



USAID | **ETHIOPIA**
FROM THE AMERICAN PEOPLE



Credit: Read M&E for USAID

MOTHER TONGUE TEACHERS' COMPETENCIES ASSESSMENT (MTTCA)

Reading for Ethiopia's Achievement Developed,
Monitoring and Evaluation (READ M&E)

2019 Report, Revised June 2020

MOTHER TONGUE TEACHERS' COMPETENCIES ASSESSMENT (MTTCA) 2019 REPORT

Reading for Ethiopia's Achievement Developed, Monitoring and Evaluation (Read M&E)

FINAL REPORT

JUNE 2020

American Institutes for Research
1000 Thomas Jefferson Street, NW
Washington, DC 20007

Submitted to:

U.S. Agency for International Development/Ethiopia
Addis Yigzaw, Contracting Officer's Representative

Under Contract No. AID-663-C-15-00001

TABLE OF CONTENTS

1. INTRODUCTION.....	1
1.1 BACKGROUND.....	1
1.2 RATIONALE FOR TEACHER ASSESSMENT IN ETHIOPIA.....	1
1.3 ASSESSMENT FRAMEWORK FOR TEACHERS' COMPETENCIES	5
1.4 PROPOSED ACTIVITIES AND WORK PLAN.....	6
1.5 SCOPE OF THE MTTCA STUDY.....	6
2. MTTCA 2019 DESIGN	8
2.1 PREPARATION AND DEVELOPMENT OF WORKSHOP MATERIALS	8
2.2 ITEM DEVELOPMENT AND ASSEMBLY OF PILOT FORMS.....	8
2.3 SAMPLING PLAN.....	10
2.4 PILOT TEST ADMINISTRATION, DATA ENTRY, AND DATA CLEANING	10
2.5 ITEM-DATA REVIEWS AND ASSEMBLY OF OPERATIONAL FORMS	11
2.6 OPERATIONAL ADMINISTRATION.....	12
2.7 DATA MANAGEMENT AND ANALYSIS	14
2.8 PRODUCTION OF REPORTS	14
3. SETTING PERFORMANCE STANDARDS FOR MOTHER TONGUE TEACHERS' COMPETENCIES ASSESSMENT	15
3.1 PROCESS FOR SETTING PERFORMANCE STANDARDS	15
3.2 SETTING PERFORMANCE LEVELS	17
3.3 SETTING CUT SCORES.....	18
4. THE MTTCA 2019 RESULTS.....	25
4.1 RESULTS FOR MTTCA TOTAL SCORES AND ITS COMPONENTS	25
4.2 TEACHER RESULTS BY PERFORMANCE STANDARDS.....	32
5. RELATIONSHIP BETWEEN TEACHERS' AND STUDENTS' PERFORMANCE	35
5.1 RELATIONSHIP BETWEEN MEAN SCORES ON EGRA AND MTTCA	35
5.2 RELATIONSHIP BETWEEN TEACHERS AND STUDENTS BY PERFORMANCE LEVELS ...	40
5.3 DECIPHERING TEACHER EFFECT	44
6. SUMMARY AND SYNTHESIS OF EXPERT OPINIONS REGARDING FACTORS ASSOCIATED WITH TEACHER AND STUDENT PERFORMANCE	49
6.1 LANGUAGE-FOCUSED DISCUSSION ON MT TEACHERS' COMPETENCIES, STUDENTS' READING OUTCOMES, AND POLICY IMPLICATIONS	49
6.2. EXPERT OPINIONS ON MT TEACHERS' COMPETENCIES AND THEIR RELATIONSHIP WITH STUDENTS' READING OUTCOMES.....	56
7. IMPLICATIONS AND RECOMMENDATIONS.....	64
7.1 KEY RESULTS OF THE MTTCA 2019.....	64
7.2 KEY INSIGHTS ON RELATIONSHIP BETWEEN TEACHERS' AND STUDENTS' PERFORMANCE.....	65
7.3 KEY INSIGHTS ON DECIPHERING THE TEACHER EFFECT IN THE CONTEXT OF HOME AND SCHOOL RESOURCES.....	65
7.4 SUMMARY OF IMPLICATIONS AND RECOMMENDATIONS	66
REFERENCES.....	67
APPENDIX A: SPECIFIC PERFORMANCE LEVEL DESCRIPTORS FOR MOTHER TONGUE TEACHERS	69

APPENDIX B: SUMMARY OF IMPLICATIONS AND RECOMMENDATIONS BASED ON QUALITATIVE DATA.....	80
---	----

FIGURES

Figure 1. Flow of Information in the Teaching–Learning–Assessment Cycle.....	2
Figure 2. Example of an Item Card.....	12
Figure 3. Flow of Activities for Setting Performance Levels.....	16
Figure 4. Flow of Activities for Setting Cut Scores	16
Figure 5. Ordered Item Booklet With Cut Scores	19
Figure 6. Example of Feedback: Agreement Data After Round 2 for Amharic	20
Figure 7. Example of Feedback: Impact Data After Round 2 for Amharic.....	20
Figure 8. Example of Feedback: Cut Score Trends by Rounds for Amharic.....	21
Figure 9. Example of Feedback: Impact Data by Round for Amharic.....	21
Figure 10. Cut Scores for All Languages After Round 4	22
Figure 11. Final Cut Scores for All Languages.....	23
Figure 12. Final Impact Data for All Languages.....	24
Figure 13. Average MTTCA Percent-Correct Total Scores by Language	26
Figure 14. Average MTTCA Percent-Correct Total Scores With Error Bars by Language.....	27
Figure 15. Average MTTCA Sub-Scores (Four Components) by Language	30
Figure 16. Average MTTCA Sub-Scores (Two Components) by Language	31
Figure 17. Percentage of Teachers Reaching MTTCA Performance Levels by Language	33
Figure 18. EGRA 2018 Results for the Oral Reading Fluency Subtask by Zones.....	37
Figure 19. MTTCA 2019 Results on Total Percent-Correct Score by Zones.....	38
Figure 20. Relationship Between Teachers' Performance Levels and Student EGRA Scores.....	41
Figure 21. Profile of EGRA School Scores for Teachers at Different Performance Levels	41
Figure 22. Percentages of Students and Teachers in Performance Levels by Language	42
Figure 23. Percentages of Students and Teachers in the Top Two Performance Levels by Language	43
Figure 24. Strength of Association Between Teacher Performance Level and Student ORF Scores in Different Contexts.....	47
Figure 25. Strength of Association of Student ORF Scores With Teacher Competencies, Home Resources, and School Resources.....	48
Figure 26. Visual Template Used for Focus Group Activity.....	57
Figure 27. Percentage of Teachers and Students That Reached Target Performance Levels	62
Figure 28. Pattern of Teacher and Student Performance	63

TABLES

Table 1. Fidelity and Outcome Indicators at Various Levels of the READ Intervention **Error! Bookmark not defined.**

Table 2. Outline of Phases for Development and Administration of MTTCA.....	6
Table 3. Number of Schools and Teachers Included in the MTTCA 2019, by Region and Zone.....	13
Table 4. General Performance Level Descriptors.....	18
Table 5. Final Cut Scores Adjusted for Maximum Points on the Test.....	23
Table 6. Experts' Interpretation of Differences Between Teacher Performances in Different Languages.....	28
Table 7. Experts' Interpretation of Differences Between Languages in Teacher Performance on Specific MTTCA Components.....	31
Table 8. Experts' Interpretation of Percentages of Teachers Reaching Targeted Performance Levels "Proficient and Advanced".....	34
Table 9. Correlations Between EGRA School Scores and MTTCA School Scores.....	35
Table 10. Experts' Interpretation of Correlations Between Teacher Competency Scores and Student Performance on ORF.....	36
Table 11. Zonal Mean Scores on the EGRA 2018 Oral Reading Fluency (ORF) Subtask.....	38
Table 12. Zonal Mean Scores for the MTTCA 2019 Total Score.....	39
Table 13. Experts' Interpretation of Differences in Performance of Teachers Across the Zones.....	40
Table 14. Experts' Interpretation of Finding 'Better Teachers = Better Students'.....	44
Table 15. Variables from Director Questionnaire Included in ANCOVA and MR.....	46
Table 16. Variables from Student Questionnaire Included in ANCOVA and MR.....	46
Table 17. Identified Factors According to Their Importance by Language Groups.....	57
Table 18. Categories of Factors Affecting Teacher Competencies Organized by Constructs.....	58
Table 19. Language Groups and Their Rating of Factors (F1 = Most Important Factor).....	59
Table 20. Factors Affecting Reading Outcomes by Language Groups Ordered by Importance.....	59
Table 21. Categorical Summary of the Factors Affecting Early Grade Reading Outcomes.....	60
Table 22. Number of Language Groups and Their Rating of Factors (F1 = Most Important Factor).....	61
Table A1. Teachers' Own Knowledge and Skills in the Language They Teach.....	70
Table A2. Subject Matter Pedagogy.....	72
Table A3. General Pedagogy.....	75

ACRONYM LIST

ANOVA	Analysis of Variance
CCA	Continuous Classroom Assessment
CTE	College of Teacher Education
CTT	Classical Test Theory
DAMIAS	Data Management and Item Analysis System
DIF	Differential Item Functioning
EGRA	Early Grade Reading Assessment
FWS	Familiar Words Reading
ILS	Initial Letter Sound
IWR	Invented Words Reading
LNR	Letter Name Recognition
MOE	Ministry of Education (Ethiopia)
MT	Mother Tongue
MTTCA	Mother Tongue Teachers' Competency Assessment
NEAEA	National Education Assessment and Examinations Agency
NLA	National Learning Assessment
ORF	Oral Reading Fluency
PLD	Performance Level Descriptor
PMP	Performance Management Plan
READ M&E	Reading for Ethiopia's Achievement Developed, Monitoring and Evaluation
READ TA	Reading for Ethiopia's Achievement Developed, Technical Assistance
READ II	Reading for Ethiopia's Achievement Developed, Institutional Improvement
READ CO	Reading for Ethiopia's Achievement Developed, Community Outreach
RSEB	Regional State Education Bureau
USAID	United States Agency for International Development

I. INTRODUCTION

This report documents the process of design, development, and implementation of the Mother Tongue Teacher Competency Assessment (MTTCA) that evaluates language skills and pedagogical knowledge and skills of teachers in the Ethiopian primary school system. The report consists of seven chapters that comprehensively address the following: (1) introduction including rationale for assessment of teachers' competencies in Ethiopia, (2) assessment design, including development of assessment frameworks, item writing, assembly of pilot forms, pilot data collection, pilot data analysis, assembly of operational forms, conducting operational data collection in seven languages and five regions, data analysis, and production of reports, (3) setting performance standards, (4) results of teacher performance data, (5) results of associations between MT teacher performance and student reading performance, (6) results of the Co-interpretation®¹ workshop with local experts, and (7) summaries, implications, and pedagogical recommendations.

I.1 BACKGROUND

The **Reading for Ethiopia's Achievement Developed (READ)** suite of projects launched by USAID has a primary goal to improve the performance of reading and writing in mother-tongue (MT) language of children in early grades in Ethiopia. **READ Monitoring and Evaluation (M&E)** was originally conceptualized as a 5-year project with the primary mandate to monitor and evaluate performance of the suite of READ projects.

The purpose of **READ Technical Assistance (TA)** project (2012-2018) is to improve reading and writing in mother-tongue language by supporting the development and implementation of MT curricula for grades 1-8 in seven targeted languages (Amharic, Afaan Oromo, Af Somali, Haddiysa, Sidaamu Afoo, Tigrigna, and Wolayttatto) used as a medium of instruction. READ TA focuses on development of MT instructional materials and professional development of teachers through the in-service programs, as well as supporting the pre-service programs for mother-tongue language teachers through the provision of technological tools to Colleges of Teacher Education (CTEs).

READ Institutional Improvement (II) provides support to the MoE and RSEBs to conduct pre-service teacher training and build the capacity of the education system to improve early grade reading and writing. **READ Community Outreach (CO)** assumes responsibility for building the capacity of parents and communities to engage them in activities that promote early-grade reading and writing, and to support school and community libraries to establish reading centers.

I.2 RATIONALE FOR TEACHER ASSESSMENT IN ETHIOPIA

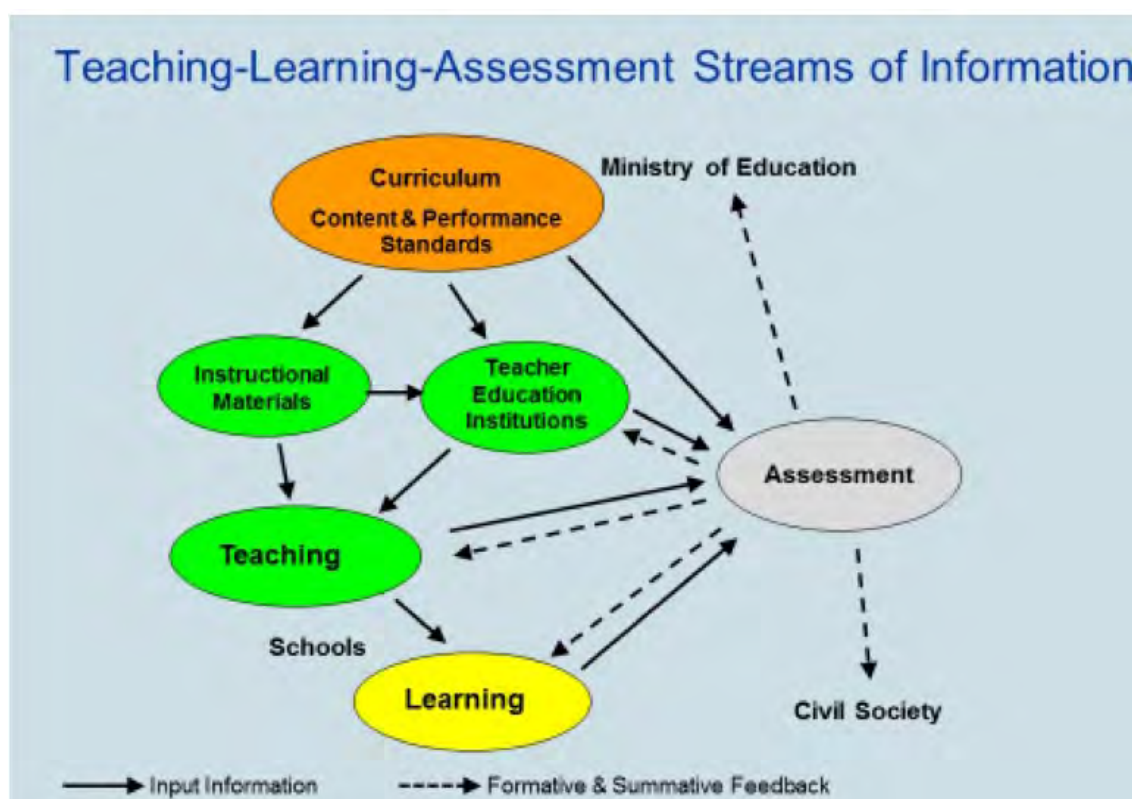
The objective of this section is to describe the rationale and work plan for development of instruments for the *assessment of teachers' competencies* and their application in the context of the READ M&E mandate.

The need to assess teachers' competencies and instructional practices in Ethiopia originates from the findings of the Early Grade Reading Assessment (EGRA) conducted under READ M&E in 2016 and 2018; general trends in international educational practice; and, more recently, a focus on teachers in Ethiopia, as expressed in the Ministry of Education's Ethiopian Educational Roadmap (MOE Ethiopia, 2018). The general flow of information in the **Teaching–Learning–Assessment** informational loop (Figure 1) shows that information stemming from curriculum is transmitted to students mainly

¹ Co-Interpretation is a collaborative process developed by AIR of making meaning out of data. It is a term used to refer to the process by which we co-interpret data and co-develop shared understanding to create meaning. It allows the people closest to the work to bring their perspectives to the interpretation of the data. It is one approach to create a roadmap for change and improvement.

through teachers; thus, teachers represent an important link in the implementation chain of educational interventions focused on improving student learning outcomes.

Figure 1. Flow of Information in the Teaching–Learning–Assessment Cycle



1.2.1. EVALUATION COMPONENTS

Although the original conception of the READ M&E project was to monitor and evaluate all three READ implementing projects, its major focus is centered on the READ TA project. READ M&E approaches its monitoring and evaluation tasks using two types of performance indicators:

1. **FIDELITY OF INTERVENTION INDICATORS:** These evaluate a project's performance against its implementation plan, which involves tracking selected Performance Management Plan (PMP) indicators to evaluate the degree to which project activities are following the original implementation plan, and whether any modifications have been made to adapt to evolving needs and priorities.
2. **OUTCOME INDICATORS:** These measure the performance of beneficiaries that are directly or indirectly targeted by an intervention. READ M&E assesses the following outcome indicators: students' reading fluency and comprehension skills in Grades 2 and 3, as measured by EGRA instruments in seven mother tongue (MT) languages.

Both types of evidence—fidelity of intervention indicators and outcome indicators—need to be *jointly analyzed* to make inferences about the factors that relate to the effectiveness of an intervention. It is also valuable to elaborate the structure of outcome indicators and envision the framework for their integration. READ outcome indicators can be organized into the following three categories, based on their proximity to the reading intervention:

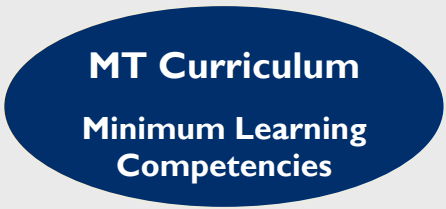
- **IMMEDIATE** outcome indicators: **teachers' competencies** in teaching reading. Because most of the READ project interventions target teachers' skills in using new MT materials to teach

reading, it is important to assess teachers' competencies that are relevant to the intervention. READ M&E proposed to develop assessments for evaluating teachers' competencies and instructional practices.

- **INTERMEDIATE** outcome indicators: *students' reading fluency and reading comprehension skills* in early grades, assessed by EGRA tasks. It is important to note that student improvements in reading are contingent on successful implementation of effective methods in teaching reading—i.e., on their exposure to the effectively implemented intervention. READ M&E proposed to evaluate EGRA results in association with both the *fidelity indicators* and *teachers' competencies*.
- **ULTIMATE** outcome indicators: *student learning outcomes* in multiple subjects in primary and upper grades. Since it is broadly accepted that reading fluency and comprehension are prerequisites for all learning, improved reading skills can be expected to be positively associated with improved learning and corresponding achievement at the ultimate outcome level. For this reason, student performance on the National Learning Assessment (NLA) in Grades 4 and 8 can be also used to evaluate the effectiveness of READ interventions.

Table I links the two types of indicators (*fidelity of intervention indicators* and *outcome indicators*) that can be assessed at each level of the READ intervention. Text highlighted in yellow applies to fidelity indicators, and text highlighted in turquoise applies to outcome indicators. From the table we see that learner outcomes are assessed through EGRA and NLA instruments, while no assessments have yet been identified for measuring teacher outcomes. This study focuses on how READ M&E addressed this lack.

Table I. Fidelity and Outcome Indicators at Various Levels of the READ Intervention

INTERVENTION FOCUS	PRODUCTS/ACTIVITIES AND EVALUATION PERSPECTIVE	
		 Fidelity indicators: The development and implementation of MT curriculum supported.
		 Fidelity indicators: Reading and writing materials appropriate for primary classrooms and pre-service and in-service teacher training developed.
		Pre-service training Fidelity indicators: Program of language-specific teaching and learning methodologies and strategies that focus on helping students learn to read and write effectively implemented. Program for support of language teaching and learning through appropriate technology and teacher aids developed. Technical support to RSEBs and MOE for the READ Institutional Improvement provided. Outcome indicators (immediate): Improved competencies and instructional practices of teachers entering profession.

	Train teachers In-service training Fidelity indicators: Training on language-specific teaching and learning methodologies and strategies that focus on helping students learn to read and write effectively delivered. Support of language teaching and learning through appropriate technology and teacher aids provided. Outcome indicators (immediate): Improved competencies and instructional practices of in-service teachers.
	Outcome indicators (intermediate): Improved student reading and writing skills, assessed by EGRA. Outcome indicators (ultimate): Improved student learning outcomes in all subjects, assessed by NLA.

1.2.2 PROPOSED ACTION ITEMS

READ M&E proposed to design and develop a package of assessments for teachers' competencies and instructional practices in teaching MT reading. The proposed package would include a test of teachers' knowledge and skills in the content they teach and in general and subject pedagogy, as well as an observational tool for assessing teachers' instructional practices. Since READ M&E already designed teacher observational tools, which were used in the READ TA Performance Evaluation study, in this study we focus on assessing MT teacher competencies.

Following the principles of best assessment practices, particularly the utilization of a standards-based interpretational framework, READ M&E also proposed to develop benchmarks for evaluating teachers' competencies by carrying out workshops to set performance standards.

As an optional line of action, READ M&E also recommended developing a data collection system that will enable longitudinal studies associating the following categories of evidence: (1) fidelity of intervention indicators, (2) teacher outcome indicators (measured by a teacher's competency test), (3) student outcome indicators (performance in reading, measured by EGRA), and (4) student outcome indicators (performance in learning, measured by NLA). Integration of these types of evidence into a single research framework would enable more complex analysis to inform policy decisions.

1.2.3 BENEFITS OF ASSESSING TEACHERS' COMPETENCIES

Since formal educational interventions are mainly delivered through teachers, assessing teachers enables the evaluation of the immediate effects of the intervention. If we only assessed student outcomes (the intermediate and ultimate effects of the intervention), we would miss a very important link in the chain of intervention effects.

Tools for assessing teachers would benefit multiple stakeholders, as they could be used for different projects in Ethiopia and, with minor adaptations, in other countries. The MOE would benefit from assessing teachers' competencies, because those assessments would provide insight into the strengths and weaknesses of teachers in different regions, which would inform policy decisions for improving in-service and pre-service teacher training.

1.3 ASSESSMENT FRAMEWORK FOR TEACHERS' COMPETENCIES

The primary objective of the teachers' competency assessment is to provide information on how well primary teachers are prepared to teach MT languages in Ethiopia. Given that the READ TA intervention was intended to affect student reading outcomes through the immediate recipients of the program—the teachers—READ M&E considers it extremely relevant to measure how much information teachers gain from the program, and whether they have all of the prerequisites required to teach reading effectively in the Ethiopian context.

READ M&E aims to acquire information about how teacher outcomes mediate the possible relationship (or lack thereof) between the intervention factors and student reading outcomes. In addition, the data can be used to create performance standards for MT teachers in Ethiopia. Eventually, this assessment of MT teachers may function as a model for measuring the competencies of MT teachers in other low- and middle-income country contexts characterized by limited resources, teachers teaching in languages students may not understand, children learning alphabetic and non-alphabetic scripts, and multilingual environments.

Based on a synthesis of the existing literature on teaching reading competencies (National Reading Panel & National Institute of Child Health and Human Development, 2000; Foorman et al., 2016), teachers' competencies (Roelofs & Sanders, 2007), and reading acquisition in alphasyllabic languages (Nag & Perfetti, 2014; Share & Daniels, 2015), READ M&E designed a plan to develop instruments and measure teachers' competencies derived from the definitions of four broad domains:

1. **KNOWLEDGE OF EARLY LITERACY TEACHING CONTENT:** Refers to what teachers need to know about teaching reading in the early grades in **the context of Ethiopia**. It will focus on teachers' knowledge of what is in a reading curriculum, what the major reading component skills are, their acquisition rates and trajectories for alphasyllabic languages (Fidel) and Roman alphabetic languages, and how these skills may manifest themselves in second (or later acquired) language learners. This domain requires definitions of these skills, an understanding of how to measure these skills, and an understanding of how to teach these skills. READ M&E incorporated the elements from this domain into the assessment component labeled "Subject Matter Pedagogy."
2. **KNOWLEDGE OF GENERAL PEDAGOGY:** Refers to what teachers need to know about general pedagogy (specific to literacy, but likely useful beyond literacy instruction) to nurture engagement with print, as well as instructional routines that are helpful for fostering literacy acquisition (for example, explicit instruction; scaffolding learning; "I do, we do, you do"; activating prior knowledge; and total physical response activities). Based on the elements from this domain, READ M&E created the assessment component labeled "General Pedagogy."
3. **KNOWLEDGE OF THE STUDENTS:** Includes knowledge about individual differences and conditions related to learning to read, such as motivation, nutrition, family background, language of the home, and potential disabilities. READ M&E integrated the elements from this domain into the component "General Pedagogy."
4. **KNOWLEDGE AND SKILLS IN THE LANGUAGE THEY ARE TEACHING:** Refers to teachers' reading comprehension, vocabulary, and grammar in the language in which the teacher is expected to teach. READ M&E included the elements from this domain into the assessment component labeled "Teachers' Language Knowledge and Skills."

I.4 PROPOSED ACTIVITIES AND WORK PLAN

The proposed activities revolve around the **objectives** of the MTTCA as a newly added component of the READ M&E project, defined as follows: to develop assessments of teachers' competencies and instructional practices in teaching reading in primary grades, to evaluate the strengths and weaknesses of MT teachers in seven Ethiopian languages in five regions, with the ultimate purpose to inform remedial activities. The **activities related to each phase** of the MTTCA project component are shown in Table 2, and in the next section we present how we integrated these activities into the project timeline, anchoring them to major milestones.

Table 2. Outline of Phases for Development and Administration of MTTCA

PHASE	DESCRIPTION	OBJECTIVE
1	Preparation	Prepare initial assessment frameworks and workshop materials
2	Item development and assembly of pilot forms	Work with MOE and content experts on item development for assessment of MT teachers in Grades 1–4 - Conduct inception meetings with MOE - Carry out item writing/reviewing workshop, including development of assessment frameworks, blueprints, and item writing specifications - Assemble pilot forms for seven languages - Design a sampling plan for piloting - Pilot administration training and data collection
3	Pilot administration	Pilot data collection by READ M&E Field Office (FO) Remotely support test scoring, data entry, and data cleaning
4	Item-data review and assembly of operational forms	Work with MOE and content experts to construct operational assessments for MT teachers in Grades 1–4 - Carry out pilot data analysis and preparation of materials for item-data reviews - Conduct content/psychometric review of all items - Assemble operational test forms - Review logistics and materials for operational administration
5	Operational administration	Carry out all stages of operational administration - Develop sampling plan - Conduct training of administrators/supervisors (by FO) - Carry out operational administration in regions (by FO) - Remotely support test scoring, data entry, and data cleaning
6	Setting performance level descriptors	Conduct Stage 1 of Setting Performance Standards: Workshop on developing conceptual definitions of performance levels (performance level descriptors, or PLDs)
7	Setting cut scores	Conduct Stage 2 of Setting Performance Standards: Workshop on setting cut scores
8	Reporting	Analyze data and produce reports on findings

I.5 SCOPE OF THE MTTCA STUDY

The Mother Tongue Teachers' Competency Assessment (MTTCA) designed by READ M&E is a tool that measures foundational teachers' competencies that have been identified in the research as aspects of teaching that are predictive of student reading success. It is not intended to be an accountability measure that evaluates individual teachers or school officers. The results should not be used to produce summative reports for individual teachers or schools, and they should not have any impact on teachers' recruitment, status, salary, or promotion.

Instead, MTTCA is designed to serve as a program evaluation instrument, providing insights into characteristics of teacher effectiveness at a systemic level. The assessment has an important formative role and is intended to provide policy and instruction experts with data they can use to design activities that lead to the improvement of teachers' effectiveness and ultimately student learning outcomes.

2. MTTCA 2019 DESIGN

2.1 PREPARATION AND DEVELOPMENT OF WORKSHOP MATERIALS

READ M&E reviewed existing reports on student performance on the EGRA 2018 and the Grade 4 NLA for information about the current status of student achievement (and therefore by implication teacher performance). The possibility of linking existing data (i.e., EGRA and NLA data) with future data (from the teacher assessments being developed) was considered as a way of providing comparative insights. READ M&E also obtained and reviewed all of the curriculum documents, syllabi, and teacher manuals to prepare for meetings and workshops.

2.2 ITEM DEVELOPMENT AND ASSEMBLY OF PILOT FORMS

Prior to the item development workshop, READ M&E conducted a series of meetings with the MOE, the National Education Assessment and Examinations Agency (NEAEA), and the United States Agency for International Development (USAID), with the goal of reviewing, modifying if necessary, and approving the project timeline and work plan. It was important at that point to agree on the shared expectations of who should be involved in each of the phases. READ M&E presented recommendations for ongoing development of assessment capacity in Ethiopia.

2.2.1 DEVELOPMENT OF ASSESSMENT FRAMEWORKS AND BLUEPRINTS

The first two days of the item development workshop were used to introduce the assessment frameworks to the MOE and RSEB officials. During these two days the READ M&E team and MT subject specialists from five regions of Ethiopia (Somali, SNNPR, Amhara, Tigray, and Oromia) worked on developing assessment frameworks to evaluate primary teachers' competencies in teaching reading. The main objective was to construct teacher tests that would measure (1) teachers' knowledge and skills in the language they teach, (2) subject-specific (MT reading) pedagogical knowledge and skills, and (3) general pedagogical knowledge and skills.

The **assessment frameworks** for teacher tests were designed and officially reviewed during the first two days of the workshop. The assessment frameworks for MT teacher tests were then used to construct blueprints and item writing tasks for the item development part of the workshop.

2.2.2 ITEM WRITING/REVIEWING WORKSHOP

In the following three days, the READ M&E technical team and a team of MT experts recruited from teacher education institutions from the five targeted regions carried out an item writing/reviewing workshop to develop items for teacher assessments.

TRAINING OF ITEM WRITERS: This workshop adopted an approach that revolved around extensive training on the principles of writing and reviewing high-quality items. While the participants of item writing workshops for student assessments are typically recruited from teachers, for the task of assessing teachers, it was necessary to recruit item writers from among current or former staff at colleges of teacher education (CTEs) and universities, MOE representatives, curriculum experts, and test development methodologists. The criteria for selecting item writers included experience, achievement, creativity, and diversity.

The training took several days, and was combined with actual item writing and peer-review practice. The training focused on the appropriate alignment of items with assessment frameworks (content standards), coverage of various **cognitive levels** (especially higher order cognitive skills such as

critical thinking and problem solving), gender fairness, and language quality. We made sure that items were based on competencies rather than textbooks. Item writers were also trained to create scoring guidelines for all items (defining correct answers and the associated point scores), including developing scoring rubrics for constructed response items (both short and extended answer item types). We focused on the following quality criteria:

- **Validity:** Alignment with curriculum and standards (link item content with student learning outcomes defined by the curriculum).
- **Reliability and discrimination:** Consistent scores in repeated applications and reliably distinguishing between weak and strong students (write items that are unambiguous, clear, and likely to show variance among students).
- **Objectivity and fairness:** Objectivity as consistency between scores obtained by different scorers (write clear scoring rubric for constructed response items), and fairness as equity between examinee groups (make sure that content of items does not favor or place at a disadvantage any group of test takers on the basis of gender, ethnic background, rural-urban location, or religion).
- **Efficiency:** Extracting as much information as possible from a test item in as little time as possible (make sure that a single test item does not take too long for a test taker to answer) and space (write items that do not occupy too much space on the page).
- **Acceptability:** Item text and graphics must be aligned with the cultural context (use themes that are acceptable regarding moral, political, religious, and other considerations).

READ M&E delivered training to item writers using a variety of materials, including a comprehensive **Item Writing Manual**, to ensure that items met industry standards of the highest technical quality. The manual included extensive examples for most of the quality principles.

ITEM AUTHORIZING: Items were mostly written during item writing workshop sessions, but some were written after the workshops and submitted to the test development team via e-mail (attached as Microsoft Word documents). A more advanced option—an online item acquisition system, functioning as part of a custom-made item banking platform—may be contemplated for future implementation.

CONTENT REVIEW: READ M&E promotes rigorous procedures for reviewing item quality. During item writing, items were peer-reviewed and then submitted to a committee for further review to ensure that they met criteria for alignment with the objective or standard being measured, and that they covered a prescribed range of cognitive complexity (according to the assessment framework). They were also reviewed for discriminatory language and content that might cause item bias, be offensive to a specific group of examinees, or inspire misconduct. Bias and sensitivity were evaluated for gender; in the future this may be extended to other examinee categories such as ethnic and religious background.

2.2.3 ASSEMBLY OF PILOT FORMS WORKSHOP

Following the week of item development, READ M&E worked with a smaller team of key subject specialists to assemble pilot test forms for the MTTCA. These pilot forms were built upon corresponding blueprints, which were equivalent to the blueprints for operational forms. Two pilot forms for each language version of the MTTCA were assembled and prepared for printing. The practical work of assembling the pilot forms was aided by item banking software.

ITEM BANKING: The process of systematically archiving item content and item statistics in either hard copy format (item cards) or electronic format is known as item banking. For MTTCA item development and the assembly of test forms, READ M&E used the electronic item banking platform FastTest, a comprehensive commercial test development software product. The advantages of an electronic item bank lie in the following functions:

- Effective browsing, search, and retrieval of item content and statistics by given criteria
- Assembling test forms by selecting items either manually or semi-automatically
- Evaluating forms assembled from selected items by producing summary tables and graphs

2.3 SAMPLING PLAN

After assembling the pilot forms, READ M&E worked closely with the MOE to develop piloting and operational sampling plans for teacher assessments. READ M&E international consultants provided a technical review and advice on the sampling processes for both piloting and operational administrations. The sampling plan for the MTTCA 2019 followed the structure of the sample for EGRA 2018: the same schools were selected in order to enable comparative analysis of teachers and student performance.

For the endline EGRA 2018 administration in Ethiopia, the study population was defined by the purpose and scope of intended generalizations. The sample size for EGRA 2018 was determined based on power analysis, considering the minimum detectable difference, which was determined in discussion with USAID and the MOE. The sampling method that AIR used in this study was *proportionate stratified cluster sampling*, with regions used as a stratum. The sampling was carried out in two stages: (1) selection of clusters (schools), and (2) selection of students within schools. Control variables for verification of sample representativeness were as follows: school size (based on the enrollment of boys and girls in targeted grades), number of male and female teachers, and teacher-student ratio.

2.4 PILOT TEST ADMINISTRATION, DATA ENTRY, AND DATA CLEANING

This stage encompasses the following activities: supervising data collection for the pilot study, scoring teacher responses on open-ended items using scoring rubrics, and entering and then cleaning data in preparation for analysis. The READ M&E field office played a key role in executing these activities, with remote support from AIR international consultants where necessary.

READ M&E also recognized the need for auditing systems to monitor the quality of standardized testing practices, which play a critical role both in the validity of test score interpretations and in the ongoing collection of validity evidence (Madaus, 1992). During data collection, READ M&E supervisors made random visits to a sample of schools to ensure quality test administration.

The next stage encompasses procedures for the digitization of examinee responses, which can be performed using scanning technology as well as through traditional procedures for manual data entry. The READ M&E project used widely accepted procedures based on scannable answer sheets and accurately digitizing test takers' responses. For pilot administration, the project used traditional procedures (i.e., manual data entry), aided by low-cost data entry software.

READ M&E trained the scoring staff (mainly teacher educators who served as developers) in the manual scoring of open-ended items. This training covered the scoring rubrics and the

implementation of quality control procedures, such as double scoring of items and the use of anchor documents, to ensure that scorers were consistent.

2.5 ITEM-DATA REVIEWS AND ASSEMBLY OF OPERATIONAL FORMS

2.5.1 ITEM ANALYSIS OF PILOT DATA

READ M&E carried out pilot-test data analysis using its proprietary Data Management and Item Analysis System (DAMIAS), which is composed of a suite of SPSS macro programs that can be customized for data analysis of students' and teachers' tests. The system is designed to work in production mode, which means it can process many test forms and items, producing rich statistical outputs in a short time. At the same time, the system enables quality control at multiple stages of data analysis. The statistical output of the system includes classical item statistics such as item difficulty, item discrimination, option and score-point distribution analysis, and differential item functioning (DIF) analysis for selected groups of examinees.

2.5.2 ITEM-DATA REVIEW

Once the pilot data analysis was complete, items were evaluated, considering their content and statistics using item cards. The item cards contained (1) the actual item, (2) information about the item content, (3) item statistics, and (4) DIF and option distribution graphs. Figure 2 shows the content of a typical item card.

READ M&E prepared the item cards for all of the pilot test items and then trained workshop participants on their use in item-data reviews. Each item was reviewed according to industry standard criteria for item quality, taking into consideration both their content characteristics and statistical performance. Each item was categorized as either (1) to be retained as is, (2) to be modified and repiloted at the next round of testing, or (3) to be discarded.

2.5.3 ASSEMBLY OF OPERATIONAL FORMS

READ M&E worked with local subject specialists and test developers on the assembly of operational forms for teacher assessments. The pool of items that was used for the assembly of operational forms came from the pilot administrations that were carried out immediately following the item development workshops. Items were also reviewed for their content and statistical quality.

Operational forms were assembled according to the blueprints developed during the item development workshop. The final operational forms were composed of about 68–70 items, yielding a total of 71–78 score points. The sections on reading comprehension, subject pedagogy, and general pedagogy were largely uniform across languages (items were originally developed in English and then translated into seven MT languages). Because vocabulary and grammar are innately specific to the language structure, the vocabulary and grammar items were developed separately for each language, and their sections were unique to each language, containing 8–10 items.

Figure 2. Example of an Item Card

28. ከሚከተሉት መካከል ደምፀልሣናዊ ግንዛቤን ለማስተማር ተገቢ የሆነው ዘዴ የትኛው ነው?

A. የአንድን ደምጽ ትዕምርት ጥቁር ስሌዳ ላይ መሳፍ

B. ደምጾችን የሚወክሉ ፊደላትን አንዲጽፉ ለተማሪዎች መግባር

C. ስማቸውን አንዲጽፉ ተማሪዎችን መጠየቅ

D. የአንድ ቃል የመነሻ ሆኖ በቃል መጥራት

COMMENTS

Asharic Teacher Competencies Pilot Admin		Domain: Subject Matter: Pedagogy	Year: 2019	Function: 0 N: 32	Accept: ☐
ItemID: AM14S12M1802	Type: M	Strand: n/a	Form: 1		Revise: ☐
TextID: n/a	Key: D	SLO: n/a	Position: 28		Reject: ☐

Options analysis							Parameters		DIF	
Option	A	B	C	D	M	Omit	Difficulty	Discrimination	N Ma	S
Percent:	32	34	33	63		0	P-Value .63	FE-Corr .30	N Fe	.27
Opt FE-Corr:	-.19	-.23	.00	.30		.00	B-Par	A-Par	ETS Cat	
Flag	I-						Flag	Flag	FavGroup	

ItemID: AM14S12M1802 Key: D

Conditional Item Means

2.6 OPERATIONAL ADMINISTRATION

2.6.1 TEST ADMINISTRATION, SCORING, AND DATA CAPTURE

READ M&E prepared the logistics and materials for operational administration. The procedures for oversight of operational administration and guidance for scoring of constructed response items were similar to those for pilot administration; however, data entry and data cleaning were conducted using technology-enhanced procedures based on scannable answer sheets and utilization of image scanners to capture data.

2.6.2 STRUCTURE OF OPERATIONAL SAMPLE

To enable longitudinal design and strengthen the power of analysis, the same schools were sampled for the 2016 EGRA and 2018 EGRA studies (to the degree possible). A total of 198 schools were assessed in both the 2016 midline and 2018 endline administrations, which constitutes 56% of the READ M&E 2018 sample. For the 2019 MTTCA study, READ M&E visited the same schools included in the 2018 EGRA sample. Teacher data were collected in a total of 342 schools that also had EGRA data, which enabled correlational analyses between student and teacher performance.

The list of zones from which READ M&E drew the 2019 MTTCA sample of schools, along with the number of schools in which data were successfully collected and the number of teachers assessed, are presented in Table 3. A map displaying the reading performance of students in assessed zones is provided in Section 5.1.

Table 3. Number of Schools and Teachers Included in the MTTCA 2019, by Region and Zone

REGION	LANGUAGE	ZONE	NO. OF SCHOOLS	NO. OF TEACHERS
Amhara	Amharic	Agew Awi	1	2
		Bahir Dar City	1	2
		Debub Gonder	6	10
		Debub Wollo	11	22
		Misrak Gojjam	10	20
		Semen Showa	8	16
		Semen Wollo	1	2
		W. Gojjam	7	13
		Amhara TOTAL	45	87
Oromia	Afaan Oromo	Arsii	13	25
		Baale	8	14
		Buunnoo Bedellee	2	4
		E. Harargee	7	14
		East Wollega	2	4
		Gujii	8	17
		I Abbaa Boora	3	6
		Naqamtee	2	4
		Shaw Bahaa	5	10
		W. Harargee	9	18
		Wallagaa Lixaa	3	6
		Oromia TOTAL	62	122
SNNPR	Haddiysa	Hadiya	46	87
	Sidamu Affo	Sidama	49	94
	Wolayttatto	Wolayta	44	86
	SNNPR TOTAL	139	267	
Somali	Aff Somali	Faafan	35	68
		Siti	9	20
		Somali TOTAL	44	88
Tigray	Tigrigna	Central	10	20
		Eastern	3	6
		Mekelle	3	6
		North West	18	38
		South East	6	12
		Southern	12	24
		Tigray TOTAL	52	106
Overall TOTAL			342	670

2.7 DATA MANAGEMENT AND ANALYSIS

All of the data analyses were conducted by READ M&E statisticians in Washington, D.C., using the following steps.

1. The SPSS-based Data Management and Item Analysis System (DAMIAS) was used to carry out data management, item analysis under classical test theory (CTT) models, and production of examinee scores necessary for standard setting and further analyses.
2. MTTCA performance was analyzed across languages and zones. In this analysis the statistical method used was analysis of variance (ANOVA).
3. Associations between student performance on EGRA 2018 and teacher performance on MTTCA 2019 were analyzed. These associations were mainly evaluated by coefficient of correlations, ANOVA, analysis of covariance (ANCOVA), and multiple regression analysis.

2.8 PRODUCTION OF REPORTS

READ M&E followed a multistep reporting procedure that included the following: (1) production of draft reports of main analyses; (2) conducting a Co-interpretation® workshop with local experts; (3) integrating the outcomes of the workshops and related qualitative analysis into the report; (4) submitting the report to the client for final feedback and completion. This approach also enables ready-to-action additional analyses and supplemental reports beyond this current scope should such be so desired by the client or other stakeholders.

3. SETTING PERFORMANCE STANDARDS FOR MOTHER TONGUE TEACHERS' COMPETENCIES ASSESSMENT

The MOE's vision for providing assessment-based support for the education of primary teachers in the regions, and especially at the zonal supervisory level, calls for conceptual and operational definitions of performance standards to be set for MT teachers in primary education. The conceptual definitions of performance standards encompass all measurable aspects of pre-service and in-service curricula and are not limited to the language-specific content scope of the MTTCA, whereas the operational definitions (cut scores) are mapped onto the actual assessment scale as defined by actual items included in the test.

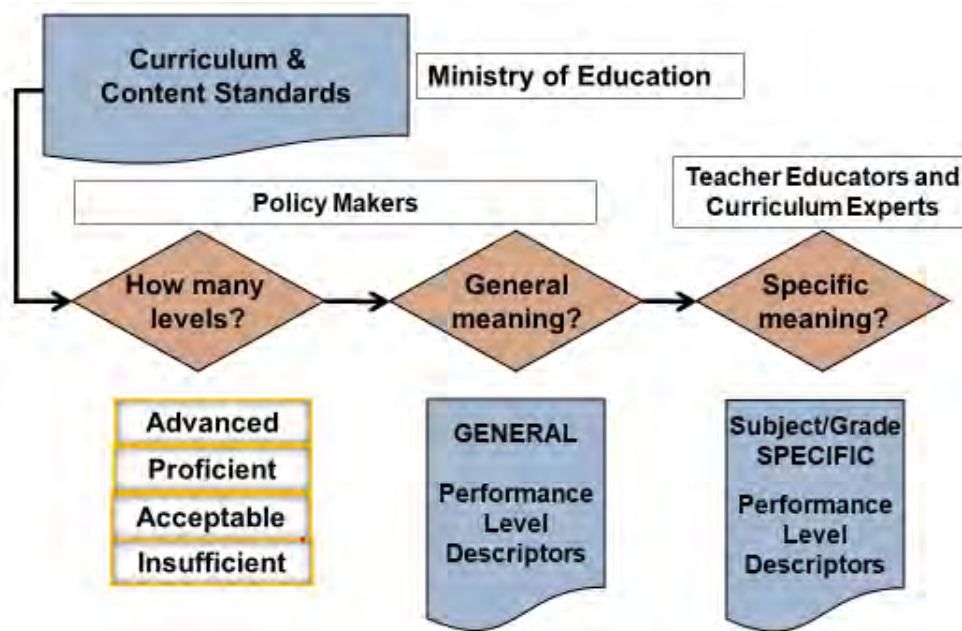
3.1 PROCESS FOR SETTING PERFORMANCE STANDARDS

SETTING PERFORMANCE STANDARDS is a process of developing *conceptual* and *operational* definitions of knowledge, skills, and competencies that teachers should possess at a specific mastery level. Traditional *content standards* are specifications of “what” teachers are expected to know and be able to do in a given area of expertise, whereas *performance standards* refer to “how much” they are expected to know and be able to do to at a particular performance level (e.g., basic, proficient, or advanced).

Conceptual and operational definitions of performance standards are a necessary component of a standards-based assessment system for evaluating teachers' performance in regard to specified achievement criteria. Incorporating performance standards into national assessment systems as descriptions of national learning targets substantially improves the interpretability of assessment results and provides a meaningful framework for monitoring educational progress at the institutional, regional, and national levels. There are **two stages** of setting performance standards: setting performance levels, and setting cut scores between the performance levels.

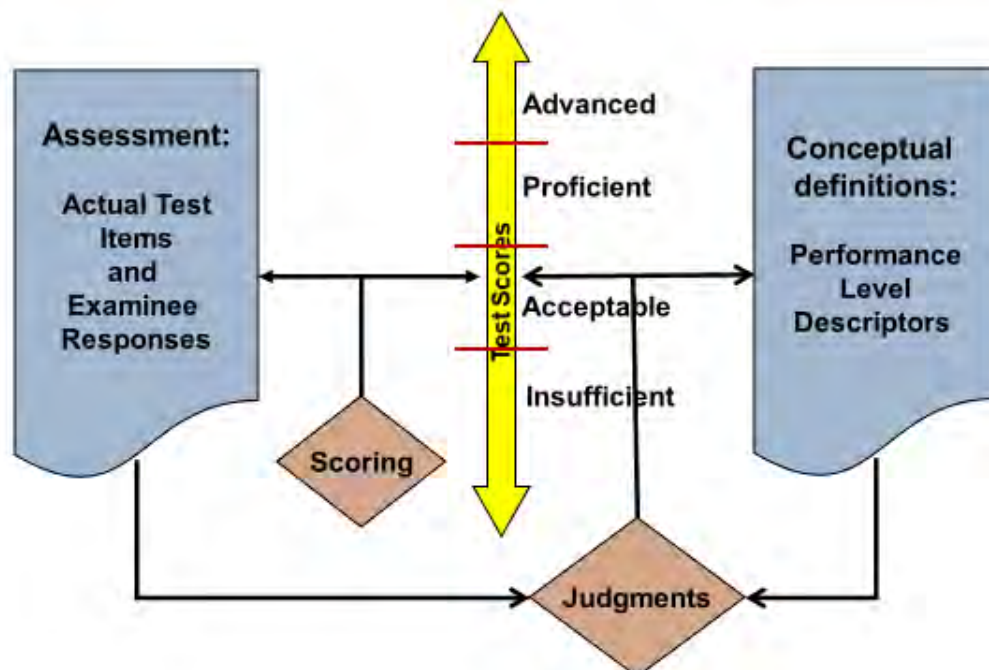
SETTING PERFORMANCE LEVELS is a procedure for conceptualizing performance levels to be used for the evaluation of teacher performance: deciding on the number and purpose of performance levels, choosing their labels, and developing general and specific descriptions of each level (performance level descriptors, or PLDs; for example, what primary MT teachers should know and be able to do in order to be considered “proficient”). This procedure is based on information drawn from official documentation on national learning standards and curricula. The procedure for setting performance levels typically employs a *focus group* method to solicit the opinions and judgments of field experts (e.g., teacher educators, supervisors, curriculum experts, MOE officials). Figure 3 summarizes the process for setting conceptual definitions of performance levels.

Figure 3. Flow of Activities for Setting Performance Levels



SETTING CUT SCORES is a procedure for establishing score ranges on an operational test scale to classify teacher performance into established levels. All methods for setting cut scores rely on obtaining judgments from subject matter experts that lead to the mapping of conceptual definitions of performance levels onto score ranges on actual assessment instruments. Those score ranges represent **operational definitions of performance levels**, and their meaning is maintained across different test forms (sets) within and across administration years through test-equating procedures. Figure 4 summarizes the process of setting cut scores.

Figure 4. Flow of Activities for Setting Cut Scores



3.2 SETTING PERFORMANCE LEVELS

The workshop for developing conceptual definitions of performance levels and PLDs for MT teachers in seven Ethiopian languages was carried out on December 11–15, 2019, in Addis Ababa. There were 58 participants, including professors and instructors from teacher education institutions, supervisors, and curriculum experts. The procedure for developing conceptual definitions of performance levels entailed the following steps.

- Making decisions about the number of levels:
 - In practice, assessment programs usually choose between three and five levels; the experts decided that four performance levels would be appropriate.
- Determining performance level labels:
 - The experts agreed on the following labels: (1) Advanced, (2) Proficient, (3) Acceptable, and (4) Insufficient.
- Developing PLDs—i.e., defining the expected performance of teachers at each level. The PLDs are a detailed specification of what teachers should know and be able to do at each performance level. There are general and specific PLDs:
 - **General PLDs:** Applicable to all primary teachers across subjects, grades, and languages.
 - **Specific PLDs:** Developed specifically for primary MT teachers, these were based on general PLDs as a common framework, and on information drawn from the official “Standards for Mother Tongue teachers grade 1–4,” developed by the MOE.

Developing the PLDs involved focus group sessions with policymakers and teacher educators representing all seven MT languages. The following is a simplified outline of the steps used to develop the specific PLDs:

- Plenary session: A discussion of what teachers should know and be able to do at each performance level.
 - Language groups: Participants were divided into seven language groups of six to eight participants in each group for individual writing of PLDs.
 - Language groups: Groups discussed individual inputs and created the group definitions of performance levels.
 - Plenary group: Table leaders presented their group results, followed by discussion and formulation of the proposed PLDs.
 - Facilitators wrote up the final PLDs in a document.
- Reviewing PLDs at the policy level (MOE and other stakeholders), leading to final approval of the conceptual definitions of MT teachers’ performance levels.

The teacher PLDs were created uniformly for all languages. Table 4 shows the **general** PLDs developed and agreed upon at the workshop, which are applicable to primary teachers in any grade or subject. The complete set of **specific** PLDs for MT teachers can be found in Appendix A. It is important to note that the development of PLDs was not restricted to the attributes that were measured in the current instruments; rather, they capture all relevant aspects of teachers’ performance deemed relevant by experts.

Table 4. General Performance Level Descriptors

INSUFFICIENT	ACCEPTABLE	PROFICIENT	ADVANCED
Teachers at this level demonstrate limited understanding of subject matter and inadequate application of pedagogical skills required for the grade level. They have low motivation and exhibit poor interpersonal skills.	Teachers at this level demonstrate partial understanding of subject matter and functional pedagogical skills required for the grade level. They show a moderate level of motivation and exhibit basic interpersonal skills.	Teachers at this level demonstrate solid understanding of subject matter. They demonstrate effective application of pedagogical skills required for the grade level and show some creativity in their implementation. They are highly motivated and exhibit effective interpersonal skills.	Teachers at this level demonstrate a comprehensive and in-depth understanding of subject matter. They demonstrate highly effective application of pedagogical skills required for the grade level and they often implement them creatively. They are exceptionally motivated, conducting research and exhibiting leadership qualities with highly effective interpersonal skills.

3.3 SETTING CUT SCORES

The workshop for setting cut scores was carried out on December 16–20, 2019, in Bishoftu. The structure of the panels was the same as for the Stage I workshop on the conceptual setting of performance levels, involving about 75% of the same experts. The setting of cut scores is a procedure for establishing score ranges on the operational test scale for the classification of teachers' test achievement into predefined levels. This procedure systematically solicits, records, and analyzes experts' judgments, drawing on their understanding of the performance level definitions, experience with the assessment (most of them participated in the development of MTTCA), and experience in evaluating teacher performance.

The process of setting cut scores is an iterative method that typically takes three to four rounds of judgments. After each round, feedback is presented to the experts to help them refine each judgment task. Feedback consists of two types of data:

- **Agreement data:** The degree to which judgments of different experts are homogeneous.
- **Impact data:** The percentage of examinees who would be classified in each performance level based on the cut scores proposed in a particular round.

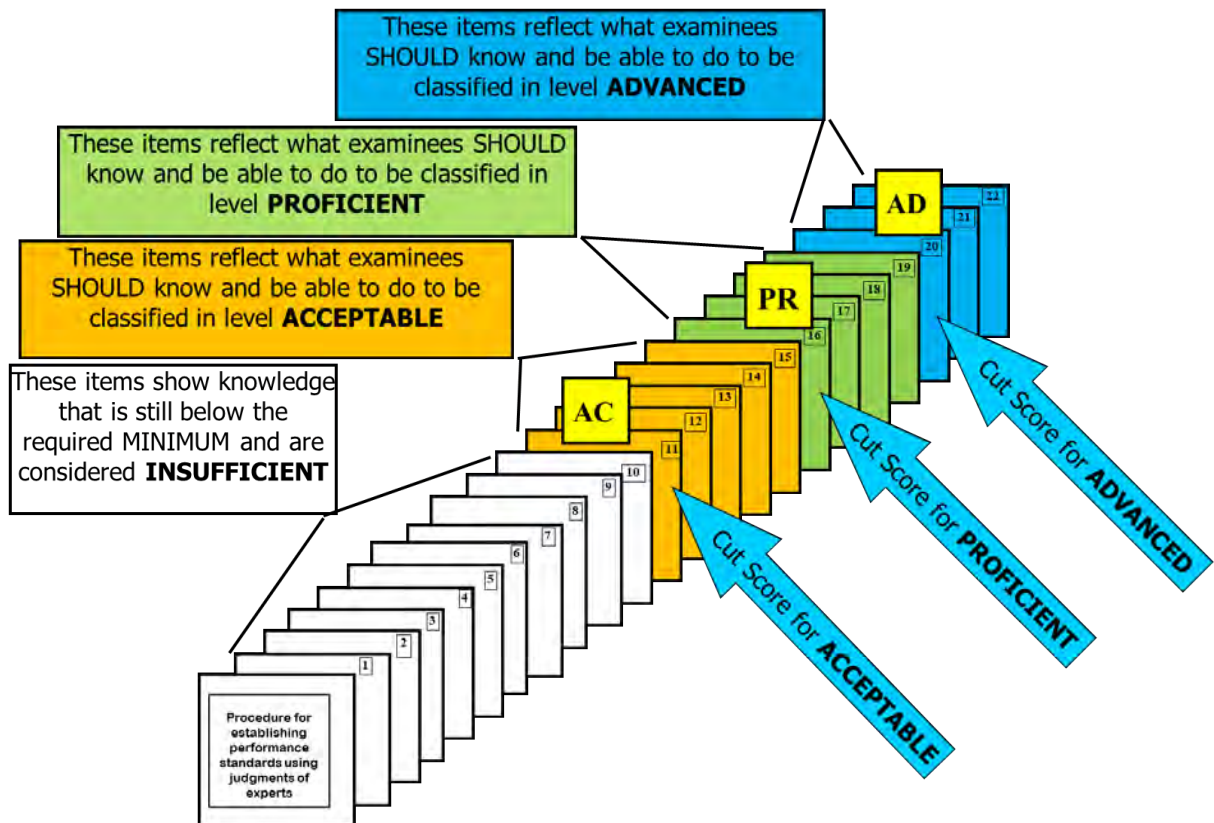
After receiving feedback, the expert groups convened discussions about the reasonableness of the agreement and impact data. Evaluation of their judgments was based on the following points:

- Does variation among the judgments of experts come from a common understanding of the standards or is it due to a difference in understanding of the standards?
- Are the percentages of examinees in performance levels consistent with the general experience of experts in teaching and evaluating primary school teachers?
- Are the cut scores and percentages of examinees of the MTTCA in performance levels showing a reasonable pattern across languages?

The discussion was followed by another round of judgments in which the panelists had another chance to adjust their judgments based on discussion and feedback. The rounds were carried out until agreement about cut scores was reached, which typically took three to four rounds.

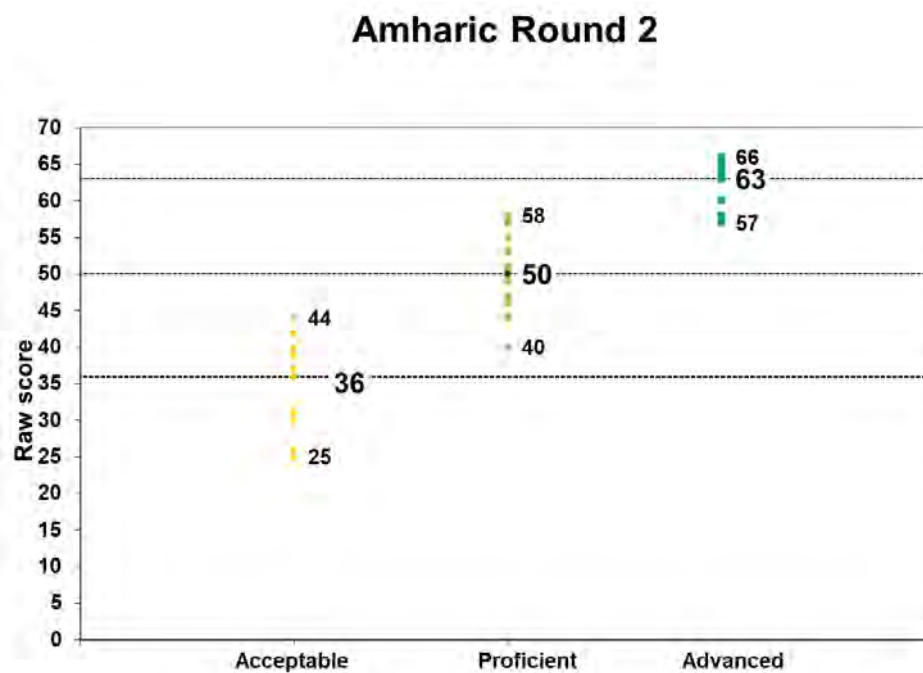
The methodology for setting cut scores for the MTTCA followed the Bookmark method (Lewis, Mitzel, & Green, 1996). In this approach, items in a test are ordered by empirical difficulty, from easiest to hardest. Panelists place a “bookmark” in the “ordered item booklet” where they estimate that an examinee at the threshold of a performance level would be expected to respond successfully to the items prior to the bookmark, and unsuccessfully to the items that follow, taking into consideration the specific performance level descriptors. The ranges of items between bookmarks represent operational definitions of performance levels. Figure 5 depicts how the ordered item booklet looked after identifying the cut scores.

Figure 5. Ordered Item Booklet With Cut Scores



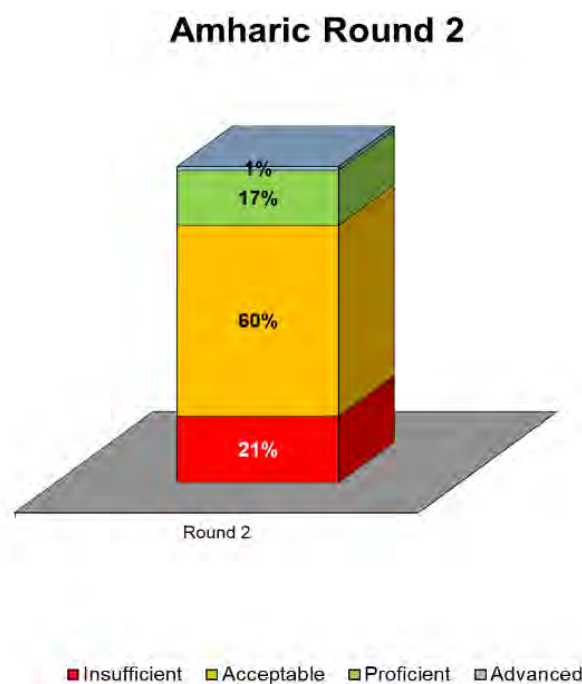
After the panelists had submitted their individual cut scores, agreement data and impact data were presented after each round. Agreement data show the spread of the individual cut scores presented by each judge, which is helpful during discussions as it helps panelists make more homogeneous judgments in future rounds. Figure 6 shows an example of agreement data for Amharic after the second round of judgments. The spread of dots represents individual judgments; the numbers indicate group minimum, maximum, and median for each point of judgments. Note that in the case shown in Figure 6, the proposed cut scores were 36 for Acceptable, 50 for Proficient, and 63 for Advanced after round 2.

Figure 6. Example of Feedback: Agreement Data After Round 2 for Amharic



After showing the proposed cut scores for the round, impact data were presented to help panelists understand how assessed teachers would be distributed among the four performance levels using the cut scores they had proposed. Impact data are represented as the percentage of teachers classified within each performance level. Figure 7 shows the impact data for round 2 for Amharic. The data show that, if the cut scores shown in Figure 6 were used, the performance of 21% of teachers would be “Insufficient,” and 60% would demonstrate “Acceptable” performance, 17% “Proficient” performance, and 1% “Advanced” performance.

Figure 7. Example of Feedback: Impact Data After Round 2 for Amharic



After rounds 2, 3, and 4, additional feedback data were presented: **cut score trends across rounds** and **impact data across rounds**. This type of data provides information about how the cut scores varied after each round. Figure 8 shows the trends of the proposed cut scores for each of the four rounds of judgments in the Amharic group. It shows that the cut score spreads decreased from round 1 to round 2, decreased their overall level from round 2 to round 3, and stabilized from round 3 to round 4.

Figure 8. Example of Feedback: Cut Score Trends by Rounds for Amha

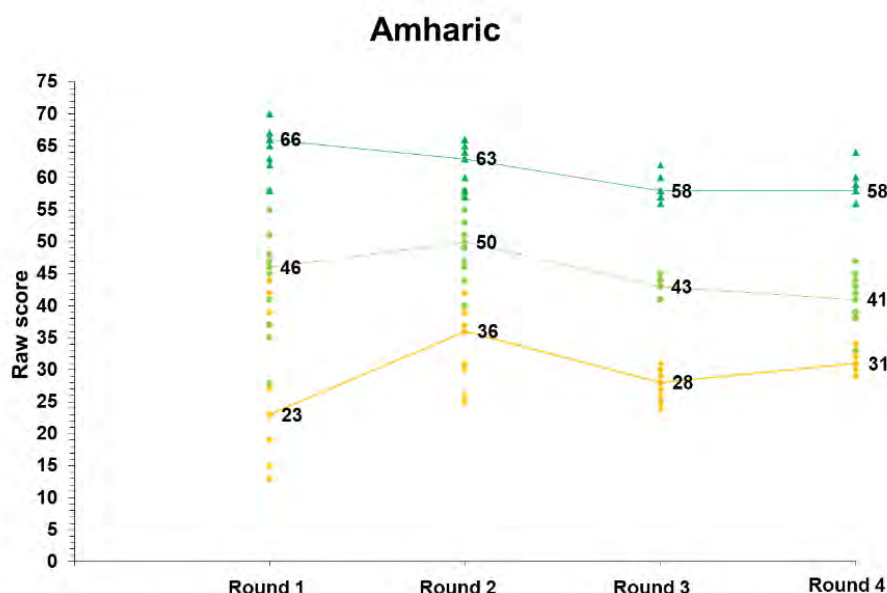
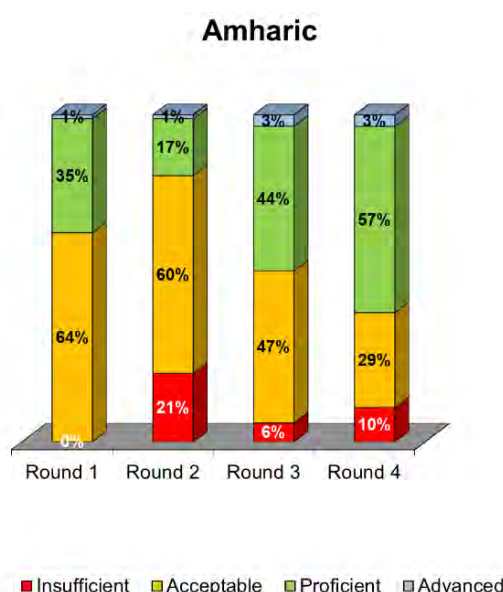


Figure 9 shows the trends for impact data by round for Amharic. These data allow panelists to evaluate whether the distribution of examinees in each performance level is reasonable, based on the panelists' experience in educating and evaluating primary school teachers, and comparing this experience with the teacher performance expectations defined in specific performance level descriptors.

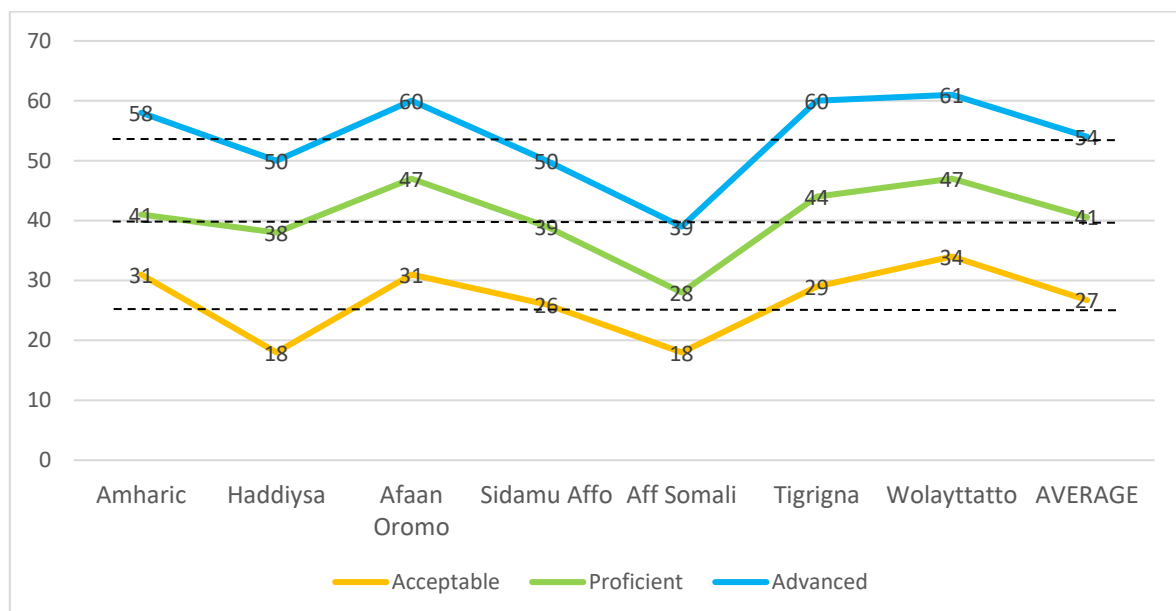
Figure 9. Example of Feedback: Impact Data by Round for Amharic



In addition to the feedback for their own language group, panelists received information about the cut scores and impact data for other language groups. This information was intended to provide insight into how the cut scores were estimated for other languages, facilitating convergence between performance standards across different languages. Because the experts agreed that teacher performance standards should be equivalent among languages, the same PLDs (developed in the Stage I workshop) were used as a reference for all seven languages in the workshop on setting cut scores. For this reason, it was expected that cut scores would reflect only minor differences in test difficulty.

However, despite extensive discussion about the need for uniformity across teacher performance standards, and despite presenting comparative feedback across languages, the language groups of experts varied in their proposed cut score levels. This was likely a result of variable teacher performance across languages and panelists' varied responses to impact data as they tried to adjust the cut scores to reflect the performance of teachers in their respective language groups. The cut scores obtained at the end of round 4 are presented in Figure 10.

Figure 10. Cut Scores for All Languages After Round 4



Because of the considerable differences between the language groups, the process for setting cut scores for the MTTCA required four rounds of judgments plus an additional moderation session. The final moderation session was conducted to fine-tune the cut scores. This session used a social and statistical moderation methodology. It was agreed that teacher performance cut scores should be determined based on the average judgments of all seven language groups from round 4.

Additional statistical adjustment was performed based on the maximum possible point score on each language form. In other words, the cut scores would use the average percent-correct score across languages, which after round 4 equaled 36 percent-correct points for Acceptable, 54 for Proficient, and 72 for Advanced. In the final moderation session, the percent-correct cut scores were set to 35 for Acceptable, 55 for Proficient, and 75 for Advanced.

Table 5. Final Cut Scores Adjusted for Maximum Points on the Test

LANGUAGE	PERFORMANCE LEVEL			MAX POINTS
	ACCEPTABLE	PROFICIENT	ADVANCED	
Amharic	27	42	57	76
Haddiysa	25	39	53	71
Afaan Oromo	27	43	59	78
Sidamu Affo	27	42	57	76
Aff Somali	25	40	54	72
Tigrigna	27	42	57	76
Wolayttatto	25	40	54	72
Pct-Correct Cut Score	35%	55%	75%	

Table 5 and Figure 11 show the final cut scores (after the final moderation session), and Figure 12 depicts the final impact data. The cut scores set during this workshop and the corresponding performance results represent the recommendations of the expert group that need to be submitted to the MOE for review and approval, after which they can become official.

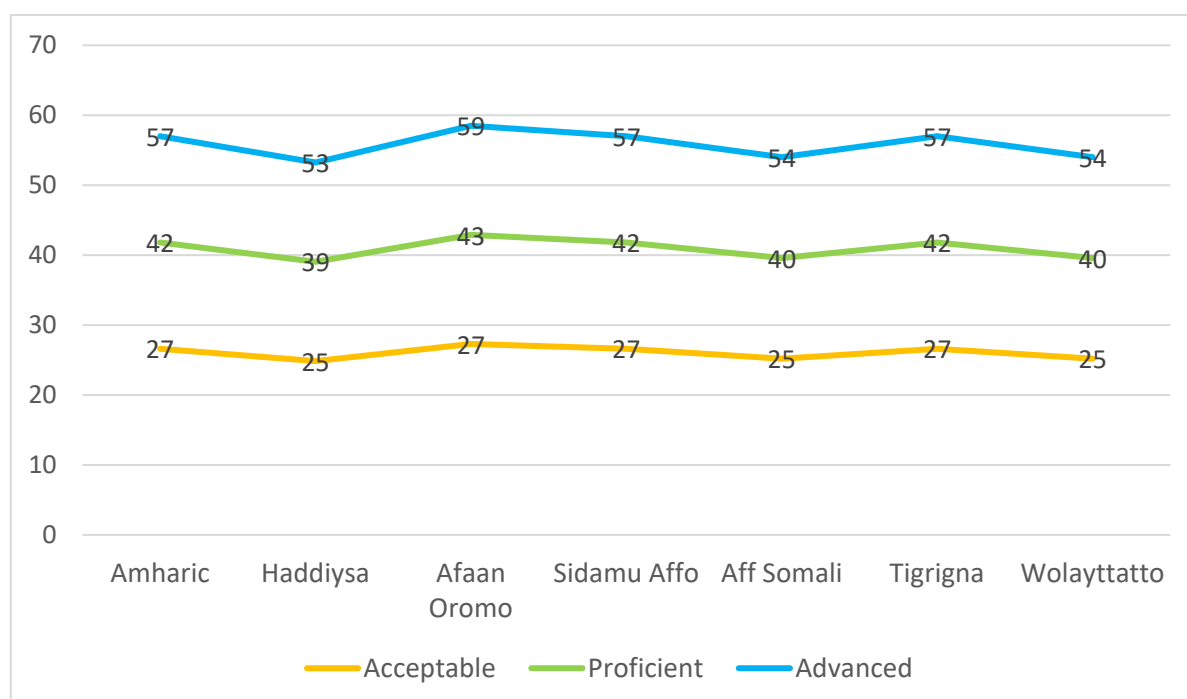
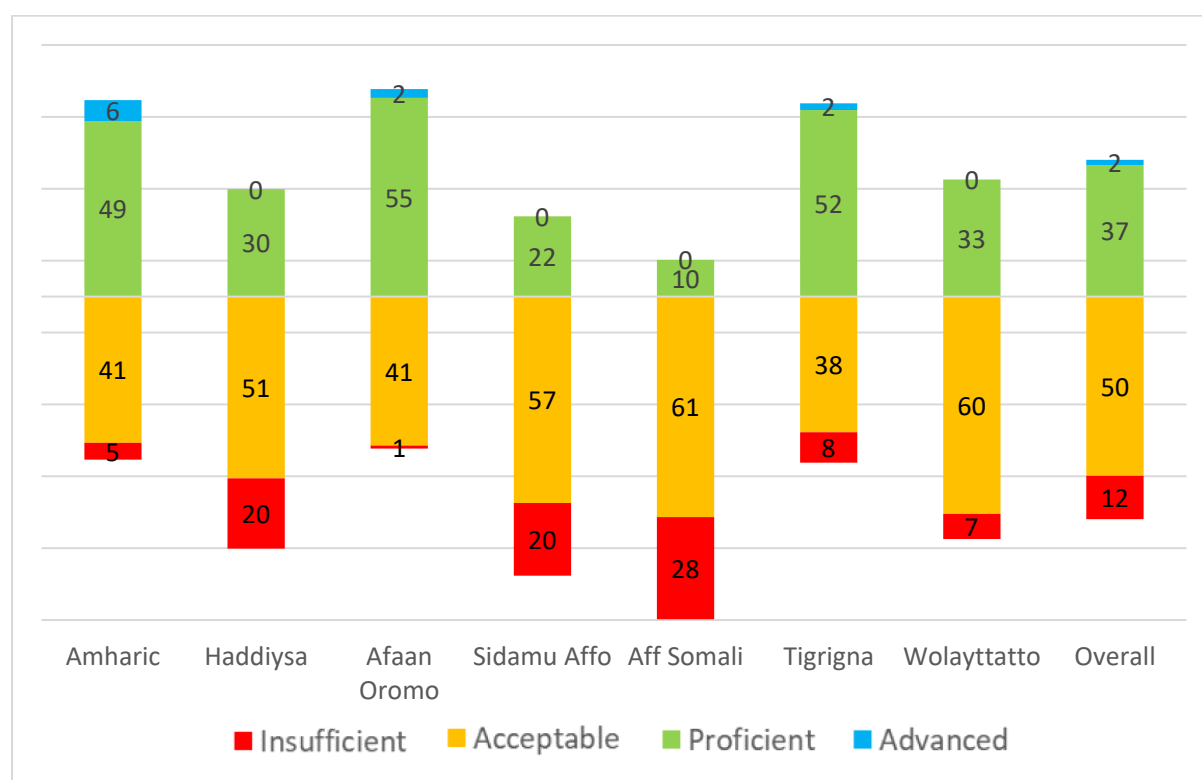
Figure 11. Final Cut Scores for All Languages

Figure 12. Final Impact Data for All Languages



4. THE MTTCA 2019 RESULTS

In this chapter we discuss the major results of the MTTCA 2019 administration, which we present using two types of reporting frameworks. The first framework uses percent-correct scores, which are relative to the maximum possible points on the MTTCA in each of the assessed languages (for the total score and for the sub-scores). Although the percent-correct score units are straightforward, transparent, and easy to understand, they do not convey enough information about the value of the results because they are affected by difficulty of the test (easy tests generate high scores and hard tests generate low scores) and the content referred to by percent-correct scores is not known. Similarly, percent-correct scores have limited comparability across languages because the difficulty of one test compared to another may be due to any of a number of reasons, including translation factors and differences in language difficulty itself affecting teacher performance on comprehension, vocabulary, and grammar measures.

The second reporting framework is based on performance standards, which addresses the limitations of the first framework by providing a means for categorizing teacher achievement into meaningful and readily interpretable performance levels. These levels are comparable across languages because they are determined based on the common conceptual definitions of teacher performance standards, whereas their operational definitions, in terms of cut scores on instruments, are developed for each language taking into consideration language specificities. The measure used in this second reporting framework is the percentage of teachers whose test achievement falls within each performance level (Insufficient, Acceptable, Proficient, and Advanced), according to the established performance standards.

Interpretations of the results presented in this and following sections are informed by discussions led at the Co-interpretation® workshop involving Ethiopian educational and assessment experts from the regions where the MTTCA was administered. The workshop took place in Adama on February 25–27, 2020, and the workshop outputs were further systemized and processed by READ M&E experts and consultants. Considering that interpretational inputs from experts in different language teams are likely to be relevant to regional policy and practice implications, we presented the experts' comments separately by language groups. This necessarily means that there are some repetitions when the same or similar interpretations are offered by teams from different languages, but preserving the regional origin of interpretations provides value for local stakeholders and policymakers.

4.1 RESULTS FOR MTTCA TOTAL SCORES AND ITS COMPONENTS

Teacher achievement on the 2019 MTTCA administration is first presented as average percent-correct scores for each of the seven languages. Results are presented for the total score on the MTTCA and for the sub-scores.

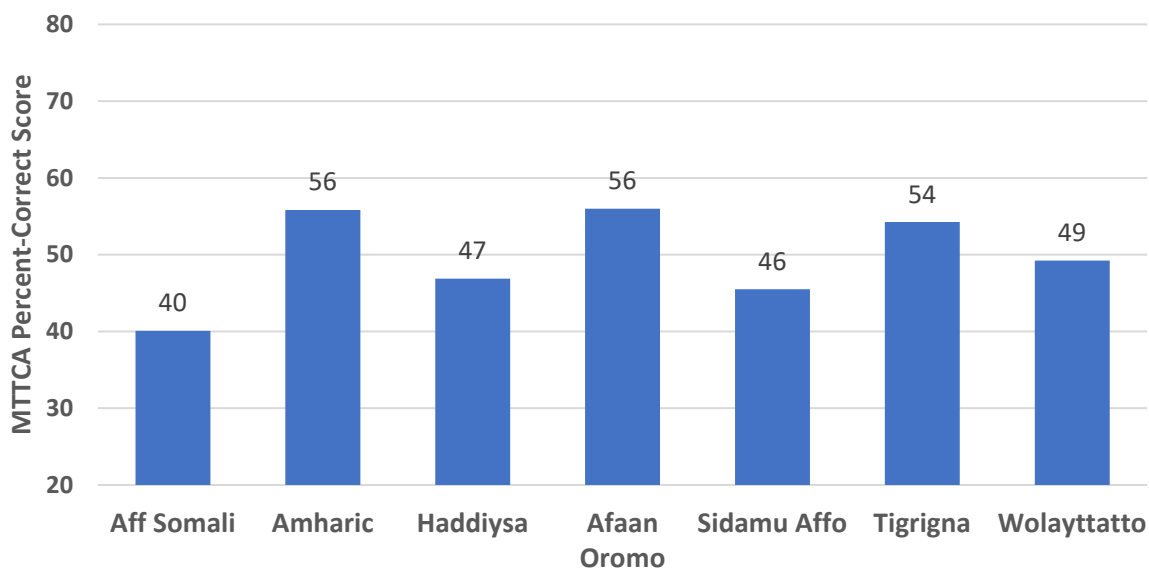
4.1.1 MTTCA TOTAL SCORES BY LANGUAGE

The means of the total percent-correct scores, as an aggregated overall performance on the competencies test for mother tongue teachers in different languages, are presented in Figure 13. The mean scores range between 56 points for Amharic and Afaan Oromo and 40 points for Aff Somali.

Statistical significance of the differences between languages was evaluated by one-way ANOVA followed by a post-hoc analysis of individual differences between pairs of languages. The overall ANOVA indicated that the variation between languages is statistically significant at the level of probability that the conclusion could be wrong (Type I error) of less than 0.1% ($p < 0.001$).

Post-hoc analysis of individual differences between languages revealed that the performance of teachers in seven assessed languages can be categorized into three relatively homogenous groups. At the top are the Amharic, Afaan Oromo, and Tigrigna teachers, in the middle group are teachers of Haddiysa, Sidamu Affo, and Wolayttatto, and in the low group are teachers of Aff Somali. The differences between averages within each of these three groups are not statistically significant, whereas differences between averages across these three groups are statistically significant.

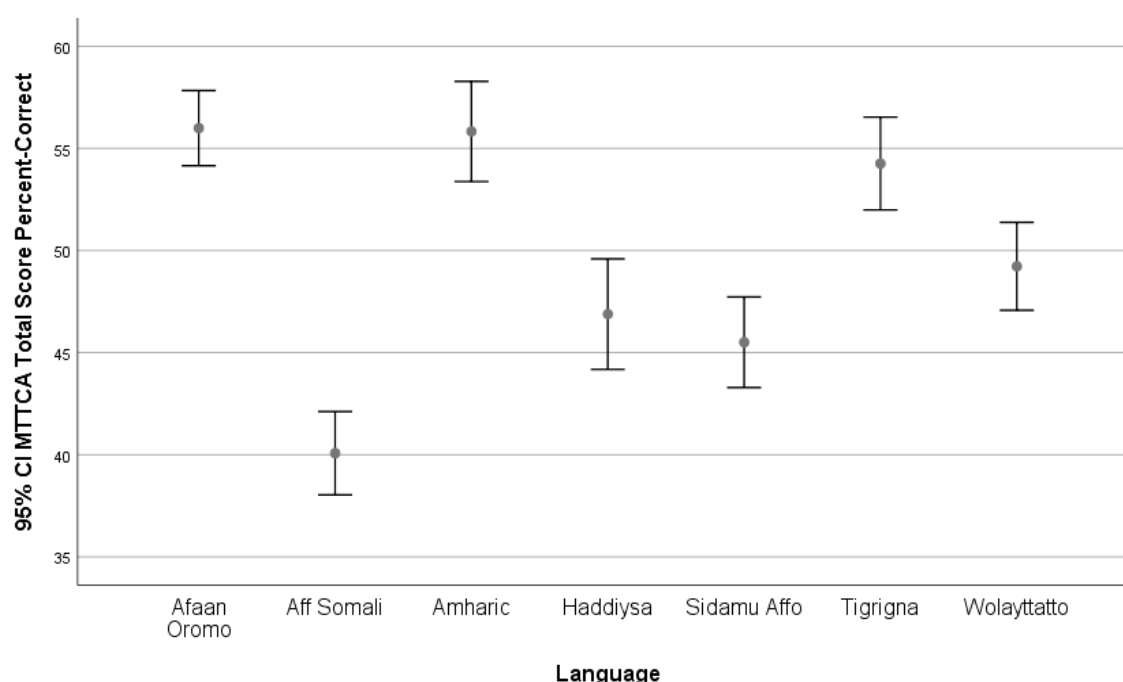
Figure 13. Average MTTCA Percent-Correct Total Scores by Language



Throughout this report, the size of the difference between two group means is expressed in terms of Cohen's d , which is commonly used to complement tests of statistical significance and interpreted as practical significance or "effect size." It expresses the size of the observed difference as a fraction of the pooled standard deviation. Cohen (1977) defined effect sizes as "small, $d = 0.2$," "medium, $d = 0.5$," and "large, $d = 0.8$." Wolf (1986) offered an educationally referenced interpretation: $0.25 =$ educationally significant (something was learned), $0.50 =$ practically/clinically significant (something substantially changed).

The difference in MTTCA mean scores between the highest and the lowest group is $d = 1.30$, which falls in the category of extremely large effect size, suggesting that there is a big performance gap between teachers coming from high-performing regions speaking Amharic and Oromo compared to Somali regions. The differences between the teachers' performance in seven MT languages are also depicted by error bars in Figure 14. The whiskers of error bars represent the confidence interval for each mean score—if the whiskers of the error bars overlap there is no significant difference, and when the whiskers do not overlap it indicates that the differences are statistically significant. The error bars clearly depict the three groups of performance on the MTTCA.

Figure 14. Average MTTCA Percent-Correct Total Scores With Error Bars by Language



Based on the results presented above, the following observations have important policy implications:

1. Overall, mother tongue teachers' performance on the competencies assessment is relatively low (barely 50 percent-correct points aggregated across all languages).
2. There are large differences between performance of teachers in different languages.
3. Three significantly different groups of languages can be identified based on the performance of MT teachers:
 - Amharic, Afaan Oromo, and Tigrigna teachers exhibit average performance around 55 percent-correct points.
 - Haddiysa, Sidamu Affo, and Wolayttatto teachers perform on average between 45 and 50 percent-correct points.
 - Aff Somali teachers exhibit an average of 40 percent-correct scores on MTTCA.

During the Co-interpretation® workshop, the team of Ethiopian experts addressed the following questions: How does the finding about differences between teacher performance in different languages fit into common experiences of the experts? What are possible explanations for these differences? What are policy implications and recommendations for actions?

Because the experts come from different regions and have insights into different background factors, their responses are clustered based on their source (language group), presented in Table 6. Please note that we have attempted to reflect in English as closely as possible what we have assumed was the intended meaning of experts.

Table 6. Experts' Interpretation of Differences Between Teacher Performances in Different Languages

EXPERT GROUP	RESPONSES
Afaan Oromo	0 Lack of reference materials in some areas more than others 0 The difference between languages in terms of the time the languages have been used for instruction for the first time 0 The prevailing misconception that knowing the language would suffice for teaching the language 0 Lack of uniformity in the training of mother tongue teachers and the frequent change in the mode of the training across regions
Aff Somali	0 Difference between languages in terms of the availability of qualified teachers in mother tongue languages 0 Mismatch between what teachers teach and their area of specialization is larger in some regions than others 0 Absence of uniformity and difference in the quality of training given to mother tongue teachers across Colleges of Teacher Education (CTEs) 0 Difference among teachers in viewing teaching as desirable profession 0 Difference among regions in terms of qualification as well as experience of mother tongue teachers 0 Reluctance of teachers to put into practice what they have learned from the trainings
Amharic	0 Difference in teachers' competencies across languages may be due to different use of the languages in mass media 0 Differences in the languages regarding the availability of government support for improving teachers' knowledge and skills 0 Difference in turnover of educational experts who participate in mother tongue language related trainings and workshops 0 Educational leadership differences among languages 0 Difference in the quality of would-be teachers and quality of training, and differences in the use of assessment methods and procedures (including continuous assessment) employed in the CTEs across languages 0 The group expected performance of Amharic teachers to be better than performance of teachers in other languages because the language is relatively well developed, has been used as medium of instruction for many years, and has several reference books 0 However, performance was the same as other languages, probably due to selection and recruitment issues for the pre-service training, implementation issues for continuous assessment in CTEs, and negative attitudes toward teaching mother tongue
Haddiysa	0 Difference in teaching experience (in some languages, teachers are newly deployed, and others have experienced teachers) 0 Teachers' commitment/readiness to improve their profession varies across language groups due to teachers' engagement in multiple activities, including business 0 Different teacher educational background/training in mother tongues across regions 0 Differences in proportion of teachers who did not graduate from CTEs engaged in teaching mother tongue languages 0 Difference in factors affecting teachers' performances, such as distance to the school, conflict, draught, famine, etc. 0 Difference in the selection, recruitment, and management of teachers 0 Differing alignment between the training teachers received and what they are expected to teach 0 Imbalance between subject pedagogy and general pedagogy in the CTE curriculum—focus on general pedagogy at the expense of subject pedagogy 0 Difference in the use of motivational strategies for teachers across languages 0 Differences in the provision of supportive supervision and feedback to teachers 0 Differing availability of mentors across languages 0 Difference in instructional leadership 0 Differing levels of teacher interest in teaching mother tongue 0 Difference in terms of the language used in the community and in the classroom

EXPERT GROUP	RESPONSES
Sidamu Affo	- o There are no adequate subject matter pedagogy courses in the CTEs in Sidama region and would-be teachers learn general pedagogy in English - o Excessive and boring repetition of content in primary textbooks in Sidama region - o Differences in recruitment process for pre-service training; lack of mechanisms for checking the authenticity of credentials for those who apply for the pre-service training in search of job - o Difference in the interest and commitment of teachers - o Differences in the quality of CTE instructors who teach would-be teachers - o Differences in the assessment mechanism employed by CTEs: reluctance, poor monitoring and supervision mechanism, and negotiation on grading are more common in some CTEs - o Differences among languages in terms of the availability of reference materials in the CTEs' libraries - o Difference in the level of motivation and professional commitment of teachers - o Difference in languages in terms of the proportion of teachers who leave the profession - o Mismatch between what teachers teach and their area of specialization - o Difference in the availability of short-term training for teachers, cascading training problem - o Difference in languages in terms of the follow-up and supervision mechanisms employed by educational leaders - o Difference in languages in terms of the availability and adequacy of textbooks and supportive materials in schools
Tigrigna	- o In some languages there is a lack of responsibility and accountability, especially during in-service training - o In some languages there is a lack of professionalism - o Difference in teachers' educational background; some mother tongue teachers may not be graduates of formal education - o Regions differ in terms of follow-up and monitoring mechanisms - o Differences between languages in the proportion of teachers graduated with mother tongue.
Wolayttatto	- o Linguistic differences (the use of Fidel & Latin scripts) - o Differences in teachers' commitment to improving their profession - o Differences among languages in terms of availability of reference materials for teachers - o Political and educational leaders across languages differ in the effort they exert in developing mother tongue languages - o Differences in teachers' selection for pre-service training; prospective teachers are selected for MT pre-service training without having interest. - o Differences in teachers' qualification (being graduate and non-graduate)—there are teachers who teach a mother tongue without having language background; even teachers with language background are not always successful, due to inadequate pedagogical knowledge about teaching MT - o Differences in teachers' commitment to improving their profession - o Languages differ in their M&E systems and instructional leadership support for teachers: There is no strong and sustainable monitoring and evaluation system at different levels of education sector. If teachers get enough support from instructional leader at school level, they will improve their skill and knowledge.

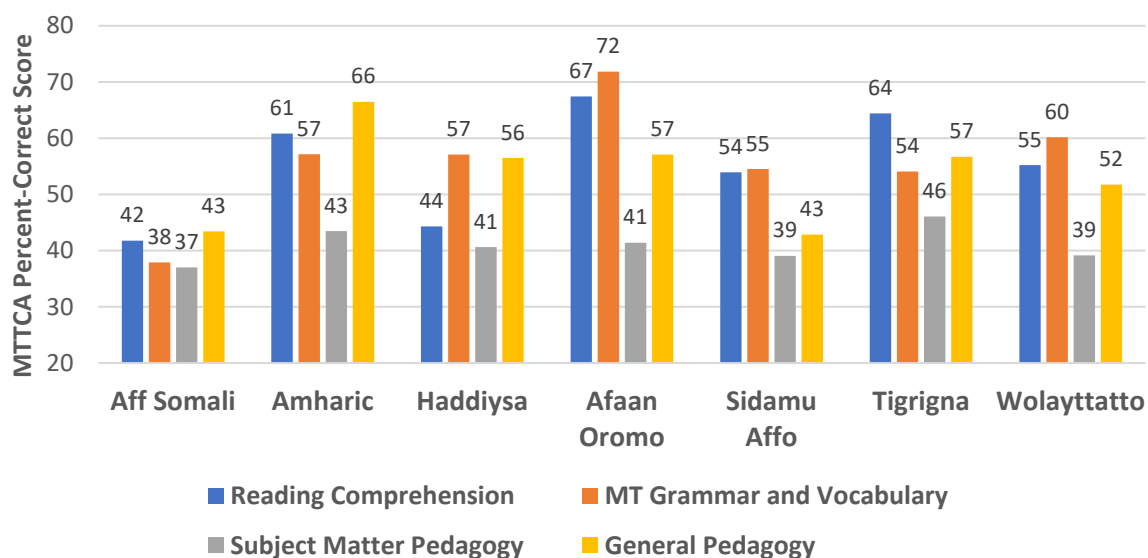
To summarize, most typical interpretation of differences in teacher competencies among languages are pointing to differences in time the languages were introduced as language of instruction, issues in teacher training, differences among regions in availability of trained teachers, differences in teachers' commitment and perception of the profession, and differences among regions in community and political support to MT teachers.

4.1.2 MTTCA SUB-SCORES BY LANGUAGE

To target strengths and weaknesses of teachers across different regions/languages, we disaggregated the total score on MTTCA into its four basic components: reading comprehension, MT grammar and vocabulary, subject matter pedagogy, and general pedagogy. Figure 15 depicts teacher performance on these basic MTTCA components:

1. Reading comprehension
2. MT grammar and vocabulary
3. Subject pedagogy
4. General pedagogy

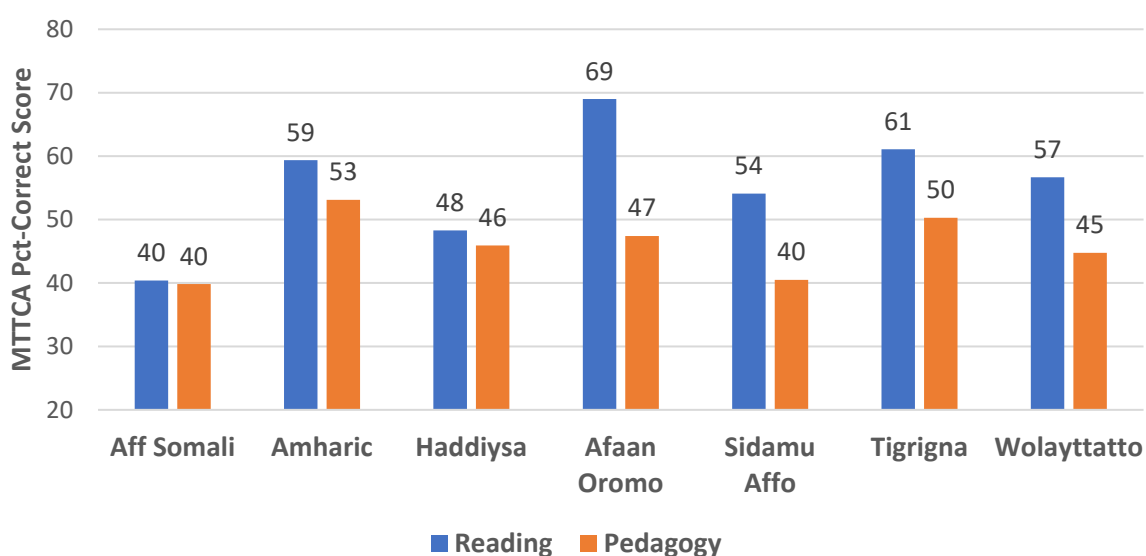
Figure 15. Average MTTCA Sub-Scores (Four Components) by Language



Since the four basic components of the MTTCA conceptually refer to two more general competency dimensions, one referring to teachers' own language skills (conveniently labeled "Reading") and the other to their teaching skills (conventionally labeled "Pedagogy"), we also present teachers' performance expressed on these two general dimensions (Figure 16). The two general dimensions are operationally defined as follows:

1. **Reading** is an aggregation of scores on reading comprehension, grammar, and vocabulary
2. **Pedagogy** as an aggregation of scores on subject pedagogy and general pedagogy

Figure 16. Average MTTCA Sub-Scores (Two Components) by Language



Based on obtained results, the following observations and comments relevant for policy considerations can be made:

1. There are substantial differences in the level of teachers' performance between the MTTCA components within languages.
2. Teacher performance on the pedagogy components (especially subject pedagogy) is substantially lower than on the reading components in most languages.
3. There are significant differences between languages in teacher performance on each of the MTTCA components.

The interpretational inputs from language experts collected during the Co-interpretation® workshop are presented in Table 7.

Table 7. Experts' Interpretation of Differences Between Languages in Teacher Performance on Specific MTTCA Components

EXPERT GROUP	RESPONSES
Afaan Oromo	0 There is only one subject matter pedagogy module in CTEs and the focus, in most cases, is on theory rather than on practical work 0 More attention is given to general pedagogy than the subject matter pedagogy
Aff Somali	0 According to the experience of the experts, mother tongue teachers are not good at subject pedagogy, but are good at reading, even higher than test scores show
Amharic	0 Teachers' score in reading does not fit to the expectation of experts 0 The duration or the length of in-service trainings is not adequate to equip teachers with new skills 0 In some cases, trainings were not offered to appropriate teachers—non-mother tongue teachers were included in many trainings organized for mother tongue teachers 0 Teachers were not interested/motivated to actively participate in trainings 0 Cascading problems of training, poor alignment between the training teachers received and what they are expected to teach 0 Lack of continuity in preparing, revising, and finalizing the curriculum and conducting trainings—they are project based, so when the project is over, the good practices decline

EXPERT GROUP	RESPONSES
Sidamu Affo	0 The result is expected as there are limited pedagogy courses in colleges; the available courses are given in English while teachers are trained to teach mother tongue 0 Teachers do not seem to be ready to learn and apply newly introduced mother tongue teaching methodology to teach reading; they prefer to present the lessons in a traditional way
Tigrigna	0 As per theoretical conception, the competencies of teachers seems to be good, but in actual practices, they may not practice the desired pedagogy, so in light of our experience, the results make sense 0 There may be different classroom practices 0 Lack of support and mentoring from principals, supervisors, and school leaders regarding teachers' professional development 0 Poor linkage between colleges and schools 0 Inability to use instructional aids and materials due to inaccessibility
Wolayttatto	0 There might be teachers who didn't take training on general pedagogy and reading competencies 0 At different levels, there are no sustainable and strong monitoring and follow-up mechanisms for ensuring teachers use of appropriate methodology 0 Some teachers may not have sufficient knowledge and skill in the new MT teaching methodology 0 Newly recruited teachers may not have mentors

The main conclusions drawn from experts' interpretations of the differences in teacher performance on specific components of MTTCA suggest that teacher training is an issue: subject matter pedagogy is insufficiently addressed at CTEs and duration and cascading nature of in-service training is not adequate to equip teachers with new skills for teaching MT.

4.2 TEACHER RESULTS BY PERFORMANCE STANDARDS

Performance standards are a necessary component of educational assessments because they provide a criterion-referenced evaluation framework for interpreting examinee performance. The purpose of performance standards is both formative and summative. Their main role is to support the MOE in developing the capacity to monitor and boost teachers' professional development progress, and to promote the quality of literacy instruction during the first four years of schooling at the system level.

There are two benefits of establishing performance standards for early-grade teachers: (1) they increase the capacity of policymakers to better support the implementation of the reading curriculum and the professional development of teachers, and (2) they create an evaluation framework for monitoring the development of teachers' competencies.

4.2.1 WHAT ARE TEACHER PERFORMANCE STANDARDS?

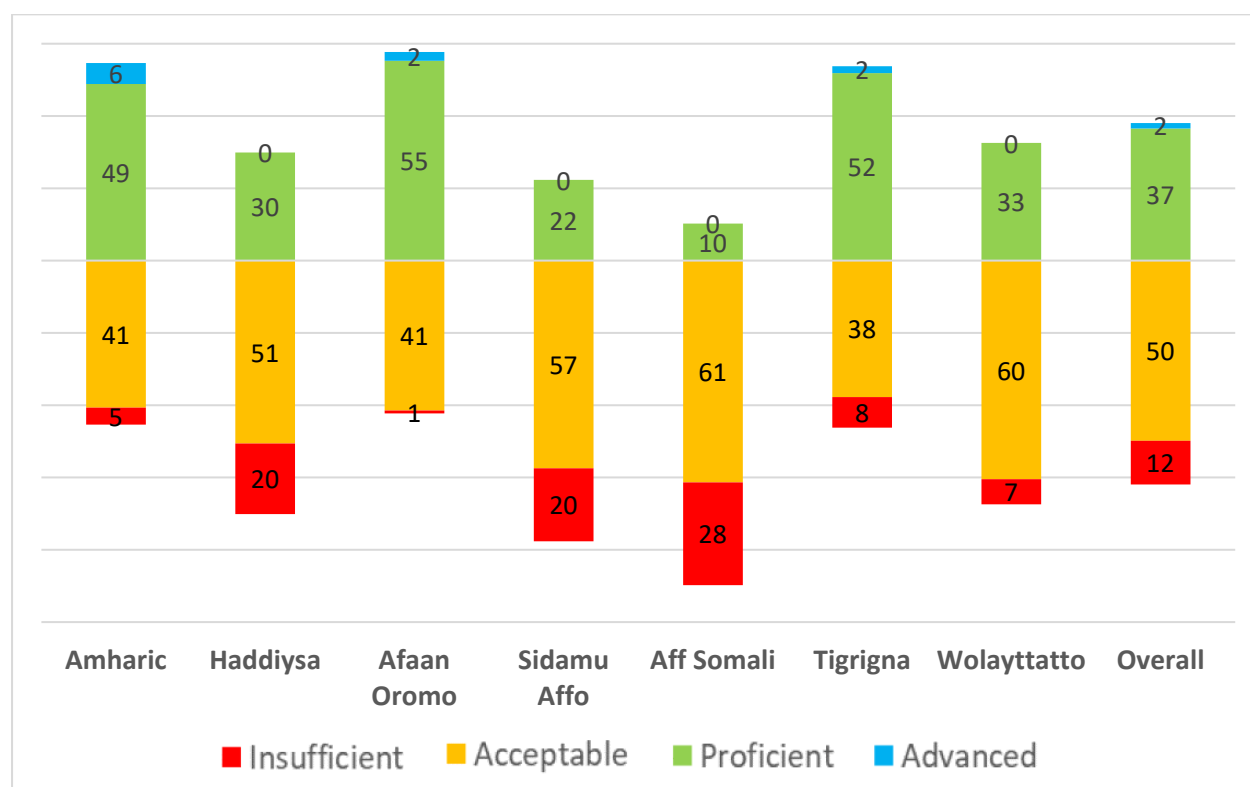
Early-grade teacher performance standards are a synthesis of conceptual and operational specifications of the **knowledge and skills teachers are expected to demonstrate** in the early grades of primary school. They are typically used to set national targets for teacher performance, first by defining the targeted competency level, and then by specifying the percentage of teachers expected to reach that level in a given time period.

4.2.2 TEACHER PERFORMANCE LEVELS BY LANGUAGE

Figure 17 depicts teacher performance expressed as the percentage of teachers in each of the four MTTCA performance levels: Insufficient, Acceptable, Proficient, and Advanced. Readers may note

that this figure is also presented in the section on setting performance standards (representing the final outcome of the cut-score-setting workshop).

Figure 17. Percentage of Teachers Reaching MTTCA Performance Levels by Language



Based on these results, the following observations and comments bearing important policy implications can be made:

1. There is a relatively low overall performance of teachers as measured by the MTTCA—when aggregated across seven languages, only 39% of teachers demonstrate the targeted levels “*Proficient or above.*”
2. There is a large variance in teacher performance between languages—for example, from 10% to 57% of teachers reaching the level “*Proficient or above.*”
3. In three languages, more than 50% of teachers are at “*Proficient or above*” (Amharic, Afaan Oromo, and Tigrigna); in three languages, this percentage is between 22% and 33% (Sidamu, Haddiysa, and Wolayttatto); and in one language, this percentage is 10% (Somali).

The interpretational inputs from language experts about the differences among teachers from different languages/regions in reaching MT teachers’ competency standards, which were collected during the Co-interpretation® workshop, are presented in Table 8.

Table 8. Experts' Interpretation of Percentages of Teachers Reaching Targeted Performance Levels "Proficient and Advanced"

EXPERT GROUP	RESPONSES
Aff Somali	- o The performance level of teachers in the region described in the report appears to be appropriate because of the availability of less qualified MT teachers, lack of sufficient experience to teach mother tongue, poor background of teachers, and shortage of in-service trainings - o There are more unqualified MT teachers in the region - o There is imbalance between demand and supply of MT teachers in the region
Amharic	- o Teacher' score in reading does not match the expectations of the experts - o Since several experienced teachers were trained based on the old mother tongue curriculum in CTEs, they need in-service/refreshment trainings/workshop to update their knowledge and skills with the new methodology - o In some cases, trainings were not offered to appropriate teachers—non-MT teachers were included in many trainings organized for mother tongue teachers - o Teachers were not interested/motivated to actively participate in trainings - o Poor alignment between the training teachers received and what they are expected to teach - o Lack of continuity in preparing, revising, and finalizing the curriculum and conducting trainings. Trainings are project based, so when project ends, use of good practices start to decline
Sidamu Affo	- o Some teachers are not well qualified to act as mother tongue teachers - o There are teachers who obtained diploma or certificate without going through proper training
Tigrigna	- o The finding of this research is found to be similar with our experience as experts and teaching at colleges and universities - o Teachers may not properly understand objectives, expected outcomes, curriculum, and minimum learning competencies (MLC) of the grade level - o Frequent in-service training provided for Tigrigna teachers
Wolayttatto	- o Less commitment of teachers to update themselves in continuous professional development (CPD), short-term training, etc. - o The absence of strong follow-up system at different levels - o The absence of different incentive packages - o Lack of reference materials

To summarize the opinions of experts from different languages, similarly to explanations for differences in MTTCA total scores and sub-scores, the differences among languages in percentage of teachers reaching targeted performance levels were mainly attributed to various aspects of teacher training, teacher motivation, and systemic support.

5. RELATIONSHIP BETWEEN TEACHERS' AND STUDENTS' PERFORMANCE

This chapter investigates the associations between teachers' and students' performance across schools that were included in the sample. We are interested in correlations between two indicators of school performance:

1. Teacher performance on the MTTCA, expressed either as average percent-correct scores on the MTTCA, or as the percentages of teachers reaching MTTCA performance levels established through standard setting.
2. Student performance on EGRA, expressed either as average EGRA sub-task scores, or as the percentages of students reaching EGRA performance levels established through benchmarking.

This chapter addresses the following questions: Do schools with higher teacher performance also demonstrate higher student performance, and vice versa? If so, is this relationship statistically significant, and what is the strength of the association? The answers to these questions are of considerable practical relevance, and may inform important policy decisions.

5.1 RELATIONSHIP BETWEEN MEAN SCORES ON EGRA AND MTTCA

On the student side, the following EGRA subtasks were included in this analysis: Familiar Words Reading (FWR), Invented Words Reading (IWR), Passage Words Reading (PWR/ORF), and Reading Comprehension (RC). On the teacher side, the following MTTCA scores were included in this analysis: total score and sub-scores (reading comprehension; grammar and vocabulary; subject pedagogy; and general pedagogy), all expressed in percent-correct metrics.

5.1.1 STUDENT EGRA SCORES VERSUS TEACHER MTTCA SCORES

Table 9 presents the coefficients of correlation between average school scores obtained from the students' 2018 EGRA and teachers' 2019 MTTCA administration. There were 342 schools where both students' and teacher's data were collected, which represented a pool of schools for computation of correlations between students' and teachers' performance. For each school we calculated the average student score on EGRA subtasks, as well as the average teacher scores on the MTTCA and its components. The school EGRA scores were then correlated with school MTTCA scores across schools as a source of variation.

Table 9. Correlations Between EGRA School Scores and MTTCA School Scores

N=342 SCHOOLS	MTTCA TOTAL SCORE	MTTCA READING COMPREHENSION	MTTCA GRAMMAR & VOCABULARY	MTTCA SUBJECT PEDAGOGY	MTTCA GENERAL PEDAGOGY
EGRA FWR	.30**	.29**	.07	.19**	.28**
EGRA IWR	.12*	.07	-.06	.10	.16**
EGRA ORF	.23**	.22**	.10*	.12*	.21**
EGRA RC	.24**	.20**	.16**	.10	.24**

Note: ** significant at $p < 0.01$, * significant at $p < 0.05$

Based on these results, the following observations and comments relevant for policy implications can be made:

1. There are statistically significant correlations between average teacher MTTCA scores and student EGRA scores across schools.
2. The strongest correlations between the MTTCA total score and EGRA were observed for the following subtasks: familiar words reading (0.30), oral reading fluency (0.23), and reading comprehension (0.24).
3. When MTTCA was analyzed by components, the strongest relationship with EGRA subtasks was found for teachers' Reading Comprehension and General Pedagogy scores.
4. Subject matter pedagogy shows a relatively low relationship with students' EGRA performance, likely because of low variance of teacher scores caused by the "floor effect," i.e., low teachers' scores on subject pedagogy.

The interpretational inputs from language experts about the correlations between teacher performance and student performance collected during the Co-interpretation® workshop are presented in Table 10.

Table 10. Experts' Interpretation of Correlations Between Teacher Competency Scores and Student Performance on ORF

EXPERT GROUP	RESPONSES
Afaan Oromo	0 The frequent change of the teaching materials might have created confusion and low performance of students on EGRA 0 Appropriate attention has not been given to subject matter pedagogy and general pedagogy in CTEs
Somali	0 The correlation of teacher performance and student performance matches the experts' common experience
Haddiysa	0 The finding does not show the correlation clearly
Sidamu Affo	0 There are no adequate subject matter pedagogy courses in the CTEs and would-be teachers learn general pedagogy in English 0 Excessive and boring repetition of content in primary textbooks
Tigrigna	0 There is no significant difference between the performances of teachers and students
Wolayttatto	0 The correlation indicates that students taught by better teachers performed better

As an overall impression, it seems that the obtained results were aligned with experts' expectations, and the correlations between teacher competencies and student performance seemed natural to them. Thus, rather than correlation itself, their comments were more pertaining to the low end of teacher and student performance reflecting their belief that the results were due to changes in instructional materials, their poor quality, and poor training for teachers.

5.1.2 MTTCA TOTAL SCORES AND EGRA ORF BY ZONES

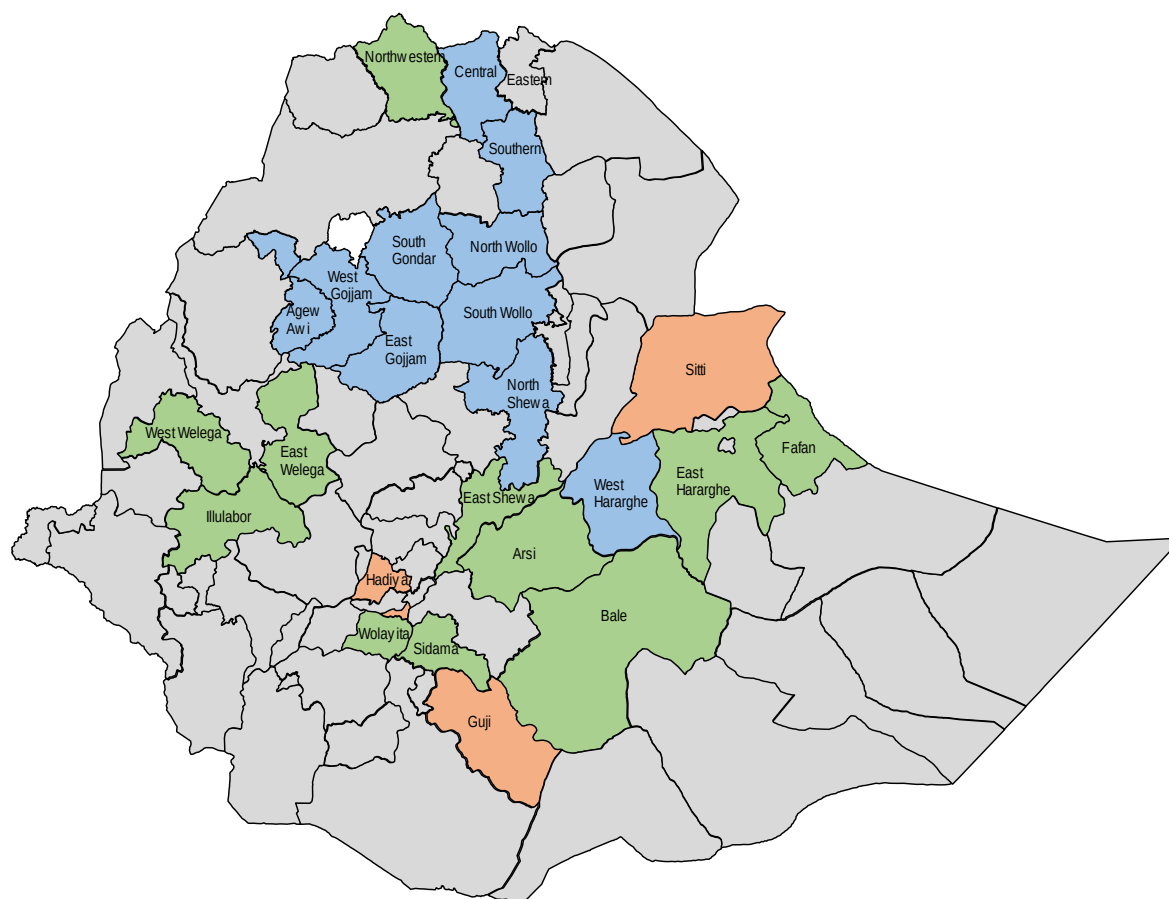
Presenting MTTCA results by smaller geographical units (zones) may help to inform policy and instructional support at the local level. In this section we present the MTTCA 2019 total score results and the EGRA 2018 results for the main subtask (ORF), disaggregated by participatory zones in which both assessments were administered. To enhance understanding of the geographical distribution of MTTCA and EGRA performance, READ M&E presents the zonal results in the form of a map, in which the zones are divided into three categories: those performing around the national mean (color coded green), those performing substantially above the national mean (color coded

blue), and those performing substantially below the national mean (color coded orange). Zonal mean scores are classified as either substantially above or below the national mean based on an effect size (Cohen's d) of 0.20. If the zonal mean score is above the national average with an effect size of at least 0.20, it is classified as *substantially above*; if the zonal mean score is below the national average with an effect size of at least 0.20, it is classified as *substantially below*.

Important note: non-representative sampling of schools was used within zones. For this reason, results should be interpreted cautiously and used primarily for descriptive and formative purposes. The results should not be used for evaluative comparisons between languages, or for high-stakes evaluations of the performance of local educational officials.

Figure 18 and Table 11 present the results for participating zones on the ORF subtask, and Figure 19 and Table 8 show the zonal results for the MTTCA total score.

Figure 18. EGRA 2018 Results for the Oral Reading Fluency Subtask by Zones

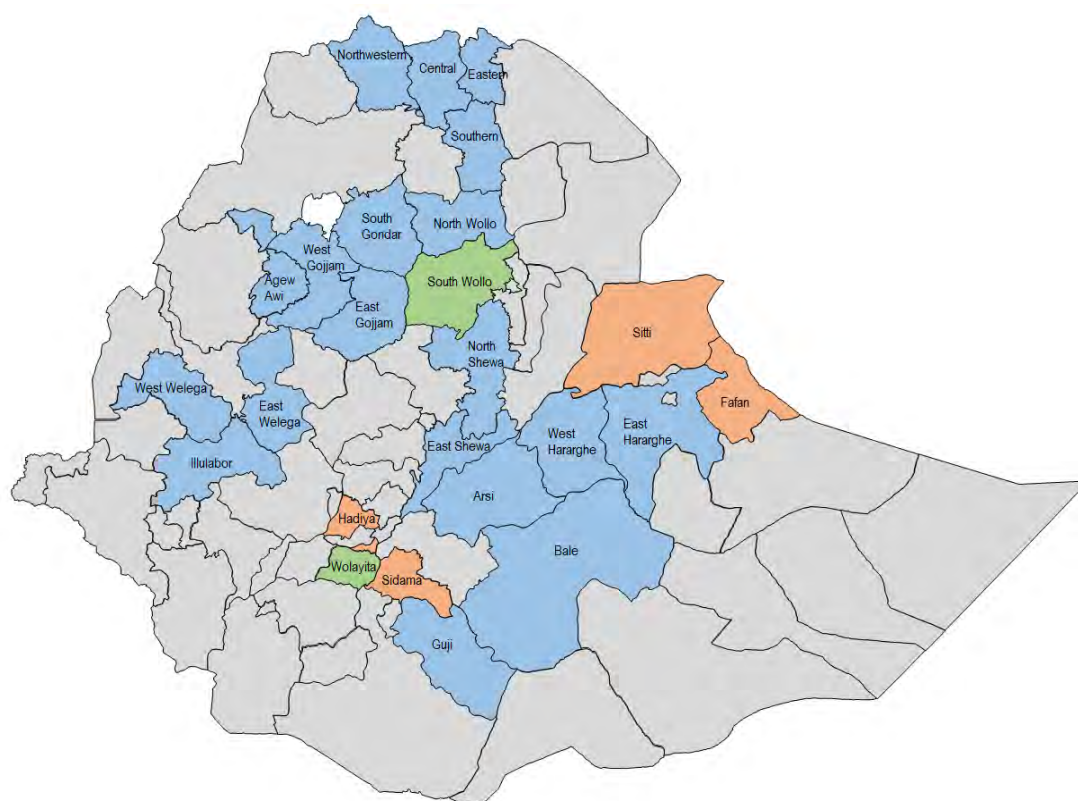


Note. Blue = above the national mean; green = around the national mean; orange = below the national mean. Categorization is based on an effect size (Cohen's d) of 0.20.

Table 11. Zonal Mean Scores on the EGRA 2018 Oral Reading Fluency (ORF) Subtask

Above the National Mean		Around the National Mean		Below the National Mean	
Zone	ORF Mean	Zone	ORF Mean	Zone	ORF Mean
Agew Awi	42.0	North West (Tigray)	19.2	Sitti	12.5
North Shewa	38.7	Baale	17.8	Guji	11.9
West Gojjam	33.0	National Mean	17.5	Hadiya	9.2
South Wollo	31.5	Arsii	17.5		
East Gojjam	31.1	West Wollega	16.9		
South Gondar	26.9	Wolayta	16.3		
Central (Tigray)	22.3	Faafan	15.7		
North Wollo	22.2	East Welega	15.5		
Southern (Tigray)	21.9	East Harargee	15.3		
West Hararghe	21.7	Sidama	15.3		
		East Shewa	14.4		
		Illubabor	14.3		

Figure 19. MTTCA 2019 Results on Total Percent-Correct Score by Zones



Note. Blue = above the national mean; green = around the national mean; orange = below the national mean. Categorization is based on an effect size (Cohen's *d*) of 0.20.

Table 12. Zonal Mean Scores for the MTTCA 2019 Total Score

ABOVE THE NATIONAL MEAN		AROUND THE NATIONAL MEAN		BELOW THE NATIONAL MEAN	
ZONE	MEAN	ZONE	MEAN	ZONE	MEAN
Shaw Bahaa	65	South Wollo	50.3	Hadiya	46.9
East Wollega	62.2	National Mean	50.1	Sidama	45.5
Bahir Dar City	62.0	Wolayta	49.2	Faafan	41.5
Eastern	61.4			Siti	35.1
Semen Wollo	59.2				
Buunnoo Bedellee	59.1				
South Gondar	57.7				
Wallagaa Lixaa	56.6				
Misrak Gojjam	56.4				
Agew Awi	55.9				
Central	55.9				
Semen Showa	55.5				
Arsii	54.9				
East Hararghe	54.7				
Gujji	54.3				
North West (Tigray)	54.0				
West Hararghe	52.9				
Mekelle	52.7				
Baale	52.7				

As the EGRA student results were correlated with the MTTCA teacher results across schools, the zonal results also suggest a general correspondence between mother tongue teachers' competencies and student reading performance (rank correlation between zonal averages on MTTCA and ORF is 0.30). These results can help Regional State Education Bureaus (RSEBs) to provide customized intervention in the zones and areas below the national mean.

Relevant questions that were addressed in the Co-interpretation® workshop: How does the finding about differences between performance of teachers across the zones fit into the common experience of local experts? What are possible explanations for these zonal differences? What recommendations can be formulated for education bureaus? Based on these results, what improvements should be set as targets for the next 5 years for the zones “around the national mean” and those “below the national mean”?

The interpretational inputs from language experts about the differences between zones in both student reading performance and MT teacher performance collected during the Co-interpretation® workshop are presented in Table 13.

Table 13. Experts' Interpretation of Differences in Performance of Teachers Across the Zones

EXPERT GROUP	RESPONSES
Aff Somali	0 According to experts' experience, Somali teachers' performance is below the national mean because of the existing challenges that have already been mentioned 0 The context and different situations of the zones might have contributed to zonal differences in MTTC assessment results
Amharic	0 The presence of better qualified and experienced teachers and relatively better supervision and leadership in towns might have contributed to zonal differences
Haddiysa	0 Adequacy of teaching learning materials 0 Difference in socioeconomic status 0 Poor reading habits 0 Poor monitoring and evaluation strategy 0 Experience variation 0 Lack of concern from stakeholders towards MT languages 0 The zonal scores of MTTC show almost all mother tongue teachers performed above the national mean
Sidamu Affo	0 Cascading approach to training for actual teachers is ineffective 0 There is no strong monitoring and supporting system that all stakeholders use in an organized way 0 Intervention covers limited areas
Tigrigna	0 The finding indicated that the result is similar for all zones (no significant differences) 0 There may be different classroom practices 0 Lack of support and mentoring from principals, supervisors, and school leaders 0 Poor linkage between colleges and schools 0 Inability to use instructional aids and materials due to inaccessibility
Wolayttatto	0 Difference in interest and commitment among teachers

To summarize, the most prevailing expert opinions about zonal differences in teacher competencies are pointing to contextual differences affecting teacher performance, mainly socioeconomic status, availability of teaching resources, and tendency that more qualified and experienced teachers concentrate in urban areas.

5.2 RELATIONSHIP BETWEEN TEACHERS AND STUDENTS BY PERFORMANCE LEVELS

In this section we first present the average EGRA student scores in relation to performance levels of their teachers. Second, we comparatively present student and teacher performance expressed in terms of performance levels—that is, the percentages of students in each of the EGRA performance levels, and the percentages of teachers in each of the MTCCA performance levels.

5.2.1 TEACHER PERFORMANCE LEVELS VERSUS STUDENT EGRA SCORES

Our primary aim is to investigate whether schools with high-competency teachers also demonstrate good student performance. We classified teachers' performance into different performance levels, based on MTTC standards, and then calculated the average student EGRA scores in schools with teachers of different performance levels (Figure 20).

We also present the same data in the form of profiles showing the average student EGRA profiles for schools where teachers of different performance levels work (Figure 21).

Figure 20. Relationship Between Teachers' Performance Levels and Student EGRA Scores

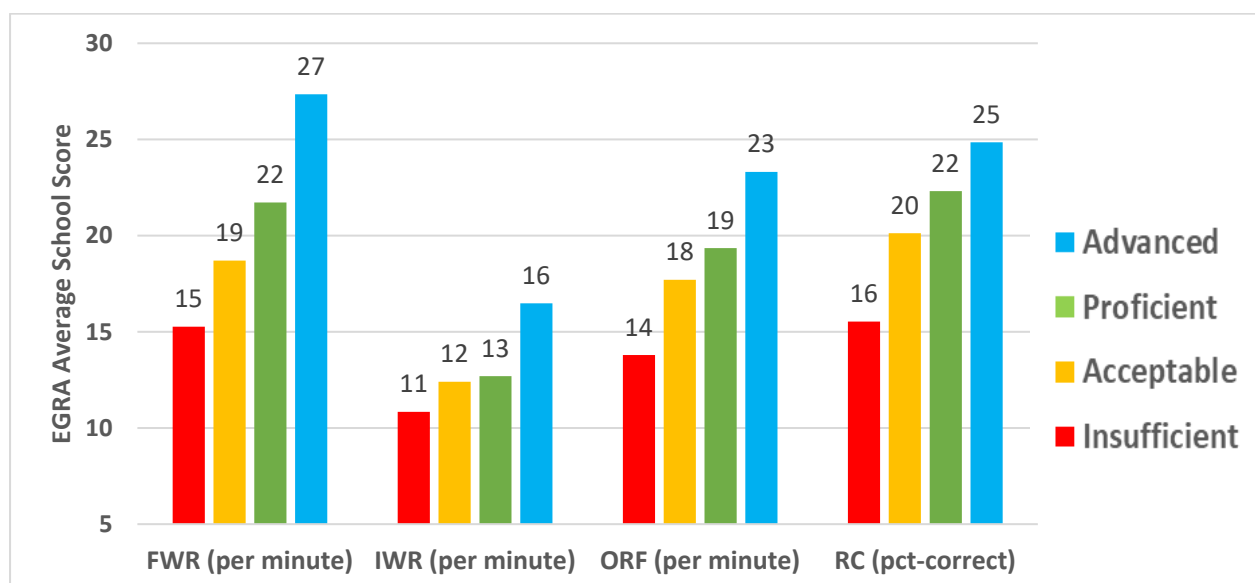
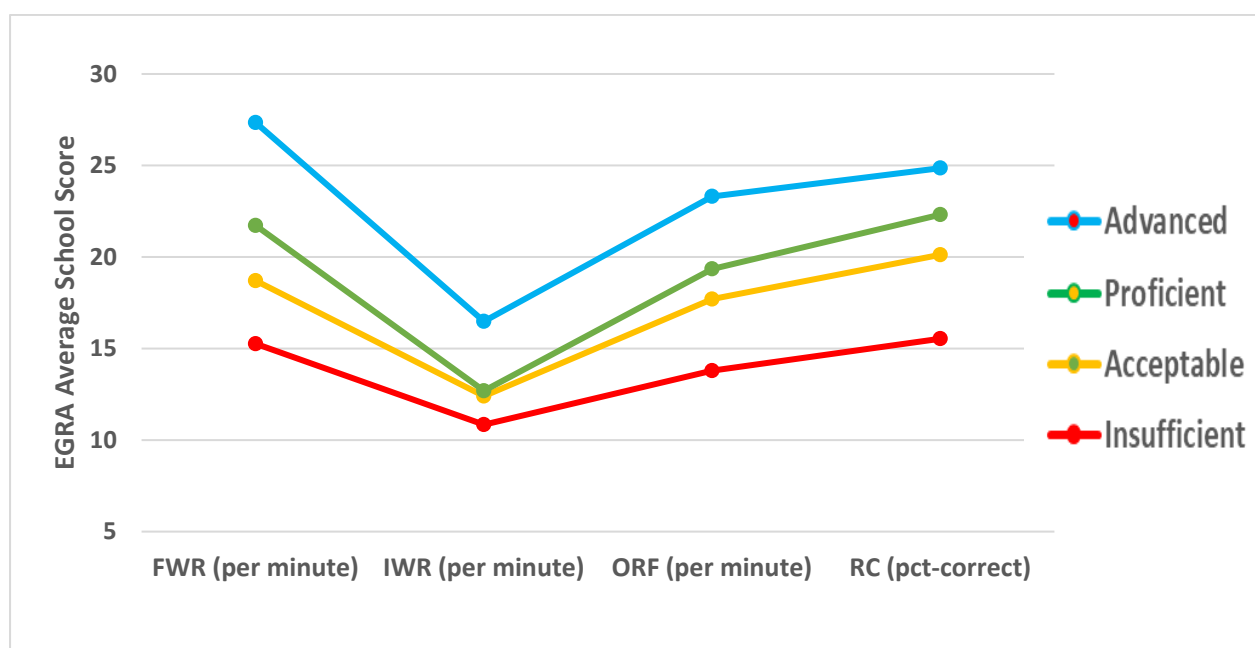


Figure 21. Profile of EGRA School Scores for Teachers at Different Performance Levels



It can be observed that students perform better on EGRA reading tasks in the schools where teachers demonstrate higher performance levels. All of the differences in EGRA scores between students taught by teachers at different performance levels were statistically significant, exhibiting large effect sizes. For example, the difference on ORF and RC between students taught by teachers performing at the Insufficient level has an effect size of 0.95, and the difference for students being taught by teachers at the Advanced level has an effect size of 0.83—both are very large effect sizes. Even the difference between students taught by teachers at the Insufficient level and students taught by teachers at the Acceptable level exhibits an effect size of 0.38 for ORF and 0.43 for RC, which are both considered moderately large effect sizes.

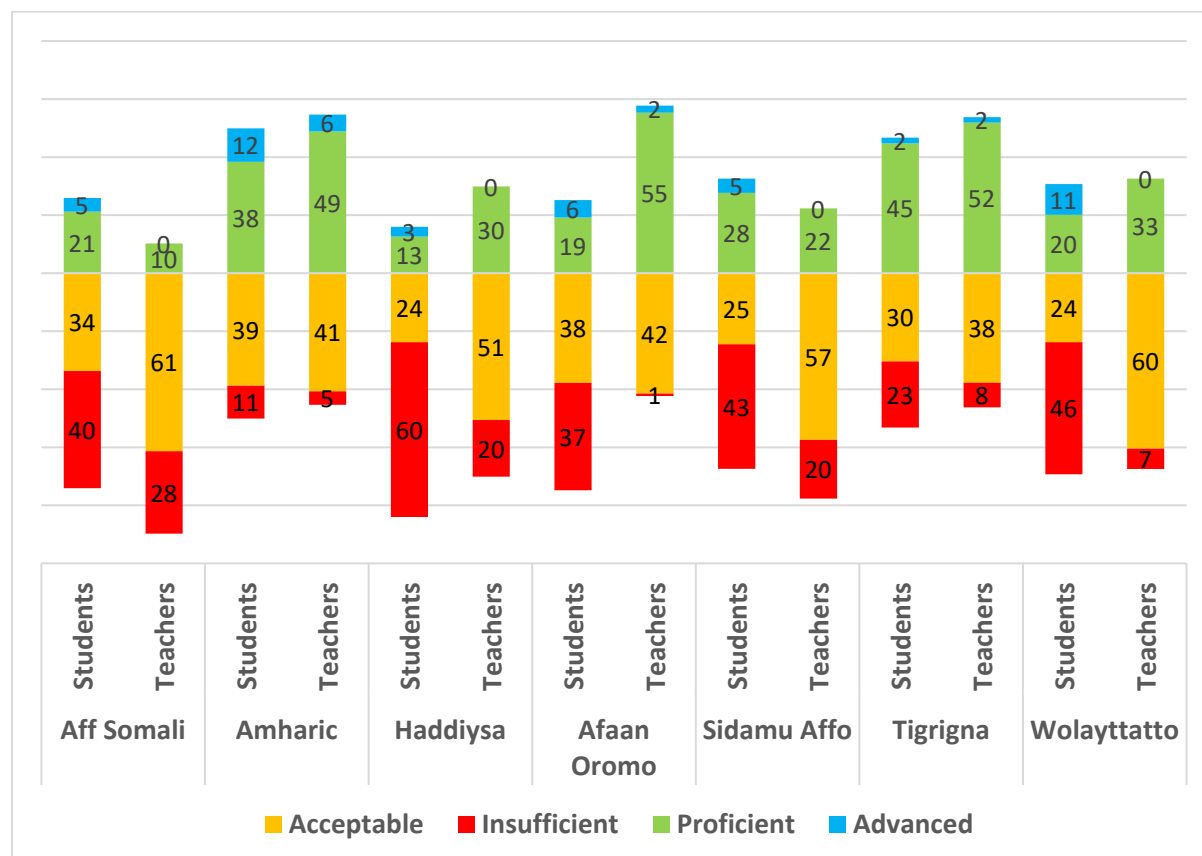
Thus, if we could improve factors that influence the academic performance of both teachers and students, such as school resources, to elevate the teacher performance level from Insufficient to Acceptable, it would likely be associated with a significant boost in student performance. Since the correlational nature of our study design does not allow for causal interpretations, it is essential to extend the research quest towards identifying factors that affect both teacher and student performance to enable the most effective targets for policy decisions. At this point, the following observations and comments relevant for policy implications can be made:

1. There is a strong correspondence between teachers' performance and students' performance in their schools. It is evident that students perform better in schools with better teachers, though this evidence is not enough to draw causal attribution.
2. The differences are statistically and practically significant, exhibiting effect sizes (Cohen's d) that are considered very large in the context of educational research (Cohen, 1977).

5.2.2 TEACHER MTTCA PERFORMANCE LEVELS VERSUS STUDENT EGRA PERFORMANCE LEVELS ACROSS LANGUAGES

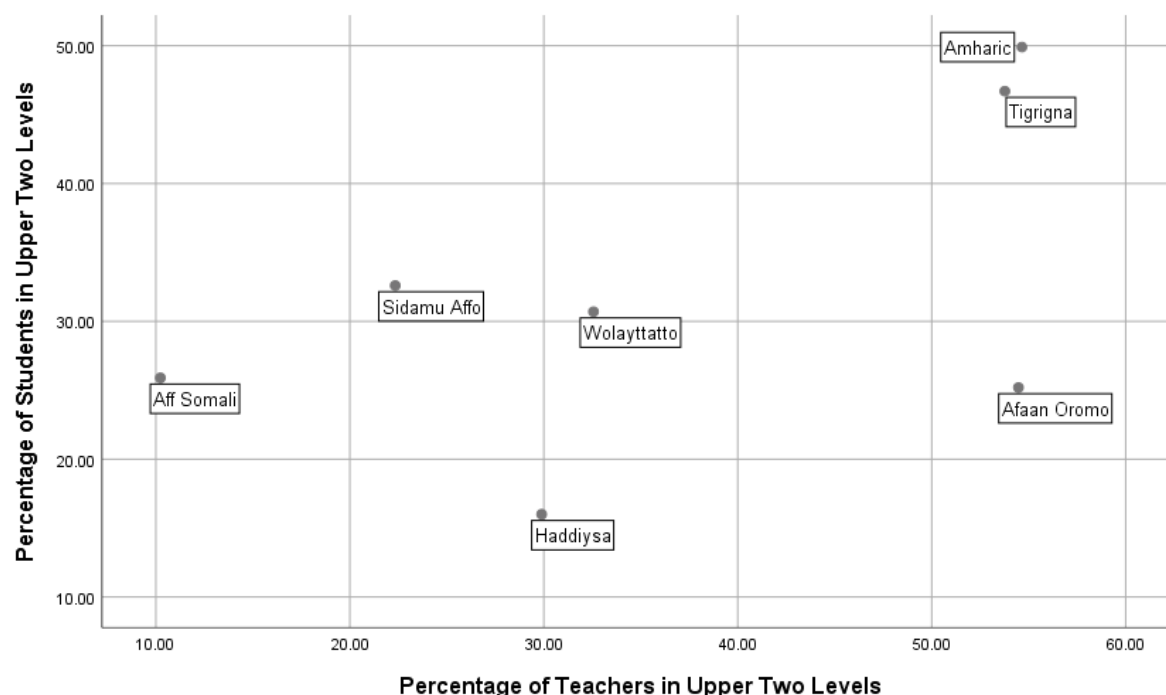
In this section we present a comparison of the distributions of the percentages of teachers in different MTTCA performance levels and the percentages of students in different EGRA performance levels in the seven Ethiopian MT languages (Figure 22). A relatively strong correspondence between distributions of teachers in performance levels on MTTCA and students in performance levels on EGRA can be observed. In most cases, the languages with high-performing teachers also exhibit high student performance.

Figure 22. Percentages of Students and Teachers in Performance Levels by Language



To further evaluate the correspondence between teacher and student performance across different languages, we created a scatterplot depicting the relationship between the percentages of teachers and students classified as performing in the top two performance levels (Figure 23). Associated with this scatterplot, the coefficient of correlation between student and teacher performance across languages equals 0.56, which is further strong evidence of this relationship.

Figure 23. Percentages of Students and Teachers in the Top Two Performance Levels by Language



To summarize, the following observations and comments relevant for policy implications can be made:

1. There is a relatively strong correspondence between teacher and student performance across languages, yielding a significant correlation of 0.56.
2. Amharic and Tigrigna are the languages with the highest performers for both students and teachers (percentages of “*Proficient* or above” approximates 50% for both teachers and students).
3. In most of the other languages it can be seen that teacher performance is relatively aligned with student performance. For example, in the Wolayttatto language, slightly more than 30% of both teachers and students reached the top two performance levels. In the Sidamu Affo language, the percentage of teachers in the upper two performance levels is 22% and the percentage of students in the corresponding top two levels is 33%. In Haddiysa, 30% of teachers reached the desired performance level as compared to 16% of students. 26% of students in Aff Somali language reached the desired reading performance level despite only 10% of teachers at the desired level of proficiency.
4. In Afaan Oromo, on the other hand, a large disparity between teacher and student performance can be observed, where high teacher performance (55% “*Proficient* or above”) is not accompanied by high student performance (only 25% of students at the targeted two performance levels).

Relevant questions that were addressed in the Co-interpretation® workshop: How does the finding about “better teachers = better students” fit into the common experience of local experts? What are plausible explanations of these associations? What are policy implications and actionable recommendations? The inputs from language experts about the associations between MT teachers’ competency performance and student reading performance collected during the Co-interpretation® workshop are presented in Table 14.

Table 14. Experts’ Interpretation of Finding ‘Better Teachers = Better Students’

EXPERT GROUP	RESPONSES
Afaan Oromo	○ We observe that, if the teachers are efficient enough, the students will also be efficient enough, since there is a direct correlation between teachers’ competencies and the students result ○ The teacher has a big influence on students’ performance by way of transferring knowledge and modeling different situations for them
Aff Somali	○ Better students are the results of better teachers. The higher the teachers’ competencies, the better students’ performance. Low EGRA scores of students in our region can be partly explained by low MTTCA results of teachers in the region.
Amharic	○ It fits with experts’ experience, because in most cases better teachers produce better students. They have a great role in shaping students’ potential and in creating a better learning environment. ○ Better teachers exert more efforts to make their students better achievers. Teachers’ capacity affects students’ performance (either positively or negatively). The reason for misalignment could be the presence of other factors that affect students’ performance.
Haddiysa	○ In real situation, people expect that the performance of the teachers has a direct influence on students’ performance. But, in the case of Aff Somali and Afaan Oromo the finding indicates an inverse relationship (students’ performance is better than the teachers’ in Aff Somali and teachers’ performance is better than students’ in Afaan Oromo.)
Tigrigna	○ There may be direct impact of the teachers towards students’ performance ○ A strong relationship exists between the performance of teachers and students; better teachers will deliver subject matter in a better way and support students to work hard and perform better
Wolayttatto	○ Better teachers are committed, cooperative, update themselves, and share experience with their students

To summarize, virtually all experts consider that quality of teachers is substantially related to student performance and that equation “better teachers = better students” is observable in schools. They even ascribe causal interpretation to this relationship indicating that “better teachers produce better students”, not being concerned that some contextual factors (such as resources) might be a determinant of both teacher and student performance. The discussion about these factors that may inflate a correlation between teacher and student performance was convened in the form of focus group activity and its outcomes are presented in the next chapter.

5.3 DECIPHERING TEACHER EFFECT

In previous sections of this chapter we presented the results of bivariate analyses of the relationship between teacher and student performance. The results clearly show that there is a statistically and practically significant association between primary teachers’ competencies measured by MTTCA and student reading performance measured by EGRA. However, this evidence is not yet enough to enable inferences about the teacher effect size, i.e., how much student performance can be attributed to the quality of teachers. Considering the correlational nature of this research, and absence of experimental manipulation that would enable causal attribution of teacher effect, further

multivariate analyses are a step toward ruling out other competing factors that contribute to student performance (mainly school and home resources typically associated with student learning outcomes).

Two types of multivariate analyses were used to evaluate the associations between teacher and student performance, analysis of covariance (ANCOVA) and multiple regression (MR) analysis. In ANCOVA, the school and home resource variables are introduced as covariates; in other words, they are statistically held constant when analyzing the relationship between teachers' competencies and student reading performance in their schools. In this way, we are analyzing the association between teacher and student performance in a hypothetical situation where school and home resources are the same.

In MR, unlike ANCOVA, we are not isolating resource variables to focus on the pure association between teacher and student performance; rather, we evaluate all the variables, including teacher performance and school and home resources, as factors associated with student performance. In this way, we can estimate and compare the degree to which all of these variables are uniquely associated with student performance through considering their overlap or collinearity. Stepwise MR selects a combination of resource variables, including teacher competencies, that are significantly related to student reading performance, removing from the equation those variables that are non-significant or redundant/overlapping. MR analysis provides regression coefficients indicating how the predictor variables should be weighted in a linear combination to enable the best prediction of the criterion variable. By comparing the standardized regression coefficients, we can evaluate a degree of unique association between each of the analyzed factors (teacher competencies, school resources, and home resources as predictors) and student reading performance.

The variables reflecting school resources and home resources associated with student performance were based on data from the school Director Questionnaire (DQ) and Student Questionnaire (SQ), respectively, collected in the same schools as a part of the EGRA 2018 administration. Selection of resource variables to be included in the ANCOVA and MR analysis was based on the univariate associations between these resource variables and student reading performance that were statistically significant in the analyses conducted on EGRA 2018 data.

The variables from DQ reflecting school resources that were included in our analyses are presented in Table 15. For descriptive purposes, the overall percentages of directors' responses in each of the response categories are also presented (note that in a few instances the response categories with low frequencies were merged).

Table 15. Variables from Director Questionnaire Included in ANCOVA and MR

QUESTIONS TO SCHOOL DIRECTORS		OVERALL PERCENTAGES OF RESPONSES			
How often are daily lesson plans reviewed?	Once a year or semester	Monthly / bimonthly	Weekly	Daily	
	9.2%	7.5%	71.3%	12.0%	
Do all MT teachers in your school have: Teacher guide for the grade they teach?			No	Yes	
			15.0%	85.0%	
Are the supplementary reading materials accessible for Grade 1–4 students?			No	Yes	
			41.9%	51.1%	
Does your school have a functioning library or a reading room?		No, or for teachers only		Yes, for students and teachers	
		29.8%		70.2%	
Is the library or reading room used by Grade 2 and Grade 3 students?			No	Yes	
			56.69%	43.4%	
How many of MT teachers have diploma	None	5 or less	5 to 10	Over 10	
	7.1%	73.1%	15.6%	4.2%	
How many MT teachers are there in your school?	None	5 or less	5 to 10	Over 10	
	1.5%	72.2%	21.2%	5.1%	

Similarly, the variables from the SQ reflecting both home and school resources that were included in our analyses are presented in Table 16. Student responses to these questions were first aggregated within schools to obtain a percentage of students giving the targeted response within each school. Then, these school data were merged with the MTTCA data so that each teacher was associated with a percentage of student responses within his/her school. For descriptive purposes, the overall averages of these student-based variables selected for our analysis are presented in Table 16.

Table 16. Variables from Student Questionnaire Included in ANCOVA and MR

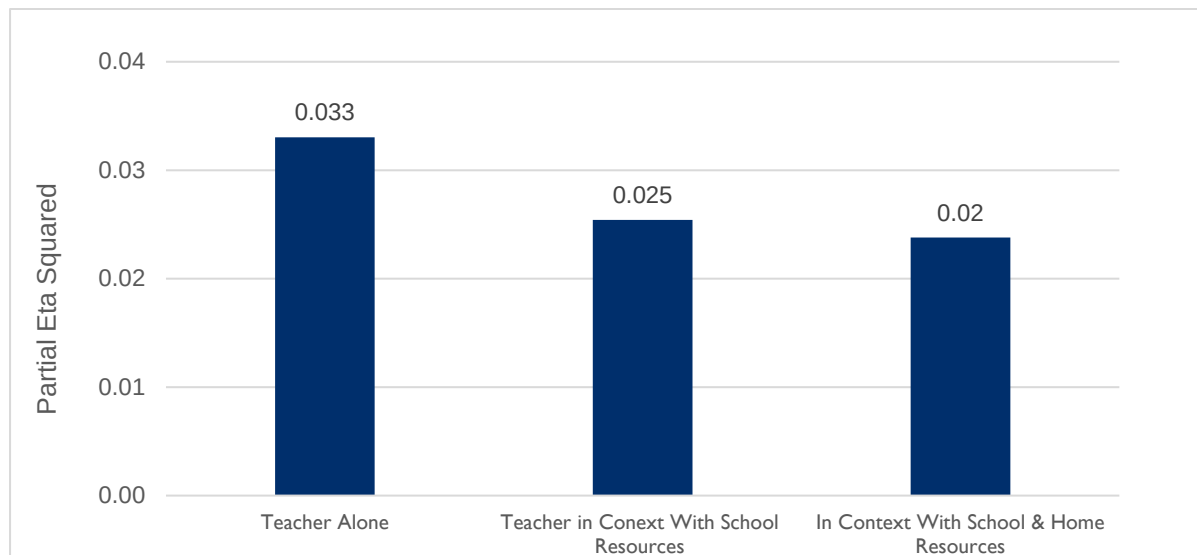
VARIABLE FROM STUDENT QUESTIONNAIRE EXPRESSED AS % OF STUDENTS PER SCHOOL	OVERALL AVERAGE
Percentage of students that have MT textbook	84.8
Percentage of students that read books other than MT textbook	47.6
Percentage of students that borrow supplementary reading materials	19.6
Percentage of students reporting that school has a library	52.1
Percentage of students reporting there is a reading corner in the classroom	21.4
Percentage of students having family members that can read and write	81.7
Percentage of students receiving help at home while reading	75.3
Percentage of students having enough time at home to read the textbook	84.5

5.3.1 ASSOCIATION BETWEEN STUDENT AND TEACHER PERFORMANCE WITH RESOURCE VARIABLES HELD CONSTANT

In the analysis of covariance (ANCOVA), the indicators of **school and home resources** were used as covariates when evaluating the differences in reading performance between students coming from schools with teachers of different levels of competencies. In other words, we evaluated the **associations between teacher and student performance** by statistically **holding the resource variables constant**.

The results show that the degree of association between MT teachers' competencies and student reading performance decreases when we include school and home resource variables; however, the relationship still stays highly significant at $p < 0.01$. This means that even in the context of equal school and home resources, there is still a significant difference in reading performance between students in schools with teachers of different levels of competencies.

Figure 24. Strength of Association Between Teacher Performance Level and Student ORF Scores in Different Contexts



5.3.2 ASSOCIATION OF STUDENT PERFORMANCE WITH TEACHER COMPETENCIES AND SCHOOL AND HOME RESOURCES

In the regression analysis we evaluated the relationship between student reading performance as a criterion variable and a set of 15 school and home resource variables (described above) plus a teacher competency score. By simultaneously including these variables in multiple regression analysis we are able to observe the unique contribution of each of the analyzed “predictors” and student reading performance as a criterion. Using the stepwise technique of regression analysis, we enabled automatic selection of those variables that have significant unique relationship with student performance.

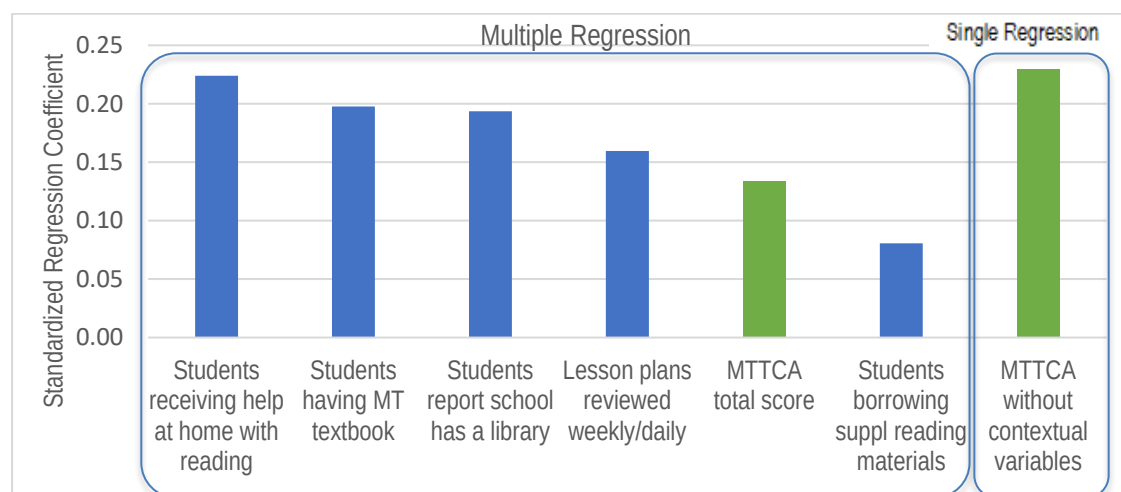
The main findings of the MR analysis are presented in Figure 25 below, which depicts the predictor variables (teacher competency score, school resource variables, and home resource variables) that demonstrate significant unique associations with student reading performance. In other words, each of these variables has a significant individual association with student performance, even when the effects of all other variables are partialled out. Out of 16 predictor variables only six survived as having a significant unique contribution to explaining the variance of student reading performance (listed in further text).

Thus, we evaluated how **Teacher Competencies** are associated with student reading performance in the context of **Home Resources** indicators and **School Resources** indicators. It appears that teacher competencies are still highly statistically and practically associated with student performance; however, the strength of association for home and school resources appears to be somewhat larger than for teacher competencies. For comparison, the green bar on the right side of Figure 25 shows the strength of association between MTTCA scores and student average ORF scores when analyzed by **single regression** without contextual variables, in which case relatively strong association of

teachers' competencies with student reading performance is likely mixed the with effect of other factors (i.e., resources) that affect both teachers' and students' performance. However, when we conduct **multiple regression** analysis including school and home resources as additional predictors, the unique contribution of teacher competencies (with resource variables partialled out) to explanation of variance in student performance appears smaller (green bar), but still remains highly statistically significant ($p < 0.001$).

Although teacher competency score is one of the factors significantly associated with student reading performance, it is interesting to observe that the variables that have the strongest association with student performance are the home resource variables "Receiving help at home with reading" and "Having a mother tongue textbook." The variables with the next strongest association with student performance are two school resource variables "School having a library" and "Frequency of reviews of daily lesson plans" and the teacher effect is somewhat lower than these home and school resource variables.

Figure 25. Strength of Association of Student ORF Scores With Teacher Competencies, Home Resources, and School Resources



It should be noted that this analysis was conducted on overall data, but it can be also done for regions separately to see how these factors are associated with student reading performance locally, so that RSEBs could prioritize activities at the local level.

To summarize, school resources and school environment, as well as home environment and resources, together with teacher competencies, are significantly associated with the average reading performance of students at the school level. The policy implication of this finding is that student reading performance may be improved by the following activities: increasing home resources (e.g., parental training, community engagement), increasing school resources (e.g., building libraries, providing supplemental readings), and increasing level of teacher competencies (e.g., continuous professional development).

The findings presented in this section are in line with ample research on the home literacy environment. These home literacy environment factors are found to be the strongest predictors of individual differences in pre-reading scores, which in turn are the strongest predictor of later reading scores (Senechal & Le Fevre, 2002; Foy & Mann, 2003; Phillips & Lonigan, 2009). The implication of this is that the home environment lays the foundation, after which the teacher and school role takes over.

6. SUMMARY AND SYNTHESIS OF EXPERT OPINIONS REGARDING FACTORS ASSOCIATED WITH TEACHER AND STUDENT PERFORMANCE

The Co-Interpretation® process noted above involved presenting results of initial quantitative analyses conducted with the data sets from assessments of students (EGRA) and teachers (MTTCA) to experts closest to the work, to bring their perspectives to the interpretation of and to inform further analyses on available data. This process enabled participants to develop shared understanding of and to strengthen the completion of this report and, correspondingly, the meaningfulness and utility of the results. The two quantitative analyses presented in Chapters 5 and the two in Chapter 6 were conducted post-workshop and were informed by the contributions of the experts who participated in the workshop.

In addition to reacting to and further informing initial quantitative analyses, the experts produced further qualitative data themselves as part of the Co-Interpretation® process. This chapter presents a synthesis of the qualitative data collected using focus groups during the Co-Interpretation® workshop conducted in Adama in February 2020. The qualitative data reflect the opinions of the participating Ethiopian language experts (teacher educators from universities and CTEs) about the three main questions: 1) What factors influence teacher performance? 2) What factors influence student performance? and 3) What is the relationship between teacher competencies and students' reading outcomes? It is important to note that the experts' opinions were collected to qualitatively reflect their experiences and views about the subject themes, inspired by the results of teachers' and students' performance, rather than to provide quantitative statistical inferences from their opinions. From that perspective, the core information presented in this chapter is not derived from quantitative data collected by MTTCA or EGRA; rather, these are results of the synthesis of qualitative data collected from Ethiopian experts to complement the quantitative analyses.

Throughout this chapter, the results are presented in a way that preserves their language and regional origin to provide information valuable for regional stakeholders in understanding the factors relevant for making policy decisions and action plans for improvement of teachers' performance and student learning outcomes in their specific regions. Also, this chapter was written in close collaboration with Ethiopian experts, to ensure that the organization of presented themes, the elaboration of quantitative and qualitative data, and the articulation of implications and recommendations reflect profound understanding of Ethiopian regional educational settings, and represent an invaluable contribution to mapping the factors relevant for making action plans intended to improve both teachers' and students' outcomes.

6.1 LANGUAGE-FOCUSED DISCUSSION ON MT TEACHERS' COMPETENCIES, STUDENTS' READING OUTCOMES, AND POLICY IMPLICATIONS

EGRA results across the last three test administrations reveal that the proportion of students in Ethiopia who are at the functional level of reading is not more than a third of the student population (USAID Ethiopia, 2018), and results are not improving over time despite significant investments in programs and corresponding assessments. It is hence essential to elicit expert insights to help explain this situation and find out possible factors affecting the reading outcomes of students, as well as teachers' competencies in teaching reading in early grades. The following sections present qualitative results from expert discussions regarding factors affecting the MTTCA results and students' early grade reading performance, as represented in the quantitative analyses presented by the READ M&E team. In addition, associations between teacher competencies and students' reading performance are considered.

Discussion results are presented separately by language with cumulative comparisons between language groups and occasional relevant external citations introduced as each separate language is presented, followed by an aggregate summary. The discussion is geared towards identifying common issues, associations, and implications and recommendations for a way forward.

6.1.1 AFAAN OROMO GROUP EXPERT DISCUSSION SUMMARY

The average MTTCA Percent-Correct Total Score for Afaan Oromo is 56, and of the teachers tested, 57% scored at the “*Proficient and Advanced*” level. Teacher performance on the pedagogy components (47) is substantially lower than on the reading component (69), which is the largest difference between reading and pedagogy, with an average-score gap of 22 percent-correct score points.

Given these scores, the Afaan Oromo group identified the factors that they believe, based on their collective expertise, affect teacher competencies and affect early grade reading outcomes in their language region, in order of importance. The factors cited as affecting teacher competencies were *training, leadership quality, resources, and selection*, and those cited as affecting early grade reading outcomes were *students’ family background, resources (teaching & supplementary reading materials), teacher competencies, curricular problem, and inappropriate time use* (spending more time in non-instructional time).

From these ratings it can be seen that training of teachers was considered as the most important factor affecting the competencies of teachers, followed by leadership quality, among participating Afaan Oromo experts. The average MTTCA Percent-Correct Total Score for Afaan Oromo was 56 points, which is average. This showed that the teachers’ performance in Afaan Oromo is comparatively better than in other language groups, though it is not very high itself. At the same time, over 57% of the teachers are in the targeted levels “*Proficient and Advanced*.” Despite these relatively high teachers’ scores, a large disparity between teacher and student performance can be observed. High teacher performance (57% teachers reached “*Proficient or Advanced*” levels) is not matched by high student performance. Only 25% of students reached the targeted top two performance levels, “*Proficient and Advanced*”). Hence, there is a mismatch between MT teachers’ competencies to teach reading and early grade students’ reading outcomes in the Oromia region.

This mismatch needs to be investigated to understand possible causes. The questions of interest in this explanation to help understand the existing mismatch include the following:

- Why is the percentage of students at the targeted top two performance levels lower than percentage of teachers performing at the equivalent targeted levels?
- Is the problem in performance by the students linked with teachers’ competencies or other related factors?
- What are the possible explanations for the observed percentages of teachers and students not reaching desired standards?

Although 57% of the teachers have reached the target level, the effect of the remaining 43% who failed to reach the target level cannot be disregarded. In addition, leadership quality, which was listed as the second most important factor affecting teacher competencies, needs consideration, since it has a direct association with the teaching-learning process. Participants noted that the teaching methods and learning culture most commonly observed in MT classrooms across the various regions reflected a more traditional, teacher-centered model rather than a student-centered model (READ TA, 2014). It was also mentioned that the majority of school directors seem to not be providing instructional leadership for teachers, which research indicates is an important factor (Bultosa,

Gadisa, & Garoma, 2017). If leadership fails to promote teachers' competencies or promote their practice in schools, not only will those who do not exhibit competencies not be motivated to improve, but also those who have the competencies may not be motivated to exert maximum effort to help children improve their reading performance. This can contribute to the poor performance of students, as students need concerted support among the stakeholders.

Beyond the above points, the students' family background and resources were listed as the first and second factors affecting early grade reading outcomes. It is known that availability of grade-appropriate reading materials (mother tongue teachers' guide, student mother tongue textbooks, and supplementary reading materials), existence of a school library and reading corner, and use of the library or reading corner are all positively associated with student performance in reading. The fact noted by the co-interpreting group that "teachers spending more time on non-academic issues" may adversely affect student learning regardless of teacher competencies level (i.e., possessing competencies does not work if they are not practiced), could have also been one cause for the discrepancy in performance between teachers and students. The large pupil section ratio (PSR) of 66 in grades 1–4 (MOE Ethiopia, 2018/19) could also be a possible cause for the poor performance of the students.

In summary, deploying properly trained teachers and resources and having proper leadership supports the teaching and learning of MT. The group suggested this can be achieved only if there is gap-filling teacher professional development and support system. These can help the school-level teaching and learning of the mother tongue to address the identified gaps. To this end, the following action points are recommended by the Afaan Oromo group:

- Strengthen MT teacher training and experience-sharing through CPD across the region.
- Revise and improve MT in-service and pre-service teacher training.
- Strengthen the ongoing effort to align the curriculum of the college of teacher education modules and students' teaching materials developed by the READ project.
- Supply enough teaching and learning materials.
- Direct school leadership to support teachers in the teaching learning process, instead of focusing only on administrative issues.
- Strengthen the community involvement with a sense of ownership and create conducive environment for the teaching-learning process.

6.1.2 AMHARIC GROUP EXPERT DISCUSSION SUMMARY

The average MTTCA Percent-Correct Total Score for Amharic is 56; of the teachers tested, 55% reached "*Proficient and Advanced*" levels, and teacher performance on the pedagogy components is 53 and reading components is 59.

The Amharic language group ranked the factors that most affected teacher competencies as *teacher's commitment, selection & recruitment, teacher training, and resource & educational materials* in order of importance, unlike the Afaan Oromo team, for whom training and leadership were the top two factors. Similarly, the Amharic group listed factors affecting early grade students' reading outcomes as *school facility, family support, content coverage, living standard, and absenteeism*. The listing indicates teacher commitment as the most important factor to predict teacher competencies, followed by selection and recruitment. The average MTTCA Percent-Correct Total Score of 56 points exhibits an average performance, with highest effect size among language groups in this assessment. The percentage of teachers reaching MTTCA performance levels is also 57. This shows the same result as Afaan Oromo, but the factors differ. The gap between reading (59%) and pedagogy (53%)

performance of teachers is much less. In this language, the percentage of students who reached the targeted top two performance levels is also 49% (EGRA, 2018).

From the evidence provided for the Amharic language group, we observe comparable performance between teachers and students clustered around average level. In other words, MT teachers' competencies are associated with early grade student reading outcomes. Training and teacher subject and pedagogical knowledge were not identified as most important factors affecting teacher performance; however, teacher commitment was identified as highly important for their performance. Thus, providing a response to the question "why do teachers fail to be committed?" is critically important. Teacher commitment is the emotional bond they demonstrate toward their work. Research shows that teachers with commitment are able to support student learning with innovative instructional strategies that lead to better achievement (Altun, 2017). If teachers are not committed, then their preparation and delivery of instruction is hampered which has a negative effect on student performance. As identified by the co-interpreters, student performance is affected by problems associated with school facilities and family background. When the issue of teacher commitment is added to these factors, it clearly affects student performance. If a solution to the problem of teacher commitment can be found, teachers are then able to resolve contextual problems (e.g., school resources) and improve their instructional effectiveness. In other words, it is highly likely that students will perform better through the improvement of teacher commitment. It is therefore important to explore more about teachers' level of commitment to teaching, and the gaps in the supply of school resources. Additionally, the following recommendations are given by the Amharic group:

- Strengthening teachers' commitment through creation of awareness about the profession, addressing issues of teachers' incentives, and installing strong follow-up system.
- Selection and recruitment need to be critically reviewed and adjusted to help attract better candidates.
- Teacher training, which was rated as the third factor, needs to be addressed. This could be done through aligning college courses and school curriculum, and strengthening teachers through CPD administered properly at the school level.
- The last factor, dealing with resources, also needs attention to solve holistic school-level problems.

6.1.3 TIGRAY GROUP EXPERT DISCUSSION SUMMARY

From the results of the quantitative analysis, the average MTTCA Percent-Correct Total Score for Tigrigna is 54. From the teachers tested, 54% reached "*Proficient* and *Advanced*" level, and teacher performance on the pedagogy components is 50 while reading components is 61. This indicates a low level of teachers' pedagogical knowledge compared to their reading competencies.

The Tigrigna language group identified the following factors as affecting teacher competencies: *professionalism*, *professional development*, *school environment*, and *incentives*, in order of importance. The group also listed the factors affecting the reading performance of students as *teacher competencies*, *learning environment* (provision of quality reading materials, availability of different materials to enhance student reading and class size), *school readiness* (parental support and socialization), *learning style and strategies*, and *curriculum* (availability of textbooks and teacher guide, knowledge of MLCs, and assessment and feedback). Hence, according to Tigray local experts, teacher professionalism and competencies are the factors that most affect teacher competencies. As the results indicate, the average MTTCA Percent-Correct Total Score for Tigrigna teachers is around average (54%), with 54% of teachers achieving the top two levels of performance ("*Proficient*" and "*Advanced*").

The overall oral reading fluency (ORF) of students is 21 words per minute, reading comprehension is 22%, and 47% of students achieved the two targeted levels (Reading with increasing fluency and comprehension, and Reading fluently with full comprehension). Lack of professionalism among teachers together with lack of teacher competencies were identified by the co-interpretation group as key factors affecting student reading performance. It was suggested that professional development efforts that lead to the improvement of teacher professionalism and competencies should help mitigate the problems that affect the reading outcomes of students. The group proposed that such efforts be focused on in future actions to support teachers. The following are additional recommendations from the Tigray group:

- Improve teachers' competencies, professionalism, and professional development through different modalities.
- Revisit school environment to make it appealing for teachers to exert maximum effort.
- Provide supplementary reading materials.
- Strengthen school preparedness for children to make them well prepared before enrolling in Grade 1.
- Assess incentive mechanisms and act accordingly.
- Devise mechanism of strengthening teachers' classroom practices, and ways of assessment and giving feedback and documentation.

6.1.4 SIDAMU AFFO GROUP EXPERT DISCUSSION SUMMARY

The average MTTCA Percent-Correct Total Score for Sidamu Affo is 46 and, from the teachers tested, 22% reached “*Proficient and Advanced*” level. Teacher performance on the pedagogy components is 40, and reading components is 54, which shows a problem of teacher competencies in general and on pedagogical skills in particular.

In an attempt to explain the above situation, the Sidamu Affo language group listed the factors that affected teacher competencies as *teacher training*, *pedagogical knowledge*, *technical skill* to use continuous classroom assessment, and *motivation*, in the order of importance.

Regarding factors affecting early grade reading outcomes in this language group, teacher competencies was listed as the first, followed by learning materials, school environment, community engagement, and societal traditional attitude in the context of the Sidama Zone.

Both dimensions suggest that the epicenter of the problem is teacher related, whether it is on the training or the competencies they currently have. The fact that only 22% of the teachers reached the top two performance levels shows the severity of the problem related to competencies of teachers. The EGRA 2018 study revealed that only 23% of the assessed students reached the top two levels (reading with increased fluency and comprehension 28% and reading fluently and with full comprehension 5%). It seems that there is an association between teacher performance and student performance, which would seem to suggest the need for interventions on improving teacher competencies. Therefore, the Sidamu Affo group suggest it is essential to revisit the ongoing teacher education, and devise mechanisms to help MT teachers develop their pedagogical knowledge and skills that lead to the improvement of student reading performance. Recommendations from the Sidamu Affo group include:

- Devise intervention to improve teacher competencies, and revise recruitment and selection criteria in general and for MT teachers in particular.
- Integrate modules with school curricula to help teachers develop pedagogical knowledge, and techniques of using continuous classroom assessment (CCA).

- Revisit learning materials and improve school environment.
- Strengthen community engagement.
- Create awareness among society on values of education and reading outcomes.
- Increase the motivation of teachers using various mechanisms.

6.1.5 WOLAYTTATTO GROUP EXPERT DISCUSSION SUMMARY

The situation of Wolayttatto is not different from that of Sidamu Affo. Average MTTCA Percent-Correct Total Score for Wolayttatto is 49, below the pass score of 50%. From the teachers tested, only 33% reached “*Proficient and Advanced*” level; teacher performance on the pedagogy components is 45 and reading components is 57, all of which show the existence of problems associated with teacher competencies.

During the Co-interpretation® workshop, the Wolayttatto language group identified and listed the factors that affected teacher competencies as *teacher commitment*, *teacher background*, *teaching & learning materials*, and *monitoring & evaluation*, according to their order of importance. As mentioned above in the Amharic section, teachers with commitment are able to support students with innovative instructional strategies that lead to better achievement. Thus, the lack of commitment of teachers, together with poor educational background, identified as the second most important factor, and with very few teachers reaching the target performance level, combine to drive down the reading performance of students. From the EGRA (2018) study, only 31% of students reached the targets of “*Proficient*” and “*Advanced*”. The ORF of students was 14 words per minute, with reading comprehension at 16%. The Wolayttatto group identified teacher competencies as a key factor, which is supported by MTTCA test results of 45% for pedagogy and 57% for reading as indicative of poor teacher capability. Considering the seriousness of the problem related to the teachers’ performance, this group considers that devising mechanisms that improve teacher competencies, including reading and pedagogical knowledge, is quite important. The following are among the recommendations from the Wolayttatto group:

- Improve teacher commitment and engagement through mechanisms such as engaging in meaningful, effective, and problem-solving CPD; introduce a system that reinforces those who work hard and is aligned with students’ performance, and introduce a system of accountability.
- Strengthen teacher competencies through various forms of training.
- Train a sufficient number of MT teachers with proper linguistic background and trained to teach the MT.
- Introduce a system of monitoring and evaluation.
- Promote students’ attitude towards learning.
- Supply sufficient resources and student reading and reference materials.
- Increase the number of MT reference material writers.
- Expand access to pre-school learning.

6.1.6 HADDIYSA GROUP EXPERT DISCUSSION SUMMARY

For the Haddiysa group, the average MTTCA Percent-Correct Total Score is 47, below the traditionally acceptable level of 50%. Of the teachers tested, 30% reached “*Proficient and Above*” level, with teacher performance on the pedagogy components 46 and reading components 48. All of these results indicate poor performance of teachers, implying poor MT teachers’ competencies in the Hadiya Zone.

During the Co-interpretation® workshop Haddiysa language group experts identified *recruitment and selection, training program, interest and motivation, and leadership* as the factors that affect teacher competencies. Once compiling the list, the group reflected that the list indicates that initially those who came to the teaching profession either did not have proper secondary school preparation, or the teacher education is not providing enough support to develop teacher competencies. The EGRA study in 2018 indicated that oral reading fluency (ORF) is 9 words per minute and reading comprehension (RC) is 17%, with only 16% of students reaching target levels (Proficient and Advanced).

The Hadiyyisa group's discussion is supported by external research, which indicates that an important goal of reading is the effective interaction between ideas in a text and reader (Pilten & Kuralbayeva, 2018). In the teacher education literature on the teaching of reading, the main role of the teacher is to develop the ability of students to interact with text (Mubarok & Sofiana, 2018). If the teacher demonstrates limitations in terms of competencies, interest, and motivation, it is likely to negatively affect the student learning performance and reading outcome.

Therefore, the selection and recruitment of teachers and the teacher education through which they gain initial teaching competencies needs critical review. Improving the competencies they possess through continuous professional development is also essential. In addition, the Haddiysa group suggested it would be useful to address the following recommendations:

- Improve selection and recruitment guidelines to attract the best performing students to become teachers.
- Strengthen the leadership role.
- Strengthen school preparedness before enrolling in Grade 1.
- Improve teacher competencies through revising pre-service teacher education curriculum, and providing short-term trainings based on needs.
- Provide sufficient reading materials.
- Maintain a supportive environment by having a standard class-to-student ratio and teacher-student ratio.

6.1.7 AFF SOMALI GROUP EXPERT DISCUSSION SUMMARY

The average MTTCA Percent-Correct Total Score of the Aff Somali MT teachers is 40, below 50%. Of the teachers tested, only 10% reached “Proficient or Above” level. Teacher performance on the pedagogy components is 40 and reading components is 40. These all indicate problems associated with teacher competencies.

The Aff Somali language group at the Co-interpretation® workshop identified *attitude towards teaching, educational leaders, geographic & environmental factors, and skills* as factors that affected teacher competencies, in that order of importance. These priorities presented by the group do not include teacher knowledge, despite only 10% of the MT teachers in the Somali region reached the target level. This is a critical problem to be addressed, and one that merits further consideration in terms of the key factors identified by this expert group vis-à-vis MTTCA results.

The factors cited by the Aff Somali group as most affecting early grade reading outcomes are curriculum, teacher competencies, supplementary materials, socioeconomic, and non-conducive environment (e.g., class-to-student ratio, and teacher-to-student ratio). The problem of teacher competencies combined with the huge pupil-to-teacher ratio of 104 for Grades 1–8 and 94 for Grades 1–4 (MOE Ethiopia, 2018/19) in the Somali region could lead to poor student performance. It is noted that having regular (up to 30 students) or large (between 31 and 60 students) class sizes is

associated with better reading performance than having extremely large class sizes (more than 61 students). Only 26% of the students have reached the target level, ORF is 15 words per minute, and RC is 16%, indicating association between poor teacher competencies and poor student performance.

The group indicated attitude and leadership as the most important factors affecting teacher competencies, which is aligned with research which indicates that improving student achievement requires attention of teachers, and teacher commitment affects student accomplishment (Altun, 2017). The most important factors affecting student reading were reported to be curriculum and teacher competencies. Thus, the group suggests that it can be seen that poor performance on the MTTCA, complemented with poor attitude and leadership, is causing poor student performance. Hence, to improve teacher competencies, the group suggests it is vital to review the ongoing teacher education and the continuous professional development, in order to mitigate the problems that contribute to poor attitude towards teaching and overall teacher competencies. In addition, the Aff Somali group recommends it would be worthwhile to:

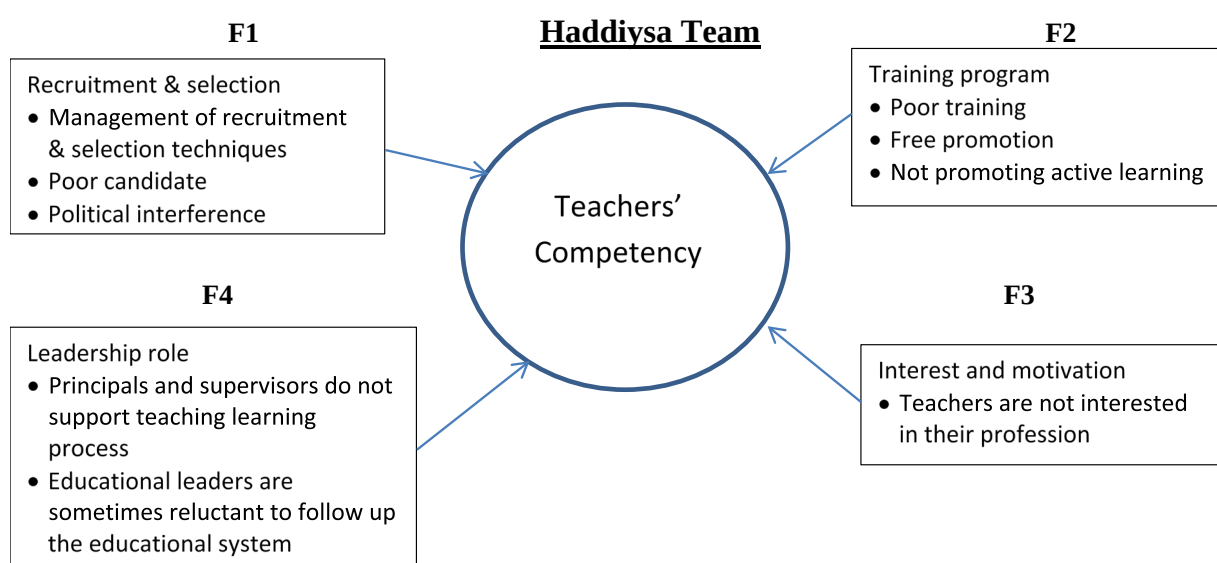
- Revise the ongoing teacher education of MT teachers, and provide CPD to develop teacher competencies.
- Promote positive attitude towards teaching.
- Assist educational leaders to be supportive of instruction.
- As geographic and environmental factors are among the factors affecting teacher competencies, teachers assigned to teach MT in remote areas should be given special incentives.
- Provide skill development packages and implement them to help teachers develop their skills.
- Provide sufficient supplementary materials.
- Maintain supportive environment that maintains a standard class-to-student ratio and teacher-student ratio.
- Revisit curricular materials and develop textbooks and related materials based on the specific characteristics of the mother tongue language and the context.
- Devise mechanisms to mitigate problems related to socioeconomic factors.
- Strengthen school readiness program (pre-primary education).
- Provide special motivation incentives for teachers assigned to teach MT in remote areas.

6.2. EXPERT OPINIONS ON MT TEACHERS' COMPETENCIES AND THEIR RELATIONSHIP WITH STUDENTS' READING OUTCOMES

Beyond co-interpreting the results from the quantitative data analysis of MTTCA 2019 and EGRA 2018, during the Co-Interpretation® workshop a focus-group activity was convened to explore and discuss two themes: (1) factors affecting teachers' competencies, and (2) factors affecting students' early grade reading outcomes. The activity was facilitated using the visual aid/template shown in Figure 1, which also represents an example of the answers produced by one of the expert groups (Haddiysa).

The two focused themes were discussed by each language group to identify potential drivers that affect the performance of both students and teachers in their regions, and to associate the results observed with student reading outcomes and real-world contexts. Each group was given the chance to articulate and prioritize the top factors that affect the early grade reading and teacher competencies components from their perspectives—five important factors for student reading, and four important factors for the teacher competencies. The factors identified by language group teams were further summarized and synthesized, outcomes of which are presented in further text.

Figure 26. Visual Template Used for Focus Group Activity



6.2.1 FACTORS AFFECTING TEACHER COMPETENCIES

According to the MTTCA total scores by language, the variation between languages is statistically significant. The post-hoc analysis of individual differences between languages also revealed that the performance of teachers in the seven assessed languages can be categorized in three relatively homogenous groups (teachers of Amharic, Afaan Oromo, and Tigrigna at the top; teachers of Haddiysa, Sidamu Affo, and Wolayttatto in the middle; and teachers of Aff Somali at the bottom).

Group responses to the prioritized factors that affect early grade reading and teacher competencies were mapped in ranked order (with 1 representing the most important factor) to perform categorical investigation and identify any potential association between the rating of the factors for each of the languages within the homogeneous groups identified by performance on MTTCA (refer to section 4.1.1), as presented in Table 17.

Table 17. Identified Factors According to Their Importance by Language Groups

LANGUAGE	FACTOR 1	FACTOR 2	FACTOR 3	FACTOR 4
Afaan Oromo	Training	Leadership quality	Resources	Selection standard
Tigrigna	Professionalism	Professional development	Conducive school environment	Incentives
Amharic	Teacher's commitment	Selection & recruitment	Teacher training	Resource & educational materials
Wolayttatto	Teacher's commitment	Teachers' background	Teaching & leaning materials	Monitoring & evaluation
Haddiysa	Recruitment & selection	Training program	Interest & motivation	Leadership role
Sidamu Affo	Teachers' training system	Lack of pedagogical knowledge	Lack of technique on using CCA	Motivational passion
Aff Somali	Attitude towards teaching	Educational leaders	Geographical & environmental factors	Skills

The factors that affect teacher competencies as stated by the language experts can be categorized into three major constructs: (1) training, professional knowledge, and development, (2) personal constructs and affects, and (3) leadership and work environment, as presented in Table 18.

Table 18. Categories of Factors Affecting Teacher Competencies Organized by Constructs

LANGUAGE GROUP	TRAINING, PROFESSIONAL KNOWLEDGE & DEVELOPMENT	PERSONAL CONSTRUCTS AND AFFECTS	LEADERSHIP AND WORK ENVIRONMENT
Afaan Oromo	• Training • Selection standard		• Leadership quality • Resources
Tigrigna	• Professionalism • Professional development		• Conducive school environment • Incentives
Amharic	• Selection & recruitment • Teacher training	• Teacher's commitment	• Resource & educational materials
Wolayttatto	• Teachers' background	• Teacher's commitment	• Teaching & leaning materials • Monitoring & evaluation
Haddiysa	• Recruitment & selection • Training program	• Interest & motivation	• Leadership role
Sidamu Affo	• Teachers' training system • Lack of pedagogical knowledge. • Lack of technique on using CCA	• Motivational passion	
Aff Somali	• Skills	• Attitude towards teaching	• Educational leaders • Geographical & environmental factors

The results from the Co-interpretation® workshop suggest participating experts perceive the most important factors affecting teacher competencies to be training, professional knowledge, and development (mentioned 13 times), followed by leadership and work environment (mentioned 10 times). Except for Sidamu Affo, whose group never mentioned leadership and work environment as a factor, the dominant factors presented among the groups are training and professional development of teachers, and the leadership and work environment. In Aff Somali, the dominant factor is leadership and work environment. The middle performing homogeneous group (Wolayttatto, Haddiysa, and Sidamu Affo) seems to offer similar issues of commitment or interest as a factor affecting teacher competencies. As outlined in Table 18, issue of teacher training, knowledge, and overall professional development are directly linked with teacher competencies. An additional common theme across all language groups except Sidamu Affo is of leadership and work environment as a factor that impacts teachers' overall competencies. These findings have important implications for policy and practice at the regional and national levels.

To further examine these three constructs, the responses were reclassified into eight subgroups which could be considered informative and actionable for both policy and practice. The training, professional knowledge, and development construct cluster was divided into three sub-groups: (1) recruitment and selection, (2) training (including background), and (3) professional development (including skills). Leadership and work environment was divided into four sub-groups: (1) educational leadership and M&E, (2) learning materials and resources, (3) school environment, and (4) incentives; while teachers' commitment, attitude, motivation is retained as is (see Table 19).

Table 19. Language Groups and Their Rating of Factors (F1 = Most Important Factor)

NO.	NAME OF CONSTRUCT CLUSTER	NAME OF SUB-CLUSTER	FREQUENCY	RATING OF FACTORS IN ORDER OF IMPORTANCE			
				F1	F2	F3	F4
1	Training, professional knowledge & development	Recruitment and selection	3	1	1		1
		Training	4	2	1	1	
		Professional development	6	1	3	1	1
2	Teachers' commitment, attitude, motivation	Teachers' commitment, attitude, motivation	5	3		1	1
3	Leadership and work environment	Educational leadership and M&E	4		2	1	1
		Learning materials and resources	3			1	2
		School environment	2			1	1
		Incentives	1			1	
Total			28	7	7	7	7

The table presents the number of language groups that identified each of the factors in this exercise. The results indicate that leadership and work environment were only identified as the first or second factor in two instances, suggesting this factor may have a subsidiary effect, rather than a primary one. Training, professional knowledge, and development are indicated as the first or second most important factor affecting teacher competencies in nine instances. This suggests the assumption that if these conditions are met, participants believe attitude and commitment could improve.

6.2.2 FACTORS AFFECTING STUDENT PERFORMANCE

Factors affecting early grade reading were also mapped to identify any themes or categories among group responses. Table 20 below presents the responses.

Table 20. Factors Affecting Reading Outcomes by Language Groups Ordered by Importance

LANGUAGE	FACTOR 1	FACTOR 2	FACTOR 3	FACTOR 4	FACTOR 5
Afaan Oromo	Student family background	Resources	Teacher competencies	Curriculum	Inappropriate time use
Tigrigna	Teacher competencies	Learning environment	School readiness	Learning style & strategies	Curriculum
Amharic	School facility	Family support	Content coverage	Standard of living	Absenteeism
Wolayttatto	Teacher competencies	Student attitude towards MT	Resources	Learner background	Access to preschool
Haddiysa	Lack of school readiness	Family background	Teacher competencies	Availability of reading materials	School environment
Sidamu Affo	Teacher competencies	Learning materials	Conducive environment	Community engagement	Societal traditional attitude
Aff Somali	Curriculum	Teacher competencies	Supplementary materials	Socio-economic	conducive environment

These factors were regrouped into three major classifications: school-related, student-related and out-of-school factors. The description of these categories is as follows:

- **School-related factors** include curriculum, teacher, leadership, and overall school environment such as conduciveness, facility, and supply of materials
- **Student related factors:** include student readiness, prior background, effort, interest, and the like
- **Out-of-school related factors:** include family background, support, and societal views

Table 21. Categorical Summary of the Factors Affecting Early Grade Reading Outcomes

LANGUAGE GROUP	SCHOOL RELATED FACTORS	STUDENT RELATED FACTORS	OUT OF SCHOOL FACTORS
Afaan Oromo	Teacher competencies Resources Curriculum Inappropriate time use		Student family background Resources
Tigrigna	Teacher competencies Learning environment Learning style & strategies Curriculum	School readiness Learning style & strategies	
Amharic	School facility Content coverage	Absenteeism	Family support Standard of living
Wolayttatto	Teacher competencies Resources	Student attitude towards MT Learner background	Resources Learner background Access to preschool
Haddiysa	Teacher competencies Availability of reading materials School environment	Lack of school readiness	Family background Availability of reading materials
Sidamu Affo	Teacher competencies Learning materials conducive environment		Learning materials Community engagement Societal traditional attitude
Aff Somali	Curriculum Teacher competencies Supplementary materials conducive environment		Supplementary materials Socioeconomic background

Through the group activity, 35 responses were gathered in relation to factors affecting early grade reading outcomes. When clustered, six of them fall within student-related factors, 14 within out-of-school factors, and 22 within school-related factors (seven of the responses were double counted; e.g. reading material can be a school-related or out-of-school-related factor). The results in this table indicate participant perceptions that the dominant category affecting early grade reading outcomes is school-related factors, followed by out-of-school factors such as family background, community engagement, and overall socioeconomic status.

Moreover, a majority of the factors above appear to be associated with teacher competencies (appropriate time use, strategies, etc.). This suggests that participants would prioritize improving teacher competencies as an important contribution to the improvement of early grade students' reading outcomes.

From the lists of factors affecting teacher competencies discussed previously (Table 19), 18 out of 28 of the listed factors have identified training, professional knowledge & development, and attitude & motivation as the most important factors affecting teacher competencies. Likewise, for the factors affecting reading outcomes (Table 21), 22 out of 42 identified factors are related to school-related competencies, and of these, 15 are related to teacher competencies and curricular and learning materials. The assumption here is that participants perceive that if teacher competencies are improved, attitude and commitment could emerge, and the most critical school-related problems could be mitigated. In addition to this, leadership was rated as a factor nine times out of the 35 responses. The teacher competencies and leadership account for close to 90% of the factors. Hence, in addition to improving teacher competencies, suggesting participants perceive it is fundamentally important to improve the skill of leadership. These findings suggest collective implications for improving selection and recruitment, training, and professional development of teachers and school leaders.

The factors related to student reading outcomes as presented above were grouped into three clusters, and further reclassified into nine sub-clusters as illustrated in Table 22 below.

Table 22. Number of Language Groups and Their Rating of Factors (F1 = Most Important Factor)

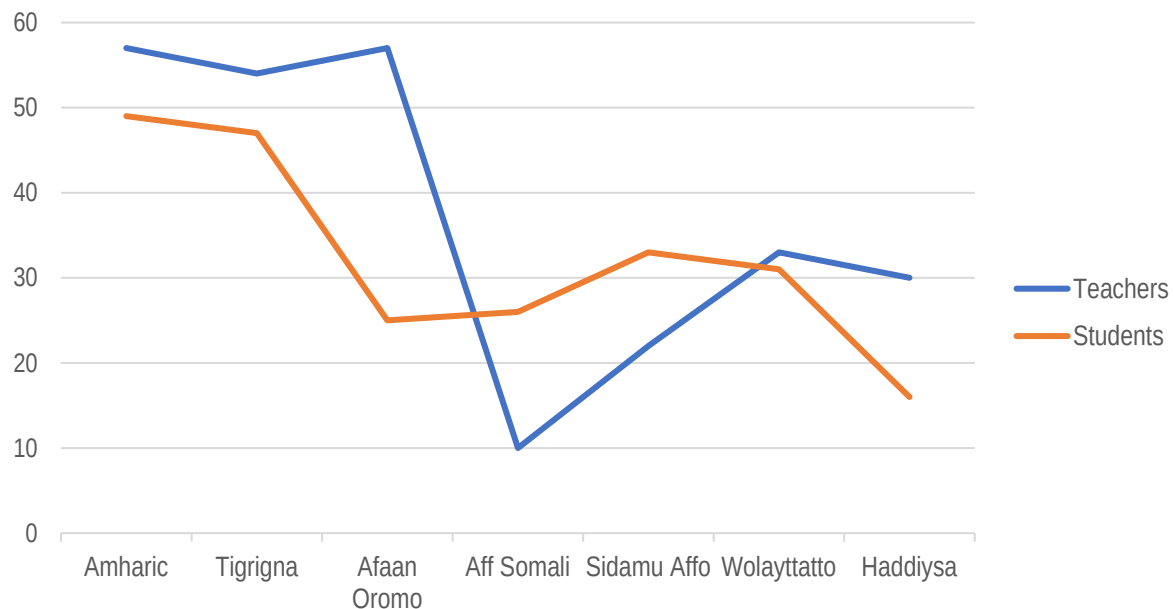
No.	Name of Cluster	Name of Sub-cluster	Frequency	Rating of Factors in Order of Importance				
				F1	F2	F3	F4	F5
1	School-related factors	Teacher competencies	8	3	1	2	1	1
		Curriculum & learning materials	7	1	1	2	2	1
		School facility & learning environment	5	1	1	1		2
		Resources	2		1	1		
2	Student-related factors	Background & preparedness	4	1		1	1	
		Attitude & behavior	2		1			1
3	Out-of-school related factors	Family & community involvement	4	1	2		1	
		Access to preschool	1					1
		Socioeconomic background & awareness	3				2	1
Total			35	7	7	7	7	7

6.2.3 RELATIONSHIP BETWEEN TEACHER AND STUDENT PERFORMANCE

The factors most frequently identified by participants in the workshop as influencing reading outcomes are teacher competencies and curriculum & learning materials. Teacher competencies was identified as factor 1 by three language groups, making it the most important cited factor. In addition, the following charts (Figure 27 and Figure 28) illustrate the pattern of teacher competencies and student performance based on MTTCA and EGRA assessments. This shows the gap in Aff Somali

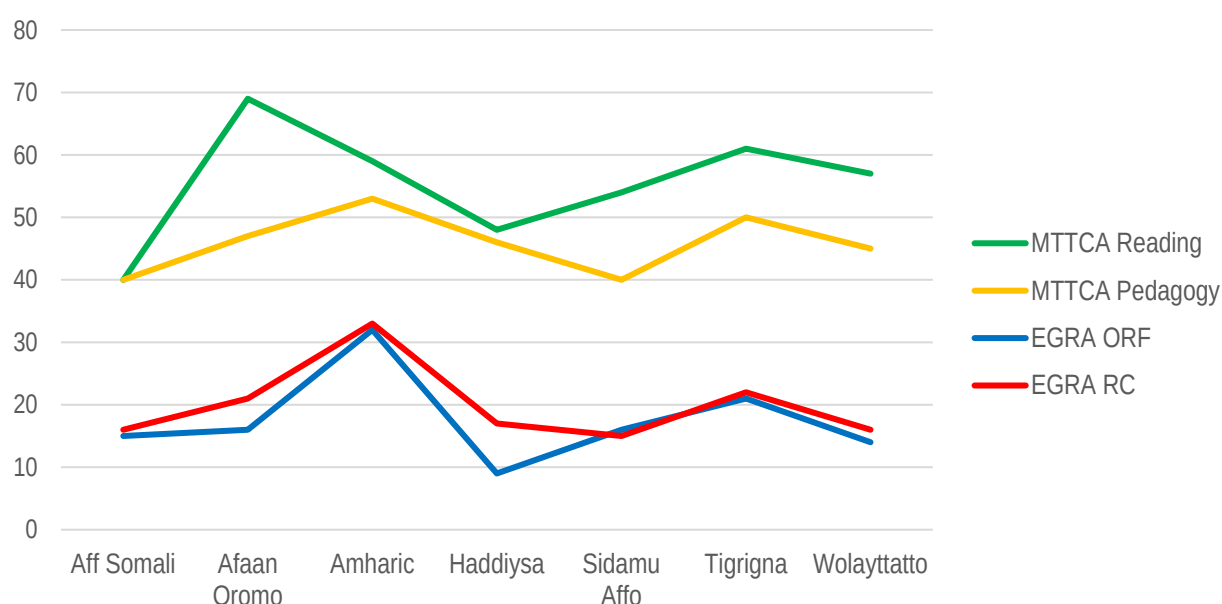
and Sidamu Affo in a sense that percentage of students reaching the top two performance levels was higher than percentage of teachers reaching the top two performance levels.

Figure 27. Percentage of Teachers and Students That Reached Target Performance Levels



The pattern of performance in reading and pedagogy of teachers is compared against the pattern in oral reading fluency and reading comprehension of the students as indicated below (note that the emphasis here is on correspondence between patterns, regardless of graph being based on different vertical axis units). It shows a similar pattern except for Afaan Oromo, at which teacher reading (reading comprehension and grammar and vocabulary) was found to be comparatively better. The overall correspondence between the patterns is indicative of the association between teachers' competencies and students' reading performance, which implies it might be worthwhile to work on improving teacher competencies.

Figure 28. Pattern of Teacher and Student Performance



Note: MTTCA in percent-correct scores, ORF in words per minute, RC in percent-correct scores

6.2.4 GENERAL EXPERT GROUP RECOMMENDATIONS FOR FURTHER CONSIDERATION

Based on the MTTCA scores and the discussions above, the following general recommendations for further consideration are offered by the participating experts:

- Teachers' competencies need to be addressed. There is a need to devise mechanisms for professional development of teachers, including revising the ongoing teacher education of MT teachers.
- Students' home background and parental and community engagement are serious factors. Hence it is important to ensure student learning will be supported by including parents and community stakeholders.
- School background with scarcity of resources and learning materials is a problem. Therefore, it would be beneficial to provide sufficient resources and learning materials:
 - Supply teachers' guides & student textbooks in 1:1 ratio
 - Prepare and utilize appropriate teaching aids
 - Provide sufficient supplemental reading materials (SRMs) in schools
 - Make the teaching learning environment attractive
 - Learners' background was a critical factor. Hence, it is important to make completing preschool a requirement in order to enroll in Grade 1, and important to expand and strengthen (1) the preparedness of students through preschool, and (2) the parent-school relationship.

7. IMPLICATIONS AND RECOMMENDATIONS

The MTTCA results offer critical evidence to inform discussions and decision-making regarding policies, professional development, strategies, and interventions to improve both teacher competencies and student reading performance in primary grades, and as a result, student learning outcomes throughout all stages of education in Ethiopia. To facilitate discussions on the MTTCA results, READ M&E organized Co-interpretation® workshop involving national experts from the regions speaking seven targeted Ethiopian languages. In this workshop, READ M&E presented both student and teacher performance results obtained in the seven languages, and carried out in-depth discussions aimed at formulating regionally contextualized interpretations of the results and recommendations for next steps and action plans.

This concluding chapter summarizes the key findings based on teacher and student performance data referring to main results of the MTTCA 2019 collected from MT teachers in seven Ethiopian languages, key insights regarding relationship between teachers' and students' performance based on the analyses of the MTTCA 2019 and EGRA 2018 data, key insights on deciphering the teacher effect in the context of other factors analyzed, and overview of corresponding implications and recommendations of these findings from READ M&E, informed by the perspectives of national experts.

Considering that the educational factors affecting teacher and student performance are contextually specific, READ M&E organized the additional recommendations from national experts offered through focus group discussions within the Co-interpretation® workshop as presented in the Appendix B.

7.1 KEY RESULTS OF THE MTTCA 2019

In this section we provide a summary of the teacher performance on the competencies assessment and key implications and recommendations for next steps that apply equally to the local languages targeted in the study. They are based on a review of the results from the 2019 administration of the assessment of MT teacher competencies:

1. There was relatively low overall performance of teachers—only 39% of teachers reaching the targeted levels “Proficient or Above.” This indicates that the focus of stakeholders should be drawn to improving teacher competencies. Therefore, MT teacher education has to be reviewed and an additional competencies enhancement package such as CPD has to be enforced.
2. Large variance in teacher performance between languages was observed, from 10% “Proficient and Above” teachers in Aff Somali to 57% “Proficient and Above” teachers in Oromia. This variation dictates focus on the teacher competencies in the context of the MT language nature and background factors relevant in each region. Hence, the teacher education and MT learning materials need to be contextually relevant, and equity of teachers' competencies in different languages sought.
3. There are substantial differences in teacher percent-correct scores between the MTTCA components, where deficiencies were observed in pedagogical knowledge and skills. Teacher performance on the pedagogy components is substantially lower than on reading components in most languages (especially subject pedagogy), with MT teachers in Oromia region demonstrating the largest difference between reading and pedagogy at 22 percent-correct score points. Thus, focus must be given to the MT-based pedagogical knowledge of teachers.

7.2 KEY INSIGHTS ON RELATIONSHIP BETWEEN TEACHERS' AND STUDENTS' PERFORMANCE

To evaluate the importance of MT teacher competencies for development of reading skills of students in early grades, the MTTCA 2019 was administered in the same schools where student reading performance was assessed by EGRA 2018. This enabled various approaches to the evaluation of associations between MT teachers' competencies and student reading skills. and the following key insights are acquired:

1. A strong positive association between student and teacher performance can be observed based on variation across the MT languages, zones, schools, and teachers.
2. In most cases, the languages where teachers are performing at a high level also have high-performing students (correlation at 0.56).
3. There is a significant correspondence between ranking of zones based on their average teacher total scores on MTTCA and ranking of zones based on student ORF performance on EGRA (rank correlation at 0.30).
4. There are significant correlations between school average scores on MTTCA and EGRA, for example, the school average for the total score on teachers' competencies test correlates with the school average for the students' score on familiar words reading, at 0.30.
5. There was a statistically significant difference in EGRA performance between students associated with categories of teachers demonstrating different MTTCA performance levels, showing that students associated with higher performing MT teachers also demonstrate higher reading skills.
6. These findings are aligned with a common notion that "better teachers = better students", however, due to correlational nature of these findings, causal attribution cannot be claimed. Thus, more research evidence targeting the factors that are associated with both teacher and student performance, is needed to allocate prioritization of intervention focus between the competencies of teachers and other factors that affect student performance.

7.3 KEY INSIGHTS ON DECIPHERING THE TEACHER EFFECT IN THE CONTEXT OF HOME AND SCHOOL RESOURCES

READ M&E stands at the position that a rigorous research-based evidence is the most relevant guideline for determining what should be at the center of the theory of change. Based on a set of univariate analyses, MT teacher competencies are found to have strong associations with the reading outcomes of students in their schools. READ M&E also conducted more sophisticated multi-variate analyses of factors associated with student early grade reading performance (analysis of covariance and multiple regression), which included into consideration three categories of variables: MT teacher competencies, home resource indicators, and school resource indicators. The key insights from these analyses are following:

1. All three categories variables (teacher competencies, school, and home resources) have a significant positive relationship with student performance.
2. The strongest association with student performance was exhibited by home resources (receiving help with reading at home), followed by school resources (school has a library) and teacher competencies, which can be also considered as a school resource indicator. Although the teacher competencies factor was ranked at the third place, the size of its unique association with student performance is significant and still relatively large.

3. This finding is in line with ample research on *home literacy environment factors*, which are found to be the strongest predictors of children pre-reading scores and later reading scores. The implication is that the home environment lays the foundation, after which the roles of school resources and teachers' competencies take effect.

7.4 SUMMARY OF IMPLICATIONS AND RECOMMENDATIONS

Combining the findings from the statistical analyses of MTTCA and EGRA data with the perspectives derived from national experts' opinions, the conclusions, implications, and recommendations for steps forward are formulated.

1. Since the advanced quantitative analysis indicated that teacher competencies play a significant role in relation to student performance, it can be recommended to provide enough of professional development opportunities for teachers to improve their competencies to teach literacy skills in the optimal way.
2. Aligned with the quantitative results, the study of experts' opinions also identified the importance of recruitment and selection of teachers for addressing the gap in teacher competencies and subsequently early grade students' reading outcomes.
3. Also, the local experts pointed out that providing support for school leaders and higher officials is important for allowing teachers to utilize their competencies and for students to engage with the learning process. The current leadership was criticized indicating its focus is more administrative than instructional. Hence, educational leadership should be considered by stakeholders as an important driver of making teacher competencies effective in practice, and substantive training support to improve school leadership is recommended.
4. Based on the quantitative research finding that home resources have strongest association with student reading performance, it is suggested to develop systems for supporting parents and communities to learn how to foster literacy development of their children.
5. In addition to quantitative results, qualitative data derived from the opinions of national experts indicated that without proper community engagement, it is difficult to mitigate observed problems in reading outcomes, as home and community factors have a highly significant effect. Therefore, culturally appropriate mechanisms should be designed to increase the engagement of parents and community members with developed guidelines and accountability indicators.
6. Based on the quantitative results, it can be recommended to increase school resources by building school libraries, providing teacher guides and other tools for teachers, providing supplemental reading materials, and motivating students to use the resources available.
7. In concurrence with this quantitative finding, national experts considered that school facilities and school environment are the most important factors affecting teacher competencies and reading performance of students. Thus, these factors should be addressed with sufficient attention in planning steps forward.

In the context of all the research findings and insights described above, the MTTCA represents an invaluable evaluation tool for assessment of teacher outcomes, which combined with the assessment of student reading and learning outcomes provides an invaluable information for guiding educational planning by stakeholders and policy makers. It can also play an important role in identifying strengths and weaknesses of the teacher cadre in targeted regions to inform actions for improvement of MT teacher competencies.

REFERENCES

- Altun, M. (2017). The Effects of Teacher Commitment on Student Achievement: A Case Study in Iraq. *International Journal of Academic Research in Business and Social Sciences*, 7, No. 11.
- Bultosa H., Gadisa T., & Garoma D. (2017). Base Line Survey of Formative Assessment Use among First Cycle Primary School Teachers: The Case of Some Selected Zones of Oromia Region, Sebeta SNE CTE, Finfine Special Zone, Oromia
- Cohen, J. (1977). *Statistical power analysis for behavioral sciences* (Rev. ed.). New York, NY: Academic Press.
- Foorman, B., Beyler, N., Borradaile, K., Coyne, M., Denton, C. A., Dimino, J., et al. (2016). *Foundational skills to support reading for understanding in kindergarten through 3rd grade* (NCEE No. 4008). National Center for Education Evaluation and Regional Assistance Working Paper. Washington, DC: U.S. Department of Education. <http://eric.ed.gov/?id=ED566956>
- Government of Ethiopia, UN Children's Fund, Save the Children, & Education Cluster. (2017). *Education in Emergency Strategic Response Plan – 2017*. Retrieved from https://reliefweb.int/sites/reliefweb.int/files/resources/ethiopoia_eie_strategic_plan_2017.pdf
- Lewis, D. M., Mitzel, H. C., & Green, D. R. (1996). Standard setting: A bookmark approach. In D. R. Green (Chair), *IRT-based standard-setting procedures utilizing behavioral anchoring*. Symposium conducted at the 1996 Council of Chief State School Officers National Conference on Large Scale Assessment, Phoenix, AZ.
- Madaus, G. F. (1992). An independent auditing mechanism for testing. *Educational Measurement: Issues and Practice*, 11, 26–31.
- MOE Ethiopia (2018/19). *Education Statistics Annual Abstract 2011 E.C.* Federal Democratic Republic of Ethiopia, Ministry of Education, October 2019, Addis Ababa.
- MOE Ethiopia (2018). *Ethiopian Education Development Roadmap (2018-30)*. Ministry of Education, Education Strategy Center (ESC).
- Mubarok, H. & Sofiana, N. (2018). Examining the Impact of Teaching Strategies and Reading Habits on Students' Reading Comprehension. *Lensa: Kajian Kebahasaan, Kesusastraan, dan Budaya*, 8, No. 2 DOI: <https://doi.org/10.26714/lensa.8.2.2018.189-202>
- Nag, S., & Perfetti, C. A. (2014). Reading and writing: Insights from the alphasyllabaries of South and Southeast Asia. *Writing Systems Research*, 6(1), 1–9. <https://doi.org/10.1080/17586801.2014.883787>
- National Reading Panel (U.S.) & National Institute of Child Health and Human Development (U.S.). (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*. Washington, DC: National Institute of Child Health and Human Development, National Institutes of Health.

- Pilten, G., & Kuralbayeva, A. (2018). Investigation of Reflecting Reading Comprehension Strategies on Teaching Environment among Preservice Classrooms. *International Electronic Journal of Elementary Education*, 10(4), 397–405. <https://doi.org/10.26822/iejee.2018438129>
- Roelofs, E., & Sanders, P. (2007). Towards a Framework for Assessing Teacher Competence. *European Journal of Vocational Training*, v40 n1 p123-139 2007.
- Share, D. L., & Daniels, P. T. (2015). Aksharas, alphasyllabaries, abugidas, alphabets and orthographic depth: Reflections on Rimzhim, Katz, & Fowler (2014). *Writing Systems Research*, 1, 17-31. doi: 10.1080/17586801.2015.1016395
- USAID Ethiopia, GRN, & MOE Ethiopia. (2015). *Results of the Early Grade Reading Benchmarking Workshop in Ethiopia*. Retrieved from <https://www.globalreadingnetwork.net/sites/default/files/eddata/EGR%20Benchmarking%20Workshop%20Ethiopia%20FINAL.pdf>
- USAID Ethiopia. (2014). READ TA Baseline Assessment Report, Teacher Professional Development and Support Systems: An Analysis of Seven Language Areas in Ethiopia, April 2014.
- USAID Ethiopia. (2018). Early Grade Reading Assessment (EGRA) 2018 Endline Report. Reading for Ethiopia's Achievement Developed Monitoring and Evaluation (READ M&E).
- Wolf, F. M. (1986). *Meta-analysis: Quantitative methods for research synthesis*. Beverly Hills, CA: Sage.

APPENDIX A: SPECIFIC PERFORMANCE LEVEL DESCRIPTORS FOR MOTHER TONGUE TEACHERS

APPENDIX A1: TEACHERS' LANGUAGE KNOWLEDGE AND SKILLS

Table A1. Teachers' Own Knowledge and Skills in the Language They Teach

STANDARD	ELEMENTS	INSUFFICIENT	ACCEPTABLE	PROFICIENT	ADVANCED
Knowledge of and skills in the mother tongue	Reading fluency and comprehension	Teachers at this level exhibit minimal understanding of a text to find out specific and main idea, infer implicit idea of a text. They rarely read with accuracy, rate, and expression.	Teachers at this level exhibit partial understanding of a text to find out specific and main idea, infer implicit idea of a text. They sometimes read with accuracy, rate, and expression.	Teachers at this level exhibit solid understanding of a text to find out specific and main idea, infer implicit idea of a text. They effectively read with accuracy, rate, and expression.	Teachers at this level exhibit a profound understanding of a text to find out specific and main idea, infer implicit idea of a text. They consistently and effectively read with accuracy, rate, and expression.
	Writing	Teachers at this level demonstrate minimal understanding of how to write sentences and paragraphs, and how to produce text in different genres using cohesive devices appropriately.	Teachers at this level demonstrate partial understanding of how to write sentences and paragraphs, and how to produce text in different genres using cohesive devices appropriately.	Teachers at this level demonstrate solid understanding of how to write sentences and paragraphs, and how to produce text in different genres using cohesive devices appropriately.	Teachers at this level demonstrate an in-depth understanding of how to write sentences and paragraphs, and how to produce text in different genres using cohesive devices appropriately.
	Listening comprehension	Teachers at this level demonstrate minimal understanding of listening to identify specific and main ideas. They rarely infer implicit ideas from different sources.	Teachers at this level demonstrate partial understanding of listening to identify specific and main ideas. They sometimes infer implicit ideas from different sources.	Teachers at this level demonstrate solid understanding of listening to identify specific and main ideas. They infer implicit ideas from different sources.	Teachers at this level demonstrate a profound understanding of listening to identify specific and main ideas. They consistently and effectively infer implicit ideas from different sources.
Knowledge of and skills in the mother tongue	Speaking fluency	Teachers at this level demonstrate minimal understanding of how to express their ideas in a logical order and maintain coherence with expected fluency and confidence.	Teachers at this level demonstrate partial understanding of how to express their ideas in logical order and maintain coherence with expected fluency and confidence.	Teachers at this level demonstrate solid understanding of how to express their ideas in a logical order and maintain coherence with expected fluency and confidence.	Teachers at this level demonstrate an in-depth understanding of how to express their ideas in a logical order and maintain coherence with expected fluency and confidence.

STANDARD	ELEMENTS	INSUFFICIENT	ACCEPTABLE	PROFICIENT	ADVANCED
	Vocabulary	Teachers at this level display minimal understanding of using different/appropriate words in different contexts and inferring the meaning of new words.	Teachers at this level display partial understanding of using different/appropriate words in different contexts and inferring the meaning of new words.	Teachers at this level display solid understanding of using different/appropriate words in different contexts and inferring the meaning of new words.	Teachers at this level display a profound understanding of using different/appropriate words in different contexts and inferring the meaning of new words.
	Grammar	Teachers at this level display minimal understanding of the rules, forms, and usage of a language.	Teachers at this level display partial understanding of the rules, forms, and usage of a language.	Teachers at this level display solid understanding of the rules, forms, and usage of a language.	Teachers at this level display a profound understanding of the rules, forms, and usage of a language.

APPENDIX A2: SUBJECT MATTER PEDAGOGY

Table A2. Subject Matter Pedagogy

STANDARD	ELEMENTS	INSUFFICIENT	ACCEPTABLE	PROFICIENT	ADVANCED
Pedagogical knowledge of and skills in mother tongue instruction	Reading	Teachers at this level are ineffective in using a variety of teaching methods to develop elements of pre-comprehension (decoding, phonemic awareness, and fluency) and early comprehension skills in students. They have low motivation to support students to read fluently and identify ideas in text.	Teachers at this level are moderately effective in using a variety of teaching methods to develop elements of pre-comprehension (decoding, phonemic awareness, and fluency) and early comprehension skills in students. They are moderately motivated to use functional teaching methods to support students to read fluently and identify ideas in text.	Teachers at this level effectively use a variety of teaching methods to develop elements of pre-comprehension (decoding, phonemic awareness, and fluency) and early comprehension skills in students. They are highly motivated to use effective teaching methods to support students to read fluently and identify ideas in text.	Teachers at this level use a variety of teaching methods in a highly effective way to develop elements of pre-comprehension (decoding, phonemic awareness, and fluency) and early comprehension skills in students. They have consistently high motivation to use effective teaching methods to support students to read fluently and identify ideas in text.
	Writing	Teachers at this level use ineffective teaching methods to teach writing skills. They have low motivation to support students to express themselves fluently.	Teachers at this level use functional teaching methods to develop students' ability to write letters and sentences that are appropriate for their grade. They are moderately motivated to use effective teaching methods to support students to express themselves fluently.	Teachers at this level effectively use a variety of teaching methods to develop students' ability to write letters and grammatically correct sentences that are appropriate for their grade. They are highly motivated to use effective teaching methods to support students to express themselves fluently.	Teachers at this level use a variety of teaching methods in a highly effective way to develop students' ability to write letters and grammatically correct sentences that are appropriate for their grade. They are highly motivated to make creative use of effective teaching methods to support students to express themselves fluently.

STANDARD	ELEMENTS	INSUFFICIENT	ACCEPTABLE	PROFICIENT	ADVANCED
Pedagogical knowledge of and skills in mother tongue instruction	Listening	Teachers at this level inadequately apply a variety of techniques to identify listening texts for the grade level. They have low motivation to use effective teaching methods to comprehend meanings from listening texts.	Teachers at this level functionally apply a variety of techniques to identify listening texts for the grade level. They are moderately motivated to use effective teaching methods to support students to comprehend meanings from listening texts.	Teachers at this level effectively apply a variety of techniques to identify listening texts for the grade level. They are highly motivated to use effective teaching methods to support students to comprehend meanings from listening texts.	Teachers at this level are highly effectively in applying a variety of techniques to identify listening texts for the grade level. They are exceptionally motivated to make creative use of highly effective teaching methods to support students to comprehend meanings from listening texts.
	Speaking	Teachers at this level inadequately apply techniques to teach speaking tasks. They have low motivation to use effective teaching methods to support students in learning speaking lessons.	Teachers at this level functionally apply a variety of techniques to teach speaking tasks. They are moderately motivated to use effective teaching methods to support students in learning speaking lessons.	Teachers at this level effectively apply a variety of techniques to teach speaking tasks. They are highly motivated to use effective teaching methods to support students in learning speaking lessons.	Teachers at this level are highly effective in applying a variety of techniques to teach speaking tasks. They are exceptionally motivated to make creative use of highly effective teaching methods to support students in learning speaking lessons.
Pedagogical knowledge of and skills in mother tongue instruction	Vocabulary	Teachers at this level use ineffective teaching methods to teach vocabulary. They have low motivation to use effective teaching methods to support students to accomplish vocabulary tasks.	Teachers at this level use somewhat effective vocabulary teaching methods. They are moderately motivated to use effective teaching methods to support students to accomplish vocabulary tasks.	Teachers at this level use a variety of effective vocabulary teaching methods. They are highly motivated to use effective teaching methods to support students to accomplish vocabulary tasks.	Teachers at this level use a variety of highly effective vocabulary teaching methods. They are consistently and exceptionally motivated to make creative use of highly effective teaching methods to support students to accomplish vocabulary tasks.

STANDARD	ELEMENTS	INSUFFICIENT	ACCEPTABLE	PROFICIENT	ADVANCED
	Grammar	Teachers at this level use ineffective methods to teach grammar. They have low motivation to support students to accomplish grammar tasks.	Teachers at this level use somewhat effective methods to teach grammar. They are moderately motivated to support students to accomplish grammar tasks.	Teachers at this level use effective methods to teach grammar. They are highly motivated to support students to accomplish grammar tasks.	Teachers at this level use highly effective methods to teach grammar. They are exceptionally motivated to make creative use of highly effective teaching methods to support students to accomplish grammar tasks.
Pedagogical knowledge of and skills in mother tongue instruction	Culture	Teachers at this level have poor understanding of how to use culture in language teaching and are limited in their use of culture. Thus, they inadequately relate a society's culture to their teaching. They are ineffective in teaching a language by comparing its dialects.	Teachers at this level have partial understanding of how to use culture in language teaching. They have partial pedagogical skills in using a society's culture in their teaching. They are somewhat effective in teaching a language by comparing its dialects.	Teachers at this level have solid understanding of how to use culture in language teaching. They effectively use a society's culture in their teaching. They effectively teach a language by comparing its dialects.	Teachers at this level have an in-depth understanding of how to use culture in language teaching. They are highly effective in using a community's culture in their teaching. They are highly effectively in teaching a language by comparing its dialects.
	Literature	Teachers at this level have poor understanding of how to use literature in language teaching. They have limitations in using pedagogical skills to incorporate literary works into their language teaching. They inadequately use literary works in the teaching-learning process.	Teachers at this level have partial understanding of how to use literature in language teaching. They have some ability to use pedagogical skills to incorporate literary works into their language teaching. They have a partial ability to use literary works to teach language skills.	Teachers at this level have solid understanding of how to use literature in language teaching. They effectively use different literary works (like folktales, stories, etc.) in their teaching, and they integrate them well to teach different language skills. They create some of their own short stories and use them effectively in the teaching-learning process.	Teachers at this level have an in-depth understanding of how to use literature in language teaching. They are highly effective in using different literary works in their teaching. They are highly capable of relating literary works to different teaching occasions. They often create their own literary materials and are highly effective in using them to provide context for teaching different language skills.

APPENDIX A3: GENERAL PEDAGOGY

Table A3. General Pedagogy

STANDARD	ELEMENT	INSUFFICIENT	ACCEPTABLE	PROFICIENT	ADVANCED
Knowledge of students and how they learn	Development: physical, social, intellectual, and emotional development of students	Teachers at this level exhibit poor understanding of students' physical, social, intellectual, and emotional development and are ineffective in using this understanding in their teaching.	Teachers at this level exhibit partial understanding of students' physical, social, intellectual, and emotional development and are somewhat effective in their teaching.	Teachers at this level exhibit solid understanding of students' physical, social, intellectual, and emotional development and use this understanding effectively in their teaching.	Teachers at this level exhibit an in-depth understanding of students' physical, social, intellectual, and emotional development, and use this understanding in their teaching with high effectiveness.
	Background: linguistic, culture, religion, socioeconomic, gender, equity	Teachers at this level make poor attempts to treat students' diversity (linguistic, cultural, religious, socioeconomic, gender) respectfully and equally.	Teachers at this level attempt to treat students' diversity (linguistic, cultural, religious, socioeconomic, gender) respectfully and equally.	Teachers at this level adequately treat students' diversity (linguistic, cultural, religious, socioeconomic, gender) respectfully and equally.	Teachers at this level very effectively treat students' diversity (linguistic, cultural, religious, socioeconomic, gender) respectfully and equally.
	Needs: identifies teaching that meets the needs/styles of students with special needs	Teachers at this level exhibit inadequate skills in using an inclusive teaching style and providing customized activities for gifted students/ students with special needs.	Teachers at this level exhibit basic skills in using an inclusive teaching style and providing customized activities for gifted students/students with special needs.	Teachers at this level exhibit effective skills in using an inclusive teaching style and providing customized activities for gifted students/students with special needs.	Teachers at this level exhibit highly effective skills in using an inclusive teaching style and providing customized activities for gifted students/students with special needs.
	Style: learning style/nature of learning	Teachers at this level are inadequate in identifying different learning styles, providing different learning experiences, and engaging students to adjust to their learning styles.	Teachers at this level partially identify different learning styles, provide different learning experiences, and engage students to adjust to their learning styles.	Teachers at this level adequately identify different learning styles, provide different learning experiences, and effectively engage students to adjust to their learning styles.	Teachers at this level are highly effective in identifying different learning styles, providing different learning experiences, and effectively engaging students to adjust to their learning styles.

STANDARD	ELEMENT	INSUFFICIENT	ACCEPTABLE	PROFICIENT	ADVANCED
General pedagogical knowledge and skills	Planning	Teachers at this level have limited or no understanding of the elements of planning. They have limited ability to develop plans for engaging children in learning activities that enhance mother tongue learning.	Teachers at this level have partial understanding of the elements of planning. They develop functional plans for engaging children in learning activities that enhance mother tongue learning.	Teachers at this level have solid understanding of the elements of planning. They develop effective plans for engaging children in learning activities that enhance mother tongue learning.	Teachers at this level have an in-depth understanding of the elements of planning. They develop highly effective plans for engaging children in learning activities that enhance mother tongue learning.
	Identification and utilization of instructional materials and resources	Teachers at this level show little or no skill in identifying relevant instructional materials and technology resources. They demonstrate limited interest in preparing teaching aids and show little or no ability to use them in the teaching-learning process.	Teachers at this level show some skill in identifying relevant instructional materials and technology resources. They demonstrate moderate interest in preparing teaching aids and show a minimally functional ability to use them in the teaching-learning process.	Teachers at this level are effective in identifying relevant instructional materials and technology resources. They prepare teaching aids from locally available materials with some creativity and use them effectively in the teaching-learning process.	Teachers at this level are highly effective in identifying relevant instructional materials and technology resources. They prepare teaching aids from locally available materials with high levels of creativity and use them efficiently in the teaching-learning process.

STANDARD	ELEMENT	INSUFFICIENT	ACCEPTABLE	PROFICIENT	ADVANCED
General pedagogical knowledge and skills	Classroom management	Teachers at this level demonstrate little or no understanding of classroom management. They are incompetent in arranging furniture and room space for use, organizing different activities using available space in the classroom and school, and displaying teaching and learning materials and students' work in the classroom. They are inefficient in encouraging self-discipline and acting immediately to address problems of discipline. They show little or no ability to treat all children in a fair and consistent manner and discourage physical punishment and mental harassment of children.	Teachers at this level demonstrate a functional understanding of classroom management. They are moderately effective in arranging furniture and room space for use, organizing different activities using available space in the classroom and school, and displaying teaching and learning materials and students' work in the classroom. They demonstrate some skill in encouraging self-discipline and acting immediately to address problems of discipline. They show some ability to treat all children in a fair and consistent manner and discourage physical punishment and mental harassment of children.	Teachers at this level demonstrate solid understanding of classroom management. They are effective in arranging furniture and room space for use, organizing different activities using available space in the classroom and school, and displaying teaching and learning materials and students' work in the classroom. They are effective in encouraging self-discipline and acting immediately to address problems of discipline. They have a strong ability to treat all children in a fair and consistent manner and discourage physical punishment and mental harassment of children.	Teachers at this level demonstrate a comprehensive and in-depth understanding of classroom management. They are highly effective in arranging furniture and room space for use, organizing different activities using available space in the classroom and school, and displaying teaching and learning materials and students' work in the classroom. They are highly effective in encouraging self-discipline and acting immediately to address problems of discipline. They have an excellent ability to treat all children in a fair and consistent manner and discourage physical punishment and mental harassment of children.

STANDARD	ELEMENT	INSUFFICIENT	ACCEPTABLE	PROFICIENT	ADVANCED
General pedagogical knowledge and skills	Using a variety of assessment techniques	Teachers at this level have an inadequate understanding of the variety of assessment techniques. They are unsuccessful in assessing their students' learning and providing immediate feedback to improve their performance. They are rarely successful in keeping records of students' performance and sharing that information with parents and school management.	Teachers at this level have limited understanding of the variety of assessment techniques. They are moderately effective in assessing their students' learning and providing immediate feedback to improve their performance. They are somewhat successful in keeping records of students' performance and sharing that information with parents and school management.	Teachers at this level have an adequate understanding of the variety of assessment techniques. They are effective in assessing their students' learning and providing immediate and effective feedback to improve their performance. They are successful in keeping records of students' performance and sharing that information with parents and school management.	Teachers at this level have an in-depth understanding of the variety of assessment techniques. They are highly effective in assessing their students' learning and providing immediate feedback to improve their performance. They are remarkably successful in keeping records of students' performance and sharing that information with parents and school management.
	Interpersonal skills	Teachers at this level are barely approachable to children, colleagues, and parents. They show minimal respect and care towards students, colleagues, and parents and do not recognize or appreciate their contributions. They have a poor relationship with others and demonstrate ineffective communication skills.	Teachers at this level are somewhat approachable to children, colleagues, and parents. They show respect and care towards students, colleagues, and parents; recognize and appreciate their contributions; and collaborate with them. They are somewhat effective in communicating their ideas and understanding others.	Teachers at this level are approachable to children, colleagues, and parents. They show respect and care towards students, colleagues, and parents; recognize and appreciate their contributions; and collaborate with them. They are effective in communicating their ideas and understanding others.	Teachers at this level are easily approachable and show high levels of respect and care towards students, colleagues, and parents. They recognize and highly appreciate the contributions of students, colleagues, and parents, and collaborate with the school community. They are exceptionally effective in communicating their ideas to others and understanding others.

STANDARD	ELEMENT	INSUFFICIENT	ACCEPTABLE	PROFICIENT	ADVANCED
General pedagogical knowledge and skills	Continuing professional development research	Teachers at this level have limited understanding of continuous professional development. They are not very interested in continuous self-upgrading. They are not able to evaluate the overall teaching-learning process and identify existing problems in the classroom environment. They are not committed to conducting professional research to solve problems and improve the teaching-learning process.	Teachers at this level have partial understanding of continuous professional development. They are moderately interested in continuous self-upgrading. They have a functional ability to evaluate the overall teaching-learning process and identify existing problems in the classroom environment. They are somewhat committed to conducting professional research to solve problems and improve the teaching-learning process.	Teachers at this level are highly motivated and frequently engaged in continuous self-improvement. They are highly competent in evaluating the overall teaching-learning process and identifying existing problems in the classroom environment. They are committed to conducting professional research to solve problems and improve the teaching-learning process. They have a positive attitude towards continuing professional development.	Teachers at this level have an excellent understanding of continuous professional development. They are extraordinarily interested in continuous self-upgrading. They are exceptionally competent in evaluating the overall teaching-learning process and are highly effective in identifying existing problems in the classroom and school environment. They are extremely committed to conducting professional research to solve problems and improve the teaching-learning process.

APPENDIX B: SUMMARY OF IMPLICATIONS AND RECOMMENDATIONS BASED ON QUALITATIVE DATA

B1 AFAAN OROMO IMPLICATIONS & RECOMMENDATIONS FOR NEXT STEPS

Professional development (including training) and leadership support were reported to be the critical factors affecting teacher competencies. Ongoing teacher education, CPD, and leadership support need to be reviewed, as they would continuously affect reading performance of students. We are informed that the Bureau of Education (BoE) has started to integrate teacher education modules (14 modules have already been revised) into teacher development programs. It would therefore be wise to strengthen these modules with a focus on mother tongue instruction and revise the leadership system in such a way that school leaders can focus on instructional support to teachers in addition to their administrative duties. These changes would require the issuance of policy directives, guidelines, and strong monitoring and evaluation systems.

There are also differences in MT teachers' competencies across zones within the same language, which needs further investigation.

B2 AMHARIC IMPLICATIONS & RECOMMENDATIONS FOR NEXT STEPS

From the aforementioned recommendations in the context of the Amhara region, the following policy implications and next steps are offered.

Teacher commitment was identified as the most important factor affecting teacher competencies. This is believed to have a cascading effect, and teacher commitment has to be one focus of policy. Therefore, the issue of commitment of teachers needs to be addressed in reform activities, and these need to be incorporated into the education sector development program to guide action for the coming years. These can be achieved through strict monitoring and support, performance appraisal, and at times rewards and incentives. When teachers are committed, they make necessary effort to support their students' learning.

Teacher competencies are affected by the selection process of would-be teachers, because candidates who have poor academic achievement and poor attitude towards teaching are enrolling in CTEs. Hence, teacher selection needs to be prioritized, and accordingly improved guidelines for selection and recruitment, especially for MT, need to be issued.

Teacher competencies are also affected by pre-service and in-service trainings. The trainings are not based on a gap analysis, and there are frequent changes of training modality (i.e., linear, cluster, generalist, and specialist). There is also poor alignment of college curricula with primary school curricula.

Teacher competencies are also affected by poor school resources, negatively influencing students' motivation for learning. Therefore, school facilities need to be improved.

Though the effect of teacher competencies is evident, poor student academic performance can be partly explained by poor family support. Therefore, reading competencies outside of school need to be monitored and parent involvement and support in the learning of their children need to be strengthened.

Absenteeism was reported to be a factor affecting the student reading outcome, which ultimately affects learning of other subjects as well. Hence, strict regulation needs to be put in place and implementation monitored periodically.

B3 TIGRIGNA IMPLICATIONS & RECOMMENDATIONS FOR NEXT STEPS

Professionalism and professional development were among the key factors affecting teacher competencies in the Tigrigna language zones. Teacher competencies were also identified as the most important factor affecting student reading outcomes. Therefore, the policy emphasis should focus on improving professionalism and professional development as a way to improve teacher competencies.

The learning environment was reported to affect teacher competencies and reading performance. Therefore, improving the school environment should be another focus.

Curricular issues and preparedness of children are the other reported factors. Therefore, there is a need to systematize preschool learning with primary education to ensure that children are ready before they join Grade I, and to align preschool and primary education through curricular review and amendment.

B4 SIDAMU AFFO IMPLICATIONS & RECOMMENDATIONS FOR NEXT STEPS

Teacher education was identified as the top factor affecting teacher competencies in the Sidama zone. Teacher competencies were also identified as the first factor to affect reading outcomes of students. Hence, students' reading could be improved through the deployment of competent teachers who have developed their competencies through their initial teacher education and through CPD during their daily work (including pedagogical knowledge and skills using CCA). Therefore, careful review needs to be made of teacher education and CPD.

Low motivation of teachers and a school environment that is not conducive to effective learning reinforce one another, backed with poor community engagement. Thus, an action plan has to be designed to promote motivation of teachers, create a conducive school environment, and mobilize community members to engage in the learning of their children.

Quality of learning materials was reported to affect student reading outcomes. Materials review and the publication of additional support materials could help to solve the problem of quality of learning materials. It is therefore essential to conduct a review of the existing learning materials and improve them so that they are contextually relevant and helpful in promoting reading.

B5 WOLAYTTATTO IMPLICATIONS & RECOMMENDATIONS FOR NEXT STEPS

In the Wolayta zone lack of teacher competencies and commitment are at the center of factors affecting student reading performance and outcomes. Improving teacher competencies and commitment is, therefore, a timely call in the context of Wolayta Zone. This could also be associated with poor student attitudes toward learning. Therefore, a holistic solution is required to solve the problem. This should include reviewing the ongoing teacher education program, engaging teachers through CPD, and employing proper measurement and evaluation.

Quality of learning materials and reference materials were cited as factors affecting student reading outcomes. Hence, it would be worthwhile reviewing the existing curricular materials (textbooks and teacher guide, including reference materials), amend gaps, and make sure that adequate supplies of materials are provided to teachers.

Teacher background was another factor, calling for training a sufficient number of MT teachers with an appropriate linguistic background who are trained to teach the MT effectively.

Access to preschool learning, and the quality of preschool, determines the preparedness of children. Thus, further emphasis should be placed on preschool learning.

B6 HADDIYSA IMPLICATIONS & RECOMMENDATIONS FOR NEXT STEPS

Teacher education, including selection and recruitment, are seriously hampering teacher competencies in the Hadiya zone. Selection, recruitment, and ongoing teacher education have to be on the policy agenda. Therefore, teacher education for MT teachers needs critical investigation to make it contextually relevant and effective.

Leadership is fundamental to success in MT learning. Leadership which ignores instructional needs in schools has a negative impact on the effectiveness of teachers and student learning outcomes. Leadership in education therefore needs to be addressed and school leaders need to be selected on the basis of merit from among those who have proven competencies to improve school function and learning outcomes.

The link between preschool and primary schools needs to be defined appropriately. As a way forward, no child should be enrolled in primary school without having preschool certification .

B7 AFF SOMALI IMPLICATIONS & RECOMMENDATIONS FOR NEXT STEPS

Poor teacher competencies, attitude, and preparation together with poor leadership, were identified as major factors affecting student reading outcomes in the Somali region. As the region with the lowest MTTCA score, Student reading outcomes are unlikely to improve until these issues are resolved. Providing more effective support to teachers is critically important and it is essential to change teacher-related variables to improve teacher competencies, through which children can improve learning outcomes. Environment is a key factor affecting teacher competencies and student reading outcomes. Thus, it would be wise to devise a contextually relevant learning system rather than completely depend on a mainstream approach.