

**Learning in Multilingual Contexts:
Language Policies, Cross-linguistic Transfer, and Reading Interventions**

Pierre de Galbert

Dr. Felipe Barrera-Osorio

Dr. Sarah Dryden-Peterson

Dr. James Kim

Dr. Dana McCoy

A Thesis Presented to the Faculty of the Graduate School of Education of Harvard
University in Partial Fulfillment of the Requirements for the Degree of Doctor of
Education

2019

© 2019

Pierre de Galbert

All rights reserved

Dedication

For Inbal, Daniela, and Luca.

This work would not have been possible without your support, grounding, and love.

In memory of Sylvia Acana,

*who dedicated her life to improving children learning outcomes in Uganda, and
addressed me as Dr. Pierre, nearly a decade too early.*

“Multilingualism is the oxygen of cultures. I hope we choose oxygen.”

Ngũgĩ wa Thiong’o, Cambridge MA, April 17, 2018

Acknowledgements

To the children of Oshikuku, their parents, and my colleagues at Nuuyoma Senior Secondary School, thank you. It is during my years as a teacher that I discovered what learning in a multilingual context entailed, where I first engaged in passionate conversations about whether teaching in Oshiwambo, English, or Afrikaans would best serve my students. You will find no data from or reference to Namibia in this dissertation, but I cannot overstate the influence these relationships had on the scholar I have become and aspire to be. Nda pandula.

I want to thank my committee for guiding and supporting me through writing this dissertation: Felipe Barrera-Osorio, Sarah Dryden-Peterson, Dana McCoy, and Jimmy Kim. The four of you brought a range of perspective from different disciplines that ultimately made my work so much better.

Felipe, I am so appreciative of the time you gave me over the last six years, time spent working with data, revising manuscripts, or discussing how to best support students in your courses. When I wandered through HGSE in search of complementary disciplinary approaches, methods, and epistemologies, you were supportive even when it defied rational choice theory, at least when it comes to efficiently meeting doctoral milestones.

Sarah, I feel privileged to have worked with and learned from you over the last ten years. The rigor with which you approach your work as a teacher, scholar, and

advisor is only matched by the warmth you show to all of us around you. I learned so much from your feedback on my qualitative work during the qualifying paper. My dissertation work is much stronger thanks to your pushing me to think more about framing my studies and interpreting ambiguous findings.

Dana, I am really thankful you joined my committee. I anticipated the high quality of support you provided on my methods from having taken a class with you. Your thoughtful and detailed feedback on methods as well as writing made my analyses and framing much better.

Jimmy, I feel lucky that you invited me to serve as your TF a few years ago, which started this professional relationship. I admire the attention you bring to preparing your teaching, and the purposefulness with which you approach your research to improve children's reading outcomes. Through your feedback on my papers, you always reminded me to think about the bigger picture and larger impact of our research.

In addition to my committee members, many faculty members at HGSE helped shape my journey. First, I must acknowledge Catherine Snow and Paola Uccelli. Through taking their classes and joining their research labs, I both learned a tremendous amount about language development, and was able to receive critical feedback on my second dissertation paper. As a faculty for courses I served as a teaching fellow, or more informal mentoring, I learned a lot from Andrew Ho, Fernando Reimers, Vicky Jacobs, and Dan Koretz. I am also thankful to Eileen McGowan who served on my committee for the Qualifying Paper and always cheered me on.

Staff at HGSE made this university more than a workplace for me. In particular, I want to thank the office of student affairs (OSA): Tracie Jones for her social justice work and constantly pushing us to make this place welcoming to all, and Kevin Boehm for his good humor and making sure student-parents feel seen; the doctoral programs office: Clara Lau and Jennifer Schroeder De Vries did the thankless work that helped to keep me pushing in this process. Last, but not least, the Gutman librarians are a treasure of this institution. Marcella Flaherty, Carla Lillvik, and Carol Kentner each supported me in small ways – researching literature, citations in APA, finding the right maps and copyrights – and all the while providing a warm and supportive space on the second floor with each smile.

The doctoral student experience is a journey that only those who go through it (and their partners) can really appreciate. Thank you, Ben Piper, for serving as a model of hard work and humility both as my Teaching Fellow all those years ago and as a scholar and practitioner now. For taking part in my various writing groups along the years, thanks to Jeraul Mackey, Sarah Bruhn, Shireen Al-Adeimi, Wenjuan Qin, Anne Lamb, and Nell O'Donnell. The Mowana research group led by Sarah Dryden-Peterson brought more than an intellectual home. In particular, thank you Celia Reddick for the numerous chats about language and East Africa; Lizzie Adelman for your passion and being vulnerable about how hard research and parenting is; Vidur Chopra for your laugh and positive attitude, support in every step of this journey, and treats from the field; Zuhra Faizi for your smile and the courage you bring to the difficult work you do in a challenging context, and feeding my family such good food. Thank you Francisco Lagos, Andy de

Barros, and Mariam Dhabi. Thank you, Sonia Alves, for the laughs, lunches, and spending all those hours trying to write during our S&P writing blocks. Thank you, Preeya Mbekeani, for these six years forging a friendship built around laughs and cries caused by problem sets, births, BBQs, birthday parties, and so much more.

Funding for my dissertation came from several sources enabling me to dedicate time to research. I am thankful to have received the Harvard University Presidential Scholarship for the first five years of my doctorate. I also benefitted from the HGSE Dean's Summer Fellowship in 2015, 2016, and 2018, which allowed me to make progress toward my qualifying paper and my two empirical dissertation papers. The Additional Insights for the 21st Century 2016-2017 Fellowship from Global TIES for Children at New York University provided funding toward the third paper of this dissertation. Equally important, the fellowship included support and feedback from Andres Molano and Kate Schwartz. Finally, the Comparative International Education Society (CIES) awarded me with a travel merit grant to attend the 2018 conference and participate in the dissertation mentoring workshop where Frances Kvietok and Professor Desmond Odugu provided helpful feedback on my work.

My most intimate thanks must go to my family. My parents were my first teachers, and their difficult conversations about where I should go to school – moving between bilingual and monolingual schools – planted the seeds for my appreciation of languages and curiosity about their role in education. My mother studied and wrote about language in her field – psychology – and, consciously or subconsciously, added fuel to my interest of the role language plays in our life. My father supported our bilingual

upbringing, and observing him oscillating between inclusion and exclusion from conversations with my American family speaking what he calls a *langue de barbare* served as a model for me in the years I lived immersed in languages I understood with various levels of proficiency. My brother Edouard shares my passion for languages and his command of German or Arabic, as well as his attempts to learn many other languages, was always a source of admiration. My brother William does not share our passion for learning languages, but his love for me has always manifested itself through complimenting me and unashamedly bragging about me, to my embarrassment, to his friends, and setting unrealistically high expectations for me. My aunts Brinton and Cathy have provided continuous support in Boston through love, meals, knowledge of navigating higher education, and childcare. Shoshana took me in as a son long before this doctoral journey started and maintained her mother-in-lawing through a careful balance of pushing me along and worrying about where the next adventure might take her grandchildren.

Finally, the three most precious thanks. Inbal, thank you for everything. Your love and support in all aspects of my life is so important it seems futile to limit it to the ways in which you made this work possible. Watching you grow into such an amazing teacher over the last 5 years may be the least obvious way in which you've helped me. The countless hours you dedicate to preparing for teaching and your dedication to searching for ways to make sure each one of your student is seen and valued in the classroom, regardless of the language they speak at home or the country they come from has been an inspiration to do better. I learned from you the technical ways in which you approach that

– WIDA, SIOP, diverse children’s books, etc. – and more importantly the love you bring to your students as a teacher that I can only try to emulate. Daniela, thank you for your love and being such a curious and caring child. Your smile and huge heart are a gift. Luca, thank you for your love, laughter, and ever surprising sense of humor. You burst into our lives in my first year of the program, and you have kept going faster and higher in all ways since then.

Table of Contents

Dedication	i
Acknowledgements.....	ii
Table of Contents	viii
Abstract	xi
List of Acronyms	xiii
Introduction.....	1
Paper 1: Ideology and Pedagogy of designing and implementing Language Policies in Education: An Analytical Framework applied to Kenya and Uganda.....	12
Introduction.....	12
1.1 Nations, States, and Language in Education.....	15
1.2 Language Planning and Orientations	23
1.3 Multilingual approaches in schools	28
2. Framework to analyze education in multilingual contexts	32
2.1. Sociolinguistic context.....	33
2.2. Community engagement and support	38
2.3. Orthographies.....	39
2.4. Availability and Representativeness of Curriculum and Materials	42
2.5. Teacher Recruitment, Training and Deployment.....	44
2.6. Assessment and examination	47
3. Education Language Policies in Kenya and Uganda	49
3.1 Education systems and official language policy.....	50
3.2 Sociolinguistic context and community support.....	53
3.3 Supporting the implementation of the language policies.....	58
Conclusion	63
Paper 2: Language transfer theory and its policy implications: Exploring interdependence between Bantu languages and English in Uganda.....	66
Policy Environment in Uganda and sub-Saharan Africa	67
Language, Literacy and Cross-linguistic transfer	69

Research Questions.....	73
Sample.....	74
Measures	76
Analytic Plan.....	78
Results.....	80
Student performance on literacy tasks	80
Cross-linguistic transfer in decoding	83
Cross-linguistic transfer in reading comprehension	85
Discussion.....	89
Limitations and Future Research	93
Conclusion	95
Paper 3: Literacy Gains for All? Impact Variation of an Early Grade Reading Program in Kenya	97
1. Introduction.....	97
2. Background and Context.....	99
2.1 Literacy in the Early Grades	99
2.2 Program variation.....	101
2.3 Teacher characteristics.....	102
2.4 School contextual factors.....	104
3. Methods.....	106
3.1 PRIMR intervention and study design.....	106
3.2 Analytical Sample and baseline balance.....	108
3.3 Measures	109
3.4 Analytical Approach	110
4. Findings.....	112
4.1 Average Treatment Effect.....	112
4.2 Heterogeneity in Gains	114
4.3 Teacher-level Moderation.....	115
4.4 School-level Moderation.....	116
5. Discussion and Conclusion.....	118
Teacher Characteristics.....	119
School characteristics.....	120
Limitations	124

Conclusion	124
Conclusion	136
Appendices.....	144
Appendix A – Additional Figures and Tables for Paper 1.....	145
Appendix B – Note on the qualitative data for Paper 1	149
Appendix C – Additional Figures and Tables for Paper 2.....	156
Appendix D - Methodological Approach to Address Clustering of Students for Paper 3	161
Appendix E - Imputation of missing data for Paper 3	168
Appendix F - Additional Tables and Figures for Paper 3.....	173
Bibliography	183

Abstract

The proportion of children across the world who have access to basic education through formal schooling has never been as high as it is today. Surveys of literacy and numeracy skills in low- and middle-income countries, however, regularly highlight low levels of learning, even for young people with several years of formal schooling. Among the structural issues faced by school systems in low- and middle-income countries experiencing rapid increases in enrollment, linguistic heterogeneity of communities and countries stands out as it poses both a threat and opportunity to quality instruction.

This dissertation is organized in three papers and aims to explore: (i) ways in which language in education policies are designed, implemented, and understood; (ii) whether cross-linguistic transfer theories apply in a new context; and (iii) heterogeneity across teachers and schools of the impact in a bilingual literacy intervention.

The first paper describes the multi-faceted factors that policymakers take into account when designing language in education policies, and presents a framework for analyzing these policies at the community, regional or state-level. The second paper uses structural equation modeling to test cross-linguistic transfer in a sample of second-grade students in Uganda. Findings suggest children in second grade have developed common decoding skills across two languages, and those starting to learn in their first language can transfer reading comprehension skills to the second. The third paper in this dissertation explores the impact of a literacy intervention in Kiswahili and English in urban public and private nonformal schools in Kenya. The study exploits the random assignment of schools to an intervention to identify whether the impact of an intervention differed

across types of teachers and schools. The study finds large average gains in literacy skills in English and Kiswahili and no differences in average impact across teacher characteristics. In contrast, the school characteristics moderate the impact of the intervention on higher-level skills (word reading and comprehension). This work contributes to support the process of designing and implementing language policies in schools focused on inclusion and equity.

List of Acronyms

ADEA	Association for the Development of Education in Africa
ASEAN	Association of Southeast Asian Nations
ATE	Average Treatment Effect
BECF	Basic Education Curriculum Framework in Kenya
CFI	Comparative Fit Index
DIBELS	Dynamic Indicators of Basic Early Literacy Skills
EGDIS	Expanded Graded Intergenerational Disruption Scale
EGRA	Early-Grade Reading Assessment
ICC	Interclass Correlation
IRT	Item-Response Theory
KCPE	Kenya Certificate of Primary Education
KICD	Kenya Institute of Curriculum Development
L1	First Language
L2	Second Language
LARA	Literacy Achievement and Retention Activity in Uganda
LMICs	Low- and Middle-Income Countries
LOI	Language of Instruction
MAR	Missing at Random
MCAR	Missing Completely at Random
MI	Multiple Imputation
MLM	Multilevel Modeling
MoEST	Ministry of Education, Science and Technology in Kenya
MOI	Medium of Instruction
NCDC	National Curriculum Development Centre in Uganda
OAU	Organisation of African Unity
OLS	Ordinary Least-Squares
PRIMR	Primary Math and Reading program in Kenya
RMSEA	Root Mean Square Error
SDG	Sustainable Development Goals
SES	Socioeconomic Status
SHRP	School Health and Reading Program in Uganda
SNNP	Southern Nations, Nationalities and People's region in Ethiopia
SRMR	Standardized Root-Mean Square Residual
TaRL	Teaching at the Right Level
UIS	UNESCO Institute for Statistics
UN	United Nations
UNEB	Uganda National Examinations Board
UNESCO	United Nations Educational, Scientific and Cultural Organization
UPE	Universal Primary Education
USAID	United States Agency for International Development
USE	Universal Secondary Education

Introduction

The right to education is a core part of the *Universal Declaration of Human Rights* (United Nations, 1948) and the 1960 *Convention against Discrimination in Education* has reaffirmed this right without exclusion “based on race, colour, sex, **language**, religion, political or other opinion, national or social origin, economic condition or birth” (Singh, 2008, p.71).¹ Since these aspirational declarations were made, access to school has steadily increased, illustrated by the drop in the proportion of out-of-school children of primary school age from 28% in 1970 to 9% in 2016 (UIS, 2019). Despite these important gains, the number of children who never access school or never finish basic education cycles remains very large in some parts of the world, perpetuating or increasing inequities (Dutcher, 2001; Walter & Benson, 2012). Furthermore, access to school does not always translate into acquiring the necessary skills to promote the “full development of the human personality” envisioned in the 1948 declaration.

Recent commitments to education for all by the international community have included explicit mention of equity and quality. Goal 4 of the *Sustainable Development Goals*, for example, calls to “ensure inclusive and equitable quality education and promote lifelong learning opportunities for all” (United Nations, 2015, p.15). As another signal to the commitment to learning, the Global Alliance to Monitor Learning was set up to bring together government and civil society to support the technical and political process of measuring learning (UNESCO, 2018). This focus on equity and quality, and

¹ Emphasis on “language” is mine.

on measuring learning, has followed increasing awareness that millions of children and adolescent were not functionally literate even when spending several years enrolled in school, a situation commonly referred to as the *global learning crisis* (UNESCO, 2013; World Bank, 2018). Unsurprisingly, children most likely to be excluded from educational opportunities, drop out of school, or spend years in school without developing literacy skills are those living in most marginalized communities, including those whose languages are not used in school.

Among the structural challenges school systems face in delivering high-quality education for all children, the choice of language in the classroom presents both a threat and an opportunity. Linguistic heterogeneity is a reality in most contexts around the world, despite the strength of the *monolingual habitus* – a set of unquestioned dispositions that make us ignore the multilingual and multicultural ways of life – in which we live (Benson, 2013). At the state level, the median number of languages per country is 19, and countries in the top decile of the distribution have at least 90 language communities (Simons & Fennig, 2018). Multilingual classrooms are increasing globally because of factors such as urbanization, economic and social mobility, intermarriage, and displacement as a result of conflict (Benson & Young, 2016). School systems that ignore the linguistic diversity of children fail to deliver an equitable education. Indeed, children who speak the language of instruction at home systematically demonstrate higher achievement on learning assessments (UNESCO, 2016a). In contrast, acknowledging this diversity and enacting multilingual language policies and programs can help reduce inequities in drop-out and completion rates. In Senegal, a six-year program implemented

in 465 classrooms in six local languages resulted in a completion rate of 43% compared with 33% in school teaching in French only, with no difference on average learning outcomes at the sixth-grade national exam (World Bank, 2010). Seid (2016)) found that using local languages in Ethiopia increased the probability that students were enrolled in primary school, and that students were enrolled in the age-appropriate grade. Similarly, using the local language can have positive effects on school attendance and completion for individuals most at risk of dropout such as girls (UNESCO, 2005), students from low-income background (Abadzi, 2006) and ethnic minorities (IRIN, 2011).

Language in education policies, however, are not designed exclusively based on pedagogical considerations. The status of language(s) in official spaces such as government, courts, and schools, is decided in part based on the sociopolitical forces and state ideology. Nation-states with monolingual and monocultural visions can use language as a unifying factor, such as Bangla in Bangladesh (Hamid & Jahan, 2015), ignoring linguistic minorities, or explicitly giving them lower status. In contrast, multicultural visions of state building can provide all linguistic groups with equal status under the law, such as in South Africa where 11 languages were given official status (N. Alexander, 1997; Desai, 1995), making it possible to use them in schools as the language of instruction. In addition to ideology, important technical and logistical concerns factor into how states make decisions about education policies, such as developing multiple curricula and deploying teachers according to language skills. Globally, the number of languages used in formal school systems has increased, at least in official policies (Albaugh, 2014; Kosonen, 2017).

In this three-paper dissertation, I explore different aspects related to education language policies. The first paper provides background on the factors states consider when planning language policies and proposes a framework for analyzing education language policies. This framework includes a set of factors necessary for choosing languages to use in classrooms and to make multilingual policies successful. The two empirical papers explore some of these factors in a given setting. The second paper seeks to build on existing research on cross-linguistic transfer – the ability of children to transfer literacy skills from one language to another – in Uganda, where no studies on cross-linguistic transfer had been conducted. The third paper explores whether a successful early-grade literacy program in Kenya had differential impact across beneficiaries and contexts.

Paper 1: Analyzing Language Policies in Education

As she instructs children in her classroom to take out their notebooks, a teacher's choice of language will be the result of a complex set of factors that include her linguistic skills, those of her students, as well as the language(s) that are used in the community, the curriculum, and ultimately, the language prescribed or proscribed by the government through an official policy. National or local governments have the mandate to design and support the implementation of language in education policies, and the choice of language(s) has important implications for the ability of children to learn. Policymakers, however, consider more than pedagogical factors when awarding an official status to languages in public institutions such as schools.

The first paper in this dissertation proposes a framework for language policy analysis and includes three main sections. In the first section, I start with presenting a historical perspective on the place of language in state building and education policy. I then describe the three types of language planning policymakers engage in explicitly or implicitly – status, corpus, and acquisition – and show the orientation a government takes toward multilingualism and how policies are implemented have implications on the extent to which linguistic minorities are successfully included as full participants. Finally, I describe the range of multilingual approaches used in schools, and how their goals reflect different visions for the long-term language development of students.

In the second section, I propose a framework to analyze language policies and practices in multilingual contexts that can be used to understand the strengths and weaknesses of the current approach, or to identify issues that need to be addressed to successfully make changes. The proposed framework includes six factors: (i) the sociolinguistic context; (ii) community support and engagement; (iii) orthographies; (iv) curriculum and materials; (v) teacher recruitment, training and deployment; and (vi) assessment and examination. The framework is meant to be used in whole or in part depending on the level of analysis (local, regional or national).

The last section of the paper applies the framework to analyze the education language policies in Kenya and Uganda. This section serves two main purposes: illustrating how to apply the framework and providing important contextual information on the two countries in which the empirical studies of this dissertation are set. The analysis shows that although the two official policies are very similar – each

recommending the use of a first language for three years of school before transitioning to English – and that the linguistic make-up of the two neighboring countries is very similar, the implementation of the policies are very different. Children in Kenya are likely to be taught in Kiswahili and English, two languages that few children speak at home, whereas a large proportion of Ugandan children are taught in a home language for the first few years of school before transitioning to English. While there is important variation within each country, this difference in language use in the early years is partly due to the much higher level of institutional support for the use of local languages in Uganda, including through the development of languages and their orthography, teaching resources, and assessments.

This paper argues that the choice of language in schools has important equity implications for individual children, as well as for linguistic communities. However, the extent to which a multilingual policy serves to reduce or exacerbate inequities and the ability for individuals to fully develop is dependent on more than an official policy. Instruction in the language that children speak at home can be beneficial to the extent that the language(s) have well developed orthography, teachers proficient in these languages and trained to teach in those languages, and a set of resources (curricula, books, assessments) to support them.

Paper 2: Exploring cross-linguistic transfer in a new context

One of the elements at the center of the framework presented in Paper 1 is the set of languages present in the community and the classrooms. The number of languages, and the differential linguistic features of these languages will impact the implementation

of a multilingual policy. Specifically, the ability of children to use their skills acquired in one language to develop those in the second language depends on the languages themselves (Bialystok, Luk, & Kwan, 2005). Paper 2 explores whether children in Uganda learning in two languages in school are able to transfer skills.

One of the main arguments educators and policymakers use to promote bilingual programs is rooted in linguistic theory. The argument follows a straightforward logic: (i) children can learn better in a language they understand, including the basic skills necessary to develop literacy, and (ii) children who have developed these skills in one language can transfer them to a second or third language. While the first statement should not need much convincing, the second relies on empirically demonstrating cross-linguistic transfer theory.

Cummins (1979) first proposed that oral language and literacy skills developed in the child's first language (L1) can transfer to the second language (L2) through a process he described as linguistic interdependence. This theory, suggesting that there is a common underlying proficiency across languages, has guided much of the research in cross-linguistic transfer. Unsurprisingly, the empirical studies based on this theory developed in North America started with samples of students speaking European languages. More problematic, however, is the fact that the vast majority of studies of cross-linguistic transfer available four decades later examine the relationship between Indo-European and/or Asian languages (Melby-Lervåg & Lervåg, 2011). The second paper in this dissertation aims to add to this literature by presenting one of only a handful

of cross-linguistic studies including African languages, and the only one to my knowledge set in Uganda.

This study has several aims, including to describe the performance of Ugandan second-grade students on literacy tasks in one of two local languages – Luganda and Runyankole-Rukiga – and English, and to explore whether there is evidence of transfer of literacy skills across these languages. In this study, I use data from 3,561 second grade students assessed in oral language, decoding skills, and reading comprehension in two languages using the Early Grade Reading Assessment (EGRA). I use correlational analysis and structural equation modeling to test whether students’ basic alphabetical knowledge skill and higher-order reading comprehension skills transfer from the first to second language.

The findings suggest evidence of cross-linguistic transfer of basic decoding skills from Luganda and Runyankole-Rukiga respectively to English. In addition, there seems to be transfer of reading comprehension skills for students who are taught in the language of the community in second grade. The limitations of this study based on the reliance of cross-sectional data and a shallow set of language measures point to the need for better assessments of language skills in multilingual settings and more research on language development, including longitudinal studies. Nevertheless, the findings support the idea that the elements at the center of the framework in Paper 1 are applicable beyond European and Asian languages, as well as in contexts linguistically similar to Uganda.

Paper 3: Identifying variation in the impact of a literacy intervention

The periphery of the framework in Paper 1 includes four elements related to the implementation of multilingual language policies: teachers, orthography, curriculum and materials, and assessments. In Paper 3, I investigate a bilingual reading program in Kenya meant to support two of these aspects: teacher training and teaching materials.

In the context of the global learning crisis described above, governments have invested heavily into early grade literacy and numeracy programs, often with the technical and financial support of bilateral and multilateral funding agencies. These literacy professional development initiatives often follow a similar model that includes developing a set of teacher guides with detailed lesson plans, in-service teacher training on the use of these resources, and learning materials for students. The language in which these programs are implemented depends on the country's language policy. In Uganda, for example, the government has supported the implementation of a reading program for grades 1-3 in more than a dozen languages. In Kenya, the national reading program Tusome is focused on English and Kiswahili.

Evaluations of early grade reading programs have largely found positive impact on low-level skills such as letter naming and word reading, and small or null effects on fluency and comprehension (Kim, Lee, & Zuilkowski, 2019). This set of findings has been interpreted by some as positive – large effects on initial skills in one or two years of implementation is a success and a foundation for future growth – and by others as negative – little impact on comprehension is a failure as that is the ultimate goal of reading. One important aspect that has been understudied is the extent to which the average impact masks heterogeneity in treatment across beneficiaries and contexts.

Understanding whether these programs vary in their impact across teachers and school contexts is important for two reasons: efficient use of resources and equity. Given the limited set of resources in low- and middle-income settings, understanding the teachers best suited to benefit from a program or the school environment that enables better implementation can maximize the benefits. In addition, an equity approach to social program delivery should ensure that programs reduce, not reinforce, existing gaps in learning.

The third paper in this dissertation focuses on a literacy intervention in Kenyan urban settings called Primary Math and Reading (PRIMR). In this study, I test whether teacher characteristics – years of experience and formal training – and school context – baseline score and whether schools are public – moderate the impact of the intervention on student learning outcomes. The study was designed as a phased-in randomized control trial that allows estimates of the average treatment impact of the intervention. While the study was not designed to measure the heterogeneous impact of the intervention, the program was implemented in both public schools and nonformal private schools, and the teachers participating included teachers with a wide range of years of professional experience, with both formally trained teachers and some without pre-service training.

The findings from my analysis show that overall, the intervention had a large positive impact on student outcome, on average, with similar effect sizes than previously reported in other studies. The impact of the intervention does not vary across teacher characteristics, suggesting that this in-service training is beneficial to all teachers regardless of years of experience in the classroom or whether they attended a formal

teacher training college. In contrast, the intervention only had positive impact on high-level skills – word reading and comprehension – in nonformal private schools, and the impact was larger in schools that demonstrated higher baseline levels of learning. This finding is particularly important to explore to better understand the reasons for the growth in gaps resulting from the intervention. The study presents plausible explanations for this finding and their implications for policy, while emphasizing the limitations of the study design. First, the study finds that all teachers can benefit from this type of in-service training, whether they have undergone formal teacher training or not, and regardless of the number of years they have been teaching. The findings that students in schools that started at a low-level benefited less from the intervention on average than their peers in higher performing schools, thereby increasing the gaps, is a reminder to focus program analyses on equity. Programs should explore adapting the interventions to match the learning environment of the school in a similar way teachers should target their instruction to the level of the students.

Paper 1: Ideology and Pedagogy of designing and implementing Language Policies in Education: An Analytical Framework applied to Kenya and Uganda

Introduction

The important gains made around the world in increasing access to formal education are rarely celebrated because of the increased awareness of the low levels of learning taking place in many settings. The current state of education is characterized by a global learning crisis with an estimated 250 million children unable to read or count well whether or not they have attended school, including 125 million who are not functionally literate despite four or more years of schooling (UNESCO, 2013; World Bank, 2018). One of the important factors contributing to low levels of learning in many contexts is the mismatch between the language spoken by children at home and that of school, with children who speak the language of instruction at home systematically demonstrating higher achievement (UNESCO, 2016a).

Countries and communities aiming to provide educational learning opportunities to all their children need to carefully consider the language(s) they use to facilitate that process. While many countries include multiple languages of instruction at some stage of the formal education system, especially in the early years of primary school, it is estimated that 38% of the world's population speak a language that is not used in schools

(Walter & Benson, 2012).² The choice of language in which a child learns at home and in school has a profound impact on their individual development, both cognitive and socioemotional. Multilingual families, where parents or grandparents speak different languages or speak with different levels of fluency and comprehension, make decisions on language use when children are born. These implicit or explicit decisions are informed, among other factors, by the language proficiency of adult speakers, childcare responsibilities that lead different levels of exposure, and the utilitarian value of the languages in the communities the child will be a part of. None of the languages are inherently a better choice than another, but the intersection of the family's values, resources, and contextual factors result in a different set of opportunities to the child based on the language(s) they grow up with. The ability to build a strong connection to a grandparent or their culture may be lost; the opportunity to speak the language used in the formal economy and government may be gained; or vice versa. Similarly, multilingual communities, regions and states make implicit or explicit decisions about language use in their institutions, including schools.

Language Planning is both a theoretical field of study and a set of activities governments or local communities engage in, implicitly or explicitly, that result in policies for how and where languages are used in public spaces, including schools. The differences between two languages that give individuals more or less economic and political power over others are not inherent to the languages themselves, but rather result from the context in which they are spoken. Policymakers engaged in language planning

² Dutcher (2001) estimated that more than 220 million school-aged children spoke a language that was either not "well known" or unwritten.

in multilingual contexts are thus not only deciding how languages are used, but also which languages to develop and with what aim (Hornberger, 2006). These decisions in the education system can be made at several different levels. In the informal education sector, language choices are typically made at a very local level. In the formal sector, the focus of this paper, language policies can be designed at the central level, provide for some regional autonomy, or allow local education administrations to make decisions. Teachers, acting as “street-level bureaucrats” (Lipsky, 1969) ultimately make the final decision about what language to use for instruction once they are left with their students in the classroom. The approach to language planning taken by decision makers, as well as the implementation of these decisions have important consequences on student learning, as well as the extent to which linguistic groups are included in the larger community.

In this paper, I argue that there is no single language policy or practice that will benefit all children grouped in a classroom or school. Specifically, a tension exists in designing policies that can benefit some individuals or communities at the expense of a larger collective, especially with regards to equity. As with the choice faced by multilingual families, communities and countries face tradeoffs based on a set of contextual factors. The choice of language used in the classroom can enable or constrain student learning, increase or decrease equity in learning opportunities across groups and regions, and strengthen or weaken opportunities for national cohesion. This paper proposes a framework to analyze language policies and practices in multilingual contexts that can be used to understand the strengths and weaknesses of the current approach, or to identify issues that need to be addressed to successfully make changes. The proposed

framework includes six factors: (i) the sociolinguistic context; (ii) community engagement and support; (iii) orthographies; (iv) curriculum and materials; (v) teacher recruitment, training and deployment; and (vi) assessment and examination. The framework is meant to be used in whole or in part depending on the level of analysis (local, regional or national). As a case example, the framework is used in this paper to analyze the education language policies and contexts in Kenya and Uganda, the settings for the two empirical studies of this dissertation.

The paper starts with presenting a historical perspective on the place of language in state building and education policy to recognize the importance of non-educational factors in the choice of language use in the classroom. I then describe the types of language planning and different orientations policymakers can take in designing language policies, and the range of multilingual education programs. Next, I present the framework to analyze language in education policy in multilingual contexts. Finally, I apply the framework to understand the extent to which the language policies can be implemented effectively in Kenya and Uganda.

1.1 Nations, States, and Language in Education

The formation and evolution of nations and states are complex processes in which the role of language is key to including or excluding membership to a larger entity for certain communities. A state is a “legal and political organization, with the power to require obedience (...) from its citizens” whereas a nation is a “community of people, whose members are bound together by a sense of solidarity” (Seton-Watson, 1977, p.1). States can thus include a number of nations, or be centered around one, and nations can

exist within, across, or without states representing them.³ Over the course of a few generations in the twentieth century, a family identifying as part of the Polish nation could thus have lived in the states of Prussia, Ukraine, Lithuania, Germany, the Soviet Union, and Poland, all without moving (Lukowski & Zawadzki, 2006). Citizens of the Kurd nation mainly live across Turkey, Syria, Iraq and Iran, and while the semi-autonomous Kurd region in Iraq voted on an independence referendum in 2017, it does not hold the power of an independent state (Aldroubi, 2017). Nations within decentralized states can negotiate the role of language in public institutions, including schools, such as French speakers in the Quebec province of Canada (Riddell, 2004), several linguistic groups in states in Ethiopia (Seid, 2016), and India (Annamalai, 2001). The history of most states and nations includes similar shifts of physical boundaries, as well as waves of inclusion and exclusion of groups, which often manifests itself in part through the choice of language in its institutions. For instance, the Republic of Latvia regained its independence from the Soviet Union in 1991, and approximately 25-30 % of its population speaks Russian as a first language (Simons & Fennig, 2018). The president of Latvia recently claimed that a new law phasing out Russian from secondary schools would help “form a more cohesive society and stronger state,” to which the Kremlin responded with threats of economic sanctions to counter what they consider discrimination and “forced assimilation” (Luhn, 2018). This exemplifies how language is central to the idea of a nation or a state and can be used to include or exclude people.

³ I use this definition to clarify the difference between nations and states in their role and power to make decisions on language policies. These terms are sometimes used interchangeably, which can lead to confusion. For example, the United Nations is an association under which people are represented through the membership of states. Similarly, the United States of America includes a pledge to the flag that calls for allegiance to “one Nation under God, indivisible” (Government Printing Office, 2011).

Different ideologies can lead to visions of monolingual nation-states, where linguistic minorities are marginalized and expected to assimilate, or multilingual inclusive states, where nations share political and cultural rights.

One of the longest lasting consequences of imperialism and colonialism in the nineteenth and twentieth centuries is the use of European languages in many countries as official or de facto state languages, including in schools. While the process of nation and state building has existed for millennia, the vast majority of today's states were either first founded or gained independence within the last two hundred years⁴, often through subjugation to a European country. The European colonial project was justified in part by the *mission civilisatrice*, a moral imperative to propagate superior languages, religions and cultures (Heller, 2008). The implications of this imperialistic project in terms of language use differed between colonial powers and eras. The French Colonial language policy was an extension of the policy banning regional languages in the *métropole*, extending the exclusive use of French to public schools in its colonies starting in 1826, with an explicit goal to assimilate (Bokamba, 1991). In contrast, British rule in Africa followed a 'dual mandate' advocating for the "modernization" of Africans through the acquisition of English, as well as the right to maintain their mother tongue (Mazrui & Mazrui, 1998, p.143). During the brief period of German colonial rule in four African countries, little German was taught, in part because one dominant view included that Africans would "brutalize" the language (Pike, 1986). While European states were not

⁴ Figure A1 in the appendix shows the frequency of countries gaining sovereignty for the first time, and the latest year they gained independence using three different scales: (i) all years; (ii) post-1700; and (iii) post-1900. In total 54% of today's states first gained sovereignty and 83% most recently became independent after 1900.

the only ones to build empires spreading their language and culture, the scale of the European colonial enterprise, as well as the era during which it occurred has led to a handful of European languages having high status globally.

In addition to European languages, Arabic and Mandarin are two of the languages with highest status globally, which is reflected both in regional bodies, and the number of countries where they have a recognized status. The United Nations currently has six official languages: Arabic, Chinese, English, French, Russian, and Spanish (United Nations, 2018).⁵ Table 1 presents the number of countries or territories these languages have official or de facto status, and their number of L1 speakers globally. English has statutory or de facto recognition in 91 countries, despite the fact that over 80% of the 339 million people who speak it as a first language come from two countries. The growth of influence of English, and other high-status languages continues without the need for military force. For example, English was chosen as the official working language of the ten-country Association of Southeast Asian Nations (ASEAN) in 2009, after having been a de facto working language since its inception, despite the small proportion of the population speaking it at home (Kirkpatrick, 2017). In contrast, the European Union recognizes 24 official languages, at a cost of €1 billion per in year funding for approximately 3,000 translators and interpreters (Kelly, 2014). The contrasting approaches of these two regional bodies reflect the political will and resources needed to promote and protect languages.

⁵ Arabic was added as an official language in 1973. French and English were originally the only two working languages, but all six official languages are now also working languages in the General Assembly and Security Council. For more detail on the history of UN languages, see United Nations (2018).

Table 1. Recognition of selected languages across countries and territories

<i>Language</i>	<i>Statutory recognition</i>	<i>De facto recognition</i>	<i>Total</i>	<i>L1 users</i>
English	37	54	91	339 million
French	41	8	49	76 million
Spanish	17	6	23	427 million
Portuguese	7	2	9	202 million
Russian	5	7	12	171 million
Arabic, Classic	23	3	26	268 million*
Chinese, Mandarin	2	2	4	897 million

Note. Data from Ethnologue Global Dataset (Simons & Fennig, 2018). The dataset includes 237 territories and countries in total. * Classic Arabic has no L1 users. The 268 million figure was calculated by adding L1 speakers of 39 varieties of Arabic. Other estimates including a different set of languages or countries range from 198 million (Jenkins, 2011) to 295 million (Parkvall, as cited in “List of Languages,” n.d.).

The protection of language rights, and the promotion of access to education in the home language has become an important feature of international declarations and regional associations in the last century, with uneven implications for children entering schools. Table 2 below presents an incomplete list of United Nations declarations and conventions that protect linguistic rights or promotes the right to education in a home language. This list illustrates both how important the protection of linguistic rights is to the international community, and how often these rights have had to be explicitly protected. Regional bodies have also shown commitment to the protection of language rights or use all languages in education in Africa (ADEA, 2010; Organisation of African Unity, 1986), in Latin America (MINEDLAC, 1996), and in Europe (European Commission, 2013). These declarations range from serving to protect the legal rights of marginalized communities to active plans to promote mother-tongue based education. For example, the Organisation of African Unity (OAU) wrote an action plan “to encourage

the increased use of African languages as vehicles of instruction at all educational levels” (Phillipson, 1996). Despite the proliferation of declarations and action plans, education policies do not systematically embrace multilingualism.

Table 2. Selected international declarations and conventions promoting language rights

Year	Document	Article(s)
1948	Universal Declaration of Human Rights	2
1960	Convention against Discrimination in Education	5
1966	International Covenant on Civil and Political Rights	27
1976	Recommendation on the Development of Adult Education	22
1978	Declaration on Race and Racial Prejudice	9
1985	Declaration on the Human Rights of Individuals who are not Nationals of the Country in which they live	5
1989	Convention 169 concerning Indigenous and Tribal Peoples in Independent Countries	28
1989	Convention on the Rights of the Child	29
1990	International Convention on the Protection of the Rights of All Migrant Workers and Members of Their Families	45
1992	Declaration on the Rights of Persons belonging to National or Ethnic, Religious and Linguistic Minorities	4
1995	Declaration and Integrated Framework of Action on Education for Peace, Human Rights and Democracy	19, 29
2001	Universal Declaration on Cultural Diversity	6

Note: Table compiled from (UNESCO, 2003).

At the national level, linguistic rights are often recognized by law, but the extent to which they are protected or promoted through the education system varies. Globally, approximately 9% of the world’s languages are used in formal schools, collectively spoken as a first language by an estimated 62% of the world’s population (Walter & Benson, 2012). The trajectory from protection and official recognition to language in education policy in Senegal exemplifies this contrast. Senegal, which has used French as its official language since independence in 1960, originally recognized 6 national

languages but modified Article 1 in the 2001 constitution to recognize “every national language that will be codified⁶” (République du Sénégal, 1963, 2001). While nearly 90 percent of the population speaks one of the six languages previously recognized (Leclerc, 2013), this change opened the possibility of full inclusion as members of the state for the remaining 10 percent of Senegalese. In Senegalese public schools, however, education has until recently been conducted only in French, with the exception of a brief experiment in a handful of schools in the 1970s and one 6-year experiment in 155 schools between 2002 and 2008 (World Bank, 2010). A new national reading program called *Lecture Pour Tous* started introducing reading in three national languages for two years, in four of the fifteen regions of the country, starting in the 2017-2018 schoolyear (Rousseau, de Galbert, & Mount-Cors, 2017). While Senegal is clearly increasing linguistic protection in its institutional framework, and promoting more use of national languages in schools, most Senegalese children continue to learn in a language they do not speak at home or in their community.

Structured opportunities for education can broadly be categorized as formal or non-formal systems, which have historically served different purposes, and had different motivations for the choice of language.⁷ There are no universally accepted definitions of

⁶ Author’s translation. Codification here refers to the development of an orthography or script. This change has led to a rapid increase of national languages recognized, numbering 21 as of 2012 (République du Sénégal, 2012, p.58).

⁷ Coombs and Ahmed (1974) include three categories, adding informal education. They defined informal education as part of the unorganized and unsystematic opportunities to learn throughout an individual’s life. Given the focus of this paper on policy and planning, I limit the discussion to formal and non-formal education.

formal and non-formal education (Hoppers, 2006; Rogers, 2004), but the following definitions are often referenced and useful to understand the main contrasts:

“Formal education as used here is, of course, the highly institutionalized, chronologically graded and hierarchically structured ‘education system’, spanning lower primary school and the upper reaches of the university. (...)

Non-formal education (...) is any organised, systematic, educational activity carried out outside the framework of the formal system to provide selected types of learning to particular subgroups in the population, adults as well as children.” (Coombs and Ahmed, 1974, p.8)

While both types of education are systematic and organized, non-formal approaches are typically more flexible and can thus provide avenues for groups that are explicitly excluded from the formal system or that are disproportionately affected by drop-out or low-quality settings, such as girls, children in rural areas, children with disabilities, and those from linguistic minorities (Yasunaga, 2014). Formal education systems have historically been concerned with elites, including in colonial territories where schools focused on educating children of local chiefs or future civil servants and used the European language as a medium both during European rule and at independence – e.g. French in West Africa and South-East Asia (Bokamba, 1991; Clayton, 1995); Spanish in South America (Reimers, 2006). Non-formal education systems have, in contrast, often made use of local languages. Prior to the arrival of the French, education in Cambodia was mainly led by monks in temples, directed to boys only, and conducted in Khmer (Clayton, 1995). Non-formal education has been used to increase literacy rates in low-income countries through the use of local languages with programs targeting out-of-school youth (Katahoire, 2006) or adults (Benson, 2004; Trudell & Klaas, 2010). Large expansions of access to formal schooling have been efforts to be more inclusive of

historically underrepresented or marginalized communities, which has simultaneously meant enrolling children in a system that uses a language they do not speak at home, an inherently inequitable process.

Multilingualism at the individual level, as well as within nations and within states is clearly not a rare phenomenon. Indeed, the presence of multiple linguistic groups living within the borders of the same state is the norm, rather than the exception (UNESCO, 2003). Until the end of the twentieth century research largely treated multilingualism as an exception, resulting in flawed paradigms such as the notion as bilinguals as “two monolinguals in one brain” (Bokamba, 2015; Grosjean, 1989). The ideology of “one state, one language”, and to a certain extent “one nation, one language”, is thus clearly problematic. Auer and Lei (2007) argue multilingualism could build bridges “between different groups within the nation, bridges with groups beyond the artificial boundaries of a nation, and bridges for cross-fertilization between cultures” (p.12). I contend that the extent to which multilingualism becomes a positive force at a community- or state-level is largely a result of the orientation to language planning and the design and implementation of sound policies.

1.2 Language Planning and Orientations

Policymakers developing language policies need to engage in three interrelated types of planning: status, corpus, and acquisition (Hornberger, 2006). Status planning defines a role for languages in a nation state or a given linguistic community. Corpus planning focuses on the language itself, whether and how it is codified and the

development of its orthography. Acquisition planning includes the efforts to create or increase opportunities for individuals to learn the language(s). This section describes these three dimensions and presents orientations or approaches to planning for language in multilingual contexts.

An administrative authority can give languages a **status**, or a language can acquire status informally through its use. While labels are used inconsistently, the types of formal status given to a language by a state can be *official*, *national*, and/or *recognized*. Official languages, sometimes called statutory, are those mandated by law to be used by the government both in writing legislative and other documents, and to conduct business in institutions such as courts and parliament. National languages are typically spoken by a sizeable proportion of the population, and can be a *de facto* language of wider communication, within a region or the entire state. Some languages or linguistic groups are officially recognized in the constitution or other legal document but have no official function in the state. In any given state or region, a language could have a different combination of these categories.⁸ In addition to their formal status, languages can acquire unofficial status or power through their use and perceived value. In fact, some sociolinguists consider language status mainly as its “perceived relative value ... usually related to its social utility” (Ricento, 2009, p.5) rather than its legality conferred by the state. I consider both the legal and informal status of languages important in analyzing multilingual education policies.

⁸ Cooper (1989) uses a different set of categories: statutory, working and symbolic based on the official function of the language at the national or provincial level.

Enhancing the status of a language can result in an increase in representation of marginalized linguistic communities in public spaces and promote the vitality of the language. For instance, the authors of the new South African constitution decided to include all 11 languages as official to give equal status to linguistic communities that had been purposefully divided under apartheid (Alexander, 1997; Desai, 1995). In Ethiopia, the 1994 violent change in regime to a federal system was motivated by a vision of political and cultural autonomy and promoted the use of local language to “acknowledge ethnolinguistic diversity and equality” (Seid, 2016). This shift to a federal system and the resulting change in language in education policy seems to have led to an increased sense of belonging to the community for some groups (Ramachandran, 2017). At the individual level, bilingual learners are more likely to maintain their heritage language when it has higher status among peers and in the community (Clément, 1986; Snow & Kang, 2006; Tse, 2001). Ultimately, languages with higher status, whether acquired formally or informally, have more vitality and are less likely to be endangered.

Education systems that include multiple languages need to plan for the management of the languages themselves. **Corpus** planning includes selecting language varieties for standardization, creating or improving their orthographies, and codification. The choice of language varieties to standardize is particularly important, and sensitive, in contexts that include endangered languages, as they typically have more regional variation (Sallabank, 2010), and it is not always clear whether two groups speak a different variety of the same language, or a separate language. Orthography can be defined as “the set of conventions for writing words in a language” (Sebba, 2007, p.10).

Depending on the current level of standardization, corpus planning can involve the development of a new orthography, when the language is only spoken, improving the orthography of a language when it is inadequate, or promoting the further development of an established language, such as reviewing and approving new words in the language. In all cases, purposeful corpus planning necessitates resources and coordination.

The identification and funding of institutions responsible for corpus planning has implications for the successful implementation of multilingual education policies. The development and promotion of the French language is supported in France by the well-endowed *Académie française* established in 1635 with the initial responsibility to “give rules to the language with certainty” and recent self-proclaimed goal of “protect the French language” (Académie française, 1995). Similarly, South Africa established one institution – the Pan South African Language Board – to promote all 11 of its official languages (Desai, 1995). In other settings, corpus planning can be formally or informally governed by cultural institutions such as kingdoms, local government authorities, or community-based organizations. The extent to which these organizations have the support of the community, and the necessary resources can impact the successful promotion of the language(s) they represent.

Hornberger (2006) includes four possible goals of **acquisition** planning in her model: revival, maintenance, interlingual communication and spread. These goals can be met through language instruction outside of the formal schooling system, such as in the adult literacy programs, or through the implementation of multilingual education policies in the formal schooling system. The next section presents a range of approaches to

multilingual programs in the formal schooling system, including the explicit and implicit goals these programs reflect.

In multilingual settings, all three language planning types –status, acquisition or corpus – related to one language is connected to the planning of other languages. Ruiz (1984) introduced the concept of *orientations* toward language planning that he defined as dispositions toward the languages and their role in a society. Ruiz argued that policy planning had mainly included two orientations: language as a *problem*, and language as a *right*. An orientation toward language as a problem tends to see non-dominant⁹ languages as a barrier to national unity and integration, and links language to other social issues of linguistic minorities such as poverty or low levels of education. In contrast, seeing language as a right considers the human right of individuals and marginalized communities to have access to services in their languages.

In addition to defining these two orientations, Ruiz called for the development of a literature and conceptualization of an orientation to language planning using language as a *resource* as an alternative.

“Language planning efforts which start with the assumption that language is a resource to be managed, developed and conserved would tend to regard language-minority communities as important sources of expertise.” (Ruiz, 1984, p.28)

While Ruiz recognized the negative impact of engaging in language planning with a problem or rights orientation, he also claimed that it may be useful to have a “repertoire

⁹ I use the terms dominant language (DL) and non-dominant language (NDL) as defined by Benson and Kosonen (2013). These terms allow reference to a language in relation to another without the ambiguity of terms such as minority or indigenous language. It is possible for language to be DL in a context and NDL in another. For example, Luganda is an NDL at the national level where English is the official language, but Luganda is a DL in some parts of eastern Uganda where Lusoga speakers have had to learn in Luganda in school.

of orientations.” Orientations toward language planning should thus consider both the context and the type of planning.

Orientations toward language planning must be examined in light of the goals set by a linguistic community or nation. At the national level, the choice of languages to develop or promote reflects the goals of the country. A country giving official status to a language will not meet its goal of inclusive citizenship without an acquisition plan that ensures the long-term viability of the language, as the case of Morocco illustrate. Morocco, which established a policy of *Arabization* at independence to reduce its dependence toward France, moved to officially recognize Berber in 2000 and subsequently made it compulsory for all children in schools (Errihani, 2017). However, the top-down approach of this policy and various implementation challenges, including the small number of teachers qualified to teach in Berber, have greatly limited progress toward the implicit goal of making Berber communities and their language full participants in the country (Zouhir, 2014). This top-down approach to language planning clearly followed a problem orientation and would need to include a proper acquisition plan with a resource orientation.

1.3 Multilingual approaches in schools

The types of education programs offered to emergent bilingual and multilingual students vary in the intensity of and duration for which multiple languages are used. In addition, the types of programs differ in the goals set for the long-term literacy of students. Table 3 presents six broad types of programs ranging from full submersion in

the dominant language to two-way bilingual programs.¹⁰ The first four types of programs have a common goal for students of linguistic assimilation in the dominant language. The programs vary in the extent to which the medium of instruction is monolingual.

Submersion programs include all students in the mainstream classrooms taught using a second language (L2) as the medium of instruction with no additional support to the first language (L1) of speakers.¹¹ Pull-out programs are similar to submersion but include times in the day for students to be removed from the mainstream classroom individually or in groups to receive additional L2 instruction. Structured immersion programs group emergent bilingual students in classrooms where instruction in subjects is differentiated to meet their language needs. In transitional bilingual programs, the first few years of school are taught using the L1 of students as they develop L2 language skills. The transition to L2 instruction can be gradual across several years, or abrupt.

In contrast to these four types of programs, developmental bilingual and two-way bilingual programs aim to promote bilingualism and biliteracy in students.

Developmental bilingualism, or late exit programs, develops L1 language and literacy for a period sufficient to maintain those skills in the long term, but aims to promote academic

¹⁰ Table is adapted from Crawford (2004) and García, Kleifgen, & Falchi, (2008) to represent models outside of the United States. This table is non-exhaustive and does not include the role that a third language may have. For example, many settings include additional languages taught as subjects.

¹¹ I use the terms first language or (L1) to refer to the home language, local language, or community language that a child is first exposed to. L2 refers broadly to a language that a child is not exposed to in her community but must learn or use at school.

achievement at secondary school and university in L2. In contrast, dual immersion aims to promote biliteracy and academic achievement in both languages¹².

Table 3. Program Models for Multilingual Learners

Program	Language used in Instruction	Duration	Goals
Submersion (sink or swim)	100% L2	Throughout Primary and Secondary Education	Linguistic assimilation to L2; Monolingualism.
Pull-out (submersion plus support)	90-100% in L2; may include some L1 support	As needed	Linguistic assimilation; remedial support in L2; Monolingualism.
Structured Immersion	90-100% in L2; may include some L1 support	1-3 years	Linguistic assimilation; quick exit to mainstream education; Monolingualism.
Transitional Bilingual (early exit)	80-100% L1 initially; transition to 90-100% L2	1-3 years	Linguistic assimilation; acquisition of L2 without falling behind academically; Monolingualism.
Developmental Bilingual (late exit)	50-90% L1 initially; gradually reach 50/50 model	5-6 years	Bilingualism and biliteracy; academic achievement in L2
Two-way Bilingual (dual immersion)	Parity of languages; may start with a 90/10 model or with a 50/50	Throughout primary and/or secondary education	Bilingualism and biliteracy; academic achievement across multiple languages

Note: Table adapted from Crawford (2004), (García, Kleifgen, & Falchi, 2008), and (García, 2009, p.123).

¹² In the original model, focused on the United States, Crawford (2004) defines two-way bilingual programs in the same way as developmental bilingual, with the key difference that the latter are aimed at speakers of non-dominant languages, whereas two-way bilingual programs include English-only speakers. In this table, I modify the last type of program to contexts that promote long-term bilingualism.

While each type of program may include a number of different pedagogical approaches or components, their goals for the long-term development of students' literacy skills are an essential characteristic of the program. The main reason for including multiple languages in the education system is to serve students who enter the system without adequate (or any) knowledge of the dominant language, and are thus emergent bilingual learners. These goals of educational programs broadly support *subtractive* or *additive* bilingualism. Subtractive bilingualism is a process during which the development of a new language (L2) is done at the expense of the first language (L1). Conversely, additive bilingualism programs support L1 maintenance and thus develop both languages (García, 2009, p.116). Although this simplified model obscures the dynamic reality of language development and the extent to which realities in a classroom match the official policy or program¹³, it is helpful to identify the explicit goals put forth by policymakers when choosing a model.

Acquisition planning is at the heart of designing language in education policies. However, designing language of instruction policies to improve the educational experience of learners cannot be done without proper consideration of language corpus and status planning. The following section proposes a set of factors related to language

¹³ An important debate exists around the notion of additive and subtractive bilingualism, and the extent to which the theories and practices around bilingual education reinforce power structures of dominant-language speakers. (García, 2009) and (García, Flores, & Chu, 2011) propose to replace this simple view with a contrast between recursive and dynamic bilingualism models. They propose this to recognize that practices of emergent bilinguals are complex and that languages are not necessarily separate and autonomous, especially in non-Western multilingual contexts. Flores and Rosa (2015) explicitly argue that additive bilingualism is flawed as the concept is fueled by raciolinguistic ideology. Despite this important debate, there has not been a significant shift in the types of policies and programs that allow for a more dynamic approach to the use of languages, such as translanguaging (Vogel & Garcia, 2017).

planning necessary to ensure quality in the provision of education in a multilingual setting.

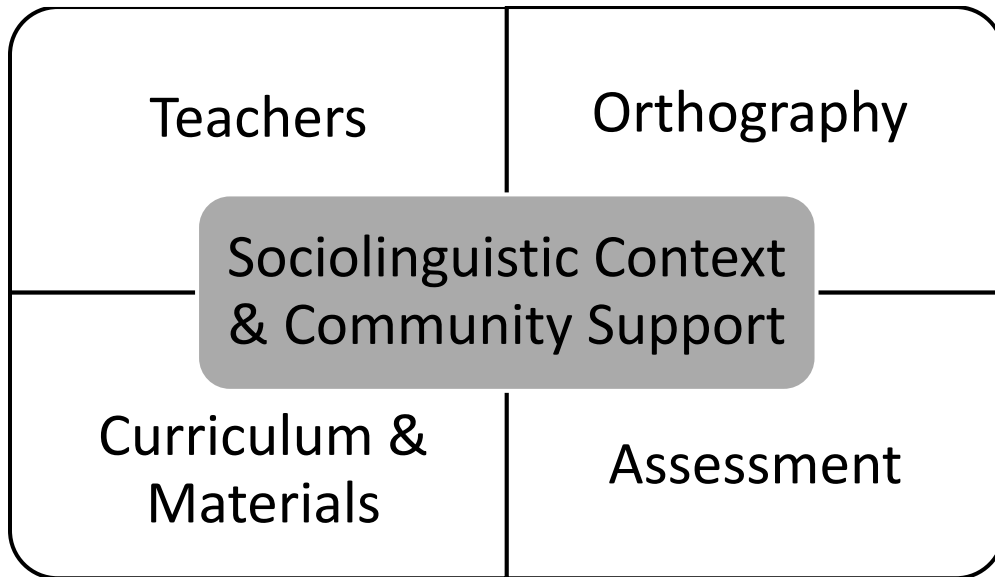
2. Framework to analyze education in multilingual contexts

As the sections above make clear, local communities, regional, and state authorities can have different motivations and goals when choosing the language used in schools. These stakeholders inform this decision in part on how they think language choice will impact the quality and relevance of the educational experience in the classroom. As Desai (1995) explains: “the merits or demerits of mother tongue education are dependent on the particular context in which learners find themselves” (p.22). More broadly, the merits or demerits of choosing any combination of languages – a mother tongue or a language foreign to the community – depends on the context of the classroom and the broader community.

In this section I propose a set of six factors necessary to understand whether the choice of language(s) used in schools in a particular setting, as well as the implementation of that policy decision, is likely to promote learning for children in a given classroom and serves the goals of the community. These factors are: (i) the sociolinguistic context; (ii) community engagement and support; (iii) orthographies; (iv) curriculum and materials; (v) teacher recruitment, training and deployment; and (vi) assessment and examination. Figure 1 presents a graphical representation of this framework. The first two factors, at the center, focus on the context in which the policy or program is considered or implemented, and are key to decide whether a multilingual policy is necessary and supported by the communities. The remaining four factors

concern the technical aspects that are necessary to ensure the quality of the education provided in the given language(s). Each section ends with some key questions that can help inform language of instruction decisions.

Figure 1. Graphical representation of the proposed framework

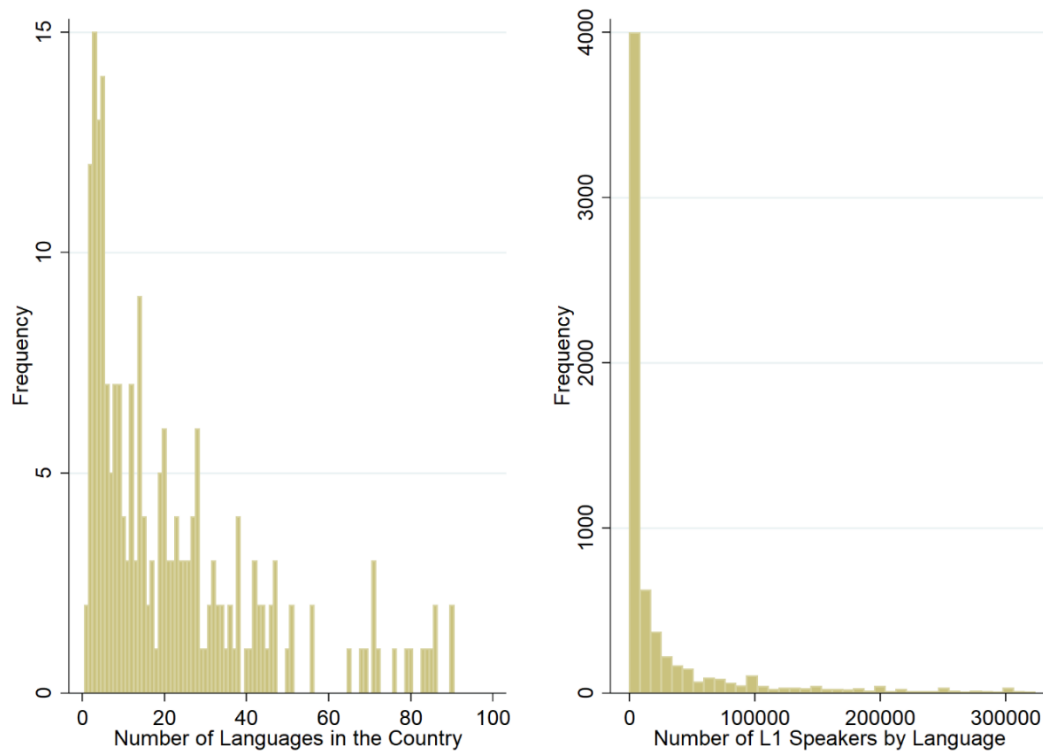


2.1. Sociolinguistic context

A deep understanding of the sociolinguistic context of a community or state is crucial to examine the language policies implemented in schools. The sociolinguistic context includes both the prevalence of languages spoken and the power languages hold in relationship with one another. The extent to which the state or communities are linguistically heterogenous is an important practical aspect of the context. From a central planning perspective, the number of language groups and the number of L1 speakers in each of these groups is an important factor in assessing the capacity to support the integration of these languages in the formal education system. Figure 2 below presents the distribution of the number of languages by country, and the number of L1 speakers in

each language. The decision to include a group of languages in the formal system must be informed in part by the capacity to support its instruction. Similarly, the best way to support a linguistic group that includes a few tens of thousands of speakers may not be through including that language in the formal education system. In contrast, linguistic communities with a small number of speakers – which is the case for the majority of linguistic groups – may choose to develop non-formal literacy programs specifically because the state is unlikely to include their language in the formal system.

Figure 2. Number of languages by country and number of L1 speakers by language

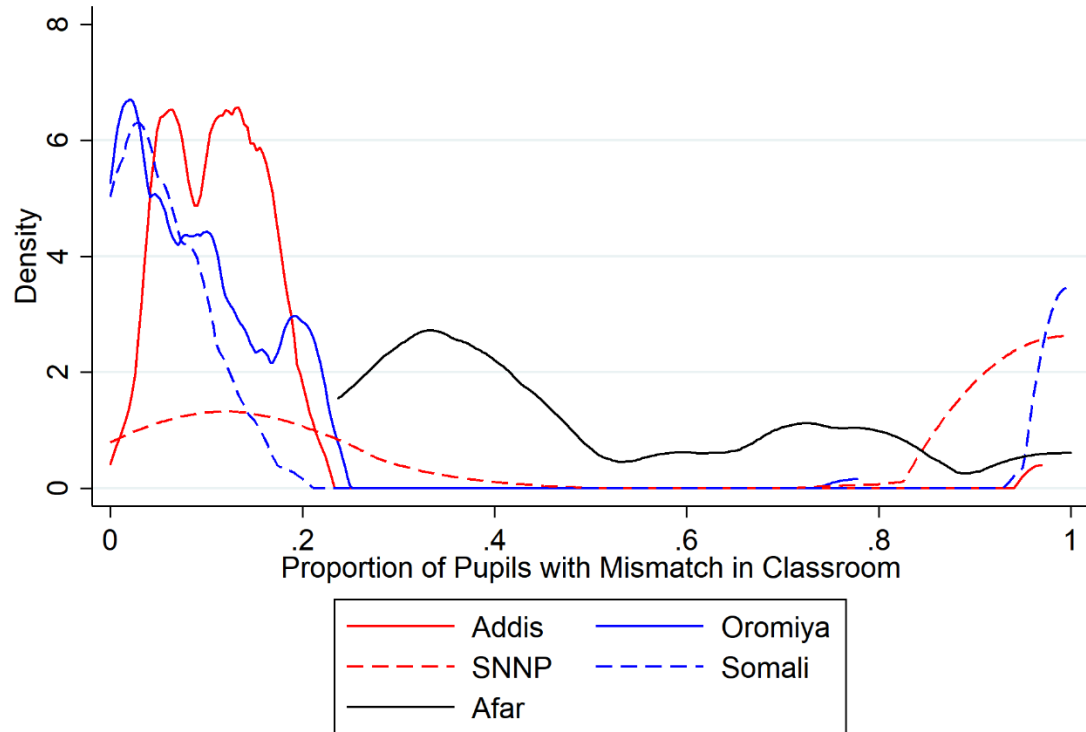


Note. Data from Ethnologue Global Dataset (Simons & Fennig, 2018). The top 10% of both distributions was dropped to improve the visual representation. There are 933 languages with more than 322,900 L1 speakers and 23 countries with more than 90 languages not represented in this figure.

In addition to the total number of languages and the size of the linguistic groups, states need to account for the relative linguistic homogeneity of communities.

Implementing a multilingual policy in a state with five linguistic communities of equal size would require a different approach and serve different goals if the five groups are geographically spread into homogenous communities than if the communities were integrated into heterogenous settings. Figure 3 shows the proportion of students in classrooms in five regions of Ethiopia that speak the official language of instruction (LOI) at home. In the Addis region, nearly all the classrooms have fewer than 20 percent of students who do not speak Amharic at home. In contrast, classrooms in Afar have at least 20% of students who do not speak the LOI at home, and that proportion varies greatly across the region. A linguistically homogenous community would lead to a classroom where all children share the same home language. Linguistically heterogenous settings, such as urban centers made more diverse through migration, can lead to classrooms with children speaking many different home languages with varying levels of linguistic proximity. For example, the average number of home languages per classroom in the Ethiopian dataset was 1.2 in Amhara, 3.2 in the Southern Nations, Nationalities and People's (SNNP) region, and 4.0 in Addis, having important consequences for the choice of LOI. The technical consequences – recruiting, training and deploying *teachers*, and developing curriculum, materials and assessments – are discussed below.

Figure 3. Proportion of pupils in the classroom who do not speak the LOI at home across 5 regions in Ethiopia.



Note. Data from Young Lives Ethiopia survey conducted in 2012-2013 in 280 4th and 5th grade classrooms across 7 regions. This figure includes data from 5 regions. SNNP is Southern Nations, Nationalities and People’s region. The sample is not meant to be representative.

Moreover, it is important to examine the linguistic context beyond a simple count of the primary language of the home, community and state. Among the factors that influence language development of multilingual children are the quantity and quality of exposure to each language¹⁴ (Hoff, 2018). Most research on multilingual language development, however, assumes a clear distinction between languages, and often a well-

¹⁴ Hoff (2018) defines higher quality exposure as hearing the language from adults who are native speakers or proficient in the language.

defined hierarchy between languages. Many people around the world, however, live in communities where multilingualism *is* the lingua franca (Desai, 1995; Di Carlo, Good, & Ojong Diba, 2019). In addition, sociolinguistic research is increasingly moving away from defining and labeling languages as discrete and with clear borders (Saraceni & Jacob, 2018). This move is motivated in part by a critical interpretation of the monolingual nation-state ideologies motivating previous research (Saraceni & Jacob, 2018), and partly on the recognition that the borders between languages and dialects are fuzzy, and mainly social constructs (Haugen, 1966). Language varieties historically exist alongside each other in a geographical space creating a dialect continuum, where two neighboring varieties differ slightly, but with difference accumulating over distances creating varieties that are no longer mutually intelligible¹⁵ (Chambers & Trudgill, 1998). Labeling a variety of the Scandinavian dialect as Norwegian or Swedish makes political and geographical sense but is arbitrary when it comes to linguistic features (ibid., p. 7). Multilingual policies thus build on or create arbitrary boundaries between dialects that must take into account the sociolinguistic factors of the communities.

Key questions:

- What language(s) do the children in this community/state speak at home and with their peers?
- What language(s) does the local administration/central government use in its institutions, and why?
- What language(s) are used in the formal and informal private sector in the community, state, and larger region, and why?

¹⁵ Chambers and Trudgill (1998) also describe a dialect continuum that exist across social classes. For example, Jamaicans of different social classes speak along a continuum between English and Creole, with no clear border distinguishing the two.

2.2. Community engagement and support

A successful language policy must account for two important factors related to communities: the level of engagement of the community in the educational process and the extent to which communities support the language choice. Parental involvement in their children's education – including communicating with children about school, supporting homework, reading with children, and participating in school events – has been shown to be positively associated with student learning outcomes (Hill & Tyson, 2009; Jeynes, 2005). Efforts to increase parental involvement in school management can also improve some measures of education quality (Gertler, Patrinos, & Rubio-Codina, 2012). A mismatch between the language of school and the language at home can create a barrier to parental engagement, such as with immigrant families or when parents have limited formal education (García Coll et al., 2002; Ji & Koblinsky, 2009). Considering education as a process children undergo year-round and all day long, rather than only through schooling, through engaging parents and communities can increase the potential for learning, especially in low-income settings where it is estimated children spend less than 20% of their annual wake time in school (Friedlander, Arshan, Zhou, & Goldenberg, 2018).

In addition to using a language conducive to parental engagement, multilingual policies must recognize the preferences of the community. Beyond a simple language mapping identifying the number of groups and speakers within these groups, the relationship between these groups should inform language choices. As Batibo (2005)

points out, people living in linguistic minorities face a dilemma between the desire to “maintain their linguistic, cultural, and ethnic identity,” and the wish to “integrate in the wider community so as to have access to education, paid jobs and interaction with the wider world” (p.55). Many studies using surveys have documented a preference from parents who speak non-dominant languages at home for the use of a dominant language in schools, such as in Botswana (Arua & Magocha, 2002), Pakistan (Mustapha, 2012), or Ghana and Burkina Faso (Beyogle, 2015). By forcing a dichotomous preference, however, these surveys do not capture the nuances of parental attitudes toward language choice. In addition, multilingual policies are often mistakenly presented or understood as successive monolingual policies, obscuring the fact that the dominant language is introduced gradually. Policies that introduce new languages in schools require both understanding parental preferences, and how they are connected to perceived power and opportunity, and to inform parents of the pedagogical advantages of using familiar languages in the school (Tembe & Norton, 2008).

Key questions:

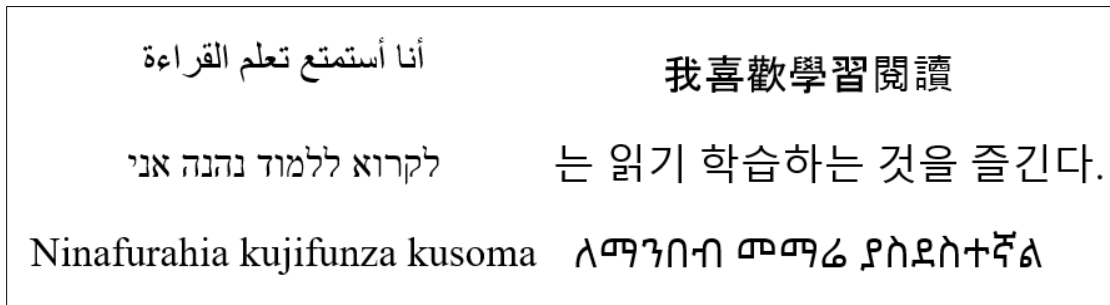
- How are adults in this community/state engaged in supporting the education system?
- What language(s) do the adults in this community want their children to learn in schools, and why?

2.3. Orthographies

Orthographies are important to understand the appropriateness of multilingual education policies in two specific ways: the adequacy of the codification of written languages used in the classroom, and the extent to which writing systems are similar

across languages. The process of developing writing systems across human history has resulted in many different writing systems, as illustrated in Figure 4. An adequate orthography has several important requirements, including phonological adequacy, simplicity, sufficient materials, and community ownership (Bradley, 2003). While the most stable and widespread world languages have had centuries or millennia to develop these, nearly half of the world's languages are not written or have writing systems not widely used (Ethnologue, n.d.), and many have only developed an orthography recently, often through the efforts of religious groups (Nabea, 2009). The process of developing new orthographies or improving existing orthographies requires both linguistic expertise of a technical nature, as well as a process in which a community “adopts” the orthography. The choice of a script (Latin or Arabic), for example, is a signal that a language resembles another, and as Cahill (2014) explains: “all orthographies are political.” From a more practical perspective, introducing curricula for certain subjects, especially science and technology, in new languages requires developing new words. Without the proper technical and political support, developing or updating an orthography can result in competing or unstable orthographies with small consequences, such as the French using the word “email” while the Quebecois use “courriel,” or more problematic issues such as instructional materials being developed using different scripts, spelling or grammatical rules. Corpus planning is thus an important component of language planning and has important consequences on the ability to implement multilingual policies.

Figure 4. Illustration of written systems



Note. Sentences all mean “I enjoy learning to read”. Languages starting from top left and going down: Arabic, Hebrew, Kiswahili, Classic Chinese, Korean, and Amharic. All languages are written horizontally here for ease of display. Arabic and Hebrew are written right to left. Chinese and Korean were traditionally written top to bottom but can also be written left to right. Kiswahili and Amharic are written left to right.

In addition to the adequacy of each language’s orthography, policies and curricula need to account for the distance between languages a child will need to learn as they progress through the system.¹⁶ The ability to transfer basic decoding skills from one language to another varies depending on the distance between these languages. When comparing three groups of bilingual children, Bialystok et al. (2005) found that those having learned Spanish and Hebrew – two alphabetic writing systems – were able to transfer decoding skills to English, but those having learned Chinese – a character-based writing system – did not. Piper and van Finkel (2017) found that the type of scripts across six languages influenced the initial reading acquisition of children in Ethiopia. Beyond basic decoding, the reading process differs across orthographies according to their transparency¹⁷ (Katz & Frost, 1992). A child in Tanzania who has learned to read in

¹⁶ The distance between two languages refers to the degree of similarities or differences between two languages, and sometimes referred to as linguistic distance (Renfrew, McMahon, & Trask, 2000). Some efforts to quantify this distance exist (e.g. Petroni & Serva, 2010) but there is no widely accepted measure.

¹⁷ Orthographic transparency refers to the extent to which the mapping between graphemes and phonemes is consistent (Sprenger-Charolles, Siegel, Jiménez, & Ziegler, 2011). For example, Spanish or Kiswahili are very transparent but English is very opaque.

Kiswahili (bottom left) will require less instruction to learn to decode English than her peer in Ethiopia who has learned Amharic first (bottom right).

Key questions:

- Are the languages used in this community/state written and are the orthographies adequate and recognized by institutions and their speakers?
- Who is responsible for developing and maintaining the corpus of languages?
- What is the linguistic distance between the languages a child would have to learn throughout their formal education?

2.4. Availability and Representativeness of Curriculum and Materials

The development of appropriate curricula and sufficient quality materials to support learning in each language is a key factor of the success of a multilingual program. The need to invest in teaching and learning materials has been widely advocated (e.g. (UNESCO, 2015b). These calls are based on the recognition that there are often few materials available in local languages (UNESCO, 2016b). A survey of reading materials appropriate for the early grades in local languages across 11 African countries found books available in 200 languages. However, there were fewer than 20 books in 141 of these languages, including 40 for which there was only 1 book (RTI International, 2015). In low- and middle- income countries, the availability of books does always result in books in student's hands. Read (2015) estimated that the textbook-to-pupil ratio at elementary schools in nine African countries ranged from 1:1 to 1:15 depending on the country and context.¹⁸ While making textbooks and teaching materials available is unlikely to have impact on learning on its own (Glewwe, Hanushek, Humpage, &

¹⁸ Urban settings had lower ratios than rural and remote areas. See Table A.27 page 240 for full estimates.

Ravina, 2011; Glewwe, Kremer, & Moulin, 2007; Sabarwal, Evans, & Marshak, 2014), these materials allow teachers to modify their pedagogy to improve learning outcomes if there are enough in the classroom (Frölich & Michaelowa, 2005).

In addition to the need for exposure to written language, the inclusion (or exclusion) of languages in the books children and teachers have access to likely reinforces linguistic hierarchies. Representation matters in many dimensions. An increase in the proportion of women members of parliament can increase women's beliefs in women's ability to govern (Alexander, 2012), as well as increase women's substantive participation in local governance (Beaman, Pande, & Cirone, 2012). It is likely harmful for children to encounter books that have primarily white male characters as reading fiction is an important mechanism through which young people develop empathy (McIntyre, 2017). In his seminal book *Decolonizing the Mind*, Ngũgĩ wa Thiong'o (1986) explains why he writes in his native Gikuyu language: "with that harmony between himself, his language and his environment as his starting point, [the Kenyan child] can learn other languages (...) without any complexes about his own language, his own self" (p.28-29). The availability of books in the languages meant to be used in school can thus help support instruction directly, as well as through increasing the perceived importance of the languages themselves.

Key questions:

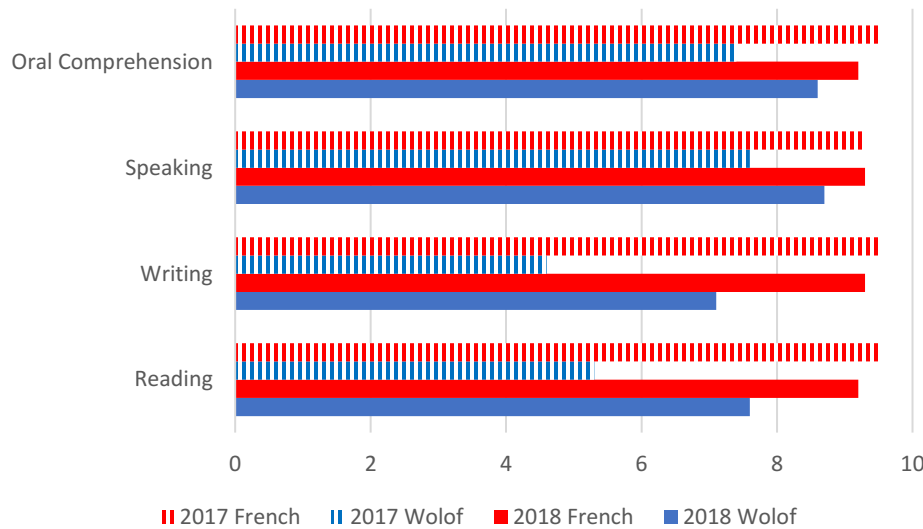
- Who is responsible for developing curricula and lesson plans? Are they available across languages of instruction?
- How many books of different levels are available across languages, and are they being used in school? What institutions are responsible for developing and distributing books?

2.5. Teacher Recruitment, Training and Deployment

A successful multilingual policy requires that classroom teachers are willing and able to teach in the language(s) that are prescribed. Teachers facing their students become the “street-level bureaucrats” (Lipsky, 1969) that make decisions around whether and how to implement an official language policy, or, as Hurdus and Lasagabaster (2018) argue, make their own policy. For example, Cincotta-Segi (2011) found in a rural mountainous area of Laos, where the monolingual official policy prescribes the use of Lao, that teachers used different approaches to teach reading in Lao, all of which included some use of Kmhmu, the language of the community. The strategies used by teachers were largely based on their proficiency in Kmhmu. The challenge to develop a cadre of teachers with both the necessary education level and the language skills to implement a multilingual policy stems in part from the cyclical nature of education systems. Teachers currently in the classroom are mainly graduates of the system themselves. For linguistic groups with few speakers, the availability of individuals having enough formal training and able to teach in that language is often limited. Even for larger linguistic groups, if the language policy in place when most teachers were students was monolingual, the confidence and ability of teachers to use their L1 in reading and writing can be limited. Figure 5 below presents how a sample of first-grade teachers in Senegal self-report their oral and written abilities in French and Wolof, the language of the community where they taught. Teachers were interviewed in 2017 prior to having ever taught using Wolof, and after one year of teaching and receiving support to teach in the language. These figures show that teachers overall have lower levels of confidence in

their Wolof abilities than in French, especially in writing and reading, but that after one year of teaching and support, the gap between the two was narrower.¹⁹

Figure 5. Teacher self-reported comfort level across language skills in Wolof and French



Note: Teachers self-reported their skills on a scale from 1 (low) to 10 (high). Figure adapted from de Galbert, Rousseau, and Mount-Cors (2018).

The process of training, recruitment and deploying teachers varies tremendously across countries, and bringing multilingual education to scale in a given state requires alignment with these processes. A comparative study of pre-service training programs across six African countries showed, for example, that the length of the training ranged from 1 to 3 years, and the time effectively spent in classrooms for a practicum ranged from 9 weeks to 1 school year, and the language used in training colleges was the

¹⁹ The same patterns were found in the Pular and Seereer speaking communities. For more detail, see the de Galbert, Rousseau & Mount-Cors (2018).

colonial language in five of the six countries despite some of the countries using local languages in the early grades (Pryor, Akyeampong, Westbrook, & Lussier, 2012). An ideal teacher professional development system is a “long-term process that includes regular opportunities and experiences planned systematically to promote growth and development in the profession” (Villegas-Reimers, 2003, p.12). Systems that include multiple languages in the classroom should integrate language specific training across teacher professional development opportunities.

The responsibility to recruit and deploy teachers varies across levels and states from highly-decentralized systems where schools are in charge to centralized systems where the ministry of education bears the duty (Bennell, 2004). One important equity issue in many contexts is the ability to deploy enough qualified teachers to rural or remote areas (Bennell & Akyeampong, 2007). Implementing more centralized policies, as has been the case in China since 2001, does not necessarily result in more equitable deployment (Han, 2013). Multilingual policies must thus make special provisions to recruit and support teachers working in linguistic communities that that live in remote or rural areas in order to avoid increasing inequities.

Key questions:

- What are the languages used as LOI and taught as subjects in teacher pre-service and in-service training?
- How are teachers recruited and deployed at each level of the system, and is language proficiency a criterion?
- What is the extent of teacher shortage and language skills across communities?

2.6. Assessment and examination

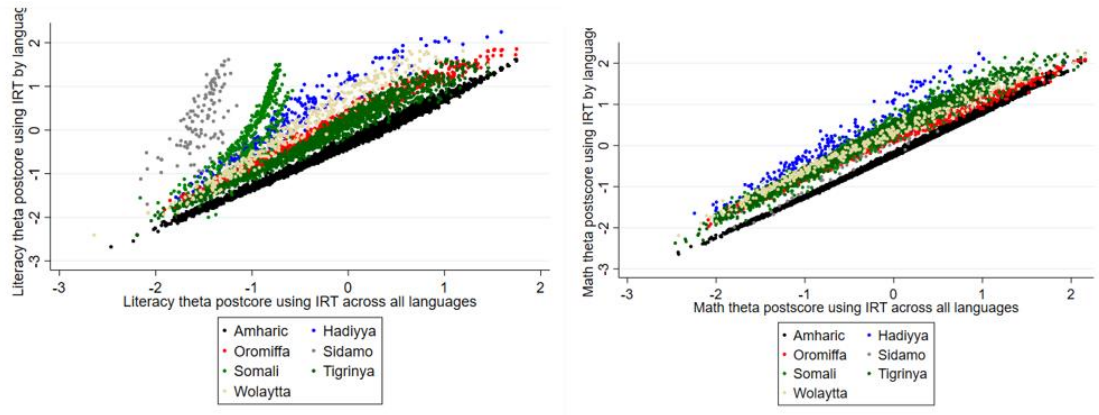
Assessing student learning takes several forms and serves different purposes for stakeholders in education systems. Formative assessment is essentially meant to provide feedback to teachers and learners about the current level of understanding of students to determine future steps, whereas summative assessment is meant to describe learning achieved at a certain point, usually the end of a course or year, for reporting to parents or policymakers (Harlen & James, 1997). An education system or a school designing a battery of assessments must therefore answer two important questions: *what are the goals of the assessments*, and *what kind of learning does it want to achieve* (Gipps, 2002)? A multilingual system using an early-exit model needs a different set of assessments than one implementing a developmental bilingual program throughout the primary cycle. In a multilingual early-exit model, a system requires a battery of assessments to measure only the most basic literacy and numeracy skills in each language, with a common set of assessments in the dominant target language for more advanced skills.

Designing assessments to measure language and literacy development can be particularly complex. Formative assessments meant to provide immediate feedback to teachers can be designed flexibly. In contrast, norm-referenced tests that help design curricular and pilot education system require complex psychometric properties. The validity of these tests is reduced when assessing second-language learners if the norms were created with monolingual students, which is often the case (García et al., 2008). Common accommodations provided when assessing students across languages include providing more time, allowing students to answer questions in their preferred language,

and translating tests (August & Hakuta, 1998; Koretz, 2008). Testing children in the language of the classroom is helpful for formative assessments. Trying to compare the performance of students across language groups to inform policy, however, can be challenging. Figure 6 presents the association between literacy and math test scores of a sample of fourth and fifth grade students in Ethiopia computed using item-response theory (IRT) within and across language groups.²⁰ The left-panel (literacy) shows a weak association between the two measures, with students identified as high-ability within the Somali or Sidamo group “dropping” well below the overall average when the items are combined across languages. In contrast, the scores in the math test are nearly identical whether the estimates are done within or across language groups. This example illustrates the need to carefully design and interpret assessments across language groups in multilingual contexts. In addition, it poses questions of equity for systems with high-stakes end-of-cycle examinations that impact whether students can advance to the next level.

²⁰ IRT is a statistical modeling method that uses items from a test to produce an estimate of a single underlying ability (literacy and math in this case) for each student (Harris, 1989; Yen & Fitzpatrick, 2006). This results in each student having a theta score with an overall mean of 0. Positive scores denote students with above-average “ability,” and negative scores indicate scores below-average.

Figure 6. Association between IRT literacy and math scores computed within and across language groups



Note: Data from the 2012-2013 Young Lives School Survey.

Key questions:

- What are the formative and summative assessments available across languages, subjects, and levels?
- In what language(s) and subjects are the end of cycle high-stakes examinations conducted?

3. Education Language Policies in Kenya and Uganda

The last section describes (i) the education systems and official language policies in Uganda and Kenya, (ii) the sociolinguistic context in the two countries, and (iii) important elements to understand the extent to which the language policies can be implemented effectively. As the lens through which the policies are analyzed is that of the state, the goal of this section is not to answer all the key questions included in the framework for all the communities in these two countries, but rather understand the explicit goals of the two language policies, and the systemic factors that can support or impede their implementation.

3.1 Education systems and official language policy

The structure of the education systems in Kenya and Uganda were similar until the 2018 introduction of the Basic Education Curriculum Framework (BECF) in Kenya. The system in Uganda follows a 7-4-2 structure with seven years of primary school, four years of junior secondary and two years of senior secondary school ((Republic of Uganda, 2008).²¹ The education system in Kenya was organized in an 8-4-4 structure with 8 years of primary education, 4 years of secondary school and four years or more of university (Republic of Kenya & UNESCO, 2012). The BECF has transformed the structure of the Kenyan system into three levels: (i) *Early Years*, which includes pre-primary and three years of primary school; (ii) *Middle School*, with 3 years of upper primary school and three years of lower secondary; and (iii) *Senior School*, with three years of upper secondary school and tertiary institutions. In both countries, the two years of pre-primary included are not compulsory and not funded by the government. Despite the differences in structures of the system, the **official** language policy in the two countries remain very similar.

In Kenya, the BECF mandates the use of the language of the catchment area as the medium for pre-primary, and the use of the “first language of the child” for literacy instruction in early primary (P1-P3) (Kenya Institute of Curriculum Development, 2017). The BECF also recommends that children be “exposed to [Kiswahili] at the earliest possible time in their schooling” and that “learners should be taught the foundational skills of reading and writing the English language at the earliest opportune time” (KICD,

²¹ Following the former British system, junior secondary school is referred to ordinary or O level, and senior secondary school is called advanced or A level.

2017, p.33). As the language suggests, this language policy provides local actors with opportunities to implement a range of language use in the classroom. This multilingual approach was not newly introduced by the BECF, as it was already the official policy stated in the 1992 syllabus (Trudell & Piper, 2014). In contrast to the policy that allows for local languages to be used for instruction in the early grades, studies conducted in Kenya have found that teachers express preference toward using English over local languages (Sure & Ogechi, 2009), observed teachers mainly using English or Kiswahili (Piper, 2010a), and found teachers justifying their use of English and Kiswahili at the expense of local languages (Dubeck, Jukes, & Okello, 2012). In general, the interpretation of the policy has resulted in English and Kiswahili being primarily used in Kenyan schools (Mose, 2017).

The policy in Uganda since 2007 is to promote the use of a “first or familiar language” as the medium for the three years of early primary, with P4 as a transition year toward the use of English (Republic of Uganda, 2006). The explicit goal of the policy is to equip children at the end of three years with “sufficient skills in English to act as the basis for developing English as the medium of instruction” for the rest of the education cycle, clearly aligning with traditional early-exit models (ibid, p.7). Similar to the BECF, language in the Ugandan curriculum provides for local decision making, such as stating that children should learn in “his/her home language or at least a language that is familiar” “wherever possible” (ibid, p.9). In contrast to Kenya, where Kiswahili and English seem to be used most often starting in the early years, local languages seem to serve as the main language of instruction in the early years. Classroom observations in

2nd grade reading lessons conducted by local officials showed that nearly 90% of teachers used the local language (Uganda National Examinations Board, 2017). Similarly, an observational study in 620 classrooms found the local language was used over 70% of the time in early primary classrooms (Piper, 2010b). In 2015, 72% and 32% of head teachers in rural and urban schools respectively participating the annual National Assessment of Progress in Education reported using a local language in the lower-primary grades.²² The NCDC conducted a small number of classroom visits in 2017 to assess the teaching of local languages in upper primary that suggested low levels of implementation (National Curriculum Development Center, n.d.). These observations suggest that schools in Uganda generally follow the policy of using local languages for three years, and English afterwards.

The education systems in Kenya and Uganda are similar to each other, and similar to those of many low- and middle-income countries, in that access to school increased drastically in recent decades following the elimination of school fees. In Kenya, the abolition of school fees at the primary level in 2003 led to a substantive increase in the completion rate, reaching nearly 80% in 2008 (Ngugi, Mumiukha, Fedha, & Ndiga, 2015). A similar policy abolishing fees for day secondary schools in 2008 has led to a small increase in access, with the net enrollment ratio reaching 33% in 2010, including large regional variation (Republic of Kenya & UNESCO, 2012). In Uganda, the elimination of school fees through the Universal Primary Education (UPE) policy in 1997 led to the number of students enrolled in primary school to nearly double from 3 million

²² Author's calculation using data from the 2015 NAPE assessment.

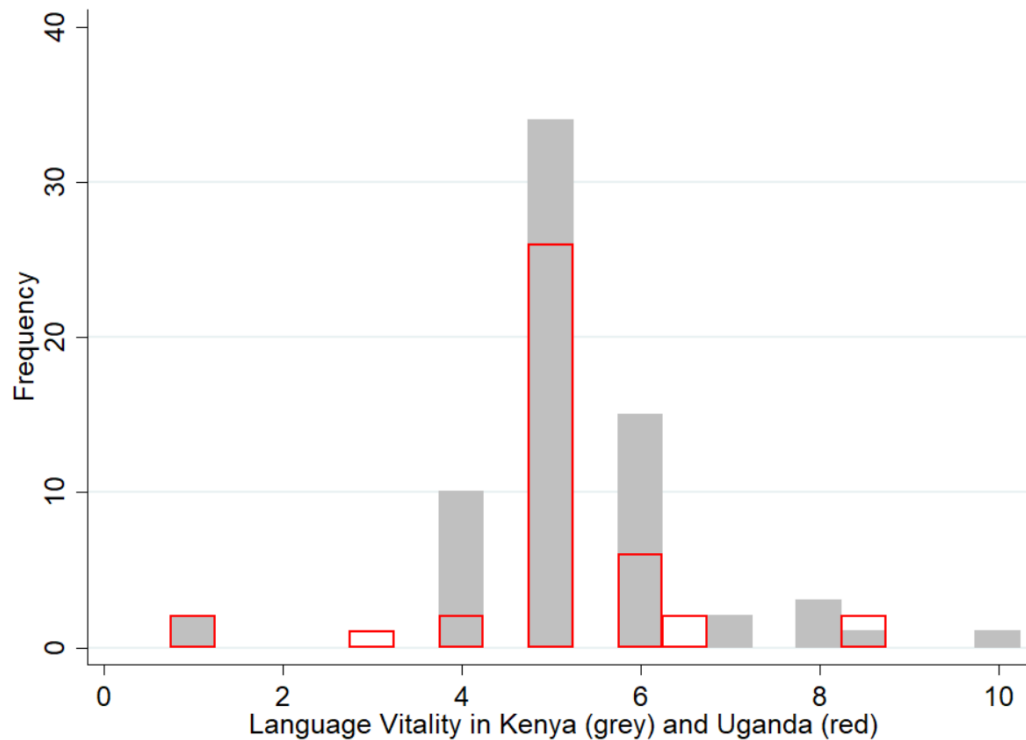
in 1996 to 5.8 million in 1998 (UIS, 2019). Following the introduction on Universal Secondary Education (USE), enrollment increased from approximately 700,000 to 1,080,000 students (54% increase) between 2006 and 2010; the transition rate from primary to secondary school also increased by 20 percentage points (UIS, 2019). These large increases in access to schooling were driven by removing entry barriers to the most marginalized groups, including linguistic minorities, thereby increasing the need to pay attention to the language spoken by children.

3.2 Sociolinguistic context and community support

Kenya and Uganda are neighboring countries located in East Africa, with a rich set of linguistic and geographical diversity. Figures A2 and A3 in the appendix include maps of the two countries with languages identified in *Ethnologue* and the families to which they belong. Kenya's population comprises 42 ethnic communities that includes three of Africa's major sociolinguistic groups: Bantu (67%), Nilotic (30%), and Cushitic (3%) (Republic of Kenya & UNESCO, 2012). The Ugandan constitution recognizes 56 indigenous communities (Republic of Uganda, 2005) and there are currently 43 living languages (Simons & Fennig, 2018). Ugandan languages are mainly Bantu (70% of the population), Nilotic (15%), or Central Sudanic (15%) in the northwest of the country, with a small number of languages belonging to the Kuliak family (Simons & Fennig, 2018; Tembe, 2008). Figure 7 below presents the distribution of vitality of languages in Kenya and Uganda according to the *Expanded Graded Intergenerational Disruption Scale* (EGDIS) meant to identify if languages are at relative risk of extinction, with values ranging from 1 (International language) to 10 (Extinct). The two distributions are

relatively similar, and the majority of languages in the two countries are scored between 4 and 6, indicating the languages are used regularly, and considered vigorous, some having abundant literature and being used in education systems. A handful of languages in both countries are threatened, or nearly extinct. The numbers presented here, however, should be taken with caution, as both the delimitation of languages and the measure of their vitality is not a precise exercise. For example, competing estimates in Kenya identify 13 languages as highly endangered (Obiero, 2008), more than twice the number identified on the EGDIS scale.

Figure 7. Vitality of Languages in Kenya and Uganda.



Note. Data from Ethnologue (2016). See table A1 in the appendix for the full description of each level.

Both Uganda and Kenya were under the colonial rule of the Great Britain from the late nineteenth century until gaining independence in 1962 and 1963 respectively, which resulted in English being an official language. The colonial experience in the two countries, however, was very different, with the British settling in large numbers in Kenya and applying direct rule, compared with a very small number of settlers in Uganda, and indirect rule (Mazrui & Mazrui, 1998). The divergences between the two countries continued after independence, with Kenya's language policies and economy creating lasting international ties and promoting English (Nabea, 2009), compared with Uganda's insularity through several brutal regime changes and a promotion of local languages (Tembe, 2008).

The place of Kiswahili in Kenya and Uganda is also very different. While Kiswahili is an official language in both countries, it plays very different roles in the lives of people, and in schools. In Uganda, Kiswahili is officially supposed to be taught as a subject in schools starting in the first grade, but is only taught in practice in some schools, usually starting in the upper primary grades (Tembe & Norton, 2011). Kiswahili is the home language few Ugandans and used daily by a small proportion of the population, mainly those living near the borders with Kenya, Tanzania and the Democratic Republic of Congo, as well as market vendors and members of the army (Tembe, 2008). Kiswahili was heavily promoted in Uganda under the military rule of Idi Amin (1971-79), which also contributed to its loss in popular support in the country since then (Mazrui & Mazrui, 1998; Namyalo & Nakayiza, 2014). In contrast, Kiswahili is used prominently in the daily lives of many Kenyans in home and schools. Kenyans living along the coast

primarily speak Kiwashili as a first language it is used as a lingua franca in the country, as well as in official capacity in some government functions (Nabea, 2009). In Kenya, Kiswahili is arguably the language of national unity as it dominates the social interactions (Samper, 2002). As a result of these two diverging experiences, Kiswahili has a much more prominent role in Kenya than in Uganda.

The extent to which there is support from the overall population today toward the use of local languages, Kiswahili, or English, as the MOI in schools is difficult to assess in either country. The Afrobarometer, which uses nationally representative samples to collect public attitudes toward democracy, governance and society (Afrobarometer, n.d.), shows that 46% of Ugandans and 65% of Kenyans believe their government is ‘better’ or ‘much better’ in its effectiveness at addressing the educational needs of the country. The data collected on language, however, do not address the language education policies directly. Opinion pieces and news articles in both Uganda (Komakech, 2014; Natukunda & Businge, 2016) and Kenya (Kiplang’at, 2014; Odhiambo, 2019) present anecdotal evidence of strong opinions for and against the use and promotion of local languages in schools and universities. The relative size of the support for the promotion of local languages in Uganda is illustrated by the petition sent to the speaker of the parliament by a large group of civil society organizations when it appeared that the secondary school reform would lead to the elimination of local language instruction (Mulumba, 2012).

When asked about the multilingual education policy in place in Uganda, parents expressed a range of opinions.²³ Some parents viewed benefits from a cultural perspective: “it has some good advantages. Because (...) [a learner] picks that foundation of that language, the culture, definitely will be protected for the future.” Others saw the positive pedagogical value of scaffolding learning through a familiar language: “you first bring that kid [on]board (...) you plant him or her in that local language, local languages he can understand, so it is good.” In contrast, other parents preferred the use of English exclusively because of its utilitarian value: “in order to prepare someone for the future, you must be knowing English.” Parents also identified some challenges of implementing such a policy such as recruiting and deploying qualified teachers: “that implies that if at all the area doesn’t have teachers, the school will not, the people will not be able to study. That one cannot work at all.”; as well as some consequences for national unity and integration of different linguistic groups: “the students who studied in Busoga, or Ankole, will not be able to get a job here, because in central here, they are speaking Luganda, so they have been studying in local language and they cannot be transferred to another area.” Importantly, however, several parents expressed an ideal situation that included literacy taught in multiple languages: “Luganda would be one of the recommended languages, then followed by English, then followed by Kiswahili, at least those three to be a proper Ugandan, he should be in a position to write those three.”

²³ Data presented from parents in this section were collected during a pilot study conducted in 2014. Appendix B describes the data collection and analyses process.

3.3 Supporting the implementation of the language policies

As described in the framework above, there are four important technical elements that need to be in place to support multilingual education policies: (i) stable orthographies; (ii) curricula and materials; (iii) systems to recruit and support teachers; and (iv) appropriate assessments systems. This section describes how these elements in Kenya and Uganda enable or inhibit the implementation of the early-exit transitional models called for by the two official language policies.

The extent to which orthographies of Kenyan and Ugandan languages is adequate for instruction in schools varies greatly. As with the vitality described in figure 7, languages in both countries range from being unwritten, to having fully developed and stable orthographies that are widely used. In Uganda, Nzogi (2011) estimated that 20 languages either did not have an orthography or had an inadequate one. There is, however, a stark contrast between the approach taken in the two countries by the government and civil society, along with funding partners, toward languages that need to develop or improve their orthographies. In Kenya, the main focus of the government has been to support literacy instruction starting in the first grade through a nation-wide reading program called Tusome that only measures progress in English and Kiswahili (Freudenberger & Davis, 2017). There has been no systematic effort to develop or improve orthographies, though some efforts have taken place to revitalize languages (e.g. the Suba language described in Obiero 2008). In contrast, the government of Uganda, along with education development partners, has led several early reading programs (e.g. School Health and Reading Program (SHRP); Literacy Achievement and Retention

Activity (LARA)) that explicitly aim to support instruction in local languages (RTI International, 2016c). In Uganda, the institutional authority to develop languages, including developing orthography, is given to *Local Language Boards* that include representatives from the government, cultural institutions such as kingdoms, and civil society. While the capacity of local language boards is sometimes limited in human and financial resources, the government, through the National Curriculum Development Center (NCDC), and its partners have invested heavily in supporting them, which has resulted in a large number of languages having developed a stable orthography in recent years (e.g. Heugh & Mulumba, 2014; Odongo, 2012; SIL LEAD, 2015).

The availability of curricula and materials that can support instruction in local languages in the two countries varies across languages. While the most prominent languages have many children's books published, there are not enough instructional materials in many others, both curricula and books aimed at supporting young children to learn how to read. For example, (Nabirye & De Schryver, 2010) found very few materials in Lusoga (in Eastern Uganda), illustrated by the fact that the only dictionary available targets individuals with more than 7 years of education, clearly not appropriate for P1-P3 children. The Kenya Institute of Curriculum Development (KICD) is the national educational research and curriculum development center in Kenya responsible for developing and approving curriculum and materials for all levels of the education system (Republic of Kenya, 2013),²⁴ and the NCDC holds the same responsibilities in Uganda. As with the development of orthographies, the differences in focus at the national-level

²⁴ KICD replaced the Kenya Institute of Education (KIE) in 2013.

have resulted in more progress toward developing instructional support in Uganda than in Kenya. Although the initial introduction of the 2007 policy in Uganda included few instructional supports in the form of local language guides and materials (Namyalo & Nakayiza, 2014), the same institutions responsible for the support to orthography development have developed curricula and instructional materials in local languages.

Of the factors described in this section, the recruitment, training and deployment of teachers to support local language instruction is the one with most similarities across Uganda and Kenya. Pre-service teacher training in Kenya is done through one of 85 Teacher Training Colleges, and admission is contingent on passing the English exam at secondary school, because it “is the language of instruction in Kenya’s education system and teachers are expected to be competent in its use” (Republic of Kenya & UNESCO, 2012, p.124). In Uganda, candidates to become primary school teachers through training in the Primary Teacher Centers (PTCs) must have completed the 4-year cycle of junior secondary level and passed English language and mathematics exams (Tembe, 2006). In both cases, the language of instruction is English, and the training received is focused on teaching in English, with no support in local languages, and limited time for practice in the classroom (Akyeampong, Lussier, Pryor, & Westbrook, 2013; Najjumba & Marshall, 2013; World Bank, 2014). In-service training is thus where support for local language instruction could take place in both countries. The support to teachers for reading instruction mirrors the national reading programs in the two countries: Tusome supports English and Kiswahili in Kenya, while SHRP and LARA support English and some local languages in Uganda.

Given the differences described above in the use of local languages in the classrooms and the institutional supports in place, it is not surprising to also note that local language assessments are much more prevalent and used in Uganda than in Kenya. The Uganda National Bureau of Education (UNEB) has adapted the Early Grade Reading Assessment (EGRA) to eleven local languages and conducted several nationally representative studies to assess early literacy skills (Uganda National Examinations Board, 2017). While UNEB does not yet have the capacity to assess early learners in all languages used in the system, the current effort presents an important signal toward the importance of monitoring learning across languages. It is also worth noting that the large citizen-led literacy assessment conducted by Uwezo in East African countries assessed children in Kenya in English and Kiswahili only, while children in Uganda were assessed in English and seven local languages (Uwezo, 2017). Paradoxically, the new basic education framework introduced in Kenya theoretically allows for greater use of local languages as it proposes to abolish the end of primary school examination. One of the important reasons parents and teachers argue for the teaching of English over local languages for their children is the fact that high-stakes examinations determine both the ability to progress in the system, and often the schools in which they will get admitted. Recognizing the detriment that high-stakes examinations pose to learning (e.g. narrowing of teaching, cheating), the new system in Kenya proposes to replace these end of cycle examinations with continuous assessment (Kaviti, 2018). It remains to be seen if this shift away from high-stakes examination in Kenya leads to opportunities for students to demonstrate their learning in languages other than English and Kiswahili.

This section described the education system and official language policies in Kenya and Uganda, as well as the two countries' linguistic history and current sociolinguistic contexts. Although Kenya and Uganda are neighbors, share a number of linguistic communities, and have similar official language in education policies, the countries' history has led to different attitudes toward English, Kiswahili and local languages. As Mazrui and Mazrui (1998) note: "Seldom have two African countries so geographically close to each other, and once ruled by the same imperial power, manifested sharper contrasting characteristics in the domain of language and political culture" (p. 159). It is also evident that despite sharing a similar official policy toward language use in the classroom, the level of institutional support and resources available are vastly different. While it is undeniable that systems in place in Uganda are more supportive of local languages, the interpretation of these findings must be done with regards to specific questions. Proponents of multilingual policies to include all linguistic communities will interpret these findings as Uganda displaying relatively more success. In contrast, if the goals prioritized are educational achievement, Kenya's 84% primary and 71% secondary school completion rates will resemble success much more than Uganda's 44% and 26% rates (UIS, 2019). In addition, 55% of Kenyans feel either "only Kenyan" or "more Kenyan than [their ethnicity]" compared with 25% of Ugandans.²⁵ In contrast, 60% of Ugandans never feel that their ethnic group is treated unfairly, compared with 43% of Kenyans. While language of instruction is only one factor contributing to the quality of instruction children receive in schools, or to the extent to which citizens

²⁵ Data from the 7th round of the Afrobarometer. 8% of Kenyans and 15% of Ugandans identify either "only" or "more" to their Ethnic group, while 36% of Kenyans and 61% of Ugandans reported identifying equally to both.

identify with the country more or less than their ethnicity, language policies and their implementation are necessarily judged according to normative views.

Conclusion

This paper had three specific aims. First, it reviewed literature on the history of language in education policy in the context of the nation and state building, dimensions of language policy and planning, and multilingual approaches in schools. Second, it proposed a framework through which language policies in education and their implementation can be interpreted as the state, regional or community level. Finally, it describes the sociolinguistic and educational context in Kenya and Uganda and compared the institutional capacity in place to support the implementation of their respective education language policies.

As explained in the first section of this paper, language policy choices in and out of schools are partly the product of ideologies of nation-state. Visions of statehood based on the one-language-one-nation ideology are more likely to exclude languages of linguistic minorities from the public space, including schools. This can mean the exclusion of groups that speak powerful languages, such as Russian speakers in Latvia, or those that speak endangered languages, such as the Batwa in Rwanda.²⁶ In contrast, visions of inclusive states, exemplified by the constitution of the *rainbow nation* of South Africa, allow for the possibility of shared political and cultural rights. These ideologies

²⁶ From a linguistic perspective, the language spoken by Batwa is sometimes considered as a dialect of Kinyarwanda (the main language spoken in Rwanda). However, the Batwa are a distinct indigenous group that lack recognition in Rwanda because of the “non-tribal Rwanda” policy instituted after the 1994 genocide (Unrepresented Nations and People of the World, 2018).

often influence the approach used in language planning, both status and acquisition, as well as its orientation, seeing language as a problem, a right, or a resource, ultimately resulting in monolingual or multilingual education policies.

The framework presented in the second section of the paper is designed for stakeholders interested in analyzing an education language policy and whether systems are in place at the local or national level to implement the policy effectively. As the analysis of the policies in Kenya and Uganda demonstrate, official language policies cannot be understood devoid of the context in which they are implemented. Although both countries recommend the use of local languages in the early years of primary school, the “street-level bureaucrats” that teachers, principals and local education officials represent in Kenya and Uganda have taken very different decisions. The support toward using local languages in Uganda is not universal, and their relatively more prominent use in classrooms is the results of important efforts by the government and civil society. The potential educational benefits for individuals – such as children learning in Luganda, a language spoken by millions of Ugandans with a stable orthography, many instructional materials and trained teachers – must be juxtaposed with the consequences at the state-level. What are the equity implications when considering the impact of the policy for children who speak Ngakarimojong, a language spoken by fewer than a million people, with scarce instructional materials and few trained teachers?²⁷

²⁷ Similar equity concerns can be asked in contexts where some children need to learn multiple scripts as others do not.

Desai (1995) stresses the importance of assessing the relative benefits mother tongue education for a child's learning with respects to the context in which instruction takes place. This paper argues that the context of the child, the school, and the community need to be analyzed to better understand the merits of the language choice both for the individual and the broader communities in which learning takes place.

Paper 2: Language transfer theory and its policy implications: Exploring interdependence between Bantu languages and English in Uganda

Countries around the world have increasingly introduced or expanded education policies that make use of multiple languages in the early years of primary school. These policies often promote the use of a language familiar to students for the first few years of schooling before transitioning to an official language used in other public institutions. While using a language familiar to the child from the time they enter school could be beneficial for several reasons, multilingual educational policies are partly motivated by research on cross-linguistic transfer, the ability to transfer skills developed in one language to a second language. This research has a strong theoretical foundation, and empirical research has largely contributed evidence of transfer of basic – if not higher-order – literacy skills. This research, however, mainly examines the transfer between European or Asian languages in high-income countries. In addition, recent theories of language development suggest a more dynamic relationship between languages including bi-directional transfer and translanguaging. As a result, policy makers in countries such as Uganda, where the study is based, make decisions on some of the important implementation options of these policies, such the number of years to use respective languages as the medium of instruction (MOI), without having contextually relevant evidence on cross-linguistic transfer or association.

This study aims to describe the performance of Ugandan second-grade students on literacy tasks in one of two local languages – Luganda and Runyankole-Rukiga – and English, and explore whether there is evidence of transfer of literacy skills across these languages.²⁸ In this study, I use data from 3,561 second grade students assessed in oral language, decoding skills, and reading comprehension in two languages using the Early Grade Reading Assessment (EGRA). I use correlational analysis and structural equation modeling to test whether students’ basic alphabetical knowledge skill and higher-order reading comprehension skills transfer from the first to second language. This study adds to our understanding of cross-linguistic transfer between Bantu languages and English. My findings support the current understanding of a positive association between literacy skills across languages found in other contexts. In addition, the findings suggest that the MOI used in school in the early years may impact students’ abilities to transfer higher-order literacy skills. Finally, this chapter argues that there is a need for further research on language practices in multilingual contexts and for better literacy assessments to understand how teachers support emergent bilinguals to build their literacy skills across languages.

Policy Environment in Uganda and sub-Saharan Africa

Uganda, a linguistically and culturally diverse country in East Africa, introduced in 2007 an education policy that promotes the use of national languages for the first three

²⁸ In this study, I use the terms first language, local language and L1 to refer to non-dominant languages used by one community within a linguistically heterogeneous state, and official language as the language with an officially recognized status in the state. In this study, Luganda and Runyankore-Rukiga are the two non-dominant Ugandan languages of interest and English is the official language. Runyankole-Rukiga is sometimes spelled Runyankore-Rukiga. In this paper, I use Runyankole-Rukiga. Occasionally, I use the abbreviation RR in tables.

years of primary school before transitioning to English, the official language (Republic of Uganda, 2006). Several other countries in sub-Saharan Africa also make use of national languages in their formal education system. The overall trend in the region in recent decades has been to integrate more languages in the education system, especially in the early years of primary school (Albaugh, 2014). One important aspect that varies across policies in different countries is the number of years national languages are used as the medium of instruction. Policies range from using only the official language in schools to full multilingual education models that allow for the use of multiple languages throughout the education system. In Togo, for example, French is used exclusively throughout the school system, whereas Eritrea uses local languages as the MOI in the primary years, and South Africa allows for the use of national languages through the entire school system (Albaugh, 2012; Broom, 2004; USAID, 2015). Early-exit transitional models, such as the one used in Uganda, support the use of a first language (L1) as MOI for two or three years before transitioning to a second language (L2) without explicitly promoting the long-term development of L1. The goal of these policies is to develop early reading skills in L1 while teaching L2 (the language used as MOI for the rest of schooling) as a subject. Crawford (2004) notes that while no single policy fits all contexts, policies like the one in place in Uganda are less effective overall than those supporting the L1 development for six years or more. Research in neighboring Kenya, which has an early-exit education policy similar to Uganda, also suggests that an early transition in that context does not allow for the development of strong enough skills in the first language (Piper, Schroeder, & Trudell, 2014).

As noted in the introduction, policies making use of national languages, and the instructional approaches based off of them, are informed in part by the growing body of literature supporting the notion of cross-linguistic transfer, suggesting that individuals can transfer skills they have learned in one language to another. Research on cross-linguistic transfer, however, has focused primarily on European and Asian languages in relatively resource-rich settings. Understanding how cross-linguistic transfer takes place in different linguistic environments and in low and middle-income countries is important to inform policies and guide practice.

Language, Literacy and Cross-linguistic transfer

The first years of formal schooling aim to develop a child's reading comprehension skills to prepare for the transition from learning to read to reading to learn, which tends to happen in the third or fourth year of school. Reading comprehension can be seen as the product of decoding and listening comprehension (Hoover & Gough, 1990) or more broadly as the combination of code-based skills and meaning-related skills (Lesaux & Marietta, 2011). This simple view is helpful to understand the two broad sets of skills necessary to develop literacy. Code-based skills include phonological awareness, phonics and fluency. Phonological awareness refers to the ability to detect and manipulate sounds spoken in language (Ehri et al., 2001). Phonics refers to the ability to connect letters to sounds and fluency is the ability to read accurately and efficiently (Lesaux & Marietta, 2011). Meaning-related skills include vocabulary and comprehension strategies. These skills, as well as other skills included in a more complex view of literacy, can be conceptualized along a continuum from *small problem spaces* to

large problem spaces, the size of problem spaces being associated with the level of difficulty to learn (Snow & Kim, 2007). *Small problem spaces* include more finite skills such as alphabetical knowledge and decoding, which are considered basic skills. In contrast, *large problem spaces* are higher-order skills and include vocabulary and meaning-making strategies. Paris (2005) presents a similar distinction between *constrained skills*, which everyone masters eventually, and *unconstrained skills*, which are distributed between people along a continuum throughout their life (e.g. vocabulary knowledge can grow in depth and breadth without limit). While much research has examined the relationship between these skills and the developmental trajectories of monolingual children, much less is known about these in plurilingual children, especially in low and middle-income countries.

Cummins (1979) first proposed that oral language and literacy skills developed in the child's first language (L1) can transfer to the second language (L2) through a process he described as linguistic interdependence. This theory, suggesting that there is a common underlying proficiency across languages, has guided the research in transfer. Most studies of transfer that examine decoding find positive correlations in alphabetical knowledge, especially when languages are similar (Dressler & Kamil, 2006). For example, Bialystok et al. (2005) found that cross-linguistic transfer in *small problem spaces* was strongly linked to the orthographic contrast – the degree of difference between the orthographies of the two languages. The greatest correlation between L1 and L2 decoding tasks in that study were for Spanish-English bilinguals followed by Hebrew-English, and then Chinese-English. Studies investigating transfer in oral language skills,

usually operationalized as vocabulary, have mixed findings but a meta-analysis suggests that there is a small positive correlation across languages (Melby-Lervåg & Lervåg, 2011). Finally, research on transfer in the large space of reading comprehension and meaning making is limited, but studies suggests a positive association between skills in two languages (Melby-Lervåg & Lervåg, 2011; Proctor, August, Snow, & Barr, 2010). The vast majority of studies of cross-linguistic transfer examine the relationship between Indo-European and/or Asian languages. In a meta-analysis including 47 studies and 52 samples, Melby-Lervåg and Lervåg (2011) found 14 samples with a first language from Asia or the Middle-East and only 1 from Africa (Herero and English). All other samples (37) had a European L1 and 51 samples had a European L2. Similarly, Proctor and Zhang-Wu (2019) searched for studies that assessed cross-linguistic relations with English and found only four language pairs that were involved in more than one study.²⁹

Current theories on the development of language and literacy for bilingual and multilingual individuals include more explicitly the dynamic relationship between languages. García and Wei (2014) argue for a single linguistic system from which individuals can access features of multiple languages through a process called *translanguaging*. This theory expands the monolingual idea of *linguaging* as a process of meaning making through language, by applying it to bilingual individuals, who make meaning and gain knowledge through the use of two or more languages (García & Wei, 2014, p.22). This theory suggests that individuals gain literacy skills irrespective of languages and apply these skills to decode and make meaning of the text from drawing

²⁹ Spanish, Chinese, French and Korean. Other languages with only one study included Chichewa, Farsi, Greek, Italian, Malay, Nahuatl, Oriya, Russian, Tamil, Urdu, and Zulu (Proctor & Zhang-Wu, 2019, p.2).

on their entire repertoire. In practice, this theory promotes instructional approaches that encourage students and teachers to draw on their knowledge across languages with flexibility (Proctor, Haring, & Silverman, 2017). This study uses cross-linguistic transfer and not translanguaging as a lens to understand literacy development because most second-grade students in the context of the study have had very little exposure to the foreign L2, in this case English, so do not have multiple languages from which to draw from.

This study aims to explore cross-linguistic transfer between one of two Bantu languages and English using data from second-grade students living in the Central and South-Western regions of Uganda, where communities mainly speak Luganda and Runyankole-Rukiga, respectively. Luganda and Runyankole-Rukiga are Bantu languages that are spoken at home by roughly one third of Uganda's population. These languages both use the Latin script and Luganda, the language with the most speakers in the country, serves as a language of wider communication, in part because it is spoken in the central region of the country where the capital city and much of the formal economy is located. In the 2002 census, Luganda was estimated to be spoken as a first language by approximately 5 million people in the Central region of Uganda, while Runyankole-Rukiga was estimated to be spoken by 3.9 million people mainly in the Southwestern part of the country (Simons & Fennig, 2018). In contrast, English was estimated to be spoken by 2.5 million people, or 7%, as a second language in the country (Lewis et al., 2016), suggesting very few children are exposed to the official language at home, especially in rural and linguistically homogenous areas.

Research Questions

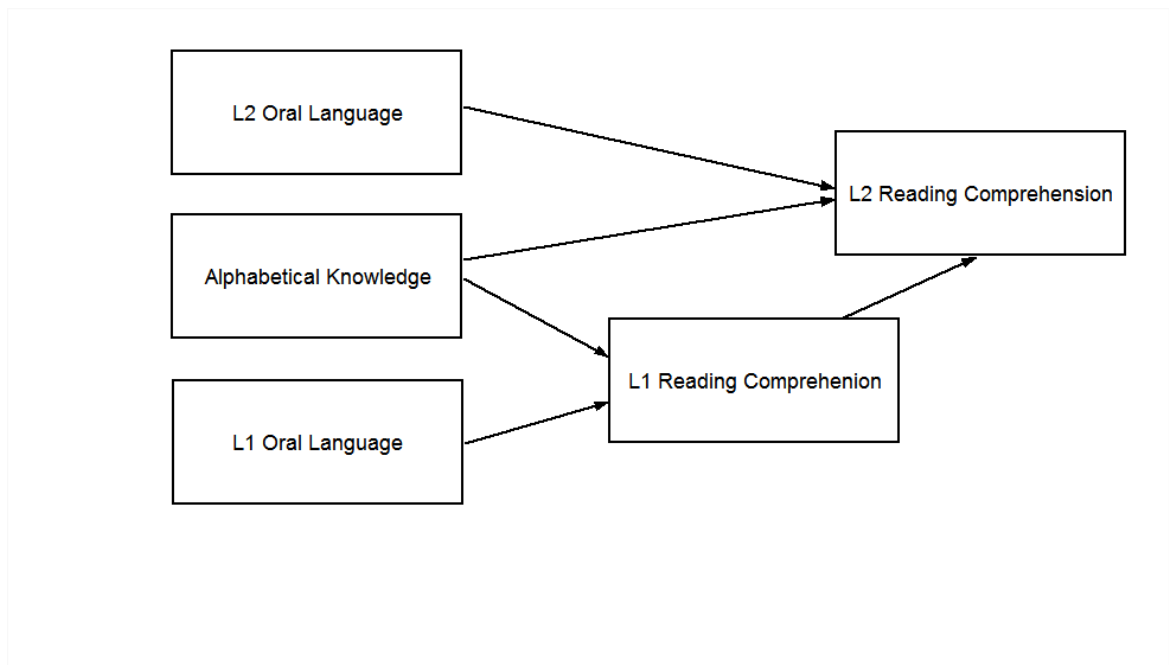
This study answers three specific research questions. The first research question aims to better understand the skills of second grade students in the sample and compare their performance across two regions, while the second and third questions explore cross-linguistic transfer:

- 1) *How do Ugandan second grade students in the Central and South-West regions compare in their performance on assessments of literacy skills in their first language and English?*
- 2) *Is there evidence of cross-linguistic transfer between Bantu languages and English in Uganda that suggests a common alphabetical knowledge skill?*
- 3) *Does a bilingual model of reading comprehension developed with English-Spanish bilinguals fit the data from Uganda in schools reporting English and Ugandan languages as the media of instruction?*

I explore transfer in *small problem spaces* by testing whether the decoding skills of students are related across languages and transfer in *large problem space* by replicating a theoretical model of reading comprehension for bilingual students proposed by Proctor et al. (2010) (see Figure 1). This model expands the simple view of reading (Hoover & Gough, 1990) from monolinguals to bilingual readers. In this model, developed and tested with Spanish-English bilinguals, individuals develop a single Alphabetical Knowledge skill across language, which predicts reading comprehension in both L1 and L2. Oral language in L1 predicts reading comprehension in L1, and oral language in L2 predicts comprehension in L2. Key to the idea of transfer in *large problem spaces*, this model proposes that reading comprehension in L1 predicts comprehension in L2. Given that the languages in this context use the same script, prior

research suggests that alphabetical knowledge should be a single construct across the first and second languages in the sample of Ugandan students (Bialystok et al., 2005; Proctor et al., 2010). Because students are in second grade, I anticipate that decoding skills will be a strong predictor of reading comprehension, and that oral language will be positively associated with reading comprehension in each language. Finally, I hypothesize that reading comprehension in the first language will partially mediate the relationship between alphabetical knowledge and reading comprehension in English, suggesting that there is some transfer in reading comprehension skills.

Figure 1. Model of Bilingual Reading Comprehension adapted from Proctor et al. (2010)



Sample

Data for this study were collected in the Central and South-West regions of Uganda by RTI International through the USAID-funded School Health and Reading

Program (SHRP) at the end of the 2014 school year.³⁰ Schools were randomly selected from a sample of representative districts in the Central and South West region as part of an evaluation of SHRP, with 94 schools having been part of the SHRP program for one year and 56 serving as control group.³¹ Thirty students were randomly selected from the second grade in each school on the day of the assessment. Participants in the study include 3,561 students enrolled in second grade in 150 schools.³² The average age of students at the time of data collection was 8 years 11 months in the Central Region and 9 years 5 months in the South-West (both with a standard deviation of 1.6 years), and students were approximately evenly split across gender (53% and 50% were boys in Central and South-West regions respectively).

As described above, the medium of instruction policy in Uganda specifies that local languages should be used in the first three years of primary schools, while English is taught as a subject. Some schools, mostly in urban areas, elect to use English as the primary MOI from the first year. In this sample, the majority of schools (66 in Central region and 70 in South-West) reported using the local language as the medium of instruction.³³ Approximately 11% of schools in the Central region reported using English as the MOI, compared with 4% in the South-West.

³⁰ The author has consulted for RTI since 2010 and supported SHRP in a limited way during the proposal stage and for analysis and report writing after the first year of implementation.

³¹ SHRP included a large number of activities over the 2013-2017 period. The main components of the reading intervention included student and teacher material development in local languages and teacher training. For more on the project or specifics of the intervention, see (RTI International, 2016c).

³² This excludes 4 students (3 in Central region and 1 in South-West) who had missing values on at least one of the eight reading measures used in the analysis.

³³ These figures were computed from interviews with Head Teachers. Students in all schools reported the local language was used by the teacher and approximately 19% of students reported their teacher also used English at times.

Table 1 presents some student and school descriptive information that shows heterogeneity across regions. Students and schools in Luganda-speaking areas generally indicated relatively higher levels of wealth, such as access to electricity or students having shoes. This is consistent with higher levels of urbanicity, which may also explain why some schools choose English as a MOI to cater for more linguistically heterogeneous classrooms. Given the process for selecting districts and schools for the intervention and study, the Luganda sample is representative of the population in the Central region and the Runyankole-Rukiga sample is representative of the population in the South-West region.

Table 1. Demographic Indicators of participating Students and Schools (%)

	Luganda	Runyankole-Rukiga
<i>Students</i>		
Attended pre-school	80.7	53.4
Missed a day of school in previous week	55.3	51.8
Brings books home from school	42.9	81.1
Student is wearing shoes	55.8	25.3
Has Electricity at home	38.8	22.8
<i>Schools</i>		
School has functioning electricity	28.4	5.6
School has functioning toilets	97.3	98.6
Report using English as the MOI	10.6	4.1

Measures

Students were assessed using an individually administered Early Grade Reading Assessment (EGRA) in English and Luganda, or English and Runyankole-Rukiga.³⁴ The

³⁴ These data do not include the order of assessment. When EGRA is administered to children in two languages, it is common practice to randomly change the order in which students are assessed. Some

EGRA instrument was developed based on the DIBELS³⁵ and its purpose is to measure some of the foundational literacy skills that readers need for beginning reading (Dubeck & Gove, 2015). Language experts and local educators developed and field tested the Ugandan versions of the instrument (RTI International, 2013). The EGRA is composed of different subtasks that each measure a different foundational reading skill (RTI International, 2016a). The present study makes use of four subtasks: oral language, letter naming, pseudoword decoding, and reading comprehension questions. Each subtask includes between 3 and 100 items that are each scored correct or incorrect.

Oral language in Luganda and Runyankole-Rukiga was measured by asking students to answer three comprehension questions after listening to a short passage, and scored on a scale of zero to three.³⁶ Oral language in English was measured using twenty questions identifying receptive vocabulary like human body parts (e.g. *mouth, back, elbow*), words from the school environment (e.g. *pencil, book, ground*) and spatial words (e.g. *on the paper, behind you*) resulting in a score from zero to twenty. While both tasks measure oral language, receptive vocabulary is a subset of oral comprehension and these two measures are not perfectly comparable.

Decoding skills were measured in the same way in English, Luganda and Runyankole-Rukiga, using two EGRA subtasks. **Letter naming** was measured through a

evidence has shown, however, that there is no difference in findings based on the order of assessment (e.g. Piper, 2010a).

³⁵ The Dynamic Indicators of Basic Early Literacy Skills (DIBELS) are a set of procedures that were widely used in the United States to measure skills in English for English speakers related to the five “big ideas” in beginning reading. For more detail, see <https://dibels.uoregon.edu/assessment/dibels>

³⁶ For example, students were read a story about a child who likes to read and then asked to answer the question “What does Nalule like to do?”

timed task with students having 1 minute to identify up to 100 letters in lower and upper-case randomly ordered with letters most frequently used the language having a higher probability of being displayed. The *pseudoword* subtask included 50 nonsense words displayed in a grid and students had 1 minute to orally decode them. In both cases, a stop-rule was triggered if students were unable to identify the first 10 letters or the first 5 pseudowords, where the assessment would be discontinued, and these students would receive a score of 0 on the subtask. Although the individual items for letter naming and pseudowords are different across languages, the subtasks are developed in a procedure that designed to make the scores comparable.³⁷

Reading comprehension was measured after students were given a timed task to read a written passage out loud for 1 minute. Five comprehension questions were orally asked to students after they read the passage, providing a comprehension score ranging from zero to five based on the number of correct answers. Questions were only asked to students able to read the relevant part of the passage, and those students who were not asked the questions received a 0 for those items.

Analytic Plan

Following the assumption that alphabetic knowledge was a common factor across languages, I used confirmatory factor analysis with alphabetic knowledge as a latent variable predicting letter naming and pseudoword decoding in both languages. Testing

³⁷ The letter name subtask, for example, is developed by (i) identifying the frequency of letters or graphemes in existing texts in the language; (ii) creating a list of 100 letters with frequency of repetition proportional to the frequency children are exposed to the letter; and (iii) randomly ordering the letters in a grid for the assessment. (for more detail, see RTI 2016b).

for adequate fit of this latent construct addressed the research question about cross-linguistic transfer in *small problem spaces*. To address whether cross-linguistic transfer happens in *larger problem spaces* – reading comprehension – I replicated the structural model proposed by Proctor et. al (2010). Specifically, the model fit and the practical and statistical significance of the path from reading comprehension in first to second language serve to answer whether comprehension skills in L1 transfer to L2. Finally, I test whether the results hold in the Central region for two sets of schools, those who report using Luganda as the medium of instruction and those who report using English.

I used Mplus to estimate the different models. One of the main assumptions of structural equation modeling is univariate and multivariate normality of endogenous variables. In my structural model, the two endogenous variables – reading comprehension in L1 and in L2 – are not normally distributed because of the large number of students unable to answer reading comprehension questions. These findings are similar to those from other EGRA assessments in Uganda and other countries in the region (see for example Piper, 2010b; RTI International, 2010; USAID, 2014). *Figures A1, A2, A3 and A4* in the appendix display the distribution of reading comprehension tasks, as well as kernel density curves overlaid with normal distribution. In addition, the assumption of bivariate normality is rejected using a Doornik-Hansen test (Luganda: $\chi^2 [4] = 8,280.24$, $p < 0.001$; Runyankole-Rukiga: $\chi^2 [4] = 14,635.48$, $p < 0.001$). Given the non-normality of the two endogenous variables, I use the maximum likelihood MLR estimator in Mplus. Used in conjunction with the COMPLEX option, this estimator is robust to non-normality

and non-independence of observations due to the clustering of students at the school level (Muthén & Muthén, 1998-2010, page533).

Model fit indices help to measure how well the theoretical models fit the data. Following the recommendations of Hu and Bentler (1999) and Kline (2015), the fit of the models will be evaluated using a variety of goodness of fit measures, including chi-square, root mean square error (RMSEA), comparative fit index (CFI) and the standardized root-mean-square residual (SRMR). Criteria for adequate fit are as follows: RMSEA $<.08$; CFI $>.90$; SRMR $<.10$ (Hu & Bentler, 1999; Kline 2015). Chi-square should be close to zero and its associated p-value above 0.05, but the measure is highly sensitive to sample size. I look at the significance of indirect paths in order to test the statistical significance of the mediation of reading comprehension in L1. In order to test whether the theoretical path between the two reading comprehension measures is supported by the data, I look at the significance of the parameter estimate. I estimate the models in each region separately instead of doing multi-group analysis across regions because the items in Luganda and Runyankole-Rukiga are different. Finally, I use multi-group analysis in the Central region to compare estimates for students in schools that use English as the MOI with those in schools that use Luganda.

Results

Student performance on literacy tasks

Tables 2 and 3 display descriptive statistics of the eight measures for the samples in Central and South-West regions respectively. Student performance overall was strong in oral language and quite low in reading skills. In the Central region, student average

scores are similar across languages, while students in the South-West tended to do slightly better in Runyankole-Rukiga than in English, on average. Given the linguistic and orthographic differences, no strong inference should be made from these comparisons. Oral language means were slightly above fifty percent in both languages in the Central region. In the South-West, students answered almost 2 of 3 oral comprehension questions correctly in L1 compared with fewer than 9 out of 20 in English. In contrast, decoding and reading comprehension scores were very low. More than half of students assessed were unable to decode a pseudoword or answer one of the five comprehension questions correctly in L1 and L2 in either region. Letter knowledge in the Central region was slightly higher, with an average of 9.3 and 8.2 letters identified in Luganda and English respectively (standard deviations of 11.3 and 11.6 respectively). Averages in the South-West were similar but students showed less variability in letter naming tasks. One important comparison shows that students in the South-West performed lower in English than their peers in the Central region on the oral language and reading comprehension subtasks ($p < .001$).³⁸

Internal consistency coefficients (alphas) are included for the entire sample and restricted samples including only students able to read at least one word in the L1 passage. These suggest high levels of reliability for tasks with at least 20 items (ranging from 0.73 in L2 Oral language in RR to 0.98 for L2 Letter knowledge in Luganda) and relatively low levels of internal consistency (ranging from 0.33 to 0.66) for L1 Oral

³⁸ Differences in performance on L2 letter and pseudoword subtasks were not statistically significant.

language (3 items) and L1 reading comprehension (5 items).³⁹ L2 reading comprehension had an acceptable level of internal consistency despite having only 5 items (above 0.7 in both regions).

Table 2. Descriptive statistics and internal consistency of the eight measures in Luganda (L1) and English (L2)

	Number of items	Mean	Median	SD	Range	Alpha a	Alpha a*
L1 Oral Language	3	1.59	2	0.81	0-3	0.33	0.27
L1 Letter knowledge	100	9.30	6	11.32	0-81	0.98	0.97
L1 Pseudoword	50	4.68	0	7.92	0-45	0.95	0.91
L1 Reading Comp.	5	0.42	0	0.92	0-5	0.56	0.56
L2 Oral Language	20	10.40	10	3.47	0-20	0.80	0.77
L2 Letter knowledge	100	8.20	4	11.72	0-96	0.98	0.97
L2 Pseudoword	50	4.41	0	8.07	0-49	0.94	0.90
L2 Reading Comp.	5	0.42	0	0.97	0-5	0.75	0.76
Observations		1709				1709	549

Note: Alpha * in the last column is estimated using the restricted sample of students including only students able to read at least one word in the L1 passage

Table 3. Descriptive statistics and internal consistency of the eight measures in Runyankole-Rukiga (L1) and English (L2)

	Number of items	Mean	Median	SD	Range	Alpha	Alpha *
L1 Oral Language	3	1.95	2	0.90	0-3	0.40	0.36
L1 Letter knowledge	100	9.83	9	6.84	0-52	0.93	0.94
L1 Pseudoword	50	6.19	0	8.92	0-41	0.93	0.91
L1 Reading Comp.	5	0.54	0	0.85	0-3	0.66	0.66
L2 Oral Language	20	8.93	9	3.03	0-18	0.73	0.69
L2 Letter knowledge	100	8.90	9	7.40	0-44	0.92	0.93
L2 Pseudoword	50	4.19	0	6.92	0-43	0.93	0.90
L2 Reading Comp.	5	0.12	0	0.42	0-4	0.70	0.68
Observations		1852				1852	822

Note: Alpha * in the last column is estimated using the restricted sample of students including only students able to read at least one word in the L1 passage

³⁹ Reliability for L1 Oral language was particularly low. Sensitivity checks were done using individual items instead of the sum score in the structural model did not result in substantive differences.

Cross-linguistic transfer in decoding

Tables 4 and 5 display Pearson correlations across all subtasks in Central and South-West regions respectively. Correlations were all positive and statistically significant, and overall stronger in the Central region than in the South-West. Decoding measures were strongly correlated both within languages and across languages in both regions. Letter knowledge across language had the strongest correlation coefficient in the Central region ($r = .81, p < .001$) and was also strong in the South-West ($r = .65, p < .001$). L1 Letter knowledge was also highly correlated with L1 Pseudoword ($r = .58$ and $r = .43, p < .001$) and L2 Pseudoword ($r = .64$ and $r = .38, p < .001$). These suggest that a common latent decoding measurement model would likely fit well. Correlations across other measures were also positive and statistically significant. The weakest correlations were generally between L1 Oral Language and all other subtasks, which is likely due to the low reliability of that subtask. Other notably strong correlations included those between the two comprehension tasks ($r = .55$ and $r = .35, p < .001$) and between oral language and comprehension in English ($r = .51$ and $r = .38, p < .001$).

Table 4 – Pearson Correlation Coefficients In Central Region, N=1709

	1.	2.	3.	4.	5.	6.	7.
Luganda							
1. L1 Reading Comp.							
2. L1 Oral Language	0.19***						
3. L1 Letter knowledge	0.49***	0.15***					
4. L1 Pseudoword	0.76***	0.17***	0.58***				
English							
5. L2 Reading Comp.	0.55***	0.14***	0.55***	0.56***			
6. L2 Oral Language	0.34***	0.19***	0.44***	0.37***	0.51***		
7. L2 Pseudoword	0.66***	0.16***	0.64***	0.77***	0.73***	0.46***	
8. L2 Letter knowledge	0.41***	0.13***	0.81***	0.51***	0.53***	0.41***	0.62***

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 5 – Pearson Correlation Coefficients In South-West Region, N=1852

	1.	2.	3.	4.	5.	6.	7.
Runyankole-Rukiga							
1. L1 Reading Comp.							
2. L1 Oral Language	0.21***						
3. L1 Letter knowledge	0.36***	0.13***					
4. L1 Pseudoword	0.73***	0.19***	0.43***				
English							
5. L2 Reading Comp.	0.35***	0.13***	0.15***	0.35***			
6. L2 Oral Language	0.33***	0.31***	0.21***	0.33***	0.38***		
7. L2 Pseudoword	0.68***	0.19***	0.38***	0.77***	0.44***	0.38***	
8. L2 Letter knowledge	0.23***	0.10***	0.65***	0.29***	0.09***	0.10***	0.26***

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Confirmatory factor analysis models in each region included a single Alphabetical Knowledge latent variable comprised of letter naming and pseudoword decoding in each language. The error terms on letter naming subtasks were allowed to covary to acknowledge the similarity of the tasks across two languages and to reflect the large correlation between these subtasks.⁴⁰ In the Central region, all four indicators loaded well

⁴⁰ I did not let the pseudoword measures covary to avoid fitting a perfectly identified model. A sensitivity analysis replacing the covariance between letter naming subtasks with pseudoword made no substantive difference. In fact, standardized loadings were larger in the South-West sample, and model fit was similar.

onto the latent Alphabetical Knowledge variable, with standardized loadings ranging from 0.66 to 0.94. The model fit the data adequately ($\chi^2 [1, N=1709] = 17.13, p < .001, RMSEA = .097, SRMR = 0.009, CFI = .995$). In the South-West, standardized loadings were very high on pseudoword in both languages (0.83 and 0.93) and lower but still adequate on letter naming tasks (0.46 and 0.31). The model also fit the data well ($\chi^2 [1, N=1852] = .03, p = .87, RMSEA < .001, SRMR = 0.001, CFI > .999$). Tables A1 and A2 in the appendix present the full results from the measurement models in the full and restricted samples, where students unable to read one word in the L1 passage were excluded. This analysis testing whether four decoding skills – skills in the *small problem space* – load onto a common factor suggests that students in this sample, for the most part exclusively taught in L1, are able to decode across languages similarly. These results confirm that decoding skills transfer from Bantu languages to English and that decoding represents a single construct for second grade students in this sample.

Cross-linguistic transfer in reading comprehension

In order to test cross-linguistic transfer in reading comprehension I fit the model proposed by Proctor et al. (2010) with a common Alphabetical Knowledge as an observed variable represented by the average of the four scores on letter and pseudoword subtasks. Oral language and reading comprehension scores in each language are also included in the model as observed scores. Figures 2 and 3 present the standardized results of the models for Central and South-West regions respectively. The model fit was adequate in both regions (Central: $\chi^2 [2, N=1709] = 2.05, p = .36, RMSEA = .004, CFI >$

.999, SRMR = .006; South-West: $\chi^2 [2, N=1852] = 42.32, p < .001$, RMSEA = .104, CFI = .959, SRMR = .021).⁴¹

Figure 2 – Results from structural equation modeling of a model predicting Luganda and English Reading Comprehension ($\chi^2 [2, N=1709] = 2.05, p < .36$). Results are standardized. *** $p < .001$

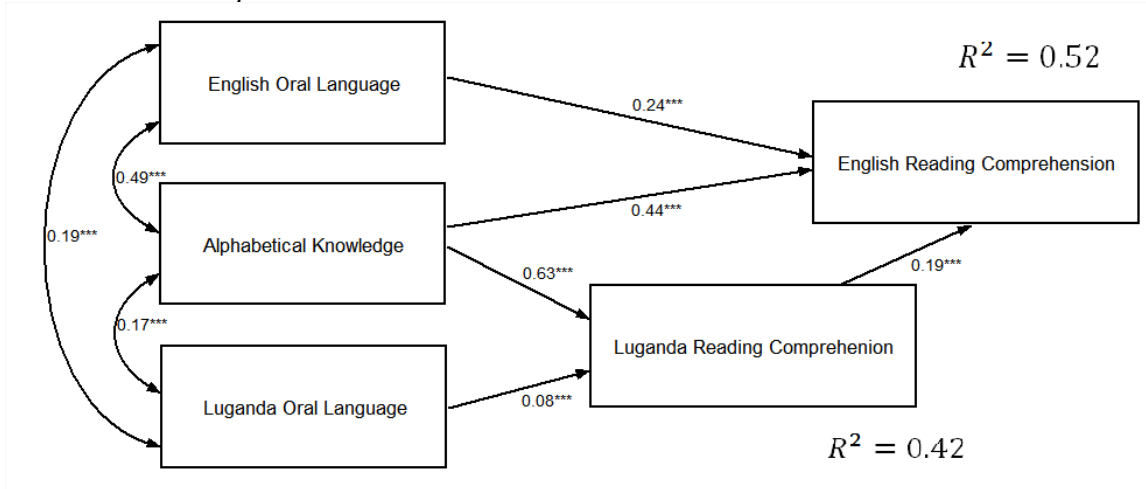
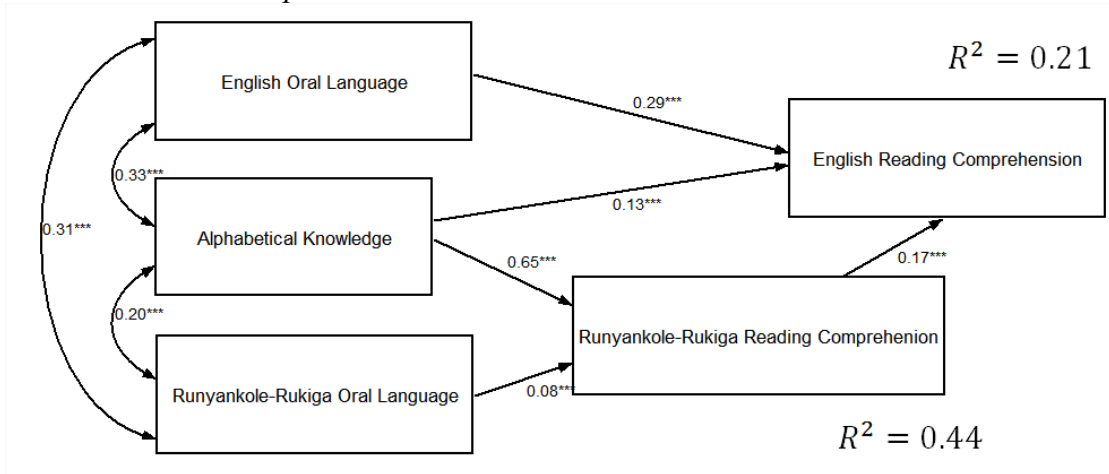


Figure 3 – Results from structural equation modeling of a model predicting Runyankole-Rukiga and English Reading Comprehension ($\chi^2 [2, N=1852] = 42.32, p < .001$). Results are standardized. *** $p < .001$



⁴¹ All model fit statistics were in line with the guidelines with the exception of RMSEA in South-West.

The direction of all paths was positive, in accordance with the hypotheses. The standardized coefficients for alphabetical knowledge suggested strong association with L1 reading comprehension (Central: $\hat{\beta} = .63$ ($SE = .03$), $p < .001$; South-West: $\hat{\beta} = .65$ ($SE = .02$), $p < .001$). The magnitude of the relationship with L2 reading comprehension was not as strong, especially in the South-West region (Central: $\hat{\beta} = .44$ ($SE = .05$), $p < .001$; South-West: $\hat{\beta} = .13$ ($SE = .04$), $p < .001$). The path from L1 to L2 reading comprehension was moderate and statistically significant in both regions (Central: $\hat{\beta} = .19$ ($SE = .06$), $p < .001$; South-West: $\hat{\beta} = .17$ ($SE = .03$), $p < .001$). This is consistent with Proctor et al. (2010) and suggests that cross-linguistic transfer from L1 to L2 is indeed occurring in reading comprehension. With the exception of the L2 in the South-West region, standardized coefficients were larger for Alphabetical Knowledge than for Oral Language, which is consistent with the literature suggesting that decoding skills are relatively more important in explaining variation than language in lower primary students (Hoover & Gough, 1990).

Tables A3, A4, A5, and A6 in the appendix report full results of the structural models in both regions for full and restricted samples. In both regions, the indirect paths from Alphabetical Knowledge to reading comprehension in L2 are positive and statistically significant. In Central region, the contribution from the indirect paths results in similar magnitude of the total effects from Alphabetical Knowledge to reading comprehension in L1 and L2 (L1: $\hat{\beta} = .63$ ($SE = .03$), $p < .001$; L2: $\hat{\beta} = .56$ ($SE = .03$), $p < .001$). In the South-West, the indirect contribution of decoding to reading comprehension in English is important but Alphabetical Knowledge remains a

stronger predictor of reading comprehension in L1 than L2 (L1: $\hat{\beta} = .65$ ($SE = .02$), $p < .001$; L2: $\hat{\beta} = .24$ ($SE = .05$), $p < .001$). These results confirm that reading comprehension in L1 partially mediates the effect of Alphabetical Knowledge on reading comprehension in English in both regions.

In the Central region, I use multi-group analysis to test whether the results are the same across two groups of schools based on the MOI reported by the Head Teacher. In order to assess whether separating the two groups significantly adds to the overall model, I use the Satorra-Bentler Scale Chi-Square test, which takes into account the non-normality of the data (Satorra & Bentler, 2001). The test shows that the model in which the paths are allowed to freely vary across language groups fits significantly better than the model in which the paths are constrained across groups ($\chi^2 [2] = 5.10$, $p = .08$). For students who were taught primarily in Luganda, the magnitude and substantive results reported above all hold, which is unsurprising given they were the vast majority of the sample ($N=1454$). Their peers ($N=255$) taught primarily in English, however, do not show the same cross-linguistic transfer (see Table A7). While the association between reading comprehension in L1 and L2 is positive, the magnitude of the direct effect is small and not significant at the .05 alpha level ($\hat{\beta} = .10$ ($SE = .06$), $p = .07$). This supports the idea that the theoretical direction of the transfer is from L1 to L2, where instruction in the first language is required to support development of the second language. For students who speak Luganda at home but are not explicitly taught in that language, their reading comprehension in Luganda does not support English reading comprehension development.

Discussion

This study sought to describe the performance of second grade students in a small set of literacy skills across two languages, and to test whether cross-linguistic transfer as observed between European and Asian languages also emerged between Bantu languages and English in Uganda. Based on existing research, I hypothesized that decoding skills would load onto a common alphabetical knowledge construct across languages and that reading comprehension in the first language would predict reading comprehension in the second language holding decoding and oral language in English constant. This study found that in both regions decoding skills represented a common construct across two languages, in accordance with the hypothesis. In addition, the association between first- and second-language reading comprehension was positive and statistically significant, suggesting that there is some cross-linguistic transfer of reading comprehension skills occurring as early as second grade. Current research has shown that children are consistently able to transfer lower-order literacy skills such as decoding, especially when orthographies across languages are similar, and this study confirms this relationship in a country and across languages rarely represented in the literature. Finally, this study found that the smaller subset of students in schools in the Central region who attend schools where English is the medium of instruction starting in the first grade did not show transfer of skills in *large problem spaces*.

Within the large set of policies that allow for the use of multiple languages in the classroom, two of the most important implementational decisions are the number of years that learners' L1s should continue to be taught and the timing of the transition to L2. The

findings of this study suggest that these decisions have an important impact on instruction as they determine which skills children will be taught to transfer explicitly, and which skills will only be taught in L2. Findings from this study focused on second-grade students support the foundational theories of cross-linguistic transfer used to promote using L1 in the early years of schooling. The ability of students to use the reading skills explicitly taught in Luganda or Runyankole-Rukiga to decode letters and non-words in English is a testament to their ability to transfer skills in *small problem spaces*. Further, despite the limited proportion of students able to read enough of the text to comprehend, students with higher levels of comprehension in their first language were able to make more meaning of the text in English, on average. However, students who attended schools where the MOI was English were not able to build on L1 reading skills to improve English reading comprehension. This finding, however, should be taken with caution, as the schools that use English as the MOI are often serving a community with different characteristics, and may be using different instructional approaches as a result of the MOI and linguistic diversity in their classrooms.

While the national MOI policy in Uganda promotes the use of local languages in the early years, schools and district officials have decision-making power regarding the language chosen. This is consistent with recommendations Crawford (2004) makes in the United States: that the full range of bilingual education policies – ranging from full immersion in L2 to dual immersion programs – should be considered to accommodate for individual and contextual differences. Schools in the Central region that chose English are likely to serve linguistically heterogenous communities. While the choice of English

in these contexts may seem most appropriate from a practical perspective, specialists would suggest that there are options that would provide L1 foundations for literacy and learning. For example, Benson and Young (2016) propose strategies to help schools build on language skills students bring with them to multilingual classrooms. While recommended practices vary based on the specific context of the school, these include using multilingual teachers, organizing multi-grade classrooms according to language groups, and supporting children in developing literacy skills in the different languages spoken in the community.

It may seem counterproductive to encourage schools in low-resource settings to develop learners' literacy skills in L1 at a time when the focus of international policymakers is on the low levels of skills developed by students in a single dominant language. This recommendation, however, is consistent with an emerging body of research suggesting the transfer of skills between languages may be bi-directional. Cummins (1979) hypothesized that the transfer of skills between languages was subject to individuals reaching a "threshold" level in the first language. In a recent meta-analysis on studies focused on bilingual students, Prevoo and colleagues (2016) found evidence of a positive relationship between L1 oral language and L2 reading, and between L2 oral language and L1 early literacy skills. Kim and Piper (2018) used longitudinal data from students in urban settings in Kenya and found evidence of transfer from English to Kiswahili and from Kiswahili to English. Children living in linguistically heterogeneous communities bring a different set of skills and resources, and face different challenges than their peers in relatively homogeneous rural settings. Translanguaging theory

suggests that these students may be developing language and literacy skills building on their exposure to multiple languages and drawing simultaneously from their entire repertoire (García & Wei, 2014). Practically, understanding that relationship is necessary to design optimal policies and to develop appropriate teaching strategies, curricula, and interventions. In Uganda, where children are sometimes punished for using their home language in part to promote their English development (Namyalo & Nakayiza, 2014), officially introducing pedagogical approaches that promote translanguaging will need to be preceded by important conversations including teachers and local communities.

While this study confirms the notion that using an L1 as the MOI in Uganda is not harmful to the development of L2 literacy from a psycholinguistic perspective, and likely beneficial, using a familiar language in school has other potential benefits. Using a familiar language enable students to participate more, as observed in classrooms in Kenya (Dubeck et al., 2012), which allows for more student-centered learning and less rote memorization. When the language of the school and home are the same, students tend to perform better on national assessments (Heugh, Benson, Bogale, & Mekonnen, 2007) and remain in school longer (Seid, 2016; World Bank, 2010). This also allows parents to be more involved through oral language development and, when also literate, reading to their children, both of which have been found to improve student reading outcomes (Hess & Holloway, 1984). Using an L1 in school in linguistically diverse settings has many potential benefits, but the timing and instructional approaches to transitioning to an L2, if this is deemed necessary, require a deep understanding of the relationship between the languages involved.

Limitations and Future Research

The evidence presented in this study presents four important limitations. Given the cross-sectional nature of this dataset, it is not possible to conclude that the skills children developed in English were transferred from the literacy instruction they received in the first language or whether these were developed simultaneously through schooling and other exposure to print. Indeed, the positive correlation in all subtasks suggests that children performing higher on any subtask were simply overall higher academic achievers, regardless of problem space or language. Proctor and Zhang-Wu (2019) find that most studies, including the current study, are not able to determine the direction of the transfer of skills, and as a result prefer using the term cross-linguistic association when directionality is not tested. Longitudinal data collection (e.g. Kim & Piper, 2018) can help to understand the learning trajectories of children learning in contexts that implement multilingual education policies and contrast these with those of children learning under monolingual policies.

In addition, the reliability, as well as the depth of measures currently available is problematic, especially oral language in L1 and reading comprehension. The current focus of early literacy intervention programs is heavily biased toward oral reading fluency, as exemplified by the proposed benchmarks for signaling “improved reading ability” under USAID’s education strategy (USAID, 2011). As educational development efforts across low and middle-income countries naturally shift their focus from the early grades toward upper primary school, governments and policy makers should increase efforts to develop more reliable measures of oral language. In addition, reading

comprehension is a complex construct and measures currently available are limited in depth.

Overall low levels of literacy skills in this sample poses both a statistical and theoretical limitation. The important floor effects due to low levels of literacy skills in a large proportion of the sample reduce the variance of the measures and thus likely attenuate the relations between variables. It is therefore possible that the lack of relationship found in schools using English as the MOI may be due to the small proportion of students with non-zero reading comprehension levels. This study thus cannot provide information on levels of transfer in settings where the majority of children are able to read fluently in one or both languages.

Finally, an understanding of the mechanisms of cross-linguistic transfer or relationships across languages is crucial to informing instruction during and after a transition to using L2 as the MOI. Uncovering mechanisms that support transfer in reading skills is more likely to occur through qualitative research including observing teachers in plurilingual contexts and emergent bilingual students. Without classroom observations that explore which languages are used in practice, and how they are used, student performance cannot be fully understood. Proctor and Zhang-Wu (2019) argue that the limited qualitative and mixed-methods studies currently available provide evidence for how complicated and dynamic the process of language development is, and that future research should adopt explanatory sequential mixed-methods frameworks. In the context where learners have some proficiency in two or more languages, pedagogical approaches

that promote *translanguaging* are likely to support the development of literacy skills in dominant languages because children use their entire repertoire to make meaning of texts.

Conclusion

Governments in linguistically diverse settings around the world are increasingly introducing or expanding multilingual education policies. Cross-linguistic transfer is complex and literacy instruction ought to be developed specifically for each language, accompanied by explicit teaching to promote transfer between L1 and L2. Understanding whether and how cross-linguistic transfer occurs between Bantu languages and English, specifically, is important both theoretically and practically (Genesee, Geva, Dressler, & Kamil, 2006.). This is the first study to my knowledge to explore cross-linguistic transfer in Uganda, and one of few exploring transfer from an African language to a colonial language. The choice of language as the MOI is not simply a pedagogical choice, and can present challenges for low and middle-income countries with rich linguistic diversity. As such, Uganda's choice to promote non-dominant languages in the early years is commendable. Linguistically diverse countries that are exclusively using an official dominant language as the MOI in schools should explore ways to include languages familiar to students in the formal school system, especially in the early years of literacy development. Similarly, countries such as Uganda that currently promote early-exit models should explore how multilingual classrooms can be expanded to include the entire primary school system, maximizing the potential of cross-linguistic transfer to develop multiliterate citizens.

In addition, research is needed to further understand multilingual literacy development of children in low and middle-income settings, especially using longitudinal data and qualitative explanatory frameworks. Finally, it is imperative to develop better literacy measures in local languages to support policy making, research, curriculum development, and ultimately instruction.

Paper 3: Literacy Gains for All? Impact Variation of an Early Grade Reading Program in Kenya

1. Introduction

Following large gains in access to primary school in low- and middle-income countries, international education policies have largely shifted focus to improving learning outcomes, resulting in a large number of interventions targeting reading instruction in the early years of school. These programs often comprise similar elements: providing multiple days of training for teachers and detailed instructional material targeting the five components of reading instruction identified as essential by the National Reading Panel in the United States (NICHD, 2000). Many of these programs have shown success in increasing basic skills such as decoding, and some have increased reading comprehension levels of young children (Kim, Boyle, Zuilkowski, & Nakamura, 2016). Evaluations of these interventions, however, typically focus on the average treatment effect of the program, and rarely examine under what conditions the programs work best, or whether all subgroups benefited. Specifically, research is lacking on whether positive impact is dependent on teachers having a prior set of skills or experience, or a conducive school environment. Understanding individual- and context-level characteristics that facilitate or hinder the effectiveness of these programs is essential when considering scale-up within countries or replication in different settings.

This study examines teacher- and school-level moderating factors of a literacy intervention in Kenyan primary schools. The Primary Math and Reading (PRIMR)

initiative was implemented in urban schools during the 2012 and 2013 academic years. I use data from the PRIMR evaluation to explore whether the intervention was successful in increasing student literacy in English and Kiswahili across types of teachers and schools. The main outcomes of interest in this study are students' scores in reading words in passages and comprehension, hereafter referred to as *higher-level literacy skills*, in contrast to sound and letter recognition, or *lower-level literacy skills*. At the teacher level, I test whether the impact differed for teachers with and without formal training and varying levels of experience. At the school level, I estimate the difference in impacts in public schools and nonformal private schools, and across baseline performance levels. Findings suggest that the intervention was highly effective in increasing student ability in both lower- and higher-level literacy skills in English and Kiswahili, on average, but that gains in learning were heterogenous. The effects were similar in classrooms taught by novice and experienced teachers, and whether or not teachers had formal training. The program was only effective, however, in increasing higher-level skills in nonformal private schools, and was not found to be effective in increasing these in public schools. Students in schools with higher performance at baseline benefitted more from the intervention than their peers in schools at the bottom of the distribution. Schools performing in the bottom 10 percent at baseline benefitted from the intervention in English but made no additional gains in Kiswahili, on average.

This paper begins with a description of the context in which literacy interventions in the early grades are taking place and a framework to explore teacher and school-level moderating factors in evaluations of literacy interventions. Methods and findings sections are followed by a discussion of the implications of heterogenous effects for

future research and evaluation of education interventions, as well as the design of literacy interventions in low-resourced contexts.

2. Background and Context

2.1 Literacy in the Early Grades

The adoption of the *Millennium Development Goals* by the international community led to policies aimed to achieve its explicit goal of universal primary education by 2015 (UNESCO, 2015a; United Nations, 2013). These policies were largely successful in increasing access to formal schooling, reducing the number of children out of school to 53 million in 2011 from 102 in 2000 (United Nations, 2013). However, in spite of, or because of the large increase in access, learning in the early grades in low- and middle-income countries remains low in what has been coined the *global learning crisis* (UNESCO, 2013; World Bank, 2018). Banerjee and Duflo (2011) highlight that 60 percent of children in grade 4 in India could not read a simple story, 80 percent of their peers in grade 3 in Pakistan could not read a short paragraph while 27 percent of grade 5 students in Kenya could not read a simple paragraph.

The recent changes in access and learning in primary school in Kenya are illustrative of that of many low- and middle-income countries (LMICs) in the first two decades of this century. Kenya introduced free primary education (FPE) in 2003, which resulted in a large increase in access to primary education, with the net enrollment ratio reaching 96% in 2013 and primary school completion 80% in 2012 (Ministry of Education, Science and Technology, 2014). However, learning outcomes, specifically literacy and numeracy levels, remain low in Kenya (Piper, Jepkemei, & Kibukho, 2015; Uwezo, 2015).

Following the increase in available evidence that access to formal school was not increasing basic skills for all, the *learning crisis* has become the focus of many international education organizations (Save the Children, 2016; UNESCO, 2014). The United States Agency for International Development (USAID) made improving reading skills in the early grades the first, and arguably most visible, of three goals in its education strategy (USAID, 2011), which was followed by the implementation of Early Grade Reading projects in many low- and middle-income countries, including Kenya. PRIMR was implemented as a collaboration between the Kenyan Ministry of Education, Science and Technology (MoEST) and USAID for three school years between 2012 and 2014.⁴² The broad goals of the program were to support literacy and numeracy instruction in the early grades. The literacy initiative included four major components: (i) providing each student with instructional booklets; (ii) training teachers on effective instructional strategies; (iii) providing teachers with detailed instructional activities; and (iv) providing teachers with support from coaches⁴³ (RTI International, 2016b). Evaluations of PRIMR have reported large learning gains across all outcomes measured but have largely focused on the average treatment effect (Piper et al., 2015; Piper & Zuilkowski, 2015; RTI International, 2016b). Following the success of PRIMR, the Kenyan government launched a national reading program called Tusome using the intervention as a blueprint for improving literacy skills.

⁴² The school year in Kenya follows the calendar year starting in January.

⁴³ In public schools, the coaches were TAC tutors (public servants responsible to provide instructional support to teachers). In private schools, coaches were hired and trained by the project team. For more on the coaching intervention see Piper and Zuilkowski (2015).

2.2 Program variation

A growing number of researchers and policymakers have made an argument to change the focus of program evaluations from “what works” on average to better understanding mediation – the mechanisms that lead to positive impact – and moderation – the conditions under which policies are more or less effective (e.g.: Peterson, 2016; Weiss, Bloom, & Brock, 2013). Understanding moderating factors of a program is especially important for two reasons: efficient use of resources and equity. Given a limited set of resources to implement or scale-up a program, understanding the individuals best suited to benefit from a program or an environment that enables better implementation can maximize the benefits. In addition, an equity approach to social program delivery should ensure that programs reduce, not reinforce, existing gaps in outcomes of interest.

Weiss et al. (2013) propose a framework to explore variation in the effect of a program that includes three sources: *treatment contrast*, *client characteristics*, and *program context*. *Treatment contrast* considers the range of alternatives that individuals could benefit from in the absence of the program. For example, the impact of attending the Perry Preschool program in the late 1960s in Michigan when few alternatives existed was very large (Heckman, Moon, Pinto, Savelyev, & Yavitz, 2010) compared with the current impact of being admitted to a Head Start program, as 60% of non-participants found an alternative center (U.S. Department of Health and Human Services, Administration for Children and Families, 2010). Feller and colleagues (2016) find positive impact of Head Start on receptive vocabulary for children who would have gone to home-based daycare but not those who would have gone to another center-based

program, a result of different *treatment contrasts* as well as *program contexts*. *Individual characteristics* of beneficiaries also lead to differential impacts, as Bloom and Weiland (2015) found that Head Start was more effective for children who entered to the program with limited exposure to English. These different sources of variation have policy implications for governments responsible with investing limited resources. In this paper I focus on variation across client characteristics and program context, primarily because the study design did not include variation in the contrast.

PRIMR primarily consisted of teacher training supported by materials and coaching across a set of mostly urban schools. While the ultimate goal of the program was to improve student outcomes, I consider teachers as the clients of the program and focus on two teacher characteristics that may influence whether and how training and resources would be most useful: teacher training and experience. I consider schools as the context in which the program is implemented and examine whether the effect varies across the type of school – public or nonformal private schools – and the quality of the learning environment at baseline. At the teacher level, understanding whether a program is more or less effective for types of teachers could help focus resources on some teachers or modify the training accordingly. Understanding whether the program varies across contexts has important equity implications.

2.3 Teacher characteristics

As cornerstones of the learning experience of children, teachers have been the focus of much education research. Teachers and teaching quality have been found to influence short-term outcomes such as student test scores (Nye, Konstantopoulos, & Hedges, 2004) and long-term outcomes such as attending university or salaries (Chetty,

Friedman, & Rockoff, 2011). The importance of teachers on student learning may be even more important in LMICs where teachers can compensate for lack of resources or literate adults in children's households (Ngware, Oketch, & Ezeh, 2011). Identifying characteristics of effective teachers, however, remains a challenge, both methodologically (as teacher characteristics are not randomly distributed across schools) and because findings on the association of teacher characteristics and student outcomes are not consistent. In low- and middle-income countries, there is strong evidence of teacher subject knowledge being positively associated with student learning, and weaker evidence that years of experience and formal training are positively associated with student outcomes (Glewwe et al., 2011; Hanushek, 2003). Altinok (2013) finds that teacher knowledge in reading and mathematics varies across 14 countries across Africa, and that these characteristics are highly correlated with experience and education.

Despite research suggesting that teacher characteristics are directly predictive of student learning outcomes, evidence on whether these characteristics *moderate* the impact of professional development interventions is weak. Interventions targeting teacher pedagogy through training and materials have shown promise when compared with "business as usual." A meta-analysis of 21 studies in LMICs found an overall effect of 0.23 standard deviations on language outcomes in the early grades, and 0.08 for math scores at the same level (Snilstveit et al., 2015). This analysis also found large heterogeneity of effects on student outcomes, including some negative treatment effects. While this analysis is helpful in assessing average treatment effects across a broad range of interventions, it concludes that no "combination or single factor in terms of intervention design, implementation or context was decisive in determining interventions'

success” (ibid, p. 207). One key aspect that has not been sufficiently explored is whether teachers with more experience or higher formal training respond better to interventions. In the United States, one study found that teachers with fewer years of experience or those with better perception of the quality of the training are more responsive to training (Downer, Locasale-Crouch, Hamre, & Pianta, 2009).

There is no clear *a priori* direction in the differential impact of an intervention such as PRIMR across teacher characteristics in the Kenyan context. Training in the Primary Teacher Training Colleges in Kenya lasts two years, but the main focus of the training to teach reading is content, not methods (Akyeampong et al., 2013), and the time spent for practicum in classrooms is less 9 weeks (World Bank, 2014). As such, an intervention based on pedagogical training coupled with scripted lessons is likely novel for all teachers, trained and untrained. If more experienced teachers respond less to professional development, novice teachers could benefit more from PRIMR. However, veteran teachers who embrace the scripted approach may be better equipped to use the skills and resources provided by the intervention.

2.4 School contextual factors

One unique aspect of PRIMR was its simultaneous implementation in formal public schools and nonformal private schools. The share of LMIC students attending private school has increased drastically in recent years, doubling from 11% to 22% between 1990 and 2010 in low-income countries and while growing from 8% to 12% in middle-income countries (Baum, Lewis, Lusk-Stover, & Patrinos, 2014). This expansion has largely taken place at the primary level and has been driven by students from lower income families’ choice of small private, nonformal, or community schools that charge

low fees.⁴⁴ Private school expansion has generally helped to increase access to education for underserved populations in a range of countries including Kenya, Ghana, Pakistan, and India (Barrera-Osorio, Blakeslee, Hoover, Linden, & Raju, 2011; Tooley & Dixon, 2005). While the private school expansion has been driven in part by entrepreneurs, governments in LMICs have also contributed to this increase through partnerships with private schools aimed at increasing capacity for expansion of access to education. Learning outcomes in these nonformal private schools are generally higher than in public schools in the same communities, but it is difficult to disentangle the impact of schooling from the selection due to payment of fees, and the quality nonformal schools, with or without government support, has remained relatively low (Day Ashley et al., 2014; Rouse & Barrow, 2009). In this context, understanding whether there are differences in the impact of interventions in public and nonformal schools can help policymakers considering different types of partnerships, specifically professional development.

A literacy program that is more effective in one type of school than another could theoretically have a compensatory effect, reducing the gap between groups of students, or an accumulated advantage effect, where groups that start ahead benefit more than their peers thereby increasing existing gaps. Figure 1 presents an idealized graphical model of the two theorized situations.⁴⁵ On the left panel, students from schools with higher baseline scores benefit much more from the program than schools with low baseline

⁴⁴ The literature uses several terms for these schools, such as low-cost, low-fee, and community schools. While heterogeneous across contexts, these schools are primarily characterized by their status (private) and the relatively low fees they charge in comparison with traditional private schools serving economically advantaged families. In this paper, I use the term ‘nonformal schools’ consistent with RTI International (2016b) and other papers reporting on PRIMR.

⁴⁵ Adapted from Miller, Farkas, Vandell, and Duncan (2014).

scores, creating a larger gap between low- and high-performing schools and students. In contrast, the right panel presents a scenario where students in relatively weaker schools at baseline benefit more from the program, narrowing the gap with their peers. Reading development research has documented the prevalence of the accumulated advantages phenomenon, coined the *Matthew Effect*, where readers developing literacy skills early have access to more vocabulary leading to growing gaps in later years (Stanovich, 1986). While the Matthew effect is an individual-level phenomenon, its corollary at the school level would lead to increased inequities in education provision. Schools providing higher quality instruction at baseline could benefit more from training and resources than struggling schools, leading to growing gaps. For example, McCoy and Wolf (2018) found that an early childhood intervention in Ghana generally had larger positive impacts on children in classrooms with higher quality at baseline. In this study, I explore whether a literacy program had a similar impact across types of schools, and whether the impact was different between low- and high-performing schools.

This study specifically aims to answer the following research questions:

- *What was the average impact of PRIMR on student word reading and comprehension in Kiswahili and English?*
- *Was there evidence of treatment heterogeneity across types of schools and teachers?*
- *Do characteristics of teachers (training and experience) and schools (ownership and baseline literacy level) moderate the effect of the intervention?*

3. Methods

3.1 PRIMR intervention and study design

PRIMR was implemented using a phased-in approach with three cohorts of schools starting in academic years 2012, 2013, and 2014 respectively. The initiative was

implemented in public schools in three urban counties as well as in private nonformal schools in Nairobi county (RTI International, 2016b).⁴⁶ Schools were grouped into 33 clusters because of the coaching component of the intervention and clusters were randomly assigned to one of three cohorts that took part in PRIMR starting in 2012, 2013 and 2014 respectively. Data were collected at baseline in January 2012, midline in October 2012, and endline in October 2013. This study considers the first two cohorts as part of the Treatment group, whether they participated in one or two years of program by the endline. Table 1 presents a summary of data used for the current study.⁴⁷ At endline, 4222 students were assessed in 212 schools, compared with 4385 students from 220 schools at baseline. A random sample of schools within each zone or cluster was selected at each round of data collection (Piper et al., 2015). As a result, only 161 schools were observed both at baseline and endline, and none of the students in this sample is observed multiple times.⁴⁸ The original study was designed to estimate the average impact of PRIMR and not to create a longitudinal sample at the student level. Sample selection was driven in part by budgetary constraints, which explains sampling schools and students randomly at each time point. As a result of the evaluation design and the focus of this study on teachers and schools as moderators, I consider the school and teacher as the treatment unit. Students randomly sampled are used to estimate the impact of the intervention on the classroom or school.

⁴⁶ Public schools were located in Nairobi, Nakuru, and Thika counties.

⁴⁷ Data from the intervention are publicly available through the Education Data for Decision Making website. <https://globalreadingnetwork.net/eddata>

⁴⁸ 114 from the Treatment group and 47 from the Control group.

3.2 Analytical Sample and baseline balance

Tables 2, 3, and 4 present baseline characteristics of the treatment and control groups, differences between the two groups, statistical tests of individual characteristic differences as well as joint tests across groups. Students in the sample were approximately 7 years old and balanced across gender in both treatment and control. Indicators of household wealth used to compute the student-level SES composite were also balanced across groups. Importantly, there were no baseline differences in the main literacy outcomes of interest in this study, word reading and comprehension, on average in schools across the two groups. Although none of the differences were statistically significant, the joint orthogonality test rejected the hypothesis that the two groups were similar across student characteristics at the 0.10 level ($F = 1.68, p = 0.07$). Table 3 shows that the majority of teachers were women in both groups and there were no significant differences in teacher age, teaching experience, or whether the teacher had a good lesson plan on the day of the survey. While teachers in the treatment group were more likely to report having formal teacher training on average, the joint significance test showed no overall statistically significant differences across groups ($F = 1.18, p = 0.31$). At the school level, the two groups had similar proportions of nonformal schools and there were no statistically significant differences on a range of school resources. There were, however, statistically significant differences in the presence of electricity at the school and the mean score achieved in the prior year by students on the Kenya Certificate of Primary Education (KCPE), the national examination held at the end of the primary school cycle. Nearly 60% of schools in the treatment group had access to electricity compared with 37% in control schools, and students in the treatment group scored 17

points higher on the KCPE, on average. The joint significance test rejected that the two groups were similar at baseline across this set of characteristics at the 0.10 level ($F = 1.57, p = 0.099$).⁴⁹ In summary, there were no baseline differences across treatment and control groups in terms of individual student characteristics and a set of teacher characteristics, but there were statistically detectable differences advantaging the treatment group in two important school characteristics – presence of electricity and student KCPE scores. In order to mitigate the possible imbalance at baseline, I include baseline covariates in analytical models below.

3.3 Measures

The outcome measures for this study are student literacy assessments. Students were assessed in English and Kiswahili using an Early Grade Reading Assessment (EGRA). EGRA was developed to provide a reliable measure of different skills that contribute to literacy, and has been adapted in more than 65 countries and over 100 languages (Dubeck & Gove, 2015). Students were assessed in six EGRA subtasks at each round of data collection. This study focuses on ***word reading*** and ***reading comprehension***, the two highest level subtasks. The word reading score is a count of words correctly read from a short story, while the reading comprehension score is the number of correct comprehension questions answered after reading the passage.⁵⁰ As

⁴⁹ The conclusion from the balance check across the KCPE score is threatened by the large number of missing values at the school level. Only 118 of 160 treated schools and 42 of 60 control schools had KCPE scores at baseline.

⁵⁰ Some studies transform the word reading measure as an oral reading fluency measure by adjusting for the time it took the student to read the number of correct words. This provides a measure in *words per minute* units. I keep the measure as the raw score of number of words read. The Pearson correlation between word reading and words per minute in the sample is 0.97 and 0.99 in English and Kiswahili respectively, suggesting there should be no difference in analysis results.

Table 2 shows, students at baseline read approximately 14 and 11 words correctly in English and Kiswahili, respectively, and answered fewer than 1 out of the 5 comprehension questions correctly.⁵¹

I use four variables to test for teacher and school level moderation. Teacher experience is measured as a continuous variable representing the number of years of teaching. Formal teacher training is a dichotomous variable indicating whether the teacher reported having been formally trained for the profession.⁵² These characteristics are measured at endline. The type of school, public or nonformal, and the baseline average literacy scores are used as school-level moderators.

Finally, I created a student-level SES score using confirmatory factor analysis to reduce the nine indicators collected about material resources of students assessed, as well as a school-level indicator of resources such as water and electricity. These variables are used as covariates the models.

3.4 Analytical Approach

The identification strategy exploits random assignment of clusters to the intervention and I use an intent to treat (ITT) model. I first estimate an average treatment effect using the following equation:

$$Y_{i,c,s} = \beta_0 + \beta_1 T_s + B_i X_i + B_c X_c + B_s X_s + \varepsilon_{i,c} \quad (1)$$

⁵¹ Although students read more English words correctly than Kiswahili, it is not possible to compare these measures across languages and infer that students read English better because of orthography and grammar are different.

⁵² Teachers are considered to have formal training if they report having a P1 teacher training certificate, an S1 diploma in teacher education or a Bachelor of Education.

The outcome variable Y are four literacy skills measured at the student level i , in classroom c and school s . These four skills include English and Kiswahili word reading and comprehension scores. The outcome is regressed using simple OLS regression on T , a dichotomous indicator of treatment and X , three vectors of student, teacher and school characteristics.⁵³ β_1 is the estimate of the average treatment effect (ATE) of the intervention. Standard errors are clustered at the *classroom* level because of the large intraclass correlation at that level.⁵⁴ In order to answer the second research question, I describe the variability in gains at the school-level as a result of the treatment by creating gain scores for each outcome of interest using the simple difference between school averages at baseline and endline for schools that were randomly sampled in both data collection rounds. Finally, I answer the third research question by testing school and teacher level moderation using the following two models:

$$Y_{i,c,s} = \beta_0 + \beta_1 T_s + \beta_2 M_c + \beta_3 T * M_c + B_i X_i + B_c X_c + B_s X_s + \varepsilon_c \quad (2)$$

$$Y_{i,c,s} = \beta_0 + \beta_1 T_s + \beta_2 M_s + \beta_3 T * M_s + B_i X_i + B_c X_c + B_s X_s + \varepsilon_c \quad (3)$$

Building on the model estimating ATE, these models including a moderating variable M measured at the teacher c or school s level and an interaction term between T and M . β_3 is an estimate of the difference in treatment effect across values of M . The only

⁵³ Controls include student grade and SES, teacher training, school type, school-level resources and baseline performance on state examinations. Some of these controls are later tested as moderators.

⁵⁴ See Appendix 1 for details on clustering, ICC and a discussion of whether to use this model or a multilevel model (MLM).

difference in the two models is the level at which the moderator M is measured (school or classroom). One key assumption in this identification strategy is that moderators are either fixed characteristics, or that that treatment did not impact the moderator. Both school-level moderators are fixed at baseline and satisfy this assumption. Both teacher-level moderators could potentially be impacted by the intervention, with a migration of teachers with more (or less) experience and training from treatment to control schools, but neither characteristics seem to have been changed by the intervention.⁵⁵ I use multiple imputation (MI) to address missing data for one key control variable and one school-level moderator, and present findings with listwise deletion as well as MI.⁵⁶

4. Findings

4.1 Average Treatment Effect

Descriptive statistics collected from teachers and head teachers from treatment schools at midline and endline suggest high levels of participation and implementation. At midline, all teachers from the treatment group reported using PRIMR English and Kiswahili lessons at least 3 days per week. At endline, only 7% of teachers in the treatment group reported having attended a PRIMR training fewer than twice, with 50% reporting having attended more than 3. Similarly, 18% of head teachers at endline reported never attending or only attending one training, compared with 55% attending

⁵⁵ See table 8 in Appendix 3 for a comparison of baseline and endline teacher characteristics across treatment and control. The age, gender and experience of teachers are very similar across both groups at the two time points. Teachers report overall higher levels of formal training at endline than baseline, but there are no significant differences between the two groups.

⁵⁶ Because of the difference in samples at baseline and endline, 45% of schools did not have data on school KCPE average scores, a key control variable, and 23% had missing baseline EGRA scores, a moderator in the model. See Appendix 2 for detail on the MI approach, missingness proportion and decisions informing the imputation model.

more than 3. While there was no significant difference at baseline, 68% and 67% of students in treatment schools at endline reported having the English and Kiswahili reading textbooks at home respectively, compared with 57% and 56% for the control group (see Table 5 for more details).

Table 6 presents the results of estimating the ATE from equation 1 with five specifications for each of the four literacy outcomes. The first column for each outcome presents the unconditional treatment effect, the second and third column present the full model with and without KCPE baseline controls respectively. The fourth and fifth columns present results with mean and multiple imputation of KCPE baseline scores, respectively. I consider the fifth model the best estimate of treatment impact as it contains the largest sample, and multiple imputation of KCPE leads to estimates very close to the control model in the second column. The main findings from this analysis are that PRIMR had a large and statistically significant impact on word reading and comprehension skills in both languages.

Students who were in treatment schools at endline read on average nearly 8 English words and 6 Kiswahili words respectively more than their peers in control schools and answered 0.59 and 0.52 more comprehension questions correctly ($p < .001$ for all estimates). These translate to effect sizes ranging from 0.31 to 0.36 standard deviations, which is slightly smaller but substantively consistent with the effect sizes ranging from 0.34 to 0.40 reported in the original study (see tables 12 and 13 in RTI International, 2016b).⁵⁷ The unconditional effect is slightly larger, but estimates are

⁵⁷ Models using standardized outcomes are included in Appendix 3.

substantively robust to the inclusion of control variables, as well as different specifications of the model and imputation approaches. Table 6 also shows that attending nonformal schools is associated with higher literacy skills on average after accounting for school and family resources and treatment. Having a teacher who reports having been formally trained is not associated with significantly higher literacy outcomes. The full model in the fifth column, using all covariates and MI for the KCPE baseline score, serves as the basis for moderation analysis presented below.

4.2 Heterogeneity in Gains

Figure 2 displays the school-level gains in average student literacy scores between the beginning of the 2012 school year (baseline) and the end of the 2013 school year (endline). As expected, students in the treatment schools largely outgrew their peers in control schools, with average school scores increasing 24 English words and 17 Kiswahili words compared with 18 English words and 12 Kiswahili words respectively in control schools. Figure 2 highlights the important spread in the gains of literacy scores in both groups, especially in treatment schools. The average student score in some treatment schools grew by less than 5 words between baseline and endline, compared with a gain of 30 to 50 words on average in other schools. Similarly, the average number of comprehension questions answered correctly declined in a few schools, while it grew by 2 or 3 questions out of 5 for approximately one third of treated schools. In the absence of change, one would expect to see some variation because of measurement error and cohort effects. Measurement error would cause change in school-level scores from one year to the next because of the sampling of different student, as well as random variation in the performance of students due to assessment conditions. Cohort effects are manifested by a

group of students performing better or worse than their peers in the previous year, on average. In control schools, heterogeneity in gains are likely the combination of measurement error, cohort effects, and differences due to learning environment resulting from school and teacher characteristics. In treatment schools, these factors could also play a role, as well as heterogeneity in treatment effect. The following sections present findings from formally testing for school and teacher level moderating factors.

4.3 Teacher-level Moderation

Table 7 presents results of models estimating teacher-level moderation using equation 2. The first and third columns of each panel show that on average, formal teacher training and years of experience are overall not statistically significantly associated with learning outcomes. Having a teacher with formal training is associated with slightly higher scores in English Word and Comprehension subtasks but essentially no association with Kiswahili subtasks. Similarly, each additional year of experience of a teacher has a small, non-significant positive association with all outcomes except Kiswahili comprehension.

The second and fourth columns of each panel include the interaction terms testing for treatment effect heterogeneity. Results consistently suggest the intervention had larger effects for students in classrooms with teachers having received no formal training, but again, estimates of the difference are not statistically significant at the 0.05 level. For example, while the average treatment effect was 7.93 English words, the estimate for students taught by untrained teachers is 10.41 words compared with 5.07 for students in

classrooms with trained teachers.⁵⁸ This pattern is true across the four literacy outcomes. Similarly, students in classrooms with more experienced teachers seem to have benefited less than those in classrooms with relatively more novice teachers. The magnitude of these differences, however, is very small and only statistically significant for Kiswahili Comprehension. Overall these results suggest that the intervention was successful in all classrooms, regardless of teacher training or experience.

4.4 School-level Moderation

Table 8 presents results of models estimating school-level moderation using equation 3 and is organized similarly to table 6. The second column in each panel shows that the intervention was only effective on high-level outcomes in nonformal private schools. Including the interaction term results in negative and non-significant treatment impact estimates for students in public schools. In contrast, the magnitude of the impact in nonformal schools is larger, with students in treatment nonformal schools reading over 11 English words ($p < .01$) and 9 Kiswahili words ($p < .001$) more than their peers, on average, which is equivalent to 0.47 and 0.51 standard deviations respectively.

Higher school-level baseline scores are associated with both higher student performance at endline across all outcomes and larger treatment impact in Kiswahili. Column 3 in all panels shows the positive association across all outcomes, with a fairly consistent association. A student from a school with 1 standard deviation higher average baseline score read 6.07 more English words ($p < .001$) and answered one third more comprehension questions correctly ($p < .001$), with similar results in Kiswahili. More

⁵⁸ Computed by adding the -5.33 estimate on the interaction term.

importantly, the intervention seems to have been more beneficial to students in schools with higher baseline scores, especially in Kiswahili. PRIMR led to an increase of 2.47 more Kiswahili Words ($p < .05$) and 0.27 more Comprehension questions ($p < .001$) respectively for students in schools having scored 1 standard deviation higher at baseline on average, which is equivalent to a 0.13 and 0.19 standard deviation difference. The estimates of interaction terms for English outcomes are also positive but not statistically significant.

Figures 3 and 4 present predicted relationships between student literacy outcomes and treatment status for schools moderated by school type and baseline EGRA scores respectively. In Figure 3, I use student SES as a predictor to highlight both the positive association between family resources and student outcomes, and the average relationship between school status, treatment, and student outcomes. Across all four outcomes, students in nonformal treatment schools (solid red line) perform substantively better than their peers in public schools. The figure also shows that students in public schools (including treatment schools) perform much lower than their peers in nonformal treatment schools, and roughly the same as students in nonformal control schools. Figure 3 shows the positive association between baseline school average and endline student outcomes, the positive impact of treatment, and the difference in that association across treatment status for Kiswahili outcomes. In schools with baseline Kiswahili scores more than 1 standard deviation below the mean, the average difference between students in treatment and control schools is very small. In contrast, students in schools that started much higher than average benefited greatly from PRIMR.

To assess whether the literacy skills are different for students in low-performing treatment and control schools, I conduct post-hoc General Linear Hypothesis tests for the 10th and 25th percentile baseline scores. For Kiswahili Word, I fail to reject that there are no differences between treatment and control schools at the 10th percentile ($F(1, 314) = 1.13$; $p=0.29$). In contrast I can reject at the 25th percentile ($F(1, 314) = 4.53$; $p=0.03$) and conclude that students in treatment schools at the 25th percentile of baseline scores had higher scores than their peers in control schools. Similarly, for Kiswahili Comprehension, I fail to reject that there are no differences at the 10th percentile ($F(1, 314) = 0.17$; $p=0.68$) and can reject at the 25th percentile ($F(1, 314) = 4.34$; $p=0.04$). These findings suggest that the intervention was effective in most schools but that schools in the bottom 10 percent of the baseline distribution did not benefit from PRIMR for Kiswahili. These schools are also largely public schools, as can be seen in figure 4, which is consistent with findings above.

5. Discussion and Conclusion

This study supports findings earlier reported about the large average treatment effects of PRIMR on literacy scores in English and Kiswahili (Piper et al., 2015; RTI International, 2016b). Importantly, however, this study shows that gains in literacy outcomes was not homogenous across school type and context. The large average treatment impact on word reading and comprehension was driven by the successful implementation in nonformal schools, and no average impact was found on higher-level literacy skills in public schools. In addition, the magnitude of treatment impact in Kiswahili was larger for schools that started with a better learning environment as

measured by average student performance. In contrast, the impact did not vary across teacher characteristics. These findings raise important questions to explore in the design and evaluation of literacy interventions that follow a similar theory of change in low- and middle-income countries, especially when considering equity as an important motivation of government policies and nongovernmental initiatives, as well as how to most effectively use limited resources.

Teacher Characteristics

The finding that the intervention was effective for all teachers, formally trained and not, novice or veterans, should be a source of satisfaction both for those who designed the intervention, and for the decision to scale-up the intervention. With the large increase in student enrollment in primary school across LMICs, a common response of governments and schools has been to hire teachers without formal training or certification (Bourdon, Frölich, & Michaelowa, 2010). The evidence on the impact of these teachers on student learning is mixed, but generally points to a lack of average difference in student outcomes across teacher status (Chudgar, 2015; Muralidharan & Sundararaman, 2013). This study also shows no association between teacher training and student outcomes but complements our understanding of this relationship by showing no difference in treatment impact. In other words, teachers with and without formal training responded to the professional development opportunity in similar ways. Similarly, the impact on learning from teachers with few years of experience and veteran teachers was the same on average, suggesting they had the same response to the training and resources.

These findings help to inform the debate on the usefulness of scripted lessons in low-resource contexts. The provision of scripted lessons to support teachers is generally

seen either as a means to turn teaching into a sterile and uncreative process (Dresser, 2012) or as an important support to novice teachers (Grossman & Thompson, 2008). The extent to which these postures are correct depends heavily on the level of scripting and the extent to which these scripts are meant to serve as scaffolds that teachers can build off. A recent comparison of early grade reading teacher guides found that scripting ranged from as few as 150 words per lesson to over 3000, with an average of 836, concluding that most successful guides had shorter lessons and were overall kept simpler (Piper, Sitabkhan, Mejía, & Betts, 2018).⁵⁹ The fact that trained and untrained teachers had the same benefits from participating in PRIMR is relatively unsurprising given the low levels of pedagogical training and reading instruction support provided in Teacher Colleges in Kenya. The lack of difference across years of teaching experience is more surprising given previous research on differential impact of teacher training, which suggests that veteran teachers were similar to their novice peers, on average, in their ability to teach reading, and their need for teaching resources. Taken together, these findings point to the need for teacher training focused on reading instruction to be integrated in pre-service and in-service, and both efforts are needed to improve reading outcomes of children.

School characteristics

Findings that the intervention was only successful in improving higher-level literacy skills in nonformal schools and potentially increased the learning gaps existing at

⁵⁹ Lessons in PRIMR in first grade contained 1203 and 1549 words per lesson in Kiswahili and English respectively, well above the average.

baseline between the two types of schools present important concerns. Analysis on the impact of the intervention on lower-level literacy skills (sound and letter recognition) revealed that the intervention had large impacts in both public and private schools.⁶⁰ While students in treated public schools performed better than their peers in nontreated schools on higher-level skills, the impact of receiving treatment is no longer significant when including school and student resources as covariates.⁶¹ Given the design of the original study was not focused on estimating differential impacts, and the fact that school status and baseline scores are correlated (nonformal schools had higher average baseline scores), it is not possible to completely disentangle the effect of these moderators. Below, I consider three factors that could explain the differential impact: learning environment, accumulated advantage, and teaching at the right level.

One possibility for the lack of impact in public schools is a generally weak learning environment. The baseline school-level measures used as a moderator in this study also constitute a proxy for learning environment. Teacher absenteeism is notoriously problematic in public institutions in LMICs, averaging 20% across countries (Chaudhury, Hammer, Kremer, Muralidharan, & Rogers, 2006). Learning conditions are sometimes worse, as found in an early grade program in public schools in Mozambique where student absenteeism ranged from 52% to 62% and teacher absenteeism was above

⁶⁰ Table A9 in appendix 3 replicates the school-level moderation analysis using sound and letter recognition in English, and letter and syllable recognition in Kiswahili. These were the two lower-skills assessed at endline. Findings show no differential impact across type of school for sounds and letters, but a similar difference for syllables as in the main results of the study.

⁶¹ The original impact evaluation focused on total average treatment effect. An appendix showed the disaggregated impact across multiple groups, including school type, as the average difference between groups divided by the pooled standard deviation, without covariates (see Appendix B of Piper et al. (2016)).

30% (Burchfield, Hua, Noyes, & van de Waal, 2017). Without more implementation data, this study cannot explain the reasons for which PRIMR had such large effects in one context and no average impact in public schools. As noted earlier, the overall success of PRIMR served as a justification to rollout the national reading program Tusome to be implemented in 28,000 schools, all of which are public except for 1,000 nonformal schools (USAID, 2014). This study points to the need to investigate the impact of programs across contexts and modify them accordingly to inform policy.

Findings that students in schools with higher learning at baseline benefited more from PRIMR, thereby increasing the learning gap and reducing equity, are consistent with the accumulated advantage framework presented in figure 1. Piper et al. (2015) noted that PRIMR had a larger impact on non-poor students than poor students. This study finds that controlling for an individual's family resources, the intervention was more successful in better-performing schools. The Kiswahili outcomes represented in figure 3 display a similar pattern as the stylized figure 1 representing accumulated advantages. Unsurprisingly, the schools in first decile of baseline performance has fewer resources than school in the rest of the sample, and students in these schools came from lower-resourced homes.⁶² The fact that students in these schools were less likely to speak the same language at home and at school, that teachers were less likely to be observed regularly, and that the schools were more likely to have been closed during the academic year present avenues to further explore treatment heterogeneity. Regardless of the reasons for the lack of impact in these schools, the findings counter the renewed focus on equity

⁶² Tables 5, 6, and 7 in Appendix 3 present a comparison of student, teacher and school characteristics between the first decile and the rest of the sample.

declared in goal 4 of the SDGs (United Nations, 2015). Equity goes beyond equal opportunity and requires a normative judgement as to what degree of inequality in a distribution of outcomes a society is willing to accept (UNESCO Institute for Statistics, 2018). PRIMR was already focused on children in Kenya who did not have access to the best educational opportunities the country had to offer. It remains to be seen how Tusome and other national reading programs will impact the distribution of reading outcomes, but a clear focus on equity should ensure that the conditions for children and communities at the lower end of the economic and educational distribution are not left behind by a rich-get-richer pattern because of structural barriers to learning.

Another consideration to explain the difference in the impact across types of school is a potential mismatch between the level of learning targeted by the intervention and the level of students in some schools. A growing body of literature is looking at interventions focused on *teaching at the right level* (TaRL). These include programs that provide an additional teacher for remedial education of lower-performing students (Banerjee, Cole, Duflo, & Linden, 2007), grouping students by ability (Duflo, Dupas, & Kremer, 2011), or using computer-adapted instruction (Muralidharan, Singh, & Ganimian, 2016). A similar paradigm can explain the fact that public schools benefiting from the intervention, especially those in the lower part of the distribution, only managed to increase lower-level literacy skills. While PRIMR targeted a range of literacy skills in both English and Kiswahili, there were no provisions to implement only one part of the program based on the initial level of the students. Designing interventions that include different levels using the TaRL framework have the potential to help lower-performing students make progress they would not otherwise make when instruction is too difficult.

However, it could also lead to widening gaps if higher performing students in low-performing schools or groups are not benefiting from an instruction with higher expectations.

Limitations

Two important limitations of the approach I use result from the study not being designed to test treatment heterogeneity. First, testing moderation in a post-hoc analysis should be clearly understood as a hypothesis-generating study to inform future hypothesis-testing studies as argued by Kraemer et al. (2002). In addition, the statistical power for these tests may be inadequate, as the sampling frame of the design was meant to identify average treatment effects. As Gelman (2018) highlights, the sample sizes necessary to estimate interaction terms can be four to sixteen times larger than that needed to identify main effects. In a context where the cost of evaluating interventions is already debated, designing studies that cost substantively more to detect treatment variation may be impractical. However, if there is a good theoretical reason to believe the intervention will not have the same impact across characteristics of teachers, or across contexts, it is worth exploring this during smaller scale programs such as PRIMR.

Conclusion

Improving literacy skills of the growing number of children having access to formal education is rightly a priority of governments and international stakeholders. The large investment in developing literacy interventions and scaling up these programs as was the case in Kenya necessitates careful research to understand the factors that promote their success. While this study was able to highlight that school context moderates the effect of the PRIMR intervention, the post-hoc analytical strategy has important

limitations. Designing interventions and evaluations with moderation in mind would enable research to collect adequate variables more closely aligned with theory in samples large enough to estimate these effects with more precision.

Table 1. Sample size and collection of data

		Baseline	Endline
Time		January 2012	Sept/Oct 2013
Clusters			
	Treatment	25	25
	Control	8	7
Schools			
	Treatment	160	159
	Control	60	53
Teachers			
	Treatment	315	313
	Control	97	102
Students			
	Treatment	3195	3176
	Control	1190	1046

Table 2. Baseline Student Characteristics: description and balance

	Control	Treatment	Control-Treatment
Student is female	0.497 (0.003)	0.499 (0.003)	-0.001
Student age	6.960 (0.056)	7.060 (0.045)	-0.099
Student is in grade 2	1.503 (0.004)	1.499 (0.001)	0.004
Student has a radio at home	0.767 (0.015)	0.769 (0.009)	-0.002
Student has a phone or mobile at home	0.924 (0.013)	0.922 (0.009)	0.002
Student has electricity at home	0.669 (0.042)	0.725 (0.024)	-0.055
Student has a television at home	0.653 (0.031)	0.703 (0.021)	-0.050
Student's mother can read and write	0.925 (0.010)	0.919 (0.007)	0.006
Number of words correct - English	13.007 (1.221)	15.008 (0.692)	-2.001
Comprehension questions correct - English	0.553 (0.070)	0.614 (0.040)	-0.061
Number of words correct - Kiswahili	10.551 (0.839)	11.164 (0.479)	-0.612
Comprehension questions correct - Kiswahili	0.791 (0.067)	0.829 (0.037)	-0.038
Maximum Number of Observations	1190	3195	4385
Date		January 2012	
Test for jointly significance			
F statistic		1.68*	
P-Value		0.07	

Note. This table presents the coefficient of the OLS regression of the treatment indicator (equal 1 if treated; 0 otherwise) against each of the listed baseline characteristics. Treatment includes schools receiving the intervention at least one year. * $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$

Table 3. Baseline Teacher Characteristics: Description and Balance

	Control	Treatment	Control-Treatment
Percentage of female teachers	0.84 (0.04)	0.89 (0.02)	-0.05
Age	35.12 (1.14)	34.72 (0.60)	0.40
Teacher formally trained	0.36 (0.05)	0.46 (0.03)	-0.10*
Teaching experience (years)	12.07 (1.09)	11.26 (0.55)	0.81
Teacher observed at least monthly	0.53 (0.05)	0.49 (0.03)	0.04
Teacher presents good scheme of work	0.44 (0.05)	0.39 (0.03)	0.05
Teacher presents good lesson plan	0.31 (0.05)	0.29 (0.03)	0.01
Maximum Number of Observations	96	315	411
Date		January 2012	
Test for jointly significance			
F statistic		1.18	
P-Value		0.31	

Note. This table presents the coefficient of the OLS regression of the treatment indicator (equal 1 if treated; 0 otherwise) against each of the listed baseline characteristics. Treatment includes schools receiving the intervention at least one year. * $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$

Table 4. Baseline School Characteristics: description and balance

	Control	Treatment	Control-Treatment
School Nonformal	0.583 (0.064)	0.563 (0.039)	0.021
Head Teacher is female	0.569 (0.066)	0.454 (0.041)	0.115
Head Teacher Experience (Years)	7.203 (0.759)	7.553 (0.407)	-0.350
School has been closed since start of year	0.051 (0.029)	0.087 (0.022)	-0.037
School Mean Score on KCPE	220.155 (7.712)	237.337 (4.198)	-17.182**
Clean and Safe Water Supply	0.610 (0.064)	0.706 (0.036)	-0.096
Presence of electricity	0.373 (0.063)	0.594 (0.039)	-0.221***
Presence of Computer Room	0.085 (0.037)	0.131 (0.027)	-0.047
Presence of Library	0.339 (0.062)	0.350 (0.038)	-0.011
School Average English Word Correct	13.216 (1.260)	15.015 (0.692)	-1.799
School Average English Comprehension	0.548 (0.069)	0.613 (0.040)	-0.064
School Average Kiswahili Word Correct	10.700 (0.871)	11.167 (0.478)	-0.467
School Average Kiswahili Comprehension	0.799 (0.069)	0.828 (0.037)	-0.029
Maximum Number of Observations	60	160	220
Date		January 2012	
Test for jointly significance			
F statistic		1.57*	
P-Value		0.099	

Note. This table presents the coefficient of the OLS regression of the treatment indicator (equal 1 if treated; 0 otherwise) against each of the listed baseline characteristics. Treatment includes schools receiving the intervention at least one year. * $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$

Table 5. Implementation Data

	Treatment	Control	Difference	p-value
Student has English reading textbook at home				
Baseline	45%	43%	2%	0.24
Endline	68%	57%	11%	<.001
Student has Kiswahili reading textbook at home				
Baseline	45%	43%	2%	0.30
Endline	67%	56%	10%	<.001
Teacher reports teaching PRIMR English lessons (Midline)				
2 days per week or less	0%			
3 days per week	2%			
4-5 days per week	98%			
Teacher reports teaching PRIMR Kiswahili lessons (Midline)				
2 days per week or less	0%			
3 days per week	2%			
4-5 days per week	98%			
Teacher reports attending PRIMR workshops (Endline)				
1 or fewer	7%			
2 or 3	43%			
> 3	50%			
Head-Teacher reports attending PRIMR workshops (Endline)				
1 or fewer	18%			
2 or 3	27%			
> 3	55%			

Table 6. Average Treatment Effect

	English Word					English Comprehension					Kiswahili Word					Kiswahili Comprehension				
Treatment	9.78*** (1.88)	7.91*** (1.63)	5.53** (2.08)	7.78*** (1.65)	7.93*** (1.67)	0.69*** (0.11)	0.60*** (0.11)	0.33* (0.13)	0.59*** (0.11)	0.59*** (0.11)	6.90*** (1.27)	5.91*** (1.17)	2.83 (1.51)	5.64*** (1.17)	5.74*** (1.19)	0.54*** (0.09)	0.52*** (0.09)	0.30* (0.12)	0.50*** (0.09)	0.52*** (0.09)
Student SES		2.74*** (0.42)	1.99*** (0.50)	2.64*** (0.42)	2.67*** (0.42)		0.25*** (0.03)	0.13*** (0.03)	0.24*** (0.03)	0.25*** (0.03)		1.59*** (0.33)	0.90* (0.42)	1.47*** (0.33)	1.54*** (0.33)		0.09*** (0.02)	0.03 (0.03)	0.08** (0.02)	0.09** (0.02)
Teacher is Trained		1.37 (1.16)	2.25 (1.46)	1.31 (1.16)	1.16 (1.18)		0.16 (0.09)	0.15 (0.11)	0.16 (0.09)	0.16 (0.09)		0.11 (0.89)	0.86 (1.21)	0.02 (0.89)	0.04 (0.91)		-0.04 (0.07)	-0.00 (0.09)	-0.05 (0.07)	-0.05 (0.07)
Baseline KCPE Score			2.26*** (0.55)	2.00** (0.66)	1.28* (0.52)			0.24*** (0.04)	0.20*** (0.04)	0.11** (0.04)			2.35*** (0.44)	2.12*** (0.49)	1.35** (0.39)			0.17*** (0.03)	0.15*** (0.03)	0.09** (0.03)
School is nonformal		14.14*** (1.87)	11.10*** (2.50)	13.50*** (1.89)	13.40*** (1.90)		0.82*** (0.16)	0.51* (0.21)	0.76*** (0.16)	0.77*** (0.16)		10.48*** (1.36)	7.42*** (1.71)	9.85*** (1.32)	9.80*** (1.35)		0.71*** (0.12)	0.38** (0.14)	0.66*** (0.11)	0.66*** (0.11)
Constant	29.66*** (1.62)	19.38*** (2.49)	23.82*** (3.17)	19.44*** (2.68)	20.05*** (2.54)	0.82*** (0.09)	0.22 (0.20)	0.67** (0.24)	0.21 (0.21)	0.28 (0.20)	21.33*** (1.05)	13.28*** (1.88)	19.37*** (2.54)	14.15*** (2.02)	14.02*** (1.90)	1.23*** (0.08)	0.74*** (0.15)	1.29*** (0.20)	0.79*** (0.16)	0.79*** (0.15)
N Students	4202	4202	2316	4202	4162	4202	4202	2316	4202	4162	4202	4202	2316	4202	4162	4202	4202	2316	4202	4162
Controls	N	Y	Y	Y	Y	N	Y	Y	Y	Y	N	Y	Y	Y	Y	N	Y	Y	Y	Y
KCPE Mean Imputation	N	N	N	Y	N	N	N	N	Y	N	N	N	N	Y	N	N	N	N	Y	N
KCPE Multiple Imputation	N	N	N	N	Y	N	N	N	N	Y	N	N	N	N	Y	N	N	N	N	Y

Note. Controls also include student grade, school-level resources, and district fixed-effects.

* p<0.05 ** p<0.01 ***p<0.001

Table 7. Teacher-level Moderation

	English Word				English Comprehension				Kiswahili Word				Kiswahili Comprehension				
Treatment	7.93*** (1.67)	10.41*** (2.19)	8.01*** (1.65)	8.59*** (2.05)	0.59*** (0.11)	0.63*** (0.14)	0.59*** (0.11)	0.53*** (0.13)	5.74*** (1.19)	7.09*** (1.61)	5.77*** (1.18)	6.91*** (1.49)	0.52*** (0.09)	0.62*** (0.12)	0.52*** (0.09)	0.63*** (0.12)	
Teacher is Trained	1.16 (1.18)	4.93* (2.35)	0.93 (1.17)	0.89 (1.19)	0.16 (0.09)	0.22 (0.14)	0.14 (0.09)	0.15 (0.09)	0.04 (0.91)	2.08 (1.55)	-0.04 (0.90)	-0.11 (0.91)	-0.05 (0.07)	0.11 (0.13)	-0.06 (0.07)	-0.07 (0.07)	
Treatment X Trained		-5.33 (2.74)				-0.09 (0.17)				-2.89 (1.91)				-0.23 (0.15)			
Teacher Experience (years)			0.15 (0.08)	0.20 (0.14)			0.01 (0.01)	0.00 (0.01)			0.05 (0.06)	0.15 (0.11)			0.00 (0.00)	0.01* (0.01)	
Treatment X Experience				-0.08 (0.16)				0.01 (0.01)				-0.15 (0.13)				-0.01 (0.01)	
Constant	20.05*** (2.54)	18.97*** (2.61)	17.29*** (2.78)	17.23*** (2.80)	0.28 (0.20)	0.26 (0.20)	0.09 (0.21)	0.10 (0.21)	14.02*** (1.90)	13.44*** (1.95)	13.14*** (2.15)	13.02*** (2.18)	0.79*** (0.15)	0.74*** (0.16)	0.70*** (0.17)	0.69*** (0.17)	
N Students		4162				4162				4162				4162			
Controls		Y				Y				Y				Y			
KCPE Multiple Imputation		Y				Y				Y				Y			

Note. Controls also include student grade and SES, school-level resources, school baseline KCPE, school type, and district fixed-effects.

* p<0.05 ** p<0.01 ***p<0.001

Table 8. School-level Moderation

	English Word				English Comprehension				Kiswahili Word				Kiswahili Comprehension			
Treatment	7.94*** (1.67)	-0.76 (2.74)	7.49*** (1.53)	7.50*** (1.53)	0.59*** (0.11)	-0.01 (0.15)	0.58*** (0.10)	0.58*** (0.10)	5.75*** (1.19)	-1.64 (2.41)	5.68*** (1.16)	5.64*** (1.15)	0.52*** (0.09)	-0.14 (0.16)	0.48*** (0.09)	0.46*** (0.09)
School is Nonformal	13.44*** (1.90)	3.06 (3.63)	11.94*** (1.96)	12.19*** (2.01)	0.77*** (0.16)	0.06 (0.23)	0.56** (0.20)	0.58** (0.20)	9.80*** (1.35)	0.99 (2.88)	8.92*** (1.35)	9.59*** (1.41)	0.65*** (0.11)	-0.13 (0.21)	0.45*** (0.12)	0.53*** (0.12)
Treatment X Nonformal		11.13** (3.37)				0.76*** (0.20)				9.46*** (2.76)				0.84*** (0.19)		
Baseline School Average			6.07*** (0.80)	5.31*** (1.18)			0.34*** (0.05)	0.29*** (0.06)			4.29*** (0.66)	2.52** (0.86)			0.27*** (0.06)	0.08 (0.07)
Treatment X Baseline Avg.				1.05 (1.28)				0.08 (0.08)				2.47* (1.01)				0.27*** (0.08)
Constant	20.02*** (2.54)	28.48*** (3.31)	18.43*** (2.61)	18.33*** (2.61)	0.28 (0.20)	0.86*** (0.22)	0.20 (0.22)	0.19 (0.22)	14.01*** (1.90)	21.20*** (2.83)	13.13*** (2.03)	12.81*** (2.03)	0.79*** (0.15)	1.43*** (0.20)	0.84*** (0.17)	0.81*** (0.17)
N Students	4162	4162	3191	3191	4162	4162	3191	3191	4162	4162	3191	3191	4162	4162	3191	3191
Controls			Y				Y				Y				Y	
KCPE Multiple Imputation			Y				Y				Y				Y	

Note. Controls also include student grade and SES, school-level resources, school baseline KCPE, school type, and district fixed-effects.

* p<0.05 ** p<0.01 ***p<0.001

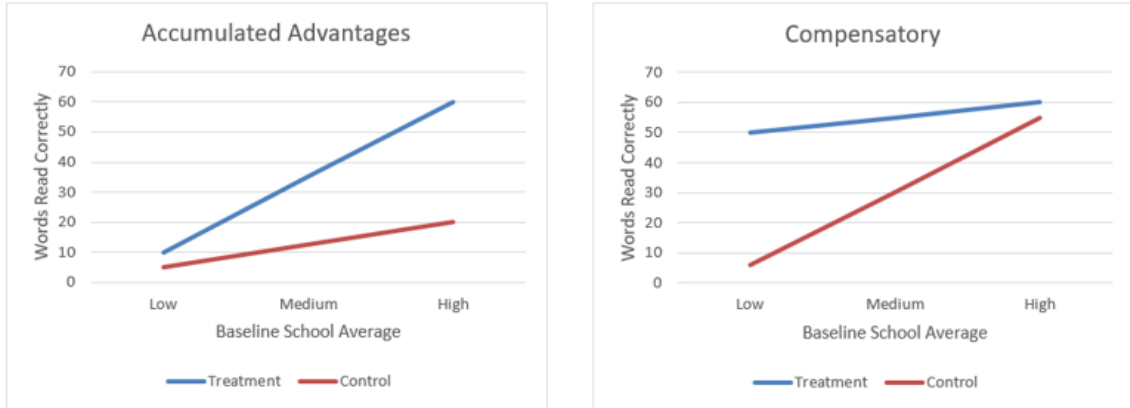
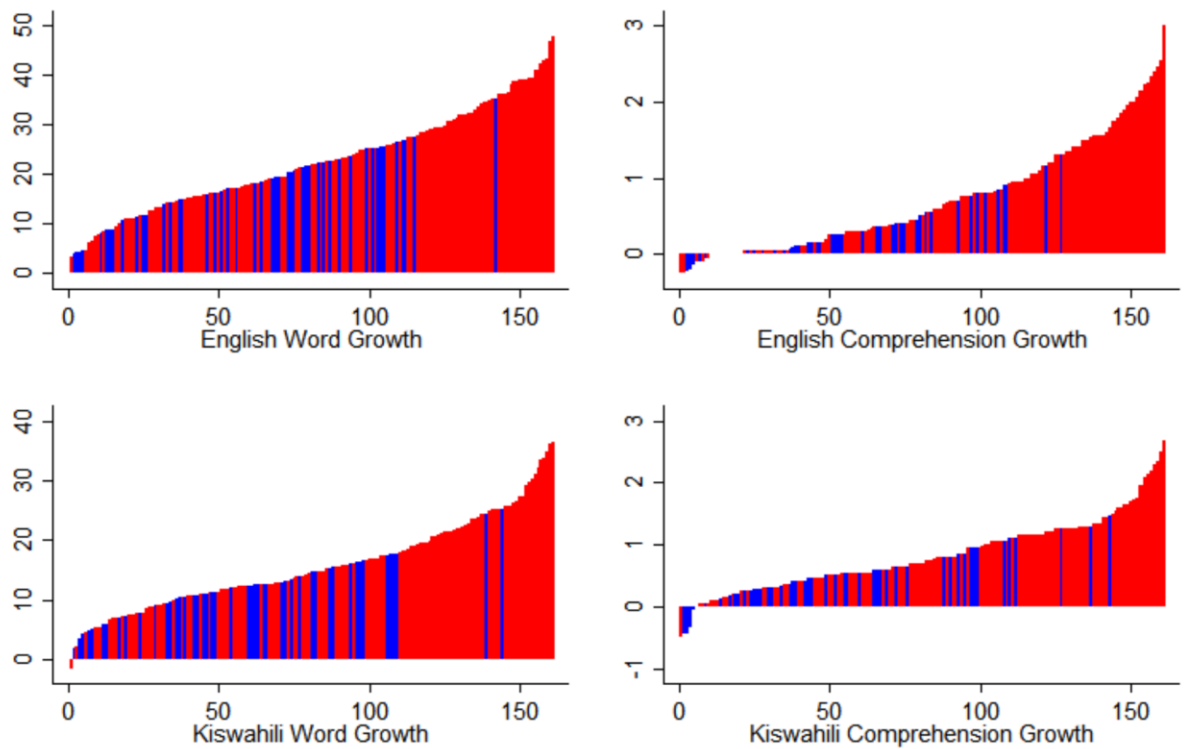
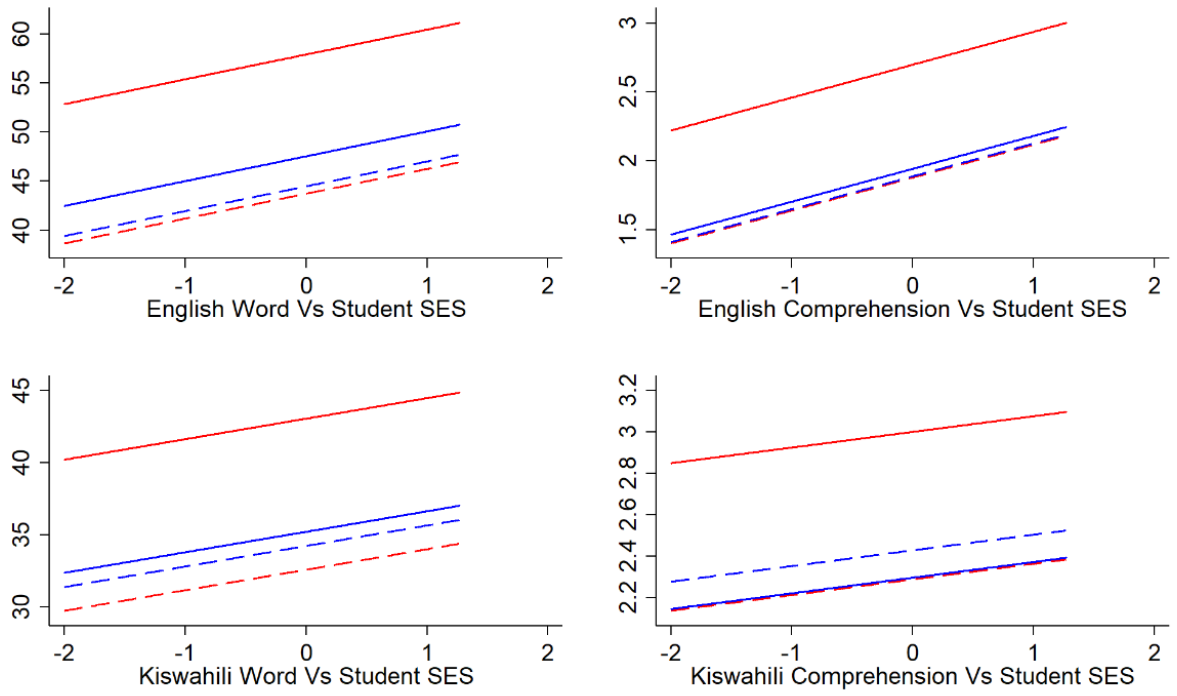


Figure 1: Idealized model of the two hypotheses around the impact of PRIMR on schools with varying baseline scores



Note: Schools Ordered by Magnitude of Growth
 Legend: Treatment Schools in Red and Control Schools in Blue

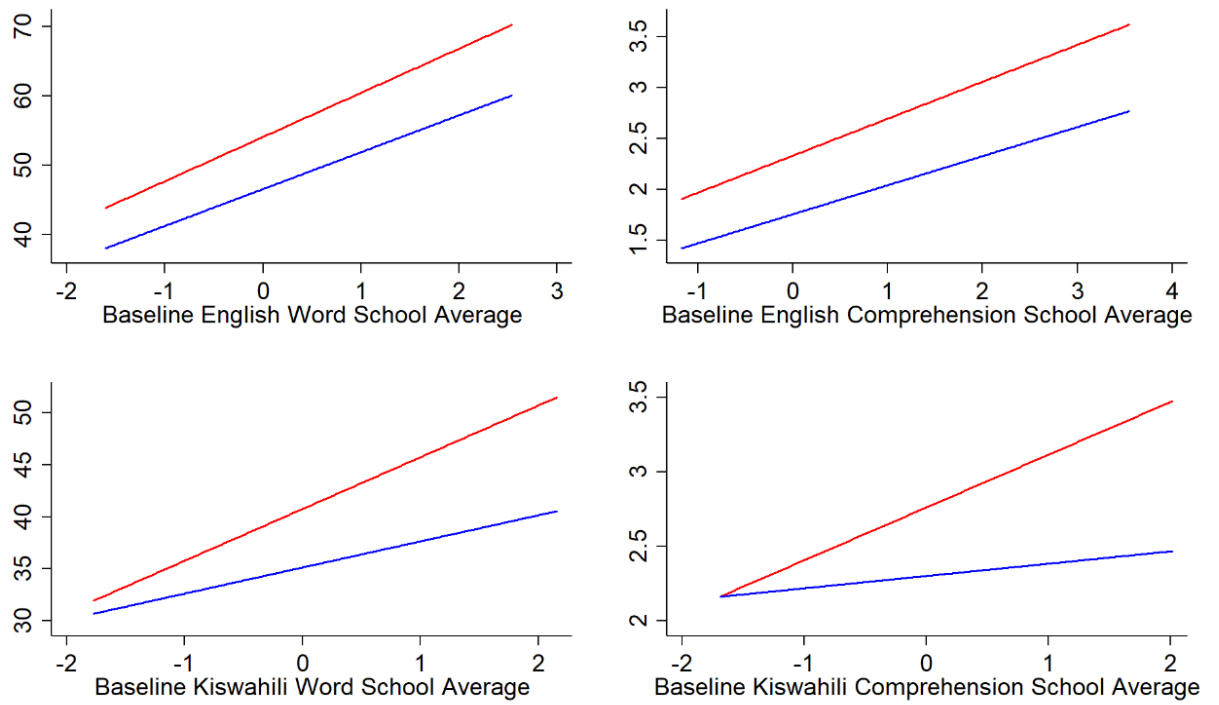
Figure 2: School level gains in literacy skills between baseline and endline



Note: Lines of best fit for grade 2 students in schools with average resources and KCPE baseline scores and untrained teachers in district 1.

Legend: Treatment Schools in Red and Control Schools in Blue
 Solid lines are nonformal schools and dashed lines are public schools

Figure 3: Predicted literacy scores of students as a function of SES by treatment status and school type.



Note: Lines of best fit for grade 2 students with average SES in schools with average resources and KCPE baseline scores and untrained teachers in district 1.
 Legend: Treatment Schools in Red and Control Schools in Blue
 Difference in Slopes in English outcomes is not significant

Figure 4: Predicted literacy scores of students as a function of school baseline literacy scores and by treatment status.

Conclusion

Education in low- and middle-income countries globally is characterized by a learning crisis illustrated by general low levels of learning and important inequities within countries. Children who speak a language at home that is not used in their classrooms are at a disadvantage compared with their peers who do not experience that misalignment. In addition to educational disparities, communities in multilingual states whose languages are not included in educational institutions face further marginalization and may not fully participate as citizens. This dissertation, focused on learning in multilingual contexts, presented a general framework for analyzing education language policies and included two empirical studies focused on early grade literacy acquisition in multilingual settings in Uganda and Kenya. At the center of the framework are two elements aimed at understanding whether a multilingual policy is necessary and supported: the sociolinguistic context and community support. The four elements at the periphery are factors that support successful implementation of a multilingual policy: teachers, assessments, orthography and curricula and materials. This conclusion employs the framework developed in the first paper to reexamine one of the central arguments of the dissertation, and expands on the need for supporting implementation of multilingual policies.

Acknowledging and Supporting Multilingualism Explicitly

One of the central claims in this dissertation is that countries should acknowledge multilingualism explicitly, and design education policies that take into account this multilingualism. This argument does not call for all countries to adopt or implement

education policies that use multiple languages as the medium of instruction in parts or all of the education system. Rather, the findings in this dissertation point to the need to question the *monolingual habitus* in which many policies are designed and implemented, especially in postcolonial settings such as Kenya and Uganda. The idea of a *linguistic habitus* builds on Bourdieu's concept of *habitus* – the set of principles and practices that individuals are socialized in and do not question – to describe the “the various language practices, including the various literacies of individuals and groups” that we are socialized into and do not question (Carrington & Luke, 1997, p.101). The *monolingual habitus* is thus a set of linguistic dispositions or practices that ignore multilingual ways of life. This monolingual habitus is evident in monolingual policies, efforts to develop one-nation-one-language states, and failures to recognize or attempts to erase linguistic diversity (Benson, 2013). The processes of subjugation or elimination of cultures and languages of minority groups has been justified by the perceived need for social and political cohesion in the nation-state (May, 2013). An alternative vision to the monolingual nation-state, one that embraces multilingualism, exists. For example, South Sudan, the world's youngest nation, gained independence in 2011 and embraced multilingualism in its constitution and its language in education policy that promotes the use of one of its 63 officially recognized languages as the medium of instruction for 3 years before transitioning to English (Spronk, 2014). Education language policies that embrace a *multilingual habitus* need not make all languages spoken in the state media of instruction. Rather, they can productively acknowledge the multilingual nature of their schools and classrooms and support pedagogies and language practices that recognize and promote linguistic diversity rather than erase it.

The first paper in this dissertation proposed a framework for analyzing language in education policies, including the need for and support toward multilingual policies. The two elements at the center of this framework –sociolinguistic context and community support for multilingual policies – can inform countries as to whether there is need and support for the use of multiple languages. The contrast between Kenya and Uganda is helpful to illustrate cases where the level of linguistic diversity – similar in the two countries – does not result in the same sociolinguistic context because of different colonial histories and the widespread use of a regional language in Kenya (Kiswahili). These differences result in variable levels of support for the implementation of multilingual policies at the institutional level.

Language policies and practices that allow for the use of multiple languages as the medium of instruction are not necessary or sufficient to support the development of a multilingual habitus. In settings that are highly linguistically heterogenous, such as the urban settings where PRIMR was implemented, it is possible to use a monolingual approach to instruction that recognizes that children come with different language skills. Hornberger and Link (2012) describe approaches that acknowledge and support multilingualism of students in classrooms that use only English as the official medium, such as a fifth grader using Spanish words in their poem shared with classmates who speak at least six other languages at home. In contrast, a policy such as the early-exit one implemented in Uganda does not guarantee support for multilingual practices. If students in the upper-primary grades are not allowed to use their first language or, as is the case in some settings, are punished for using a non-dominant language, the promotion of a first

language in the lower-primary grades does little to promote a multilingual habitus. This is the case even when students who start developing literacy skills in their first language are able to transfer these skills to English, as findings from the second paper in this dissertation suggest.

A multilingual education policy alone, however, do not necessarily translate into equal access to quality education and participation. South Sudan, for example, has been engulfed in a crisis since December 2013 that has contributed to nearly 2.5 million refugees and 2 million internally displaced (United Nations High Commission for Refugees (UNHCR), 2018) and severely impeded on the education of its citizens, with over 2 million children out of school (Unicef, 2019). The mere presence of multiple languages in a community and their official recognition is not sufficient. Similarly, acknowledging languages spoken by students in a classroom is not sufficient for countries aiming to reduce inequities and help all children thrive in their education system. For example, Mulimbi and Dryden-Peterson (2018) describe how the curriculum in Botswana promotes equal citizenship by *tolerating* individuals from ethnic minorities rather than *recognizing* the perspectives and contributions of groups.

Supporting Multilingual Education Practices

The four factors at the periphery of the framework proposed in Paper 1 concern supports needed to successfully implement monolingual or multilingual policies in a multilingual setting.⁶³ The two empirical studies in this dissertation present findings that

⁶³ The pertinence of orthography development in a monolingual policy is less salient, as non-dominant languages would likely be restricted to oral use.

relate directly to three of these components: curriculum and materials development, teacher training, and the design of assessments.

One of the important challenges education systems in low- and middle-income countries face is the recruitment, training and deployment of **teachers** who can support quality instruction. In settings that choose to use multiple languages for instruction, this challenge is compounded by the cyclical nature of educational inequities. Linguistic communities with low-levels of school completion will include smaller number of graduates who could be trained as teachers. This situation, in turn, results in a smaller number of teachers who speak the language able to be deployed to teach in those communities. In addition, teacher training is still conducted in the dominant language in many settings, including Kenya and Uganda, even when policies allow or promote the use of local languages in the classrooms where teachers are to be deployed. Further, pre-service training usually includes very little support to teaching early literacy skills in any language. As a result, in-service training such as those provided by SHRP in Uganda and PRIMR in Kenya are the primary means through which many teachers in the early grades in low- and middle-income settings receive support for literacy instruction. As the findings suggest in the third paper of this dissertation, this training can have large impact to help instruction of basic literacy skills for all teachers, regardless of their years of experience or pre-service training.

A second factor important to supporting the implementation of multilingual policies is the development of **curricula and teaching materials** in multiple languages. The settings described in this dissertation highlight the paucity of resources to support

instruction, especially in non-dominant languages. More importantly, curricula needs to be appropriate for the languages used in instruction. Recent work reviewing teacher guides in 19 early grade reading projects across 13 countries revealed important variation in the number of words per lesson (154 to 3,679) indicating different levels of scripting for teachers, as well as a large propensity from teachers to modify the lessons to meet their needs (Piper et al., 2018). This variation highlights the need for curricula that are developed locally and adapted to the needs of the language. Importantly, these guides, as well as those developed for the vast majority of literacy interventions supported by bilateral and multilateral agencies rely on monolingual reading comprehension frameworks developed for English in the United States (Kim et al., 2016). While the principles behind the simple view of reading comprehension (Hoover & Gough, 1990) as the product of decoding and oral comprehension likely apply in many languages and across contexts, the development of curricula in multilingual settings should examine this universality critically. The second paper in this dissertation, for example, suggests different findings for bilingual reading development in Central Uganda based on the language used in the classroom. This finding points to the need for different instructional approaches and potentially different materials to best support learning in urban and rural contexts because of the differing levels of linguistic homogeneity.

Finally, the two empirical papers in this dissertation highlight the need for more and better **assessments** across languages. As noted in the framework in Paper 1, formative assessments are important to support teacher instruction, and summative assessment can guide policy. Multilingual policies can only be successfully implemented

with the availability of multiple assessments in each language. While the last decade has seen a large increase in the availability of internally valid literacy measures in non-dominant languages, this increase has largely been limited to the EGRA and ASER-style measures of reading that provide very coarse measures of literacy proficiency. The increased availability of these measures must be acknowledged as a useful set of tools but must also be examined critically. The focus of instruction on an isolated set of foundational sub-skills is an important challenge that has risen with the influence of EGRA as an evaluative instrument (Bartlett, Dowd, & Jonason, 2015). This could be one of the reasons for the lack of impact from the PRIMR intervention on higher-level skills in public schools found in Paper 3 of this dissertation. With programs being evaluated mainly by their success in increasing letter naming, oral reading fluency, and comprehension, there is little incentive to focus on other important skills such as vocabulary or writing. The psychometric analysis of the oral language comprehension in Luganda and Runyankole-Rukiga in Paper 2 found very low internal validity because of the small number of items. Similarly, the reading comprehension tasks in EGRA and ASER only measure a limited range of comprehension skills. Confining the skills measured, and the languages in which they are assessed is problematic for teachers and policymakers who have larger goals for the education system.

The global commitment to promote access to quality education for all children includes a clear mandate to end inequities across several dimensions, including linguistic background of children. Through an increase in inclusive policies and numerous government and community-led initiatives, the number of languages used in formal

schooling has been increasing around the world. This increase should be seen as a progress toward reducing inequities, but should not reduce engagement with communities in making policy decisions or mask the need to support their implementation. Desai (2015) reminds us that “the merits or demerits of mother tongue education are dependent on the particular context in which learners find themselves” (p.22). The framework I propose in this dissertation expands the notion of *context* to include six elements necessary to examine the language policies that can lead to meaningful and effective instruction for all students in the classroom.

Appendices

Appendix A – Additional Figures and Tables for Paper 1

Appendix B – Note on the qualitative data for Paper 1

Appendix C – Additional Figures and Tables for Paper 2

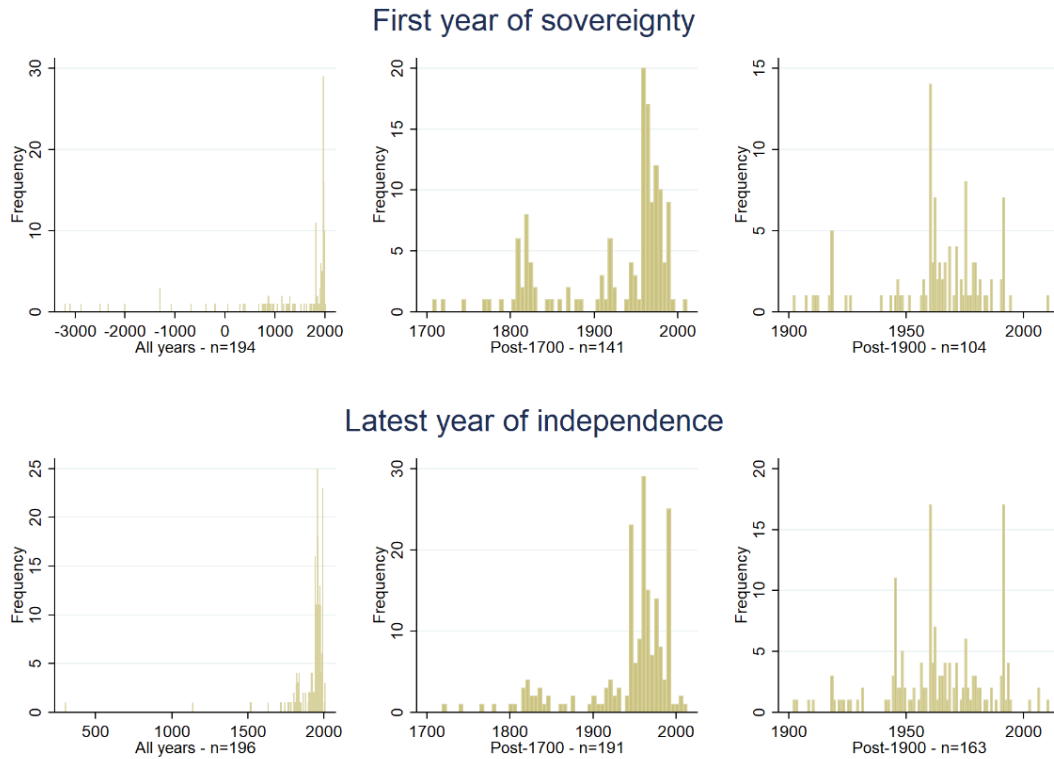
Appendix D – Methodological Approach to Address Clustering of Students for
Paper 3

Appendix E – Imputation of missing data for Paper 3

Appendix F – Additional Figures and Tables for Paper 3

Appendix A – Additional Figures and Tables for Paper 1

Figure A1. First year of sovereignty and latest year of independence for today's states.



Note: Data compiled by author using information accessed on October 16, 2018 from https://en.wikipedia.org/wiki/List_of_sovereign_states_by_date_of_formation. Negative numbers for dates indicate BC. Some of the dates in these data are imprecise (e.g. India first founded in 2000 BC), and the existence or date of independence of some states are contested (e.g. Palestine, 1988). Further, some partially recognized states are not included in this list (e.g. Western Sahara or Kosovo). My use of the data without modification is not a validation of the decisions made by the authors. My intention in creating this figure is to illustrate large trends in state formation rather than validate the status of individual state.

Figure A2. Map of Kenya and its languages.

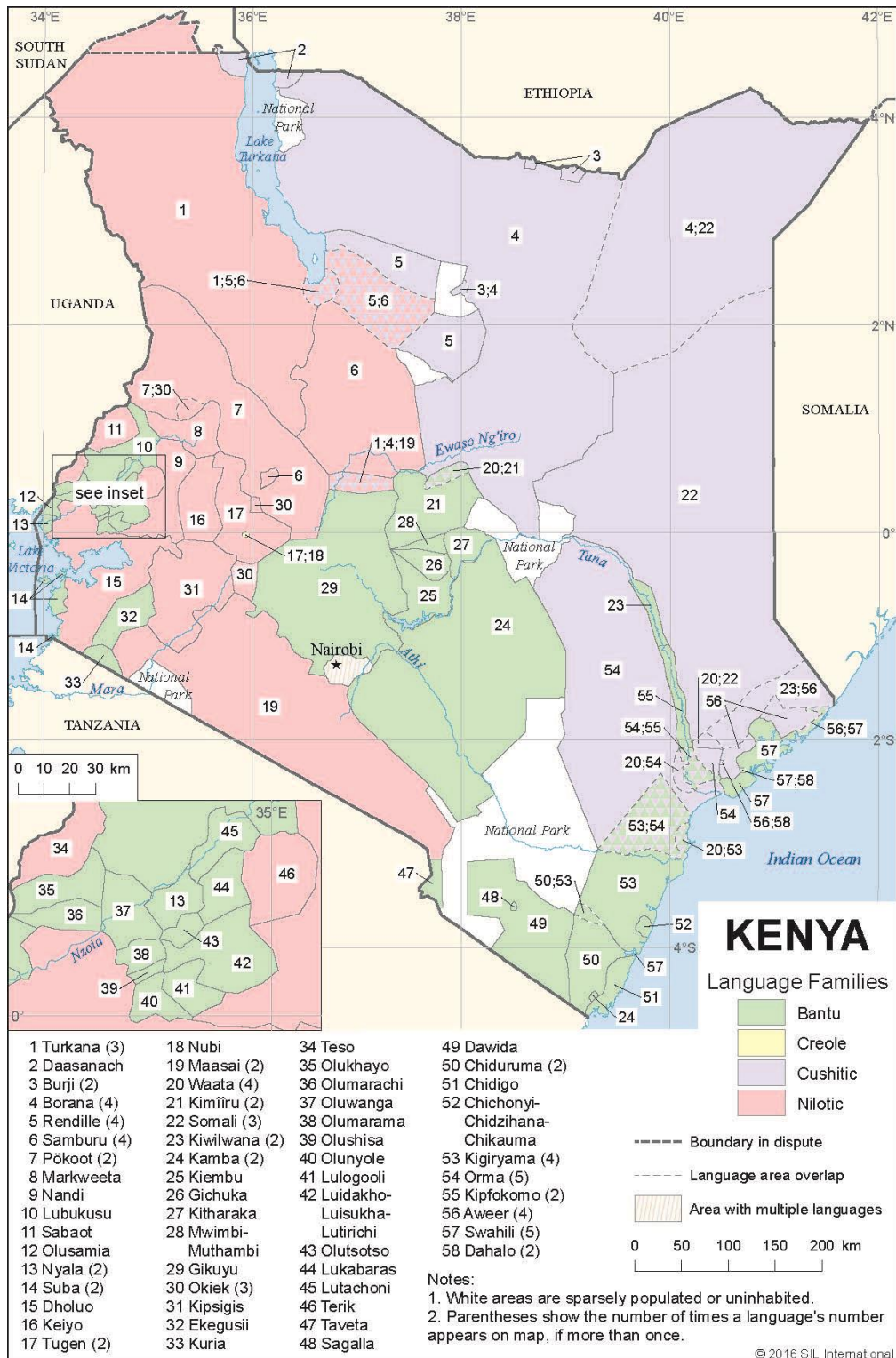


Figure A3. Map of Uganda and its languages.

Level	Label	Description
0	International	The language is widely used between nations in trade, knowledge exchange, and international policy.
1	National	The language is used in education, work, mass media, and government at the national level.
2	Provincial	The language is used in education, work, mass media, and government within major administrative subdivisions of a nation.
3	Wider Communication	The language is used in work and mass media without official status to transcend language differences across a region.
4	Educational	The language is in vigorous use, with standardization and literature being sustained through a widespread system of institutionally supported education.
5	Developing	The language is in vigorous use, with literature in a standardized form being used by some though this is not yet widespread or sustainable.
6a	Vigorous	The language is used for face-to-face communication by all generations and the situation is sustainable.
6b	Threatened	The language is used for face-to-face communication within all generations, but it is losing users.
7	Shifting	The child-bearing generation can use the language among themselves, but it is not being transmitted to children.
8a	Moribund	The only remaining active users of the language are members of the grandparent generation and older.
8b	Nearly Extinct	The only remaining users of the language are members of the grandparent generation or older who have little opportunity to use the language.
9	Dormant	The language serves as a reminder of heritage identity for an ethnic community, but no one has more than symbolic proficiency.
10	Extinct	The language is no longer used and no one retains a sense of ethnic identity associated with the language.

Table adapted from Ethnologue (n.d.).

Appendix B – Note on the qualitative data for Paper 1

As described in the main paper, the qualitative data presenting community perspectives on language of instruction in Uganda was collected as part of a study I conducted in 2014 (de Galbert, 2017). The questions relating to language of instruction were not analyzed in the published study⁶⁴. This appendix briefly describes the methods used to collect and analyze the data, as well as preliminary results. The focus groups were conducted to answer the two following research questions:

- How do parents perceive the Thematic Curriculum and the Language of Education Policy in Uganda?
- How do parents' perspectives about language relate to sociolinguistic and psycholinguistic theories of language planning?

Methods

In order to answer these research questions, I conducted 7 focus groups with 20 parents in total in communities in Central and Eastern Uganda in December 2014. The communities were selected from four different districts, including urban and rural settings⁶⁵. The languages spoken at home by participants were either Luganda or Lusoga. Participants were recruited through contacting Head Teachers who were informed of the goals of the study. The number of participants in each focus group ranged from 2 to 5, and included a total of 13 fathers and 7 mothers⁶⁶. The focus groups were guided by a semi-structured questionnaire with general questions about the *Thematic Curriculum* and language policies. These were purposefully open-ended and I asked follow-up questions

⁶⁴ The purpose of asking the questions was meant to serve as a pilot for a future larger study on parental perspectives.

⁶⁵ For more detail on the selection process, see de Galbert (2017).

⁶⁶ In some cases, the participants were legal guardians (siblings or uncles) but not biological parents. I refer to all participants as parents in this appendix.

to encourage discussion between participants. I recorded and transcribed the data from the focus groups, and analyzed them using Atlas.ti, a qualitative analysis software. I coded transcripts from the focus groups first using an etic approach focused on the two research questions. Codes included references to the thematic curriculum (e.g. positive perspective, negative perspective) and psycholinguistic and sociolinguistic theories (e.g. language transfer; language power; diglossia).

Given the small scale of this study, it is important to examine the results keeping in mind two important limiting factors. First, parents all spoke Bantu languages at home, and most spoke Luganda. As a result, there are no data from parents from a different linguistic family, which could affect their perspectives. Further, there are no data from more linguistically marginalized communities. Basoga speakers can be thought of having been linguistically dominated by Luganda, such as having to learn in Luganda in the past, but they remain a relatively better off than some linguistic communities in Uganda given their central geographical location. The second limitation comes from having conducted the focus groups in English. While some parents preferred to answer in a different language, the selection criteria included the ability to understand English. By definition, this limits the range of views presented to those of people who had the chance to learn English. Despite these two important limitations, a number of important themes emerged from the data, which could be useful to understand parental preferences regarding language of instruction.

Results: Research Question 1 – Perceptions of the Thematic Curriculum and language policy

Overall, parents seemed to identify the curriculum merely as the language of instruction policy. In expressing their perception of the use of local languages, several positive aspects, as well as important challenges were identified.

The most salient findings from the study is that parents overwhelmingly associated the *Thematic Curriculum* with the language of instruction policy⁶⁷. For example, one participant said: “The thematic curriculum is what? Is it that one about the child about learning in his language?” When they acknowledged having heard about the policy, which was not always the case, parents explained it as a policy to use local languages in schools. None of the parents who described the policy on their own included the pedagogical choices made in introducing themes for the lower primary, or the difference between subject and theme-based instruction. For example, one father described the policy as follows:

Ok, so with thematic curriculum, they told us that the government designed, I don't know how they designed it, but in Uganda here we have communities, or tribes that speak different languages. So the government said that we introduce thematic curriculum so that the children are taught in the real language that is spoken in that area.

When none of the parents recognized the name of the policy, and I described the different elements of the curriculum, the mention of language use immediately triggered a recognition of the policy introduced in 2007. In describing the use of local languages in the first three years of school, there were two main reasons parents appreciated the policy. The most frequently cited positive aspect of the policy was that children could

⁶⁷ As described in the main paper, the Thematic Curriculum included an important change in the pedagogical approach to early years of schooling, as well as a shift in the language policy.

fully understand what teachers explained starting in the first grade. For example, a mother said: *“That’s why she’s saying that it’s very good, that children learn in their language. Because they normally get everything the teacher teaches.”*

The other frequently cited positive aspect parents identified was the importance of knowing the home language in depth, in order to preserve and promote the local culture.

As one father expressed:

Yes, it has some good, advantages. Because it builds culture of a learner. If a learner is, say for example a Muganda, and he picks that foundation of that languages, the culture definitely will be protected for the future.

Parents also identified a number of challenges in implementing the policy. In some cases, these were referred to as reasons the policy should be abandoned, while others focused on the challenges that needed to be overcome. One recurrent theme of the focus group discussion was the notion that migration within Uganda was rendered challenging by the policy. Parents willing to travel and work in different areas would put their children at risk.

According to me, the policy wouldn’t be bad, but the problem is that, that policy works if a child is to study from a particular area from beginning to when he completes primary level. Because an example, we have soldiers here. With so many children, and those soldiers they are their transfers are frequent, so it brings a child in primary here, and he studies in Luganda, then after two, after a year, that soldier is transferred from here, and transferred to Bundibugyo. – Focus Group Participant

Similarly, some parents thought this policy would prevent some teachers from working in different areas, and possibly preventing children from accessing good education, either through having fewer qualified teachers, or because they would not

learn a language other than their own, limiting their access to economic opportunities in the capital city. The following two participants expressed this concern.

If at all I am a teacher, I am a good teacher, I know Luganda only and English, that implies I can't go to the northern part of Uganda, because for them they are supposed to learn in their language, which I don't understand, which I can't speak. That implies that if at all the area doesn't have teachers, the school will not, the people will not be able to study. That one cannot work at all. – Focus Group Participant

That situation, the students who studied in Busoga, or Ankole, will not be able to get a job here, because in central here, they are speaking Luganda, so they have been studying in local language and they cannot be transferred to another area. – Focus Group Participant

Research Question 2 – Parental perspectives and sociolinguistic and psycholinguistic concepts

Sociolinguistic concept

Parents overall confirmed the literature on the power of English and Luganda in the context of nationalism and globalization. While some parents viewed the importance of local languages in school, it was evident that English proficiency was of the utmost importance. One expression that emerged repeatedly to support this assertion was the concept of “dotcom world.” This was used in several focus groups to refer to the importance of technology –implicitly used to refer to as an Anglophone phenomenon – and globalization in general.

Our country is now under transition. We are trying to transform everything, because now the world is under dotcom (...) I believe English should be prioritized. Despite, the local language is good, because people learn how to be discipline, they learn, but in order to prepare someone for the future, you must be knowing English.(...) a friend of mine went to Japan, and was teaching English. So even those bigger countries, they are trying to change, give English a priority, yet they are big countries.

Parents who expressed the importance of English often viewed the policy as a choice of a national language over English, as opposed to its intention, which is to use national languages to support English in later grades.

And I see some countries who use French, like Rwanda, they are now changing to English. Yeah. I think English is international, so it's a must for a child who is growing now to learn and know that language.

Several participants also alluded to the importance of Luganda, though there were few participants who were not from the Central region, making it difficult to differentiate Luganda's power and the importance of a local language. In one focus group with Lusoga speakers, however, the power and importance of Luganda was evident.

Once I cross the river Nile, I am as good as a Muganda. I blend easily, it's the language that people use. So (...) Luganda has become pretty much the informal national language. – Focus Group Participant

This sentiment was echoed by some Baganda expressing the need for all Ugandans to learn Luganda and Swahili:

Actually, being Ugandan, Luganda would be one of the recommended languages, then followed by English, then followed by Kiswahili, at least those three to be a proper Ugandan, he should be in a position to write those three. – Focus Group Participant

Uganda is normally, is dominated actually, the central region, and majority of economic activities in the country are carried out in the central region. So every person from any other corner of the country comes and communicates thoroughly, so Luganda is very easy to learn, and it joins so many people. – Focus Group Participant

Psycholinguist concept

When asked how using local languages could help their children learn, few parents referred to the concept of cross-linguistic transfer. There were, however two notable exceptions. One parent referred to transfer of letter-knowledge and decoding when saying:

Most of the kids who don't speak English, some of them cannot read the Luganda script, but if you know to read the Luganda script, it's very easy to read the English script. – Focus Group Participant

Another parent described the learning process in child development according to the idea of threshold described by Cummins (1979), stressing the importance of developing the first language enough to subsequently promote second-language development:

That is why we are begin[ing] from the grassroots, to teach from nursery to P3, that curriculum of local languages. One, you first bring that kid [on]board, because his brain is growing. If it is growing, you plant him or her in that local language, local languages he can understand properly, so it is good. – Focus Group Participant

Appendix C – Additional Figures and Tables for Paper 2

Table A1 – Standardized and Unstandardized Coefficients for CFA

Observed Variable	β	Central N=1709			South-West N=1852		
		B	SE		β	B	SE
L2 Letter knowledge	0.68	1.00			0.46	1.00	
L2 Pseudoword	0.82	0.84	0.07		0.92	2.62	0.32
L2 Letter knowledge	0.66	0.99	0.05		0.31	0.73	0.06
L2 Pseudoword	0.94	0.98	0.09		0.83	1.82	0.17

Note. CFA = Confirmatory factor analysis; SE are standard errors from the unstandardized coefficients

Table A2 – Standardized and Unstandardized Coefficients for CFA in restricted sample

Observed Variable	β	Central N=549			South-West N=822		
		B	SE		β	B	SE
L2 Letter knowledge	0.56	1.00			0.32	1.00	
L2 Pseudoword	0.64	0.73	0.09		0.84	3.24	0.79
L2 Letter knowledge	0.56	1.08	0.09		0.19	0.67	0.13
L2 Pseudoword	0.90	1.21	0.25		0.74	2.54	0.44

Note. CFA = Confirmatory factor analysis; SE are standard errors from the unstandardized coefficients

Table A3 – Results from Structural Equation Model in Central Region (N=1709)

Observed Variable	β		B	
	L1 RC	L2 RC	L1 RC	L2 RC
Direct				
Alphabetical Knowledge	0.63 (0.03)	0.44 (0.05)	0.07 (0.01)	0.05 (0.01)
L1 Oral	0.08 (0.03)		0.09 (0.04)	
L2 Oral		0.24 (0.03)		0.07 (0.01)
L1 Reading Comprehension		0.19 (0.06)		0.20 (0.06)
Indirect				
Alphabetical Knowledge		0.12 (0.04)		.014 (.005)
Total				
Alphabetical Knowledge	0.63 (0.03)	0.56 (0.03)	0.07 (0.01)	0.06 (0.01)
R^2	0.42	0.52		

Note. RC = Reading Comprehension;

Table A4 – Results from Structural Equation Model in Central Region (N=549)

Observed Variable	β		B	
	L1 RC	L2 RC	L1 RC	L2 RC
Direct				
Alphabetical Knowledge	0.44 (0.04)	0.36 (0.04)	0.06 (0.01)	0.06 (0.01)
L1 Oral	0.13 (0.04)		0.21 (0.06)	
L2 Oral		0.35 (0.05)		0.15 (0.03)
L1 Reading Comprehension		0.18 (0.05)		0.19 (0.06)
Indirect				
Alphabetical Knowledge		0.08 (0.03)		.011 (.004)
Total				
Alphabetical Knowledge	0.44 (0.04)	0.44 (0.03)	0.06 (0.01)	0.07 (0.01)
R^2	0.22	0.46		

Note. RC = Reading Comprehension;

Table A5 – Results from Structural Equation Model in South-West Region (N=1852)

Observed Variable	β		B	
	L1 RC	L2 RC	L1 RC	L2 RC
Direct				
Alphabetical Knowledge	0.65 (0.02)	0.13 (0.04)	0.09 (.004)	0.01 (.003)
L1 Oral	0.08 (0.02)		0.07 (0.02)	
L2 Oral		0.29 (0.05)		0.04 (0.01)
L1 Reading Comprehension		0.17 (0.03)		0.09 (0.02)
Indirect				
Alphabetical Knowledge		0.11 (0.02)		.008 (.002)
Total				
Alphabetical Knowledge	0.65 (0.02)	0.24 (0.05)	0.09 (.004)	0.02 (.004)
R^2	0.44	0.21		

Note. RC = Reading Comprehension;

Table A6 – Results from Structural Equation Model in South-West Region (N=822)

Observed Variable	β		B	
	L1 RC	L2 RC	L1 RC	L2 RC
Direct				
Alphabetical Knowledge	0.41 (0.03)	0.15 (0.04)	0.06 (0.01)	0.02 (.004)
L1 Oral	0.12 (0.03)		0.13 (0.03)	
L2 Oral		0.37 (0.04)		0.08 (0.01)
L1 Reading Comprehension		0.12 (0.03)		0.08 (0.03)
Indirect				
Alphabetical Knowledge		0.05 (0.02)		.005 (.002)
Total				
Alphabetical Knowledge	0.41 (0.03)	0.20 (0.05)	0.06 (0.01)	0.02 (0.01)
R^2	0.19	0.22		

Note. RC = Reading Comprehension;

Table A7 – Standardized Results from Multi-Group Analysis in Central Region (Luganda N=1454 and English N=255)

Observed Variable	β - Luganda		β - English	
	L1 RC	L2 RC	L1 RC	L2 RC
Direct				
Alphabetical Knowledge	0.67 (0.02)	0.38 (0.05)	0.47 (0.08)	0.56 (0.07)
L1 Oral	0.05 (0.03)		0.22 (0.10)	
L2 Oral		0.18 (0.02)		0.30 (0.08)
L1 Reading Comprehension		0.28 (0.06)		0.10 (0.06)
Indirect				
Alphabetical Knowledge		0.19 (0.04)		0.05 (0.03)
Total				
Alphabetical Knowledge	0.67 (0.02)	0.57 (0.03)	0.47 (0.08)	0.61 (0.06)
R^2	0.46	0.50	0.29	0.67

Note. RC = Reading Comprehension;

Figure A1 – Kernel Density Curves of Luganda Measures in the Central Region with normal curve overlay.

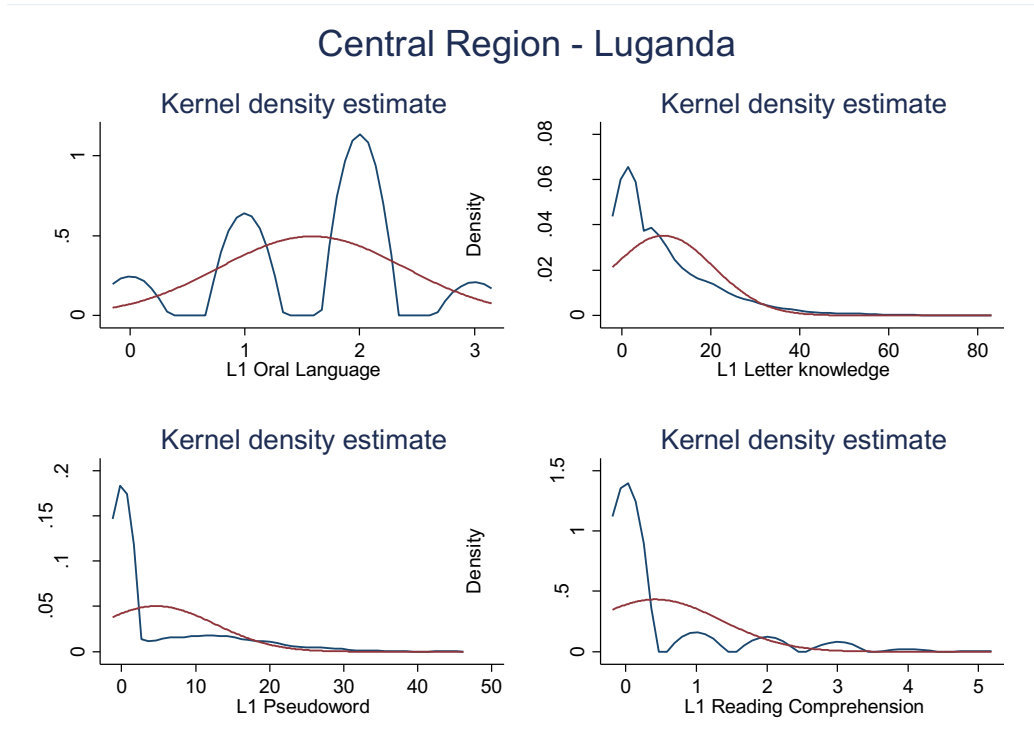


Figure A2 – Kernel Density Curves of English Measures in the Central Region with normal curve overlay.

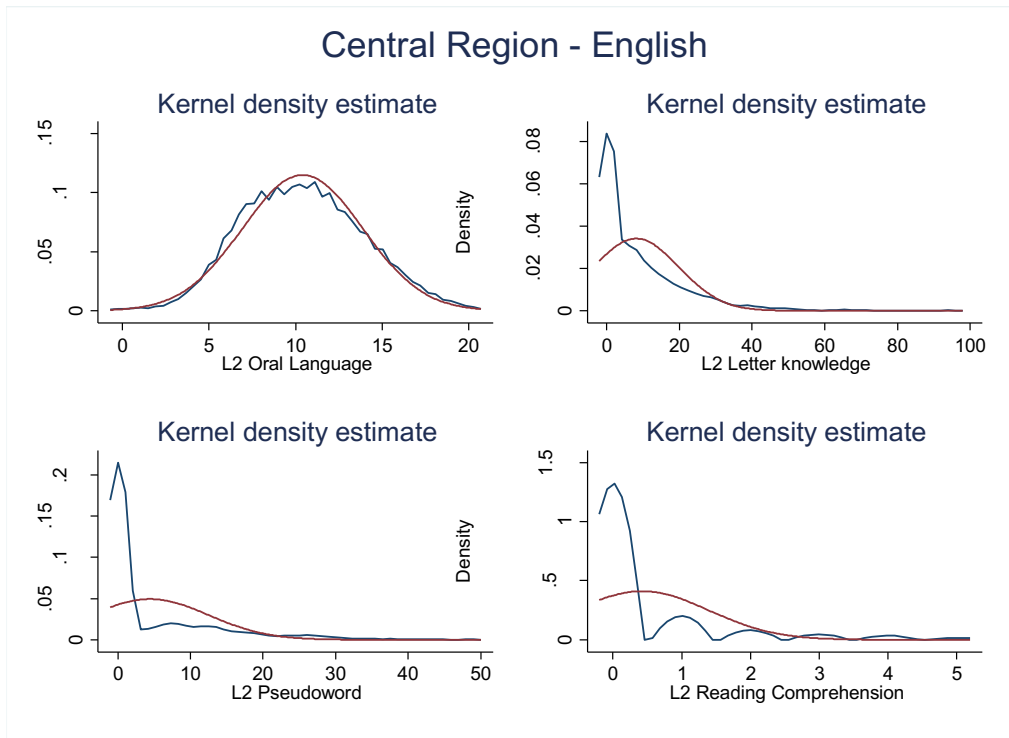


Figure A3 – Kernel Density Curves of Runyankole-Rukiga Measures in the South-West Region with normal curve overlay.

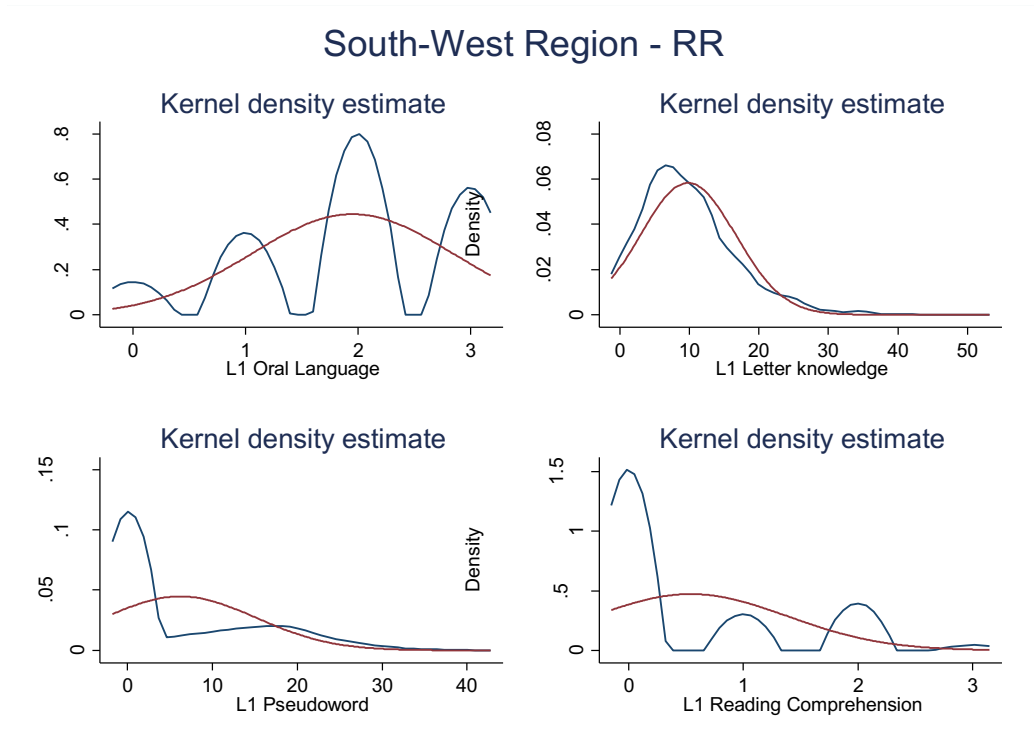
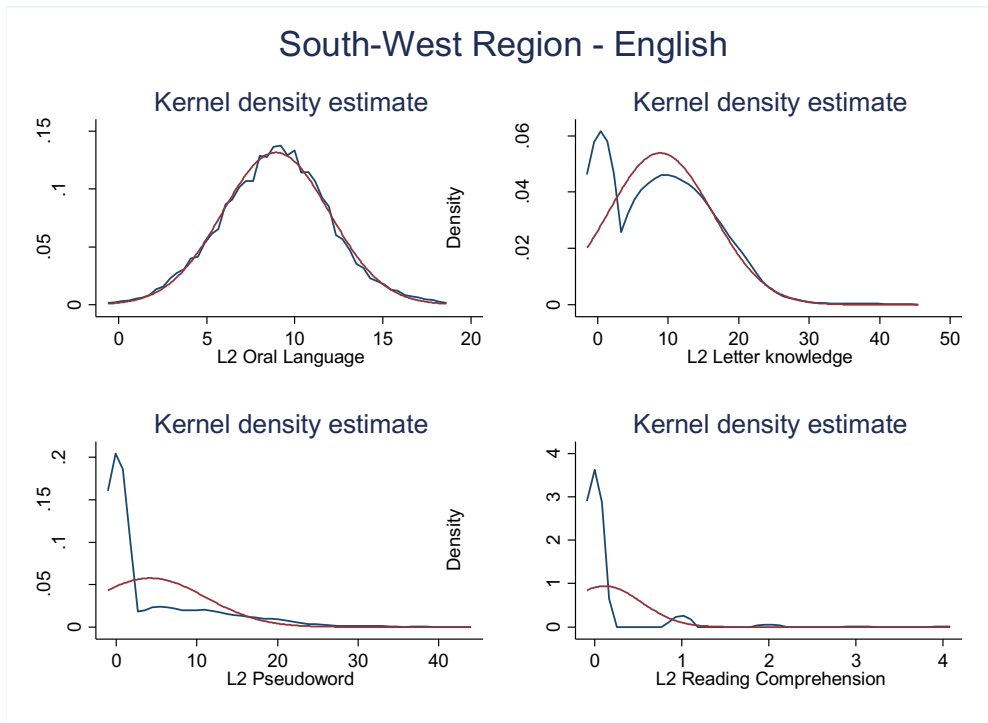


Figure A4 – Kernel Density Curves of English Measures in the South-West Region with normal curve overlay.



Appendix D - Methodological Approach to Address Clustering of Students

for Paper 3

The data in the current study were collected from students randomly sampled from grade 1 and 2 classrooms in schools identified to take part in the study. These student observations are therefore nested into classrooms, and schools, and thus are not independent. Analyzing these data using classical regression ignoring this clustering violates the Ordinary Least Square assumption that observations are independent and identically distributed (i.i.d) and could result in misleading parameter estimates (Gelman & Hill, 2006), as well as smaller standard errors (Cameron & Miller, 2015). This appendix briefly describes two common options to address this (clustering standard errors and multilevel modeling), presents analysis of the average treatment effect (ATE) of the intervention using both approaches, and argues for the simpler approach in this study.

Is it necessary to take clustering into account in this study?

The extent to which clustering of observations needs to be taken into account is largely dependent on the intraclass correlation (ICC), which is the proportion of variance in the outcome attributable to group membership. At one extreme, if the ICC is 0, then none of the variance in student reading is due to school membership, and students can be assumed to have been randomly sampled from the larger population. At the other extreme, an ICC of 1 would mean that all students within a school (or classroom) would have the same reading level, and variation in the outcome would thus be entirely dependent on school membership. In a two-level model, where students are nested either into classrooms or schools, the intraclass correlation is estimated using the following formula:

$$ICC = \frac{\sigma_{between}^2}{\sigma_{between}^2 + \sigma_{within}^2} \quad (1)$$

where $\sigma_{between}^2$ is the variance between groups (either classroom or school) and σ_{within}^2 is the variance within groups, and the sum of the two is the total variance. In the case of this study, I also consider the students to be nested into classrooms and into schools in a three-level model. As a result, I decompose the variance using the following formulas:

$$ICC_{school} = \frac{\sigma_{between-school}^2}{\sigma_{between-school}^2 + \sigma_{between-classroom}^2 + \sigma_{within}^2} \quad (2)$$

$$ICC_{classroom} = \frac{\sigma_{between-school}^2 + \sigma_{between-classroom}^2}{\sigma_{between-school}^2 + \sigma_{between-classroom}^2 + \sigma_{within}^2} \quad (3)$$

Table 1.1 below presents the intra-class correlation of six models across the four outcomes of interest in this study. The top panel presents the ICC with the outcome regressed unconditionally, including first a school-level clustering, then a classroom-level clustering, and finally with a three-level model. In the bottom panel, the outcome is regressed on the student’s grade level, with the same three clustering approaches. Controlling for grade is important given that the study was designed to sample one first-grade and one second-grade classroom in each school, which means classroom assignment should explain a large proportion of variance. In models clustering only at the school level, the ICC ranges from 0.12 to 0.25, suggesting that at least 12 percent of the variation in student outcome is explained by between-school variance. At the classroom level, the ICC ranges from 0.21 to 0.33 and 0.35 to 0.45 with and without controlling for grade, respectively. In the three-level models controlling for grade, school-level variance explains between 12 and 19 percent of the variation in student outcome, and classroom assignment explains between 21 and 33 percent of the same variation.

Table 1.1. Intraclass Correlations (ICC) at Baseline

Outcome	English Word	English Comp	Kiswahili Word	Kiswahili Comp
Unconditional				
School alone	0.19	0.17	0.13	0.12
Classroom alone	0.45	0.41	0.39	0.35
School	0	0	0	0
Classroom	0.45	0.41	0.39	0.35
Control for Grade				
School alone	0.25	0.21	0.18	0.16
Classroom alone	0.33	0.32	0.24	0.21
School	0.19	0.12	0.13	0.12
Classroom	0.33	0.32	0.24	0.21

The magnitude of the ICC in baseline data suggests that clustering should be taken into account in this study at the classroom level. Huang (2016) suggests that ignoring clustering even with moderate ICC (0.05) could lead to underestimated standard errors.

Two common approaches to take into account this correlation are (i) using robust clustered standard errors and (ii) multilevel modeling (MLM). As a result of the estimates above, in comparing the two approaches, I first cluster standard errors at the classroom level, and then use a three-level MLM.

First Approach: using clustered standard errors

In order to estimate the average treatment effect of the intervention, I estimate the following model:

$$Y_{i,c,s} = \beta_0 + \beta_1 T_s + B_i X_i + B_c X_c + B_s X_s + \varepsilon_{i,c} \quad (4)$$

In this model, the outcome variable measured at student i is regressed on a school-level Treatment indicator T , as well as a set of student, classroom and school covariates, with the standard errors ε clustered at the classroom level⁶⁸.

Table 1.2 (same as table 6 in the paper) below presents the results of this model across the four reading outcomes of interest: English Word reading, English Comprehension, Kiswahili Word reading and Kiswahili Comprehension. In each case, the models included are a simple regression with no covariates in the first column, the full model excluding KCPE baseline in the second column, the full model with no imputation in the third column, the full model with mean imputation in the fourth column, and the full model with multiple imputation in the fifth model⁶⁹.

Second Approach: using a three-level model

In order to estimate the average treatment effect of the intervention, I estimate the following model:

$$Y_{i,c,s} = \beta_{0cs} + B_{ics} X_i + \varepsilon_{ics} \quad (5)$$

$$\beta_{0cs} = \gamma_{00s} + B_{0cs} X_c + u_{0cs} \quad (6)$$

$$\gamma_{00s} = \zeta_{000} + \beta_{00s} T_s + B_{00s} X_{00s} + V_{000} \quad (7)$$

where the random components are distributed as follows:

$$\varepsilon_{ics} \sim N(0, \sigma^2)$$

⁶⁸ Student-level covariates include grade and SES; classroom-level covariates include teacher training status; school-level covariates include school type (public or private), school average KCPE score, and a composite of school resources; district-level fixed effects are also included.

⁶⁹ As noted in the body of the paper, one important baseline covariate (KCPE mean score) has a large number of missing values at baseline. As a result I use two approaches to impute the baseline score, and these are discussed briefly in the main body of the paper, and in more detail in Appendix 2.

$$u_{0cs} \sim N(0, \tau_2^2)$$

$$V_{000} \sim N(0, \tau_3^2)$$

Similar to the first approach, the student level outcome Y is regressed on a vector of student-level covariates. In contrast to the model above, the classroom-level mean has a random intercept β_{0cs} , itself regressed on a set of classroom-level covariates. Finally, the school-level mean γ_{00s} is regressed on both the Treatment indicator T and a set of school-level covariates.

Table 1.3 below presents the same set of estimates as table X.2 using the three-level model presented in this section.

Comparing the two models

As the main goal of this appendix is to justify choosing one approach, as opposed to interpreting the findings, I restrict this discussion to similarities and differences across estimates. The main takeaway of the comparison between the two sets of estimates is that they are remarkably similar. One notable difference is that standard errors tend to be slightly smaller in the first approach, resulting in the treatment parameter estimate for Kiswahili comprehension using the full model without imputation being statistically significant using the first approach and not the second, despite having the same point estimate.

Overall, the point estimates on the Treatment indicator are substantively identical. More importantly, the trends on this indicator are stable across models including different sets of covariates (not shown in tables below) and to the different approaches to imputation. The patterns are also similar when fitting models that include moderation to answer substantive questions of interest. Using MLM does allow to account for clustering at both the classroom and school at the same time, whereas robust clustered standard errors cannot. However, this does not result in a large magnitude difference in the standard errors. As a result, I see no statistical justification to choosing one model over the other. I therefore turn to non-statistical arguments to choose the first approach over MLM.

While the choice between the two approaches are often based on discipline-specific practices, there are substantive reasons to choose MLM over robust standard errors. One advantage using an MLM approach is the ability to examine the relationship between individual-level variables and the outcome, and whether this relationship differs across schools. In the case of an impact evaluation, this would be especially interesting if the treatment assignment were done at the individual level. In the case of this study, where treatment assignment is done at the school level, and only a small part of the theory of action is through individual students (provision of textbooks to students), there is no

major substantive advantage of using MLM. The research questions in this study do not require MLM.

The two main advantages of using the first approach in this study are the simplicity of the model, and the reduced computational demand. In terms of communicating the model to a broader audience, it may be easier to use a traditional model than an MLM. In addition, an MLM is computationally more demanding, especially when taking multiple imputation into account. As a result, I will use the first approach, using clustered standard errors in a multiple regression OLS model.

Table 1.2. Average Treatment Effect - Econometric Approach

	English Word					English Comprehension					Kiswahili Word				Kiswahili Comprehension					
Treatment	9.78*** (1.88)	7.91*** (1.63)	5.53** (2.08)	7.78*** (1.65)	7.93*** (1.67)	0.69*** (0.11)	0.60*** (0.11)	0.33* (0.13)	0.59*** (0.11)	0.59*** (0.11)	6.90*** (1.27)	5.91*** (1.17)	2.83 (1.51)	5.64*** (1.17)	5.74*** (1.19)	0.54*** (0.09)	0.52*** (0.09)	0.30* (0.12)	0.50*** (0.09)	0.52*** (0.09)
Student SES		2.74*** (0.42)	1.99*** (0.50)	2.64*** (0.42)	2.67*** (0.42)		0.25*** (0.03)	0.13*** (0.03)	0.24*** (0.03)	0.25*** (0.03)		1.59*** (0.33)	0.90* (0.42)	1.47*** (0.33)	1.54*** (0.33)		0.09*** (0.02)	0.03 (0.03)	0.08** (0.02)	0.09** (0.02)
Teacher is Trained		1.37 (1.16)	2.25 (1.46)	1.31 (1.16)	1.16 (1.18)		0.16 (0.09)	0.15 (0.11)	0.16 (0.09)	0.16 (0.09)		0.11 (0.89)	0.86 (1.21)	0.02 (0.89)	0.04 (0.91)		-0.04 (0.07)	-0.00 (0.09)	-0.05 (0.07)	-0.05 (0.07)
Baseline KCPE Score			2.26*** (0.55)	2.00** (0.66)	1.28* (0.52)			0.24*** (0.04)	0.20*** (0.04)	0.11** (0.04)			2.35*** (0.44)	2.12*** (0.49)	1.35** (0.39)			0.17*** (0.03)	0.15*** (0.03)	0.09** (0.03)
School is nonformal		14.14*** (1.87)	11.10*** (2.50)	13.50*** (1.89)	13.40*** (1.90)		0.82*** (0.16)	0.51* (0.21)	0.76*** (0.16)	0.77*** (0.16)		10.48*** (1.36)	7.42*** (1.71)	9.85*** (1.32)	9.80*** (1.35)		0.71*** (0.12)	0.38** (0.14)	0.66*** (0.11)	0.66*** (0.11)
Constant	29.66*** (1.62)	19.38*** (2.49)	23.82*** (3.17)	19.44*** (2.68)	20.05*** (2.54)	0.82*** (0.09)	0.22 (0.20)	0.67** (0.24)	0.21 (0.21)	0.28 (0.20)	21.33*** (1.05)	13.28*** (1.88)	19.37*** (2.54)	14.15*** (2.02)	14.02*** (1.90)	1.23*** (0.08)	0.74*** (0.15)	1.29*** (0.20)	0.79*** (0.16)	0.79*** (0.15)
N Students	4202	4202	2316	4202	4162	4202	4202	2316	4202	4162	4202	4202	2316	4202	4162	4202	4202	2316	4202	4162
Controls	N	Y	Y	Y	Y	N	Y	Y	Y	Y	N	Y	Y	Y	Y	N	Y	Y	Y	Y
KCPE Mean Imputation	N	N	N	Y	N	N	N	N	Y	N	N	N	N	Y	N	N	N	N	Y	N
KCPE Multiple Imputation	N	N	N	N	Y	N	N	N	N	Y	N	N	N	N	Y	N	N	N	N	Y

Note. Controls also include student grade, school-level resources, and district fixed-effects.

* p<0.05 ** p<0.01 ***p<0.001

Table 1.3. Average Treatment Effect MLM approach																					
Treatment	English Word					English Comprehension					Kiswahili Word					Kiswahili Comprehension					
		9.54*** (2.23)	7.80*** (1.90)	5.54* (2.42)	7.68*** (1.89)	7.91*** (1.92)	0.68*** (0.15)	0.60*** (0.14)	0.32* (0.16)	0.59*** (0.14)	0.59*** (0.14)	6.78*** (1.51)	5.89*** (1.49)	2.85 (2.01)	5.62*** (1.46)	5.82*** (1.49)	0.54*** (0.11)	0.51*** (0.11)	0.30 (0.16)	0.50*** (0.11)	0.52*** (0.12)
Student SES		1.19*** (0.35)	1.24** (0.46)	1.17*** (0.35)	1.20** (0.35)		0.14*** (0.02)	0.10** (0.03)	0.14*** (0.02)	0.15*** (0.02)		0.74* (0.29)	0.48 (0.39)	0.72* (0.29)	0.75* (0.30)		0.04 (0.02)	0.00 (0.03)	0.04 (0.02)	0.04 (0.02)	
Teacher is Trained		1.15 (1.16)	1.49 (1.48)	1.12 (1.15)	1.05 (1.17)		0.16 (0.09)	0.17 (0.11)	0.15 (0.09)	0.16 (0.09)		0.00 (0.93)	0.30 (1.26)	-0.05 (0.92)	-0.01 (0.94)		-0.01 (0.07)	0.00 (0.10)	-0.01 (0.07)	-0.01 (0.07)	
Baseline KCPE Score			2.33** (0.80)	2.10* (0.93)	0.49 (0.62)		0.24*** (0.05)	0.21** (0.07)	0.03 (0.04)			2.38*** (0.66)	2.16** (0.72)	0.58 (0.48)			0.17*** (0.05)	0.15** (0.06)	0.04 (0.04)		
School is nonformal		14.27*** (2.40)	10.78*** (3.03)	13.59*** (2.38)	13.95*** (2.42)		0.84*** (0.18)	0.52** (0.20)	0.77*** (0.17)	0.82*** (0.18)		10.58*** (1.88)	7.19** (2.52)	9.94*** (1.85)	10.29*** (1.88)		0.73*** (0.14)	0.38 (0.20)	0.68*** (0.14)	0.71*** (0.14)	
Constant		29.86*** (1.93)	19.93*** (3.29)	24.71*** (4.05)	19.92*** (3.38)	20.11*** (3.31)	0.82*** (0.13)	0.24 (0.24)	0.68* (0.27)	0.23 (0.25)	0.26 (0.24)	21.46*** (1.31)	13.52*** (2.58)	19.96*** (3.37)	14.32*** (2.63)	13.84*** (2.58)	1.24*** (0.10)	0.73*** (0.20)	1.29*** (0.26)	0.78*** (0.20)	0.74*** (0.20)
N Students	4202	4202	2316	4202	4162	4202	4202	2316	4202	4162	4202	4202	2316	4202	4162	4202	4202	2316	4202	4162	
Controls	N	Y	Y	Y	Y	N	Y	Y	Y	Y	N	Y	Y	Y	Y	N	Y	Y	Y	Y	
KCPE Mean Imputation	N	N	N	Y	N	N	N	N	Y	N	N	N	N	Y	N	N	N	N	Y	N	
KCPE Multiple Imputation	N	N	N	N	Y	N	N	N	N	Y	N	N	N	N	Y	N	N	N	N	Y	

Note. Controls also include student grade, school-level resources, and district fixed-effects.

* p<0.05 ** p<0.01 ***p<0.001

Appendix E - Imputation of missing data for Paper 3

As described in the main text of the study, the two samples of schools selected at baseline and endline were random selections of schools from the treatment and control population, resulting in a number of schools at endline missing baseline data. Of the 262 schools sampled at endline, 114 (72%) treatment and 47 (89%) control schools were also sampled at baseline respectively. In this appendix, I start with a description of the missing data in the study, I then explain the rationale for using multiple imputation (MI), discuss some of the assumptions related to MI, and finally the decisions made in implementing MI in this study.

Missing data in this study

Two school baseline variables of interest are missing in a number of schools important enough to warrant estimating models with and without imputed data. The first variable is the school KCPE average score, measuring the performance of students on end of primary school leaving exam the year prior to the intervention. This variable is included in models for two reasons: first, it represents an important control variable as a measure of student performance; second, the variable is not balanced at baseline. The second missing variable of interest in the baseline scores on the EGRA instrument. This variable is a theorized moderator to the impact of the intervention. Table 2.1 includes descriptive statistics for the two variables at baseline, and the prevalence of missing values at endline.

Approaches to missing data

Missing data can be addressed in many ways, including dropping entire observations of the analysis (listwise deletion), and assigning values to the missing variables (imputation). Listwise deletion is appropriate when data is MCAR, and when the sample is large enough that the loss of data will not affect statistical power. Single imputation is assigns one value to each missing variable. One traditional approach has been mean-imputation, which assigns missing values the average of the sample. A similar approach is regression imputation where covariates are used to predict the missing value using OLS regression and assign \hat{Y} as the missing value. However, single imputation will bias the standard errors by reducing the amount of uncertainty in the data, thus also biasing the statistical tests associated (Johnson & Young, 2011). Multiple Imputation (MI) addresses this by creating multiple copies of the datasets with different values imputed to the missing variables in each dataset⁷⁰.

⁷⁰ Another “modern” approach to missing data that addresses the problems with single imputation is full-information maximum likelihood (FIML). The main difference between MI and FIML is that FIML does not assign values to missing data and deals with missing data and parameter estimates in one step. These approaches are theoretically similar and tend to produce similar results when the number of imputed datasets is large, therefore there seems to be no major difference (Johnson & Young, 2011).

The process for multiple imputation consists of three steps: (1) imputation, (2) estimation, and (3) pooling. In the first step, M datasets are created using a chosen imputation model. The different estimations are then conducted separately using each dataset. Finally, the estimates are combined into a single multiple-imputation estimate (StataCorp, 2013).

Table 2.1 Descriptive Statistics of Missing Data

	Mean	SD	N	% Missing
KCPE Baseline				
Treatment	0.02	0.89	84	47.17%
Control	-0.32	0.93	33	37.74%
English Word				
Treatment	13.80	8.65	114	28.30%
Control	11.50	9.06	47	11.32%
English Comprehension				
Treatment	0.54	0.48	114	28.30%
Control	0.50	0.56	47	11.32%
Kiswahili Word				
Treatment	10.46	5.94	114	28.30%
Control	9.52	6.23	47	11.32%
Kiswahili Comprehension				
Treatment	0.78	0.46	114	28.30%
Control	0.73	0.52	47	11.32%

Assumptions related to MI

Two important assumptions to consider when imputing data is the reason for missing data and the proportion of missing data. In this study, the missing baseline information for the EGRA measures are undeniably missing completely at random (MCAR), meaning that the reason for missing and the value of the missing data are independent from both observed and unobserved variables. The reason for missing data for KCPE scores, in contrast, is due both to random sampling and for traditional reasons related to inability to collect data. The schools for which no data were collected at baseline are MCAR, but those for which the schools were unable or unwilling to provide the KCPE scores cannot be considered MCAR. For the purpose of this study, I consider these missing at random (MAR), which assumes that the missingness is unrelated to the value of the missing variable, but maybe related to other characteristics of the school. As a result, I present

estimates of the model in three ways: without controlling for KCPE, including KCPE with a smaller sample, and with KCPE scores imputed.

As presented in table 2.1 above, the level of missingness in this study ranges from 11% for baseline EGRA scores in the control group to 47% for KCPE scores in the treatment group. While these proportions are higher than what the typical 5-10% range observed in large surveys, Johnson and Young (2011) note that MI performs well even when missing data approaches 50%, especially with a larger M .

Imputation model

One important decision in MI is the selection of variables for the imputation model. In order to avoid bias in the analysis model, White, Royston, and Wood (2011) recommend to include all covariates from all analyses models, as well as other predictors of the missing variable. These authors also recommend imputing baseline data only using other baseline variables to respect the independence of baseline variables from randomization (White et al., 2011). Based on these recommendations and given that the missing values in this case are at the school level baseline variables, I use the following school-level covariates in the imputation model: six measures of school resources⁷¹, school status, head teacher characteristics⁷², and geographical district.

Another decision related to MI is the number of imputations to perform to obtain stable and unbiased estimates. The two main factors to consider when deciding on M is the size of the sample and the proportion of missing data. Larger sample and lower proportions of missing data require fewer imputations. While many MI estimations use 5 imputations, Johnson and Young (2011) recommend using an M of 25 with larger proportions of missing data, a recommendation I follow given the relatively large proportion of missing data in my sample.

In order to implement MI in this study I use the built-in `mi` command in Stata 13 (StataCorp, 2013).

Table 2.2 below presents the results of the average treatment impact using the following model:

$$Y_{i,c,s} = \beta_0 + \beta_1 T_s + B_i X_i + B_c X_c + B_s X_s + \varepsilon_{i,c} \quad (1)$$

⁷¹ Presence of safe water, electricity, feeding program, girls' washroom, computer room, and library.

⁷² Sex and whether the head teacher reported having received school management training.

In this model, the outcome variable measured at student i is regressed on a school-level Treatment indicator T , as well as a set of student, classroom and school covariates, with the standard errors ε clustered at the classroom level⁷³.

In the first column, the model only includes the Treatment indicator. The second column presents estimation of the final model excluding the school mean KCPE score. The third column includes the full model without imputation, the fourth column presents the full model with mean imputation, and the fifth column presents results using multiple imputation described above.

⁷³ Student-level covariates include grade and SES; classroom-level covariates include teacher training status; school-level covariates include school type (public or private), school average KCPE score and a composite of school resources; district-level fixed effects are also included.

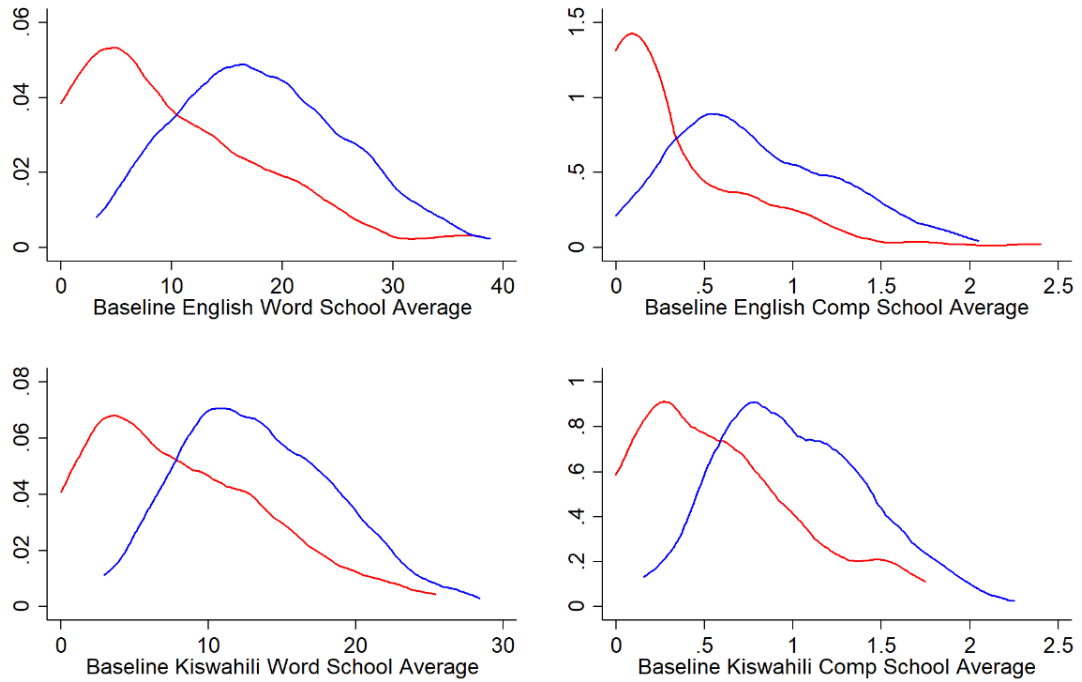
Table 1.2. Average Treatment Effect - Econometric Approach

	English Word					English Comprehension					Kiswahili Word					Kiswahili Comprehension				
	Treatment	9.78*** (1.88)	7.91*** (1.63)	5.53** (2.08)	7.78*** (1.65)	7.93*** (1.67)	0.69*** (0.11)	0.60*** (0.11)	0.33* (0.13)	0.59*** (0.11)	0.59*** (0.11)	6.90*** (1.27)	5.91*** (1.17)	2.83 (1.51)	5.64*** (1.17)	5.74*** (1.19)	0.54*** (0.09)	0.52*** (0.09)	0.30* (0.12)	0.50*** (0.09)
Student SES		2.74*** (0.42)	1.99*** (0.50)	2.64*** (0.42)	2.67*** (0.42)		0.25*** (0.03)	0.13*** (0.03)	0.24*** (0.03)	0.25*** (0.03)		1.59*** (0.33)	0.90* (0.42)	1.47*** (0.33)	1.54*** (0.33)		0.09*** (0.02)	0.03 (0.03)	0.08** (0.02)	0.09** (0.02)
Teacher is Trained		1.37 (1.16)	2.25 (1.46)	1.31 (1.16)	1.16 (1.18)		0.16 (0.09)	0.15 (0.11)	0.16 (0.09)	0.16 (0.09)		0.11 (0.89)	0.86 (1.21)	0.02 (0.89)	0.04 (0.91)		-0.04 (0.07)	-0.00 (0.09)	-0.05 (0.07)	-0.05 (0.07)
Baseline KCPE Score			2.26*** (0.55)	2.00** (0.66)	1.28* (0.52)			0.24*** (0.04)	0.20*** (0.04)	0.11** (0.04)			2.35*** (0.44)	2.12*** (0.49)	1.35** (0.39)			0.17*** (0.03)	0.15*** (0.03)	0.09** (0.03)
School is nonformal		14.14*** (1.87)	11.10*** (2.50)	13.50*** (1.89)	13.40*** (1.90)		0.82*** (0.16)	0.51* (0.21)	0.76*** (0.16)	0.77*** (0.16)		10.48*** (1.36)	7.42*** (1.71)	9.85*** (1.32)	9.80*** (1.35)		0.71*** (0.12)	0.38** (0.14)	0.66*** (0.11)	0.66*** (0.11)
Constant	29.66*** (1.62)	19.38*** (2.49)	23.82*** (3.17)	19.44*** (2.68)	20.05*** (2.54)	0.82*** (0.09)	0.22 (0.20)	0.67** (0.24)	0.21 (0.21)	0.28 (0.20)	21.33*** (1.05)	13.28*** (1.88)	19.37*** (2.54)	14.15*** (2.02)	14.02*** (1.90)	1.23*** (0.08)	0.74*** (0.15)	1.29*** (0.20)	0.79*** (0.16)	0.79*** (0.15)
N Students	4202	4202	2316	4202	4162	4202	4202	2316	4202	4162	4202	4202	2316	4202	4162	4202	4202	2316	4202	4162
Controls	N	Y	Y	Y	Y	N	Y	Y	Y	Y	N	Y	Y	Y	Y	N	Y	Y	Y	Y
KCPE Mean Imputation	N	N	N	Y	N	N	N	N	Y	N	N	N	N	Y	N	N	N	N	Y	N
KCPE Multiple Imputation	N	N	N	N	Y	N	N	N	N	Y	N	N	N	N	Y	N	N	N	N	Y

Note. Controls also include student grade, school-level resources, and district fixed-effects.

* p<0.05 ** p<0.01 ***p<0.001

Appendix F - Additional Tables and Figures for Paper 3



Note: Kernel Density
Legend: Public Schools in Red and Nonformal Schools in Blue

Figure A1: Kernel density of baseline school average English and Kiswahili scores by school type

Table A1. Baseline Student Characteristics: description and balance across school type

	Public School			Nonformal			Double Difference
	Control	Treatment	Control-Treatment	Control	Treatment	Control-Treatment	
Student is female	0.498 (0.002)	0.498 (0.004)	0.000	0.497 (0.005)	0.500 (0.005)	-0.003	-0.003
Student age	7.114 (0.088)	7.347 (0.069)	-0.233	6.849 (0.066)	6.835 (0.048)	0.014	0.247
Student is in grade 2	1.500 (0.000)	1.500 (0.000)	0.000	1.506 (0.008)	1.498 (0.001)	0.008	0.008
Student has a radio at home	0.790 (0.027)	0.776 (0.015)	0.014	0.751 (0.016)	0.764 (0.011)	-0.013	-0.027
Student has a phone or mobile at home	0.919 (0.016)	0.911 (0.013)	0.008	0.927 (0.020)	0.930 (0.012)	-0.003	-0.011
Student has electricity at home	0.402 (0.066)	0.550 (0.043)	-0.147	0.863 (0.022)	0.862 (0.015)	0.001	0.148
Student has a television at home	0.477 (0.046)	0.562 (0.037)	-0.085	0.780 (0.027)	0.814 (0.015)	-0.034	0.051
Student's mother can read and write	0.909 (0.020)	0.901 (0.011)	0.008	0.935 (0.010)	0.932 (0.008)	0.003	-0.005
Number of words correct - English	5.686 (0.910)	11.266 (1.118)	-5.580	18.319 (1.452)	17.926 (0.741)	0.393	5.973
Comprehension questions correct - English	0.175 (0.051)	0.399 (0.059)	-0.224	0.821 (0.089)	0.781 (0.047)	0.041	0.265
Number of words correct - Kiswahili	5.556 (0.744)	8.997 (0.788)	-3.441	14.171 (0.962)	12.853 (0.528)	1.318	4.759
Comprehension questions correct - Kiswahili	0.389 (0.059)	0.642 (0.058)	-0.253	1.080 (0.077)	0.975 (0.042)	0.105	0.358
Maximum Number of Observations	500	1400		690	1795		4385
Date				January 2012			

Note. This table presents the coefficient of the OLS regression of the treatment indicator (equal 1 if treated; 0 otherwise) against each of the listed baseline characteristics. Treatment includes schools receiving the intervention at least one year.

Table A2. Baseline Student Characteristics: description and balance for students in schools above and below the English word scholl average median

	School Below Median			School Above Median			Double Difference
	Control	Treatment	Control-Treatment	Control	Treatment	Control-Treatment	
Student is female	0.500 (0.005)	0.496 (0.006)	0.004	0.494 (0.004)	0.502 (0.003)	-0.008	-0.012
Student age	7.003 (0.073)	7.244 (0.069)	-0.241	6.900 (0.087)	6.897 (0.053)	0.003	0.244
Student is in grade 2	1.499 (0.001)	1.499 (0.001)	-0.000	1.510 (0.010)	1.499 (0.001)	0.011	0.011
Student has a radio at home	0.772 (0.021)	0.763 (0.015)	0.010	0.760 (0.019)	0.775 (0.011)	-0.015	-0.025
Student has a phone or mobile at home	0.928 (0.012)	0.885 (0.015)	0.043	0.918 (0.027)	0.954 (0.009)	-0.036	-0.079
Student has electricity at home	0.522 (0.058)	0.576 (0.041)	-0.055	0.881 (0.026)	0.855 (0.018)	0.025	0.08
Student has a television at home	0.531 (0.037)	0.545 (0.033)	-0.014	0.828 (0.032)	0.843 (0.015)	-0.015	-0.001
Student's mother can read and write	0.918 (0.015)	0.902 (0.011)	0.015	0.934 (0.014)	0.933 (0.007)	0.001	-0.014
Number of words correct - English	6.156 (0.652)	7.481 (0.505)	-1.325	22.814 (1.164)	21.660 (0.622)	1.154	2.479
Comprehension questions correct - English	0.185 (0.030)	0.237 (0.029)	-0.052	1.068 (0.088)	0.946 (0.046)	0.122	0.174
Number of words correct - Kiswahili	6.130 (0.596)	6.342 (0.427)	-0.212	16.867 (0.844)	15.425 (0.462)	1.442	1.654
Comprehension questions correct - Kiswahili	0.450 (0.050)	0.482 (0.036)	-0.031	1.272 (0.075)	1.134 (0.039)	0.138	0.169
Maximum Number of Observations	700	1499		490	1696		4385
Date				January 2012			

Note. This table presents the coefficient of the OLS regression of the treatment indicator (equal 1 if treated; 0 otherwise) against each of the listed baseline characteristics. Treatment includes schools receiving the intervention at least one year.

Table A3. Baseline Student Characteristics: description and balance for students in classrooms with and without trained teachers

	Teacher Trained			Teacher not Trained			Double Difference
	Control	Treatment	Control-Treatment	Control	Treatment	Control-Treatment	
Student is female	0.497 (0.003)	0.500 (0.004)	-0.003	0.489 (0.007)	0.498 (0.004)	-0.009	-0.006
Student age	6.953 (0.127)	7.262 (0.068)	-0.309	7.013 (0.088)	6.883 (0.059)	0.130	0.439
Student is in grade 2	1.471 (0.060)	1.554 (0.026)	-0.084	1.516 (0.035)	1.460 (0.022)	0.056	0.14
Student has a radio at home	0.791 (0.026)	0.775 (0.013)	0.016	0.767 (0.019)	0.768 (0.012)	-0.001	-0.017
Student has a phone or mobile at home	0.878 (0.024)	0.924 (0.011)	-0.046	0.930 (0.015)	0.920 (0.012)	0.011	0.057
Student has electricity at home	0.577 (0.081)	0.673 (0.035)	-0.096	0.641 (0.053)	0.771 (0.027)	-0.13	-0.034
Student has a television at home	0.601 (0.065)	0.670 (0.030)	-0.069	0.625 (0.035)	0.732 (0.025)	-0.107	-0.038
Student's mother can read and write	0.921 (0.020)	0.922 (0.008)	-0.000	0.925 (0.013)	0.920 (0.009)	0.004	0.004
Number of words correct - English	11.906 (2.562)	15.914 (1.165)	-4.008	11.247 (1.442)	14.482 (0.897)	-3.236	0.772
Comprehension questions correct - English	0.539 (0.149)	0.668 (0.070)	-0.128	0.468 (0.086)	0.577 (0.049)	-0.109	0.019
Number of words correct - Kiswahili	9.609 (1.691)	12.029 (0.824)	-2.420	9.423 (1.071)	10.608 (0.629)	-1.185	1.235
Comprehension questions correct - Kiswahili	0.701 (0.141)	0.880 (0.064)	-0.179	0.721 (0.086)	0.795 (0.049)	-0.074	0.105
Maximum Number of Observations	340	1442		620	1683		4385
Date				January 2012			

Note. This table presents the coefficient of the OLS regression of the treatment indicator (equal 1 if treated; 0 otherwise) against each of the listed baseline characteristics. Treatment includes schools receiving the intervention at least one year.

Table A4. Baseline Student Characteristics: description and balance for students in classrooms with teachers above and below median experience

	Below 7 years of experience			Above 7 years of experience			Double Difference
	Control	Treatment	Control-Treatment	Control	Treatment	Control-Treatment	
Student is female	0.501 (0.004)	0.494 (0.004)	0.007	0.491 (0.005)	0.505 (0.005)	-0.014	-0.021
Student age	6.952 (0.071)	7.130 (0.065)	-0.179	6.975 (0.102)	6.978 (0.062)	-0.003	0.176
Student is in grade 2	1.465 (0.026)	1.462 (0.020)	0.004	1.568 (0.043)	1.542 (0.023)	0.026	0.022
Student has a radio at home	0.760 (0.018)	0.761 (0.013)	-0.001	0.780 (0.024)	0.778 (0.011)	0.002	0.003
Student has a phone or mobile at home	0.934 (0.012)	0.929 (0.011)	0.006	0.906 (0.027)	0.914 (0.013)	-0.008	-0.014
Student has electricity at home	0.625 (0.053)	0.678 (0.033)	-0.053	0.745 (0.054)	0.779 (0.028)	-0.033	0.02
Student has a television at home	0.626 (0.039)	0.662 (0.029)	-0.036	0.699 (0.044)	0.751 (0.024)	-0.052	-0.016
Student's mother can read and write	0.922 (0.014)	0.914 (0.009)	0.008	0.929 (0.015)	0.924 (0.009)	0.005	-0.003
Number of words correct - English	11.061 (1.464)	12.669 (0.979)	-1.607	16.318 (1.935)	17.698 (1.007)	-1.380	0.227
Comprehension questions correct - English	0.425 (0.078)	0.490 (0.054)	-0.065	0.768 (0.122)	0.756 (0.061)	0.012	0.077
Number of words correct - Kiswahili	8.791 (1.038)	9.881 (0.692)	-1.090	13.552 (1.280)	12.639 (0.725)	0.913	2.003
Comprehension questions correct - Kiswahili	0.643 (0.080)	0.723 (0.052)	-0.080	1.041 (0.105)	0.951 (0.058)	0.091	0.171
Maximum Number of Observations	750	1709		440	1486		4385
Date	January 2012						

Note. This table presents the coefficient of the OLS regression of the treatment indicator (equal 1 if treated; 0 otherwise) against each of the listed baseline characteristics. Treatment includes schools receiving the intervention at least one year.

Table A5. Baseline Student Characteristics: Bottom 10th percentile comparison

	Top 90%	Bottom 10%	Difference
Student is female	0.499 (0.003)	0.493 (0.008)	0.006
Student age	6.993 (0.036)	7.407 (0.151)	-0.414***
Student is in grade 2	1.500 (0.001)	1.500 (0.000)	0.000
Student has a radio at home	0.765 (0.008)	0.799 (0.028)	-0.033
Student has a phone or mobile at home	0.933 (0.007)	0.821 (0.029)	0.112***
Student has electricity at home	0.769 (0.018)	0.156 (0.029)	0.613***
Student has a television at home	0.737 (0.016)	0.242 (0.028)	0.495***
Student's mother can read and write	0.927 (0.005)	0.854 (0.026)	0.074***
Speaks the same language at school and home	0.255 (0.016)	0.107 (0.032)	0.148***
Number of words correct - English	15.829 (0.590)	1.595 (0.438)	14.233***
Comprehension questions correct - English	0.659 (0.035)	0.007 (0.004)	0.652***
Number of words correct - Kiswahili	12.028 (0.394)	1.271 (0.165)	10.756***
Comprehension questions correct - Kiswahili	0.898 (0.031)	0.067 (0.012)	0.831***
Maximum Number of Observations	3965	420	4385
Date		January 2012	

Note. This table presents the coefficient of the OLS regression of an indicator equal 1 if school is in bottom 10 percent at baseline and 0 otherwise, against each of the listed baseline characteristics. Kiswahili Word Reading is used to identify 10th percentile * $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$

Table A6. Baseline Teacher Characteristics: Bottom 10th percentile comparison

	Top 90%	Bottom 10%	Difference
Percentage of female teachers	0.896 (0.016)	0.732 (0.070)	0.165***
Age	34.202 (0.553)	40.605 (1.711)	-6.403***
Teacher formally trained	0.430 (0.026)	0.525 (0.080)	-0.095
Teaching experience (years)	10.966 (0.512)	15.744 (1.646)	-4.778***
Teacher observed at least monthly	0.538 (0.026)	0.146 (0.056)	0.391***
Teacher presents good scheme of work	0.404 (0.026)	0.425 (0.079)	-0.021
Teacher presents good lesson plan	0.289 (0.024)	0.375 (0.078)	-0.086
Maximum Number of Observations	367	41	411
Date	January 2012		

Note. This table presents the coefficient of the OLS regression of an indicator equal 1 if school is in bottom 10 percent at baseline and 0 otherwise, against each of the listed baseline characteristics. Kiswahili Word Reading is used to identify 10th percentile
 *p<0.10 **p<0.05 ***p<0.01

Table A7. Baseline School Characteristics: bottom 10th percentile comparison

	Top 90%	Bottom 10%	Difference
School Nonformal	0.628 (0.034)	0.000 (0.000)	0.628***
Head Teacher is female	0.524 (0.036)	0.143 (0.078)	0.381***
Head Teacher Experience (Years)	7.209 (0.368)	9.762 (1.370)	-2.553**
School has been closed since start of year	0.066 (0.018)	0.190 (0.088)	-0.125**
School Mean Score on KCPE	235.397 (4.121)	215.814 (7.254)	19.584*
Clean and Safe Water Supply	0.712 (0.032)	0.381 (0.109)	0.331***
Presence of electricity	0.581 (0.035)	0.095 (0.066)	0.486***
Presence of Computer Room	0.126 (0.024)	0.048 (0.048)	0.079
Presence of Library	0.369 (0.034)	0.143 (0.078)	0.226**
School Average English Word Correct	15.888 (0.596)	1.595 (0.438)	14.293***
School Average English Comprehension	0.657 (0.035)	0.007 (0.004)	0.650***
School Average Kiswahili Word Correct	12.071 (0.399)	1.271 (0.165)	10.799***
School Average Kiswahili Comprehension	0.899 (0.031)	0.067 (0.012)	0.832***
Maximum Number of Observations	60	160	220
Date	January 2012		

Note. This table presents the coefficient of the OLS regression of an indicator equal 1 if school is in bottom 10 percent at baseline and 0 otherwise, against each of the listed baseline characteristics. Kiswahili Word Reading is used to identify 10th percentile
 *p<0.10 **p<0.05 ***p<0.01

Table A8. Baseline and Endline Teacher Characteristics

	Control	Treatment	Control-Treatment
Baseline			
Percentage of female teachers	0.842 -0.038	0.891 -0.018	-0.049
Age	35.121 (1.145)	34.721 (0.605)	0.400
Teacher formally trained	0.362 (0.050)	0.463 (0.028)	-0.101*
Teaching experience (years)	12.068 (1.092)	11.261 (0.552)	0.808
N	95	313	386
F statistic		1.452	
Endline			
Percentage of female teachers	0.843 (0.036)	0.885 (0.018)	-0.042
Age	35.284 -1.055	35.495 -0.575	-0.211
Teacher formally trained	0.578 (0.049)	0.613 (0.028)	-0.035
Teaching experience (years)	11.206 -0.928	11.572 -0.513	-0.366
N	102	313	415
F statistic		0.449	

Note. This table presents the coefficient of the OLS regression of the treatment indicator (equal 1 if treated; 0 otherwise) against each of the listed baseline characteristics. Treatment includes schools receiving the intervention at least one year. * $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$

Table A9. School-level Moderation on sound, letter, and syllable

	English Sound				English Letter				Kiswahili Letter				Kiswahili Syllable			
Treatment	2.89*** (0.30)	1.89** (0.61)	3.22*** (0.31)	3.22*** (0.30)	27.32*** (2.27)	19.99*** (3.99)	29.22*** (2.40)	29.10*** (2.39)	22.92*** (2.54)	16.87*** (4.65)	25.12*** (2.76)	25.04*** (2.72)	12.01*** (1.98)	1.81 (4.23)	12.35*** (2.01)	12.05*** (1.99)
School is Nonformal	0.65 (0.34)	-0.54 (0.73)	0.59 (0.42)	0.70 (0.42)	13.03*** (3.00)	4.28 (5.14)	13.90*** (3.71)	14.28*** (3.77)	12.58*** (3.02)	5.37 (5.77)	12.75*** (3.50)	13.93*** (3.56)	10.74*** (2.56)	-1.43 (5.13)	8.21** (2.80)	9.30** (2.81)
Treatment X Nonformal		1.27 (0.69)				9.39 (4.81)				7.74 (5.53)				13.06** (4.80)		
Baseline School Average			0.16 (0.15)	-0.18 (0.25)			1.80 (1.14)	0.94 (1.62)			3.88** (1.35)	0.76 (2.23)			5.54*** (1.14)	2.96 (1.52)
Treatment X Baseline Avg.				0.46 (0.27)				1.33 (1.88)				4.37 (2.34)				3.74* (1.79)
Constant	2.79*** (0.47)	3.76*** (0.71)	2.71*** (0.59)	2.66*** (0.60)	20.29*** (4.18)	27.42*** (5.21)	14.81** (5.36)	14.77** (5.39)	22.99*** (4.10)	28.87*** (5.64)	19.83*** (4.80)	19.28*** (4.88)	27.08*** (3.21)	37.01*** (4.84)	25.44*** (3.55)	25.05*** (3.52)
N Students	4162	4162	3191	3191	4162	4162	3191	3191	4162	4162	3191	3191	4162	4162	3191	3191
Controls			Y				Y				Y				Y	
KCPE Multiple Imputation			Y				Y				Y				Y	

Note. Controls also include student grade and SES, school-level resources, school baseline KCPE, school type, and district fixed-effects.

* p<0.05 ** p<0.01 ***p<0.001

Bibliography

- Abadzi, H. (2006). *Efficient learning for the poor: insights from the frontier of cognitive neuroscience*. <https://doi.org/10.1596/978-0-8213-6688-2>
- Académie française. (1995). *Statuts et règlements*. Retrieved from http://academie-francaise.fr/sites/academie-francaise.fr/files/statuts_af_0.pdf
- ADEA. (2010). *Policy guide on the integration of African languages and cultures into education systems*. Retrieved from Association for the Development of Education in Africa (ADEA) website: <http://unesdoc.unesco.org/images/0021/002156/215634e.pdf>
- Afrobarometer. (n.d.). Our history. Retrieved January 29, 2019, from <http://www.afrobarometer.org/about/our-history>
- Akyeampong, K., Lussier, K., Pryor, J., & Westbrook, J. (2013). Improving teaching and learning of basic maths and reading in Africa: Does teacher preparation count? *International Journal of Educational Development, 33*(3), 272–282. <https://doi.org/10.1016/j.ijedudev.2012.09.006>
- Albaugh, E. A. (2012). *Language policies in African states—updated, January 2012*. Brunswick, ME: Bowdoin College.
- Albaugh, E. A. (2014). *State-building and multilingual education in Africa*. Cambridge, United Kingdom: Cambridge University Press.
- Aldroubi, M. (2017, September 12). Iraqi parliament rejects Kurdistan independence referendum. *The National*. Retrieved from

<https://www.thenational.ae/world/mena/iraqi-parliament-rejects-kurdistan-independence-referendum-1.627743>

Alexander, A. C. (2012). Change in women's descriptive representation and the belief in women's ability to govern: a virtuous cycle. *Politics & Gender*, 8(04), 437–464.

<https://doi.org/10.1017/S1743923X12000487>

Alexander, N. (1997). Language policy and planning in the new South Africa. *African Sociological Review*, 1(1), 82–92.

Altinok, N. (2013). *The impact of teacher knowledge on student achievement in 14 sub-Saharan African countries*. Paris: UNESCO.

Annamalai, E. (2001). *Managing multilingualism in India: Political and linguistic manifestations*. New Delhi: SAGE Publications.

Arua, A. E., & Magocha, K. (2002). Patterns of language use and language preference of some children and their parents in Botswana. *Journal of Multilingual and Multicultural Development*, 23(6), 449–461.

<https://doi.org/10.1080/01434630208666479>

Auer, P., & Wei, L. (2007). *Handbook of multilingualism and multilingual communication* (Vol. 5). Berlin, Germany: De Gruyter.

August, D., & Hakuta, K. (1998). *Educating language-minority children*. Washington, D.C: National Academy Press.

Banerjee, A. V., Cole, S., Duflo, E., & Linden, L. (2007). Remediating education: Evidence from two randomized experiments in India. *The Quarterly Journal of Economics*, 122(3), 1235–1264.

- Banerjee, A. V., & Duflo, E. (2011). Why aren't children learning. *Development Outreach*, 13(1), 36–44.
- Barrera-Osorio, F., Blakeslee, D. S., Hoover, M., Linden, L. L., & Raju, D. (2011). Expanding educational opportunities in remote parts of the world: evidence from a RCT of a public private partnership in Pakistan. *Third Institute for the Study of Labor (IZA) Workshop, "Child Labor in Developing Countries."* Presented at the Mexico City, Mexico. Retrieved from http://www.iza.org/conference_files/childl2011/blakeslee_d6783.pdf
- Bartlett, L., Dowd, A. J., & Jonason, C. (2015). Problematizing early grade reading: Should the post-2015 agenda treasure what is measured? *International Journal of Educational Development*, 40, 308–314. <https://doi.org/10.1016/j.ijedudev.2014.10.002>
- Batibo, H. (2005). *Language decline and death in Africa: Causes, consequences, and challenges* (Vol. 132). Clevedon, UK: Multilingual Matters.
- Baum, D., Lewis, L., Lusk-Stover, O., & Patrinos, H. (2014). What matters most for engaging the private sector in education. *World Bank Framework Paper*. Retrieved from <https://wdronline.worldbank.org/handle/10986/21756>
- Beaman, L., Pande, R., & Cirone, A. (2012). Politics as a male domain and empowerment in India. In S. Franceschet, M. L. Krook, & J. Piscopo, *The impact of gender quotas* (Vol. 13). New York, NY: Oxford University Press.
- Bennell, P. (2004). *Teacher motivation and incentives in sub-Saharan Africa and Asia*. 60.

- Bennell, P., & Akyeampong, K. (2007). *Teacher motivation in sub-Saharan Africa and south Asia*. Retrieved from https://www.researchgate.net/profile/Kwame_Akyeampong/publication/44838477_Teacher_motivation_in_Sub-Saharan_Africa_and_South_Asia/links/0a85e53c3d8bbacf8a000000.pdf
- Benson, C. (2004). Bilingual schooling in Mozambique and Bolivia: from experimentation to implementation. *Language Policy*, 3(1), 47–66. <https://doi.org/10.1023/B:LPOL.0000017725.62093.66>
- Benson, C. (2013). Towards adopting a multilingual habitus in educational development. In C. Benson & K. Kosonen (Eds.), *Language issues in comparative education: inclusive teaching and learning in non-dominant languages and cultures*. Rotterdam; Boston: SensePublishers.
- Benson, C. J., & Kosonen, K. (2013). *Language issues in comparative education: inclusive teaching and learning in non-dominant languages and cultures*. Retrieved from <http://site.ebrary.com/id/10720818>
- Benson, C., & Young, C. (2016). How can mother tongue-based MLE be carried out in classrooms where three or more local languages are represented as mother tongues? In B. Trudell & C. Young (Eds.), *Good Answers to Tough Questions in Mother Tongue-Based Multilingual Education* (p. 7). SIL International.
- Beyogle, R. (2015). Language contact and language attitudes in two Dagara-speaking border communities in Burkina Faso and Ghana. In E. Zsiga, O. T. Boyer, & R. Kramer, *Languages in Africa*. Washington, D.C: Georgetown University Press.

- Bialystok, E., Luk, G., & Kwan, E. (2005). Bilingualism, biliteracy, and learning to read: Interactions among languages and writing systems. *Scientific Studies of Reading*, 9(1), 43–61.
- Bloom, H. S., & Weiland, C. (2015). *Quantifying variation in Head Start effects on young children's cognitive and socio-emotional skills using data from the National Head Start Impact Study*. MDRC.
- Bokamba, E. G. (1991). French colonial language policies in Africa and their legacies. In D. Marshall, *Language planning: Essays in honor of Joshua A. Fishman, Volume 3* (pp. 175–214). John Benjamins Publishing.
- Bokamba, E. G. (2015). Multilingualism as a Sociolinguistic Phenomenon: Evidence from Africa. In E. Zsiga, O. T. Boyer, & R. Kramer, *Languages in Africa*. Washington, D.C: Georgetown University Press.
- Bourdon, J., Frölich, M., & Michaelowa, K. (2010). Teacher shortages, teacher contracts and their effect on education in Africa. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 173(1), 93–116. <https://doi.org/10.1111/j.1467-985X.2009.00601.x>
- Bradley, D. (2003). *Issues in orthography development and reform*. Presented at the Language Development, Language Revitalization and Multilingual Education in Minority Communities in Asia, Bangkok, Thailand. Retrieved from http://palaso.org/ldc/plenary_papers/david_bradley.pdf
- Broom, Y. (2004). Reading English in multilingual South African primary schools. *International Journal of Bilingual Education and Bilingualism*, 7(6), 506–528. <https://doi.org/10.1080/13670050408667828>

- Burchfield, S., Hua, H., Noyes, D., & van de Waal, W. (2017). Improving early grade reading outcomes: *Aprender a ler* in Mozambique. *New Directions for Child and Adolescent Development*, 2017(155), 117–130. <https://doi.org/10.1002/cad.20189>
- Cahill, M. (2014). Non-linguistic factors in orthographies. In Cahill, Michael & K. Rice (Eds.), *Developing Orthographies for Unwritten Languages*. Dallas, Texas: SIL International.
- Cameron, A. C., & Miller, D. L. (2015). A practitioner's guide to cluster-robust inference. *Journal of Human Resources*, 50(2), 317–372. <https://doi.org/10.3368/jhr.50.2.317>
- Carrington, V., & Luke, A. (1997). Literacy and Bourdieu's sociological theory: A reframing. *Language and Education*, 11(2), 96–112. <https://doi.org/10.1080/09500789708666721>
- Chambers, J. K., & Trudgill, P. (1998). *Dialectology*. Cambridge, United Kingdom: Cambridge University Press.
- Chaudhury, N., Hammer, J., Kremer, M., Muralidharan, K., & Rogers, F. H. (2006). Missing in action: teacher and health worker absence in developing countries. *The Journal of Economic Perspectives*, 20(1), 91–116.
- Chetty, R., Friedman, J. N., & Rockoff, J. E. (2011). *The long-term impacts of teachers: Teacher value-added and student outcomes in adulthood*. Retrieved from National Bureau of Economic Research website: <http://www.nber.org/papers/w17699>
- Chudgar, A. (2015). Association between contract teachers and student learning in five francophone African countries. *Comparative Education Review*, 59(2), 261–288. <https://doi.org/10.1086/680351>

- Cincotta-Segi, A. (2011). Talking in, talking around and talking about the L2: three literacy teaching responses to L2 medium of instruction in the Lao PDR. *Compare: A Journal of Comparative and International Education*, 41(2), 195–209. <https://doi.org/10.1080/03057925.2011.547285>
- Clayton, T. (1995). Restriction or resistance? French colonial educational development in Cambodia. *Educational Policy Analysis Archives*, 14.
- Clément, R. (1986). Second language proficiency and acculturation: an investigation of the effects of language status and individual characteristics. *Journal of Language and Social Psychology*, 5(4), 271–290. <https://doi.org/10.1177/0261927X8600500403>
- Coombs, P. H., & Ahmed, M. (1974). *Attacking rural poverty: How nonformal education can help* (No. 10091). Essex, CT: International Council for Educational Development.
- Crawford, J. (2004). *Educating English learners: Language diversity in the classroom* (5th ed.). Los Angeles, CA: Bilingual Education Services.
- Cummins, J. (1979). Linguistic interdependence and the educational development of bilingual children. *Review of Educational Research*, 49(2), 222–251. <https://doi.org/10.3102/00346543049002222>
- Day Ashley, L., Mcloughlin, C., Aslam, M., Engel, J., Wales, J., Rawal, S., ... Rose, P. (2014). *The role and impact of private schools in developing countries: A rigorous review of the evidence. Final report*. Department for International Development.

- de Galbert, P. (2017). *Beyond primary school completion: Exploring how a public-private partnership in Uganda supported secondary school*. Retrieved from <https://dash.harvard.edu/handle/1/33797214>
- de Galbert, P., Rousseau, M., & Mount-Cors, M. F. (2018). *Lecture pour tous Senegal. EGRA CI analysis report*.
- Desai, Z. (1995). The evolution of a post-apartheid language policy in South Africa: An on-going site of struggle. *European Journal of Intercultural Studies*, 5(3), 18–25. <https://doi.org/10.1080/0952391950050303>
- Di Carlo, P., Good, J., & Ojong Diba, R. (2019). Multilingualism in rural Africa. In *Oxford Research Encyclopedia of Linguistics*. <https://doi.org/10.1093/acrefore/9780199384655.013.227>
- Downer, J. T., Locasale-Crouch, J., Hamre, B., & Pianta, R. (2009). Teacher characteristics associated with responsiveness and exposure to consultation and online professional development resources. *Early Education and Development*, 20(3), 431–455.
- Dresser, R. (2012). The impact of scripted literacy instruction on teachers and students. *Issues in Teacher Education*, 21(1), 18.
- Dressler, C., & Kamil, M. L. (2006). First-and Second-Language Literacy. In D. August & T. Shanahan, *Developing literacy in second-language learners: Report of the national literacy panel on language minority children and youth*. Retrieved from <http://psycnet.apa.org/psycinfo/2006-10122-009>
- Dubeck, M. M., & Gove, A. (2015). The Early Grade Reading Assessment (EGRA): Its theoretical foundation, purpose, and limitations. *International Journal of*

Educational Development, 40, 315–322.

<https://doi.org/10.1016/j.ijedudev.2014.11.004>

Dubeck, M. M., Jukes, M. C. H., & Okello, G. (2012). Early primary literacy instruction in Kenya. *Comparative Education Review*, 56(1), 48–68.

<https://doi.org/10.1086/660693>

Duflo, E., Dupas, P., & Kremer, M. (2011). Peer effects, teacher incentives, and the impact of tracking: Evidence from a randomized evaluation in Kenya. *The American Economic Review*, 101(5), 1739–1774.

Dutcher, N. (2001). *Expanding educational opportunity in linguistically diverse societies*.

Retrieved from <https://eric.ed.gov/?id=ED466099>

Ehri, L. C., Nunes, S. R., Willows, D. M., Schuster, B. V., Yaghoub-Zadeh, Z., &

Shanahan, T. (2001). Phonemic awareness instruction helps children learn to read: Evidence from the National Reading Panel’s meta-analysis. *Reading Research Quarterly*, 36(3), 250–287.

Errihani, M. (2017). English education policy and practice in Morocco. In R. Kirkpatrick, *English language education policy in the Middle East and North Africa* (pp. 115–131). New York, NY: Springer Publishing International.

Ethnologue. (n.d.). How many languages in the world are unwritten? Retrieved January 31, 2019, from Ethnologue website: <https://www.ethnologue.com/enterprise-faq/how-many-languages-world-are-unwritten-0>

European Commission. (2013, September 26). *Frequently asked questions on languages in Europe*. Retrieved from http://europa.eu/rapid/press-release_MEMO-13-825_en.htm

- Feller, A., Grindal, T., Miratrix, L., & Page, L. C. (2016). Compared to what? Variation in the impacts of early childhood education by alternative care type. *The Annals of Applied Statistics*, *10*(3), 1245–1285. <https://doi.org/10.1214/16-AOAS910>
- Freudenberger, E., & Davis, J. (2017). *Tusome external evaluation—Midline report*. Retrieved from MSI, A Tetra Tech Company website:
https://pdf.usaid.gov/pdf_docs/PA00MS6J.pdf
- Friedlander, E. W., Arshan, N., Zhou, S., & Goldenberg, C. (2018). Lifewide or school-only learning: Approaches to addressing the developing world’s learning crisis. *American Educational Research Journal*, *35*. <https://doi.org/0002831218792841>
- Frölich, M., & Michaelowa, K. (2005). *Peer effects and textbooks in primary education: evidence from francophone sub-Saharan Africa*. 61.
- García Coll, C., Akiba, D., Palacios, N., Bailey, B., Silver, R., DiMartino, L., & Chin, C. (2002). Parental involvement in children’s education: Lessons from three immigrant groups. *Parenting*, *2*(3), 303–324.
https://doi.org/10.1207/S15327922PAR0203_05
- García, O. (2009). *Bilingual education in the 21st century: a global perspective*. Malden, MA: Wiley-Blackwell.
- García, O., Flores, N., & Chu, H. (2011). Extending bilingualism in U.S. secondary education: New variations. *International Multilingual Research Journal*, *5*(1), 1–18. <https://doi.org/10.1080/19313152.2011.539486>
- García, O., Kleifgen, J. A., & Falchi, L. (2008). *From English language learners to emergent bilinguals*. Retrieved from

<http://deimos3.apple.com/WebObjects/Core.woa/DownloadTrackPreview/tc.columbia.edu.1463244836.01463261407.1463634935.pdf>

García, O., & Wei, L. (2014). Language, bilingualism and education. In *Translanguaging: Language, bilingualism and education* (pp. 46–62). Retrieved from http://link.springer.com/chapter/10.1057/9781137385765_4

Gelman, A. (2018, March 15). You need 16 times the sample size to estimate an interaction than to estimate a main effect. Retrieved June 18, 2018, from Statistical Modeling, Causal Inference, and Social Science website: <http://andrewgelman.com/2018/03/15/need-16-times-sample-size-estimate-interaction-estimate-main-effect/>

Gelman, A., & Hill, J. (2006). *Data analysis using regression and multilevel/hierarchical models*. Cambridge, United Kingdom: Cambridge university press.

Genesee, F., Geva, E., Dressler, C., & Kamil, M. (2006). Synthesis: Cross-linguistic relationships. In D. August & T. Shanahan, *Developing literacy in second-language learners: Report of the national literacy panel on language-minority children and youth*. (pp. 153–174). Retrieved from <http://www.psych.mcgill.ca/perpg/fac/genesee/23.pdf>

Gertler, P. J., Patrinos, H. A., & Rubio-Codina, M. (2012). Empowering parents to improve education: Evidence from rural Mexico. *Journal of Development Economics*, 99(1), 68–79.

Gipps, C. V. (2002). *Beyond testing: Towards a theory of educational assessment*. New York, NY: Routledge.

- Glewwe, P., Hanushek, E. A., Humpage, S. D., & Ravina, R. (2011). *School resources and educational outcomes in developing countries: A review of the literature from 1990 to 2010*. Retrieved from National Bureau of Economic Research website: <http://www.nber.org/papers/w17554>
- Glewwe, P., Kremer, M., & Moulin, S. (2007). *Many children left behind? Textbooks and test scores in Kenya*. Retrieved from National Bureau of Economic Research website: <http://www.nber.org/papers/w13300>
- Government Printing Office. *United States Code Title 4 - FLAG AND SEAL, SEAT OF GOVERNMENT, AND THE STATES.* , (2011).
- Grosjean, F. (1989). Neurolinguists, beware! The bilingual is not two monolinguals in one person. *Brain and Language*, 36(1), 3–15. [https://doi.org/10.1016/0093-934X\(89\)90048-5](https://doi.org/10.1016/0093-934X(89)90048-5)
- Grossman, P., & Thompson, C. (2008). Learning from curriculum materials: Scaffolds for new teachers? *Teaching and Teacher Education*, 24(8), 2014–2026. <https://doi.org/10.1016/j.tate.2008.05.002>
- Hamid, M. O., & Jahan, I. (2015). Language, identity, and social divides: Medium of instruction debates in Bangladeshi print media. *Comparative Education Review*, 59(1), 75–101. <https://doi.org/10.1086/679192>
- Han, L. (2013). Is centralized teacher deployment more equitable? Evidence from rural China. *China Economic Review*, 24, 65–76. <https://doi.org/10.1016/j.chieco.2012.10.001>
- Hanushek, E. A. (2003). The failure of input-based schooling policies. *The Economic Journal*, 113(485), F64–F98.

- Harlen, W., & James, M. (1997). Assessment and learning: Differences and relationships between formative and summative assessment. *Assessment in Education: Principles, Policy & Practice*, 4(3), 365–379.
<https://doi.org/10.1080/0969594970040304>
- Harris, D. (1989). Comparison-of 1-, 2-, and 3-parameter IRT models. *Educational Measurement: Issues and Practice*, 14(64), 05.
- Haugen, E. (1966). Dialect, language, nation. *American Anthropologist*, 68(4), 922–935.
<https://doi.org/10.1525/aa.1966.68.4.02a00040>
- Heckman, J. J., Moon, S. H., Pinto, R., Savelyev, P. A., & Yavitz, A. (2010). The rate of return to the HighScope Perry Preschool Program. *Journal of Public Economics*, 94(1–2), 114–128. <https://doi.org/10.1016/j.jpubeco.2009.11.001>
- Heller, M. (2008). Language and the nation-state: Challenges to sociolinguistic theory and practice. *Journal of Sociolinguistics*, 12(4), 504–524.
<https://doi.org/10.1111/j.1467-9841.2008.00373.x>
- Hess, R. D., & Holloway, S. D. (1984). Family and school as educational-institutions. *Review of Child Development Research*, 7, 179–222.
- Heugh, K., Benson, C., Bogale, B., & Mekonnen, A. G. Y. (2007). Final report: Study on medium of instruction in primary schools in Ethiopia. *Commissioned by the Ministry of Education, Ethiopia*.
- Heugh, K., & Mulumba, B. M. (2014). *Implementing local languages medium education in the early primary curriculum of Ugandan schools: A Literacy and Basic Education (LBE) intervention in six districts in North and North West Uganda*.

- Hill, N. E., & Tyson, D. F. (2009). Parental involvement in middle school: A meta-analytic assessment of the strategies that promote achievement. *Developmental Psychology, 45*(3), 740–763. <https://doi.org/10.1037/a0015362>
- Hoff, E. (2018). Bilingual Development in Children of Immigrant Families. *Child Development Perspectives, 12*(2), 80–86. <https://doi.org/10.1111/cdep.12262>
- Hoover, W. A., & Gough, P. B. (1990). The simple view of reading. *Reading and Writing, 2*(2), 127–160.
- Hoppers, W. (2006). *Non-formal education and basic education reform: A conceptual review*. Retrieved from <https://eric.ed.gov/?id=ED495405>
- Hornberger, N. H. (2006). Frameworks and models in language policy and planning research. In *An Introduction to Language Policy: Theory and Method*. Malden, MA: Blackwell.
- Hornberger, N. H., & Link, H. (2012). Translanguaging and transnational literacies in multilingual classrooms: a biliteracy lens. *International Journal of Bilingual Education and Bilingualism, 15*(3), 261–278. <https://doi.org/10.1080/13670050.2012.658016>
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6*(1), 1–55.
- Huang, F. L. (2016). Alternatives to multilevel modeling for the analysis of clustered data. *The Journal of Experimental Education, 84*(1), 175–196. <https://doi.org/10.1080/00220973.2014.952397>

- Hurdus, J., & Lasagabaster, D. (2018). Mappings and policies-in-the-making: Toward an ontogenetic understanding of language-in-education policy. *Language and Education, 32*(3), 227–241. <https://doi.org/10.1080/09500782.2018.1434786>
- IRIN. (2011, November). Ethnic minorities face higher school drop-out risk. Retrieved April 12, 2017, from <http://www.irinnews.org/report/94139/bangladesh-ethnic-minorities-face-higher-school-drop-out-risk>
- Jenkins, O. B. (2011). Population Analysis of the Arabic Languages. Retrieved November 20, 2018, from <http://strategyleader.org/articles/arabicpercent.html>
- Jeynes, W. H. (2005). A meta-analysis of the relation of parental involvement to urban elementary school student academic achievement. *Urban Education, 40*(3), 237–269. <https://doi.org/10.1177/0042085905274540>
- Ji, C. S., & Koblinsky, S. A. (2009). Parent involvement in children’s education: An exploratory study of urban, Chinese immigrant families. *Urban Education, 44*(6), 687–709. <https://doi.org/10.1177/0042085908322706>
- Johnson, D. R., & Young, R. (2011). Toward Best Practices in Analyzing Datasets with Missing Data: Comparisons and Recommendations. *Journal of Marriage and Family, 73*(5), 926–945. <https://doi.org/10.1111/j.1741-3737.2011.00861.x>
- Katahoire, A. (2006). *Effective literacy programs. Case studies in non formal education*. Libreville, Gabon: Association for the Development of Education in Africa (ADEA).
- Katz, L., & Frost, R. (1992). The reading process is different for different orthographies: The orthographic depth hypothesis. In C. Fowler, *Speech Research Status Report* (pp. 147–160). New Haven, CT: Haskins Laboratories.

- Kaviti, L. (2018). *The new curriculum of education in Kenya: A linguistic and education paradigm shift*. 12.
- Kelly, M. (2014). Recent developments in EU language policy. *European Journal of Language Policy*, 6(1), 119–50.
- Kenya Institute of Curriculum Development. (2017). *Basic education curriculum framework*. Retrieved from <https://kicd.ac.ke/wp-content/plugins/pdfjs-viewer-shortcode/pdfjs/web/viewer.php?file=https%3A%2F%2Fkicd.ac.ke%2Fwp-content%2Fuploads%2F2018%2F03%2FCURRICULUM-FRAMEWORK-FINAL-2017.pdf&download=false&print=false&openfile=false>
- Kim, Young-Suk G., Lee, H., & Zuilkowski, S. S. (2019). Impact of literacy interventions on reading skills in low- and middle-income countries: A meta-analysis. *Child Development*, 0(0). <https://doi.org/10.1111/cdev.13204>
- Kim, Young-Suk Grace, & Piper, B. (2019). Cross-language transfer of reading skills: an empirical investigation of bidirectionality and the influence of instructional environments. *Reading and Writing*, 32(4), 839–871. <https://doi.org/10.1007/s11145-018-9889-7>
- Kim, Y.-S., Boyle, H., Zuilkowski, S. S., & Nakamura, P. (2016). *Landscape report on early grade literacy*. Washington, D.C: USAID.
- Kiplang'at, J. (2014, January 27). Teachers and parents oppose new rule on language in schools. *Daily Nation*. Retrieved from <http://www.nation.co.ke/news/Teachers-fight-rule-on-using-local-language-in-classrooms/-/1056/2163022/-/jsehr0/-/index.html>

- Kirkpatrick, A. (2017). Language education policy among the Association of Southeast Asian Nations (ASEAN). *European Journal of Language Policy*, 9(1), 7–25.
<https://doi.org/10.3828/ejlp.2017.2>
- Kline, R. B. (2011). *Principles and practice of structural equation modeling* (3rd ed). New York: Guilford Press.
- Komakech, M. (2014, February 14). Thematic curriculum is not failing so give it a chance. *Daily Monitor*. Retrieved from
<http://www.monitor.co.ug/OpEd/Commentary/Thematic-curriculum-is-not-failing-so-give-it-a-chance/-/689364/2205312/-/u7hqih/-/index.html>
- Koretz, D. M. (2008). *Measuring up : What educational testing really tells us*. Cambridge, MA: Harvard University Press.
- Kosonen, K. (2017). *Language of instruction in Southeast Asia* [Background Paper]. UNESCO.
- Kraemer, H. C., Wilson, G. T., Fairburn, C. G., & Agras, W. S. (2002). Mediators and moderators of treatment effects in randomized clinical trials. *Archives of General Psychiatry*, 59(10), 877. <https://doi.org/10.1001/archpsyc.59.10.877>
- Leclerc, J. (2013). Sénégal. In *L'aménagement linguistique dans le monde*. Retrieved from <http://www.axl.cefan.ulaval.ca/afrique/senegal.htm>
- Lesaux, N. K., & Marietta, S. H. (2011). *Making assessment matter: using test results to differentiate reading instruction*. United States: Guilford Press.
- Lipsky, M. (1969). *Toward a theory of street-level bureaucracy*. Madison: Institute for Research on Poverty, University of Wisconsin.

- List of languages by number of native speakers. (n.d.). In *Wikipedia*. Retrieved January 23, 2019, from http://en.wikipedia.org/wiki/List_of_languages_by_number_of_native_speakers
- Luhn, A. (2018, April 3). Moscow threatens sanctions against Latvia over removal of Russian from secondary schools. *The Telegraph*. Retrieved from <https://www.telegraph.co.uk/news/2018/04/03/moscow-threatens-sanctions-against-latvia-removal-russian-secondary/>
- Lukowski, J., & Zawadzki, H. (2006). *A concise history of Poland*. Cambridge, United Kingdom: Cambridge University Press.
- May, S. (2013). *Language and minority rights: Ethnicity, nationalism and the politics of language* (2nd ed.). <https://doi.org/10.4324/9780203832547>
- Mazrui, A. A., & Mazrui, A. M. (1998). *The Power of Babel: Language and governance in the African experience*. University of Chicago Press.
- McCoy, D. C., & Wolf, S. (2018). Changes in classroom quality predict Ghanaian preschoolers' gains in academic and social-emotional skills. *Developmental Psychology*, 54(8), 1582–1599. <https://doi.org/10.1037/dev0000546>
- McIntyre, J. C. (2017). *Gender and race in children's picture books: A tragedy in three studies*. Harvard University, Cambridge, MA.
- Melby-Lervåg, M., & Lervåg, 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. <https://doi.org/10.1111/j.1467-9817.2010.01477.x>

- Miller, E. B., Farkas, G., Vandell, D. L., & Duncan, G. J. (2014). Do the effects of head start vary by parental preacademic stimulation? *Child Development, 85*(4), 1385–1400. <https://doi.org/10.1111/cdev.12233>
- MINEDLAC. (1996). *Final report*. Retrieved from UNESCO website: <http://unesdoc.unesco.org/images/0011/001121/112197Eo.pdf>
- Ministry of Education, Science and Technology. (2014). *Education for All 2015 National Review Report: Kenya*.
- Mose, P. N. (2017). Language-in-education policy in Kenya: Intention, interpretation, implementation. *Nordic Journal of African Studies, 26*(3), 17.
- Mulimbi, B., & Dryden-Peterson, S. (2018). Responses to cultural diversity in Botswana's schools: links between national policy, school actions and students' civic equality. *Journal of Curriculum Studies, 50*(3), 364–386.
- Mulumba, B. M. (2012, October 10). *Petition absence of local languages in the proposed ordinary secondary level curriculum*.
- Muralidharan, K., Singh, A., & Ganimian, A. J. (2016). *Disrupting education? Experimental evidence on technology-aided instruction in India*.
- Muralidharan, K., & Sundararaman, V. (2013). *Contract Teachers: Experimental Evidence from India* (No. w19440). <https://doi.org/10.3386/w19440>
- Mustapha, Z. (2012). *Parents' choices of language as the medium of instruction in schools*. Retrieved from http://www.aserpakistan.org/document/aser/2012/notes_aser_2012/Zubeida_Mustafa.pdf

- Muthén, L. K., & Muthén, B. O. (1998-2010). *Mplus user's guide. sixth edition*. Los Angeles, CA: Muthén & Muthén.
- Nabea, W. (2009). Language policy in Kenya: Negotiation with hegemony. *The Journal of Pan African Studies*, 3(1), 18.
- Nabirye, M., & De Schryver, G.-M. (2010). The monolingual Lusoga dictionary faced with demands from a new user category. *Lexikos*, 20(1), 326–350.
- Najjumba, I. M., & Marshall, J. H. (2013). *Improving learning in Uganda, volume 2: Problematic curriculum areas and teacher effectiveness: insights from national assessments*. Retrieved from <http://elibrary.worldbank.org/doi/book/10.1596/978-0-8213-9850-0>
- Namyalo, S., & Nakayiza, J. (2014). Dilemmas in implementing language rights in multilingual Uganda. *Current Issues in Language Planning*, 1–16.
<https://doi.org/10.1080/14664208.2014.987425>
- National Curriculum Development Center. (n.d.). *Status of teaching of and learning local language at upper primary in Uganda*.
- National Institute of Child Health and Human Development [US]. (2000). *Report of the national reading panel: Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: Reports of the subgroups*. Retrieved from <https://www.nichd.nih.gov/publications/pubs/nrp/Documents/report.pdf>
- Natukunda, C., & Businge, C. (2016, March 28). Mother tongue: A dying language of the heart. *New Vision*. Retrieved from

http://www.newvision.co.ug/new_vision/news/1420699/mother-tongue-dying-language-heart

- Ngugi, M., Mumiukha, C., Fedha, F., & Ndiga, B. (2015). Universal primary education in Kenya: advancement and challenges. *Journal of Education and Practice*, 10.
- Ngware, M. W., Oketch, M., & Ezech, A. C. (2011). Quality of primary education inputs in urban schools: Evidence from Nairobi. *Education and Urban Society*, 43(1), 91–116.
- Nye, B., Konstantopoulos, S., & Hedges, L. V. (2004). How large are teacher effects? *Educational Evaluation and Policy Analysis*, 26(3), 237–257.
- Nzogi, R. (2011). *Orthography development at SIL: The interface of linguistics, literacy principles, and community involvement*. Presented at the Multilingual Education Network (MLen) Uganda, Kampala, Uganda.
- Obiero, O. J. (2008). Evaluating language revitalization in Kenya: The contradictory face and place of the local community factor. *Nordic Journal of African Studies*, 17(4), 247–268.
- Odhiambo, T. (2019, January 19). It's time our schools, universities started teaching local. *Daily Nation*. Retrieved from <https://www.nation.co.ke/lifestyle/weekend/It-is-time-schools-universities-start-teaching-local-languages/1220-4941568-ic9o82/index.html>
- Odongo, J. (2012). *SIL led Kumam orthography workshop report*.
- Organisation of African Unity. (1986). *The language plan of action for Africa*. Retrieved from <http://www.bisharat.net/Documents/OAU-LPA-86.htm>

- Paris, S. G. (2005). Reinterpreting the development of reading skills. *Reading Research Quarterly*, 40(2), 184–202. <https://doi.org/10.1598/RRQ.40.2.3>
- Peterson, A. (2016). Getting ‘What Works’ working: Building blocks for the integration of experimental and improvement science. *International Journal of Research & Method in Education*, 39(3), 299–313.
<https://doi.org/10.1080/1743727X.2016.1170114>
- Petroni, F., & Serva, M. (2010). Measures of lexical distance between languages. *Physica A: Statistical Mechanics and Its Applications*, 389(11), 2280–2283.
<https://doi.org/10.1016/j.physa.2010.02.004>
- Phillipson, R. (1996). Linguistic imperialism: African perspectives. *ELT Journal*, 50(2), 160–167.
- Pike, C. (1986). History and imagination: Swahili literature and resistance to German language imperialism in Tanzania, 1885-1910. *The International Journal of African Historical Studies*, 19(2), 201–233. <https://doi.org/10.2307/219426>
- Piper, B. (2010a). *Kenya early grade reading assessment findings report*. Research Triangle Park, NC: RTI International.
- Piper, B. (2010b). *Uganda early grade reading assessment findings report: Literacy acquisition and mother tongue*. Research Triangle Park, NC: RTI International.
- Piper, B., Jepkemei, E., & Kibukho, K. (2015). Pro-poor PRIMR: Improving early literacy skills for children from low-income families in Kenya. *Africa Education Review*, 12(1), 67–87. <https://doi.org/10.1080/18146627.2015.1036566>
- Piper, B., Schroeder, L., & Trudell, B. (2011). *Oral Reading Fluency and Comprehension in Kenya: Reading Acquisition in a Multilingual Environment*.

- Piper, B., Sitabkhan, Y., Mejía, J., & Betts, K. (2018). *Effectiveness of teachers' guides in the global south: Scripting, learning outcomes, and classroom utilization*.
<https://doi.org/10.3768/rtipress.2018.op.0053.1805>
- Piper, B., & van Ginkel, A. J. (2017). Reading the script: How the scripts and writing systems of Ethiopian languages relate to letter and word identification. *Writing Systems Research*, 9(1), 36–59. <https://doi.org/10.1080/17586801.2016.1220354>
- Piper, B., & Zuilkowski, S. S. (2015). Teacher coaching in Kenya: Examining instructional support in public and nonformal schools. *Teaching and Teacher Education*, 47, 173–183. <https://doi.org/10.1016/j.tate.2015.01.001>
- Prevoo, M. J., Malda, M., Mesman, J., & van IJzendoorn, M. H. (2016). Within-and cross-language relations between oral language proficiency and school outcomes in bilingual children with an immigrant background: A meta-analytical study. *Review of Educational Research*, 86(1), 237–276.
- Proctor, C. P., August, D., Snow, C., & Barr, C. D. (2010). The interdependence continuum: A perspective on the nature of Spanish–English bilingual reading comprehension. *Bilingual Research Journal*, 33(1), 5–20.
<https://doi.org/10.1080/15235881003733209>
- Proctor, C. P., Haring, J. R., & Silverman, R. D. (2017). Linguistic interdependence between Spanish language and English language and reading: A longitudinal exploration from second through fifth grade. *Bilingual Research Journal*, 1–20.
<https://doi.org/10.1080/15235882.2017.1383949>
- Proctor, C. P., & Zhang-Wu, Q. (2019). Cross-linguistic relations among bilingual and biliterate learners. In V. Grøver, P. Uccelli, M. Rowe, & E. Lieven, *Learning*

through language: Towards an educationally informed theory of language learning (pp. 218–234). Cambridge: Cambridge University Press.

- Pryor, J., Akyeampong, K., Westbrook, J., & Lussier, K. (2012). Rethinking teacher preparation and professional development in Africa: an analysis of the curriculum of teacher education in the teaching of early reading and mathematics. *Curriculum Journal*, 23(4), 409–502. <https://doi.org/10.1080/09585176.2012.747725>
- Ramachandran, R. (2017). Language use in education and human capital formation: Evidence from the Ethiopian educational reform. *World Development*, 98, 195–213. <https://doi.org/10.1016/j.worlddev.2017.04.029>
- Read, T. (2015). *Where have all the textbooks gone?: Toward sustainable provision of teaching and learning materials in sub-Saharan Africa*.
<https://doi.org/10.1596/978-1-4648-0572-1>
- Reimers, F. (2006). Education and social progress. In V. Bulmer-Thomas, J. Coatsworth, & R. Cortes-Conde (Eds.), *The Cambridge economic history of Latin America* (pp. 427–480). <https://doi.org/10.1017/CHOL9780521812900.013>
- Renfrew, C., McMahon, A., & Trask, L. (2000). *Time depth in historical linguistics*. Cambridge: McDonald Institute for Archaeological Research.
- Republic of Kenya. *The Kenya Institute of Curriculum Development Act, 2013*. , Pub. L. No. 4 (2013).
- Republic of Kenya, & UNESCO. (2012). *Education for all end decade assessment (2001-2010)*. Retrieved from
<http://unesdoc.unesco.org/images/0022/002260/226082e.pdf>
- Republic of Uganda. (2005). *Constitution of the Republic of Uganda*.

- Republic of Uganda. (2006). *Primary School Curriculum*.
- Republic of Uganda. *THE EDUCATION (PRE-PRIMARY, PRIMARY AND POST-PRIMARY) ACT, 2008.* , (2008).
- République du Sénégal. *La Constitution du Sénégal.* , (1963).
- République du Sénégal. *CONSTITUTION DE LA REPUBLIQUE DU SENEGAL DU 22 JANVIER 2001.* , (2001).
- République du Sénégal. (2012). *Programme de Développement de l'Education et de la Formation (PDEF) 2012-2025*.
- Ricento, T. (2009). *An introduction to language policy: Theory and method*. Oxford: Blackwell.
- Riddell, T. Q. (2004). The impact of legal mobilization and judicial decisions: The case of official minority-language education policy in Canada for francophones outside Quebec. *Law & Society Review*, 38(3), 583–610. <https://doi.org/10.1111/j.0023-9216.2004.00058.x>
- Rogers, A. (Ed.). (2004). *Non-formal education: Flexible schooling or participatory education?* Hong Kong: Comparative Education Research Centre.
- Rouse, C. E., & Barrow, L. (2009). School vouchers and student achievement: Recent evidence and remaining questions. *Annual Review of Economics*, 1(1), 17–42. <https://doi.org/10.1146/annurev.economics.050708.143354>
- Rousseau, M., de Galbert, P., & Mount-Cors, M. F. (2017). *Lecture pour tous Senegal EGRA (early grade reading assessment) baseline report*.

- RTI International. (2010). *Ethiopia early grade reading assessment. Data analysis report: Language and early learning*. Research Triangle Park, NC: RTI International.
- RTI International. (2013). *USAID/Uganda School Health and Reading Program The Status of Early Grade Reading and Support to Primary School Teachers to Teach Reading in Uganda: Cluster 1 Baseline Report*. Research Triangle Park, NC: RTI International.
- RTI International. (2015). *Survey of children's reading materials in African languages in eleven countries. Final report*. Washington, D.C: U.S. Agency for International Development.
- RTI International. (2016a). *Early Grade Reading Assessment Toolkit*. Research Triangle Park, NC: RTI International.
- RTI International. (2016b). *The Primary Math and Reading (PRIMR) Initiative Endline Impact Evaluation – Revised Edition*. Research Triangle Park, NC: RTI International.
- RTI International. (2016c). *USAID/Uganda school health and reading program: Early grade reading assessment results*. Research Triangle Park, NC: RTI International.
- Ruiz, R. (1984). Orientations in language planning. *NABE: The Journal for the National Association for Bilingual Education*, 8(2), 15–34.
- Sabarwal, S., Evans, D. K., & Marshak, A. (2014). The permanent input hypothesis: The case of textbooks and (no) student learning in Sierra Leone. *World Bank Policy Research Working Paper*, (7021). Retrieved from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2491934

- Sallabank, J. (2010). Language documentation and language policy. In P. K. Austin, *Language Documentation and Description* (Vol. 7, pp. 141–171). London: SOAS.
- Samper, D. A. (2002). *Talking Sheng: The role of a hybrid language in the construction of identity and youth culture in Nairobi, Kenya* (Ph.D., University of Pennsylvania). Retrieved from <https://search.proquest.com/docview/305509897/abstract/DB70FB7289314606PQ/1>
- Saraceni, M., & Jacob, C. (2018). Revisiting borders: Named languages and de-colonization. *Language Sciences*. <https://doi.org/10.1016/j.langsci.2018.05.007>
- Satorra, A., & Bentler, P. M. (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika*, 66(4), 507–514.
- Save the Children. (2016). *Lessons in Literacy: 8 principles to ensure every last child can read*. London, UK: Save the Children.
- Sebba, M. (2007). *Spelling and society: The culture and politics of orthography around the world*. Cambridge, United Kingdom: Cambridge University Press.
- Seid, Y. (2016). Does learning in mother tongue matter? Evidence from a natural experiment in Ethiopia. *Economics of Education Review*, 55, 21–38. <https://doi.org/10.1016/j.econedurev.2016.08.006>
- Seton-Watson, H. (1977). *Nations and states*. Boulder, Colorado: Westview Press.
- SIL LEAD. (2015). *USAID/Uganda School Health and Reading Program: SIL LEAD Annual Report*. Retrieved from <https://static1.squarespace.com/static/51afb275e4b06f81432164f0/t/5665fa2bc21>

b86061401eed2/1449523755396/2015-10-13+Year+3+-+Annual+Report+-
+SIL+LEAD+-+FINAL.pdf

Simons, G. F., & Fennig, C. D. (2018). *Ethnologue: Languages of the world, twenty-first edition*. Retrieved from <https://www.ethnologue.com/>

Singh, K. (2008). *UNESCO's convention against discrimination in education (1960): Key pillar of the education for all*. (1), 16.

Snilstveit, B., Stevenson, J., Phillips, D., Vojtkova, M., Gallagher, E., Schmidt, T., ...
Eyers, J. (2015). *Interventions for improving learning outcomes and access to education in low- and middle-income countries. A systematic review* (No. 24).
London: International Initiative for Impact Evaluation (3ie).

Snow, C. E., & Kang, J. Y. (2006). Becoming bilingual, biliterate, and bicultural. In W. Damon & R. M. Lerner, *Handbook of child psychology*. Retrieved from
<http://onlinelibrary.wiley.com/doi/10.1002/9780470147658.chpsy0403/full>

Snow, C. E., & Kim, Y.-S. (2007). *Large problem spaces: The challenge of vocabulary for English language learners*. Retrieved from
<http://psycnet.apa.org/psycinfo/2006-21627-007>

Sprenger-Charolles, L., Siegel, L. S., Jiménez, J. E., & Ziegler, J. C. (2011). Prevalence and reliability of phonological, surface, and mixed profiles in dyslexia: A review of studies conducted in languages varying in orthographic depth. *Scientific Studies of Reading, 15*(6), 498–521.
<https://doi.org/10.1080/10888438.2010.524463>

- Spronk, T. (2014). Addressing the challenges of language choice in the implementation of mother-tongue based bilingual education in South Sudan. *Multilingual Education, 4*(1), 16. <https://doi.org/10.1186/s13616-014-0016-z>
- Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly, 21*(4), 360–407.
- StataCorp. (2013). *Stata 13 base reference manual*. College Station, TX: STATA press.
- Sure, K., & Ogechi, N. O. (2009). *Linguistic human rights and language policy in the Kenyan education system*. Ethiopia: OSSREA.
- Tembe, J. H. (2008). *Language Education Policy and Multilingual Literacies in Ugandan Primary Schools*. University of British Columbia, Vancouver, Canada.
- Tembe, J., & Norton, B. (2008). Promoting local languages in Ugandan primary schools: The community as stakeholder. *Canadian Modern Language Review, 65*(1), 33–60. <https://doi.org/10.3138/cmlr.65.1.33>
- Tembe, J., & Norton, B. (2011). English education, local languages and community perspectives in Uganda. In H. Coleman, *Dreams and realities: Developing countries and the English language*. London: British Council.
- Tooley, J., & Dixon, P. (2005). *Private education is good for the poor: A study of private schools serving the poor in low-income countries*. Washington, D.C: Cato Institute.
- Trudell, B., & Klaas, A. R. (2010). Distinction, integration and identity: Motivations for local language literacy in Senegalese communities. *International Journal of Educational Development, 30*(2), 121–129.

- Trudell, B., & Piper, B. (2014). Whatever the law says: Language policy implementation and early-grade literacy achievement in Kenya. *Current Issues in Language Planning, 15*(1), 4–21. <https://doi.org/10.1080/14664208.2013.856985>
- Tse, L. (2001). Resisting and reversing language shift: Heritage-language resilience among US native biliterates. *Harvard Educational Review, 71*(4), 676–709.
- Uganda National Examinations Board. (2017). *The achievement of learners in early grade reading in selected districts of Uganda: Midline report national*. Kampala, Uganda: Uganda National Examinations Board.
- UIS. (2019). *UNESCO Institute for Statistics (UIS) database*. Retrieved from <http://data.uis.unesco.org/>
- UNESCO. (2003). *Education in a multilingual world*. Paris: UNESCO.
- UNESCO. (2005). *Mother tongue-based teaching and education for girls: Advocacy brief*. Bangkok: UNESCO.
- UNESCO. (2013). *The global learning crisis*. Paris: UNESCO.
- UNESCO. (2014). *Advocacy toolkit for teachers to provide a quality education*.
- UNESCO. (2015a). *Education for all 2000-2015: Achievements and challenges*. Paris: UNESCO.
- UNESCO. (2015b). *Investing in teachers is investing in learning. Background paper for the Oslo summit on education for development*.
- UNESCO. (2016a). *Education for people and planet: Creating sustainable futures for all*. Paris: UNESCO.
- UNESCO. (2016b). *If you don't understand, how can you learn?* (No. Policy Paper 24). Paris: UNESCO.

- UNESCO. (2018). *SDG 4 data digest. Data to nurture learning*. Retrieved from UNESCO Institute for Statistics website:
<http://uis.unesco.org/sites/default/files/documents/sdg4-data-digest-data-nurture-learning-2018-en.pdf>
- UNESCO Institute for Statistics. (2018). *Handbook on measuring equity in education*. Retrieved from <http://uis.unesco.org/sites/default/files/documents/handbook-measuring-equity-education-2018-en.pdf>
- Unicef. (2019). *Education Briefing Note*. New York: Unicef.
- United Nations. *Universal Declaration of Human Rights*. , (1948).
- United Nations. (2013). *The Millennium Development Goals Report*. New York: United Nations.
- United Nations. (2015). *Draft outcome document of the United Nations summit for the adoption of the post-2015 development agenda*. New York: United Nations.
- United Nations. (2018, April). What are the official languages of the United Nations? - Ask DAG! Retrieved October 19, 2018, from <http://ask.un.org/faq/14463>
- United Nations High Commission for Refugees (UNHCR). (2018). *South Sudan Situation*. Retrieved from <https://data2.unhcr.org/en/documents/download/66950>
- U.S. Department of Health and Human Services, Administration for Children and Families. (2010). *Head start impact study. Final report*. Washington, D.C.
- USAID. (2011). *USAID Education Strategy 2011-2015*. Washington, D.C: U.S. Agency for International Development.
- USAID. (2014). *Evaluation: Tusome baseline study*. Washington, D.C: USAID.

- USAID. (2015). *Report on language of instruction in Senegal*. Washington, D.C:
- USAID.
- Uwezo. (2015). *Are our children learning? The state of education in Kenya in 2015 and beyond*. Nairobi: Twaweza East Africa.
- Uwezo. (2017). *Are our children learning? Lessons from Uwezo learning assessments from 2011 to 2015*. Retrieved from Twaweza East Africa website:
<https://www.twaweza.org/uploads/files/East%20Africa%20Report%202015.pdf>
- Villegas-Reimers, E. (2003). *Teacher professional development: An international review of the literature*. Paris: International Institute for Educational Planning.
- Vogel, S., & Garcia, O. (2017). Translanguaging. In *Oxford Research Encyclopedia of Education*. Retrieved from
<http://education.oxfordre.com/view/10.1093/acrefore/9780190264093.001.0001/acrefore-9780190264093-e-181?print=pdf>
- wa Thiong'o, N. (1986). *Decolonising the mind: The politics of language in African literature*. London: James Currey.
- Walter, S., & Benson, C. (2012). Language policy and medium of instruction in formal education. In B. Spolsky, *The Cambridge handbook of language policy* (pp. 278–300). Cambridge: Cambridge University Press.
- Weiss, M. J., Bloom, H. S., & Brock, T. (2013). *A conceptual framework for studying the sources of variation in program effects*. Retrieved from MDRC website:
http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2393091

- White, I. R., Royston, P., & Wood, A. M. (2011). Multiple imputation using chained equations: Issues and guidance for practice. *Statistics in Medicine*, 30(4), 377–399. <https://doi.org/10.1002/sim.4067>
- World Bank. (2010). *Studies on experimental bilingual education in Senegal*. Washington, D.C: World Bank.
- World Bank. (2014). *SABER country report - Kenya*.
- World Bank. (2018). *World Development Report 2018: Learning to Realize Education's Promise*. <https://doi.org/10.1596/978-1-4648-1096-1>
- Yasunaga, M. (2014). *Non-formal education as a means to meet learning needs of out-of-school children and adolescents* (p. 26). Retrieved from UNESCO Institute for Statistics website: <http://allinschool.org/wp-content/uploads/2015/01/OOSC-2014-Non-formal-education-for-OOSC-final.pdf>
- Yen, W. M., & Fitzpatrick, A. R. (2006). Item response theory. *Educational Measurement*, 4, 111–153.
- Zouhir, A. (2014). Language policy and state in Morocco: The status of Berber. *Digest of Middle East Studies*, 23(1), 37–53. <https://doi.org/10.1111/dome.12039>