multivariate analysis (such as the Heckman selection model or multivariate ordinary least squares regression analysis).
- 5d. ☐ **Exclude** if the study does not include a valid comparison condition, such as “no intervention,” “pipeline,” or “business as usual.”
- 5e. ☐ **Exclude** if the study does not include any of the following outcome measures.
- 5f. ☐ **Include** for research question 1 if the study passes all of the previous criteria.

- 6. ☐ Qualitative primary study (research question 2)

- 6a. ☐ **Exclude** if the study is not tied to specific LOI policy interventions.
- 6b. ☐ **Exclude** if the study does not report a qualitative methodology, such as interviews or focus groups.
- 6c. ☐ **Exclude** if the study does not include students.
- 6d. ☐ **Include** for research question 2 if the study passes all of the previous criteria.

- 7. ☐ **Include** for Phase 2 if any of 5f, 6d, or 7d are checked.

---

## SOURCES OF SUPPORT

---

This review is supported UK Aid through the Centre of Excellence for Development Impact and Learning (CEDIL).

---

## DECLARATIONS OF INTEREST

---

None of the proposed authors have developed impact studies that focus on the LOI transition policies.

---

## PRELIMINARY TIMEFRAME

---

Deliverable	Date to Submit
Title Registration	June 2020
Protocol	July 2020

Deliverable	Date to Submit
Draft review	May 2021
Final review	August 2021

---

## PLANS FOR UPDATING THE REVIEW

---

The authors will assess updating the review every three years if funding becomes available.

---

## ACKNOWLEDGMENTS

---

Our sincere appreciation goes to the Centre of Excellence for Development Impact and Learning (CEDIL) for the financial and technical support provided. In addition, we would like to thank Dr. Thomas de Hoop for his guidance and technical oversight of this review.

---

## AUTHOR DECLARATION

---

### Authors' responsibilities

By completing this form, you accept responsibility for preparing, maintaining and updating the review in accordance with Campbell Collaboration policy. The Campbell Collaboration will provide as much support as possible to assist with the preparation of the review.

A draft review must be submitted to the relevant Coordinating Group within two years of protocol publication. If drafts are not submitted before the agreed deadlines, or if we are unable to contact you for an extended period, the relevant Coordinating Group has the right to de-register the title or transfer the title to alternative authors. The Coordinating Group also has the right to de-register or transfer the title if it does not meet the standards of the Coordinating Group and/or the Campbell Collaboration.

You accept responsibility for maintaining the review in light of new evidence, comments and criticisms, and other developments, and updating the review at least once every five years, or, if requested, transferring responsibility for maintaining the review to others as agreed with the Coordinating Group.

### Publication in the Campbell Library

The support of the Coordinating Group in preparing your review is conditional upon your agreement to publish the protocol, finished review, and subsequent updates in the Campbell Library. The Campbell Collaboration places no restrictions on publication of the findings of a Campbell systematic review in a more abbreviated form as a journal article either before or after the publication of the monograph version in *Campbell Systematic Reviews*. Some journals, however, have restrictions that preclude publication of findings that have been, or will be, reported elsewhere and authors considering publication in such a journal should be aware of possible conflict with publication of the monograph version in *Campbell Systematic Reviews*. Publication in a journal after publication or in press status in *Campbell Systematic Reviews* should acknowledge the Campbell version and include a citation to it. Note that

systematic reviews published in *Campbell Systematic Reviews* and co-registered with the Cochrane Collaboration may have additional requirements or restrictions for co-publication. Review authors accept responsibility for meeting any co-publication requirements.

**I understand the commitment required to undertake a Campbell review, and agree to publish in the Campbell Library. Signed on behalf of the authors:**

**Form completed by: Pooja Nakamura**

**Date: 30 June 2020**



[www.CEDILprogramme.org](http://www.CEDILprogramme.org)