



USAID | PHILIPPINES
FROM THE AMERICAN PEOPLE

EdData II

DEP/AME: Philippines Analytic Support Services for Early Grade Reading (PhilEd Data II)

Component 2: Early Grade Reading Assessment
Results: A cross-language look at MTB-MLE
implementation in the Philippines

EdData II Technical and Managerial Assistance, Task Number 15
Contract Number: AID-492-M-12-00001
July 30, 2014

This publication was produced for review by the United States Agency for International Development. It was prepared by RTI International.

DEP/AME: Philippines Analytic Support Services for Early Grade Reading (PhilEd Data II)

Component 2: Early Grade Reading Assessment Results: A cross-language look at MTB-MLE implementation in the Philippines

Education Data for Decision Making (EdData II)
Task Order No. 15

Prepared for
USAID/Washington
Mitch Kirby
Senior Education Advisor, Asia Bureau
Contracting Officer's Technical Representative
Data for Education Programming/Asia and Middle East
1300 Pennsylvania Avenue NW
Washington, DC 20523

Prepared by
Sarah Pouzevara, Joseph DeStefano, Chris Cummiskey and Jennifer Pressley
3040 Cornwallis Road
Post Office Box 12194
Research Triangle Park, NC 27709-2194

RTI International is a trade name of Research Triangle Institute.

The author's views expressed in this publication do not necessarily reflect the views of the United States Agency for International Development or the United States Government.

Table of Contents

List of Tables	iv
List of Figures	iv
Abbreviations	vi
Acknowledgments.....	vii
Map of the Philippines and regions of data collection.....	viii
I. Introduction	1
A. Background.....	1
B. DepED's MTB-MLE	2
C. Study design and methodology.....	7
II. Findings	11
A. Is MTB-MLE being implemented?	11
B. What are the predominant teaching methods being used?.....	16
C. How well are children learning how to read in the MT?	20
D. How well are children acquiring L2 and L3?	39
E. What factors are associated with reading skills?	46
III. Summary and Conclusions	51
A. Summary of EGRA results by language.....	51
B. Limitations	58
C. Conclusions.....	61
D. Recommendations	62
Annex A: Summary of results from 2013 EGRA study	66
Annex B: MTB-MLE Index.....	71
Annex C: Sampling methodology	74

List of Tables

Table 1:	Information on languages of 2014 EGAs.....	2
Table 2:	EGRA instrument subtests in the Philippines.....	8
Table 3:	Study sample.....	10
Table 4:	National curriculum standards related to reading	11
Table 6:	Summary of select indicators related to MTB-MLE implementation, by language and grade	19
Table 7:	Correlations between subtests.....	32
Table 8:	Reading accuracy: Percent correct out of total attempted	35
Table 9:	Minimum number of words attempted for each comprehension question, by language.....	35
Table 10	ORF and comprehension.....	38
Table 11:	G2 teacher expectations for Filipino reading.....	39
Table 12:	Correlations between subtests.....	41
Table 13:	Summary of EGRA results: Average by subtest and grade	44
Table 14:	Summary of EGRA results: Proportion of zero scores by subtest and grade.....	45
Table 15:	Select indicators for Cebuano	52
Table 16:	Select indicators for Hiligaynon	53
Table 17:	Select indicators for Ilokano	55
Table 18:	Summary of key indicators for Maguindanaoan.....	57
Table 19:	Breakdown of classroom sample by type of classroom.....	61
Table 20:	Reading fluency and comprehension as benchmarks	65
Table C-1.	Number of Excluded schools from the 2013 BIES school list and reasons why schools were excluded.	74
Table C-2:	Sample methodology for the 2014-Four Region EGRA-SSME.....	75
Table C-5:	Counts by Region-Grade of the completed assessments/questionnaires at the student level.	76

List of Figures

Figure 1:	Scores on the MTB-MLE Implementation Index—Teachers.....	13
Figure 2:	Scores on the MTB-MLE Implementation Index—Materials	14
Figure 3:	Scores on the MTB-MLE Implementation Index—Instruction.....	15
Figure 4:	Pupil language activity during MT lesson	17
Figure 5:	Classroom configuration during MT lesson	18
Table 5:	Percent of classroom observation where no instruction was taking place	18
Figure 6:	Average percent correct responses on PA and LC tasks	21
Figure 7:	Scores on listening comprehension and phonemic awareness exercises, by grade and language	21
Figure 8:	Phonemic awareness: Distribution of scores by grade and language	22
Figure 9:	Listening comprehension: Distribution by grade and language	23
Figure 10:	Letter-sound identification scores, by region and grade	26
Figure 11:	Letter-sound identification: Distribution of scores, by grade and language	26
Figure 12:	Correct non-words per minute: Average scores by region and grade, with and without zero scores.....	28
Figure 13:	Non-word reading: Distribution of scores, by grade and language	29

Figure 14:	Correct familiar words per minute: Average scores by region and grade, with and without zero scores.....	30
Figure 15:	Familiar word reading: Distribution of scores, by grade and language	31
Figure 16:	Oral reading fluency (with and without zero scores), by language and grade.....	33
Figure 17:	Oral reading fluency: Distribution of scores, by grade and language.....	34
Figure 18:	Reading comprehension: Average percent correct by language and grade.....	36
Figure 19:	Distribution of comprehension scores (percent correct out of comprehension questions attempted) by language and grade.....	37
Figure 20:	Comprehension and oral reading fluency	38
Figure 21:	Filipino—Proportion of zero scores.....	40
Figure 22:	Filipino familiar word reading: Correct words per minute by language	40
Figure 23:	Teacher expectations for English reading acquisition	42
Figure 24:	English vocabulary: Percent zero scores by language	43
Figure 25:	English vocabulary: Percent of correct responses by language	43
Figure 26:	Summary of differences in performance by gender, across languages and subtests (percentage point difference, girls-boys).....	47
Figure 27:	Differences in oral reading associated with home factors	48
Figure 28:	Differences in average oral reading fluency associated with school-related factors.....	49
Figure 29:	Odds ratios for variables associated with being in the top 25% of readers	50
Figure A-1:	Comparison of Filipino scores across regions	66
Figure A-2:	Comparison of English scores across regions.....	66

Abbreviations

ARMM	Autonomous Region of Muslim Mindanao
cfwpm	correct familiar words per minute
clspm	correct letter sounds per minute
cnwpm	correct non-words per minute
cwpm	correct words per minute
DepED	Department of Education
EdData II	Education Data for Decision Making
EGRA	early grade reading assessment
G	grade
IRR	inter-rater reliability
KFW	Filipino language commission
LC	learning comprehension
MT	mother tongue
MTB-MLE	mother-tongue based, multilingual education
OFW	overseas foreign worker
PA	phonemic awareness
PhilEd Data	Philippines Analytic Support Services for Early Grade Reading
Phil-IRI	Philippines Informal Reading Inventory
Q	quarter
SES	socio-economic status
USAID	United States Agency for International Development

Acknowledgments

This study was carried out under the auspices of USAID's Data for Education Planning in Asia and the Middle East (DEP/AME) task order to the Education Data for Decision Making (EdData II) Blanket Purchase Agreement for which RTI International is the contractor.

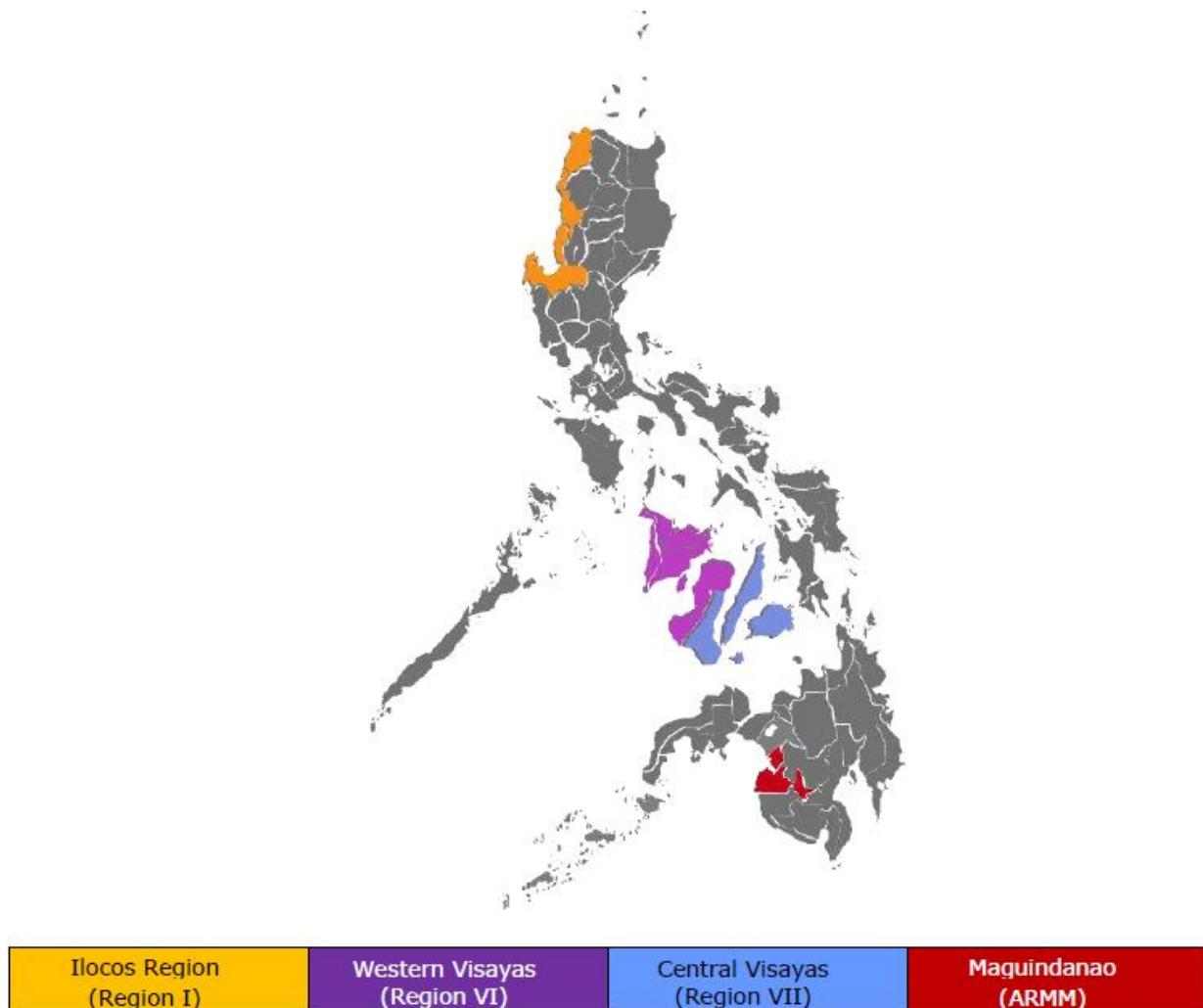
The work in the Philippines however could not have been completed without the interest, support and commitment of the Education Office of USAID/Philippines, in particular Robert Burch, Mirshariff Tillah and Marie Antoinette Reyes-Hayles. Likewise, the Philippines Department of Education provided the leadership and committed the time and effort necessary to develop the instruments, collect the data and analyze and interpret the results of this study. In particular we wish to acknowledge Undersecretary for Programs and Projects Dr. Dina S. Ocampo and the staff within her office who are committed to improving reading outcomes in the Philippines. This especially includes Dr. Marilyn D. Dimaano, Director of the Bureau of Elementary Education (BEE), Dr. Rosalina J. Villaneza, Focal Person for MTB-MLE in BEE, and Jocelyn Tuguinayo, Senior Education Program Specialist in BEE. The Regional Directors and other Department of Education staff at the regional level in Regions I, VI, VII and ARMM made invaluable contributions to the development of assessment and data collection instruments in each regional language and provided the person-power and logistical support necessary to organize all the field data collection in their respective regions, in particular Dina Bonao, Nerio Benito Eseo, Nena Meñoza and Saada Tubing. We also thank the Regional Directors Tersita M. Velasco, John Arnold S. Siena, Carmelita T. Dulangon, and Jamar M. Kulayan.

In addition to the contributions of DepED, the field work would not have been possible without the collaboration of our partner Taylor Nelson Sofres (TNS), in particular Yazmin Tolentino, Joel Flores and Mao Ocampo provided leadership on the part of TNS in the development of the instruments, training of assessors and management of field operations. TNS field staff gamely traveled far and wide to supervise and support data collection in the four regions.

Our other partner on all the EdData II work in the Philippines has been SEAMEO INNOTECH. For this study SEAMEO helped organize the final policy dialogue workshop.

Lastly, and perhaps most importantly, we acknowledge the tireless efforts of Rufino Jamilano, RTI International's in-country task coordinator for all EdData II activities. Rufino's constant diligence and commitment made it possible to collaborate productively with all the players involved in this and previous activities.

Map of the Philippines and regions of data collection



Shaded areas designate the areas where data collection for this report took place. The region boundaries and areas where the languages are spoken may be different than the boundaries represented on this map.

I. Introduction

A. Background

To read and understand a simple text is one of the most fundamental skills a child can learn; in 2014 it should be the responsibility of a functioning primary school system to provide at least sufficient reading skills to the children it serves. Evidence indicates that learning to read both *early* and at a sufficient *rate* are essential for learning to read well, and consequently understanding what is read. Acquiring literacy becomes more difficult as learners grow older; those who do not learn to read in the first few grades are more likely to repeat and eventually drop out (Early Grade Reading Assessment [EGRA] Toolkit, RTI International, 2009).

The Government of the Philippines has placed high priority on improving literacy and is currently undertaking comprehensive reform initiatives including curricular reforms and implementation of a mother-tongue based multilingual education (MTB-MLE) program in 19 national languages. In support of these reforms, the United States Agency for International Development (USAID)/Philippines and the Philippines Department of Education (DepED) initiated the USAID-funded *Strengthening Information for Education Policy, Planning, and Management in the Philippines* project (PhilEd Data I) in April 2012. During 2012 and 2013, PhilEd Data provided technical assistance and support to several activities, one of which was the implementation of a national, sample-based evaluation of reading in the early grades of elementary education. The principal outcomes of this component were:

- EGRA instruments were adapted to English (Grade 3), Filipino (Grade 3), and Ilokano (Grade 1).
- The instruments were used to collect data from a nationally representative sample in Grade 3 (6 regions, 2,400 students) and a regionally-representative sample in Grade 1 (1 region, 500 students).
- Results were presented to and discussed with a range of stakeholders in July 2013.

The results provided baseline data on Grade 3 students' reading achievement and language readiness under the previous (prior to adopting mother tongue-based instruction) national curriculum. PhilEd Data I also explored the relationship of these skills to selected factors in students' school and home environment. This evidence can be used to inform policy and the implementation of appropriate interventions in schools and classrooms across the country, as well as to monitor the effect of MTB-MLE in years to come. The regional language EGRA for Ilokano was the first rigorous and large-scale assessment of student learning in one of the mother-tongue languages. The data on Grade 1 progress towards learning to read in the mother tongue is a starting point from which DepED may develop benchmarks and standards for reading progress in the mother tongue and improve teaching methods and materials based on evidence of the effectiveness of existing inputs.

It is in this context that the **2014 Philippines EGRA study** (PhilEd Data II) is expanding the knowledge base for the MTB-MLE program by conducting a comparable exercise in three additional mother-tongue languages—Cebuano¹, Hiligaynon, and Maguindanaoan (**Table 1**). Another representative set of data was collected for Ilokano, which will provide additional

¹ The local name is Sinugbanong Binisaya. For convenience, we will refer to the language as Cebuano.

information on whether quality of MT instruction is improving with implementation experience, and how much children improve their reading skills from Grade 1 to Grade 2. The 2014 study used the same EGRA instrument as 2013—with appropriate adaptions—to observe and measure reading skills in five key areas:

- Alphabetic principle
- Phonemic awareness
- Vocabulary
- Fluency
- Comprehension

The 2014 fieldwork is sponsored by USAID/Washington under the Education Data for Decision-Making (EdData II) program, and technical assistance is provided by RTI International and a local survey research partner, TNS Global².

Table 1: Information on languages of 2014 EGRAs

Region #	Region Name	Language	Region population**	Est. speakers nationwide***
ARMM	Autonomous Region in Muslim Mindanao	Maguindanaoan	3,256,140	1,100,000
I	Ilocos (Northern Luzon)	Ilokano	4,748,372	6,920,000
VI	Western Visayas	Hiligaynon	7,102,438	5,770,000
VII	Central Visayas	Sinugbuhanong Binisaya or Cebuano	6,800,180	15,800,000

** Census data from 2010, National Statistical Coordination Board. http://www.nscb.gov.ph/secstat/d_popn.asp

*** Data from www.ethnologue.com, based on 2000 census or SIL data

B. DepED's MTB-MLE

The current MTB-MLE policy follows decades of experimentation with different language policies and pilot programs each aiming to find the right balance between protecting and valuing indigenous languages (mother tongues), while also allowing Filipinos from across the nation to communicate using a common language (Filipino) and at the same time creating opportunities for international opportunity through development of fluency in an international language (English).³ These experiments revealed varied yet generally positive results supporting the notion that children learn best when instruction is in the language they are most familiar with, while also showing this does not inhibit their ability to progressively learn other languages.

In July 2009 DepED issued Order No. 74: “Institutionalizing Mother Tongue-Based Multilingual Education (MTB-MLE).” The Order requested the use of the learner’s mother

² For more information, contact Sarah Pouzevara, RTI International (soucez@rti.org).

³ For example, First Iloilo Experiment (1948–1954); the Cebu Experiment (pre-1960s); the Antique Experiment (1952); the First Rizal Experiment (1953–1959); another Rizal experiment (1960–1966); the First Language Component Bridging Program (FLC-BP) Pilot Project in Ifugao (1986–1993); the Lubuagan Multilingual Education Program (1998–); the Lingua Franca Project (1999–2001); the Culture-Responsive Curriculum for Indigenous People–Third Elementary Education Project (2003–2007); the Double Exposure in Mathematics Initiative of Region IV-B (2004–2007); and others.

tongue in improving learning outcomes and promoting Education for All—a tenet also highlighted in President Benigno Aquino's 10-Point Education Agenda when he took office in 2010. The Philippine DepED cites four forms of child development fortified by its MLE program: Language, Cognitive, Socio-cultural, and Academic (Department of Education, Order No. 16, 2012). This marks a shift from an exclusively bilingual orientation to a multilingual one, in which literacy in multiple languages is a key goal, as well as a departure from previous orientations where the mother tongue was seen as merely a bridge to learning other languages.⁴

Following the release of DepED Order 74, a Strategic Plan for implementation of the MLE program was completed in 2010. The Strategic Plan laid out specific activities to be undertaken, including advocacy, pre- and in-service teacher training, materials development, policy development, resource mobilization, and assessment, monitoring and evaluation.⁵ In school year (SY) 2011–2012, over 900 schools across the country began using one of 12 major languages as media of instruction (MOIs) in Grade 1. In SY 2012–2013, these pioneer schools proceeded to do the same in Grade 2; simultaneously, all public schools began implementing the program in Grade 1, as the pioneer schools had done the year before. In SY 2013–2014, at least 7 languages were added as MOIs in Grade 1, while the first 12 languages were being used in Grade 2 classrooms across the country and in the Grade 3 classrooms of pioneer schools. In SY 2014–2015, all public schools will be implementing one of several mother tongues in Grades K–3, both as an MOI and as a separate language subject.

While there are broad policy guidelines and support for training and materials development from the central level, regions also have a great deal of autonomy in the level of support that they provide to schools through the regional MTB-MLE coordinators. For example, the efforts that DepED has undertaken since 2012 to expand MTB-MLE nationwide include:

- Curriculum reforms
- Developing a corps of trainers to deliver in-service training
- Establishment of regional and division mother-tongue coordinators
- Writing teachers' guides and learner materials for mother-tongue literacy and distributing them to schools
- Developing EGRA in the languages of instruction
- Piloting EGRAs in schools and training teachers on EGRA
- Providing training (EGRA-ARATA [Assessment Results Associated Teaching activities]) to teachers in all regions, March 2013

The use of EGRA as a classroom based tool, a regional monitoring tool, or a standardized national assessment tool is still being negotiated in the context of a national assessment framework which is still in the process of development at the time of this publication.

The following sections of this report highlight the ways in which MTB-MLE has been implemented thus far in the four languages of assessment.

⁴ For more information see Firth McEachern's December 31, 2013, PhilEd Data report, Local languages, literacy, and assessment in the Philippines: Implications for the Early Grade Reading Assessment instrument development and use (available at www.eddataglobal.org/documents/index.cfm?fuseaction=pubDetail&id=560).

⁵ Republic of the Philippines, Department of Education. Mother Tongue Based Multilingual Education Strategic Plan. 13 February 2010

USAID/Basa Pilipinas

In Region I and Region VII the USAID/Basa Pilipinas program has been supporting MTB-MLE through a variety of inputs, and using EGRA as a measurement tool. In the period prior to this data collection, all Grade 1 and 2 teachers in the target divisions of La Union and Cebu Province received training in Enhanced Reading Instruction. The trainings reviewed in depth the domains of literacy, and teachers were oriented on why and how to approach the tasks of oral language development, phonological awareness, and vocabulary development.

Questioning tactics were explored as a component of effective classroom practice. Teachers developed materials for use in the classroom, including read alouds and pedagogical tools. This training was done in the first quarter of the school year and also included a discussion between DepED leadership and teachers on the K to 12 curriculum and MTB-MLE. In November 2013, Grade 3 teachers were also trained in aspects of teaching reading and writing, and they received read-aloud books in the mother tongues for use in the classroom.

All trainings incorporated demonstration videos and addressed strategies for assessing student performance in the classroom. The trainings covered use of the mother tongue, but not exclusively. The content of the trainings was designed to apply to all three languages (mother tongue, English and Filipino) as well as to bridging strategies. During the training of trainer activities, teachers in other divisions in the two regions were involved; therefore they have also had some initial exposure to the content.⁶

The sample sizes used in this study do not allow us to isolate the effect of the USAID/Basa program, and indicators specific to the program were not included in the survey. It is nevertheless important to consider the degree to which each region has been exposed to specific reading training programs, materials, and the EGRA instrument and methodology. The following sections of this report describe region-by-region implementation of MTB-MLE, as provided by the regional MTB-MLE coordinators. Section II, further below, will describe the findings from survey instruments (teacher questionnaires and classroom observations), which were designed to measure elements of implementation of MTB-MLE. Finally, Section III provides information, where available, on indicators that show a statistically significant association with reading scores.

Region I (Ilocos Region) – Ilokano

The Ilocos region in northwestern Luzon includes four provinces, Ilocos Norte, Ilocos Sur, La Union, and Pangasinan. The language, Ilokano, is also spoken outside of Ilocos region, and within the region some other language groups are prevalent, particularly Pangasinan in the Pangasinan province. Within the region, words used, and pronunciation, can vary between the north and south.

The region has a well-developed MTB-MLE



⁶ Description provided by USAID/Basa Pilipinas (Karen Cassidy)

program, and has been pioneering mother-tongue instruction in Ilokano since 2011. The region was the site of the first regional EGRA conducted through USAID/Philippines and DepED, with the support of RTI International, in 2013.

In 2012, prior to the full implementation of the MTB, all concerned teachers were trained. The region likewise initiated trainings for the Education Program Supervisors and Public Schools District Supervisors to keep them abreast of feature of the K-12 Basic Education Curriculum. During implementation of the MTB-MLE, it was found that the Ilokano orthography and other instructional materials (Teacher's Guides, Learner's Materials and Curriculum Guides) proved to be of great help in teaching MTB-MLE. All Grade 1 and 2 teachers have been distributed materials (Teacher's Guides, Learner's Materials) produced by DepEd. Teachers also constructed and made teacher-made instructional materials such as big books and short stories.

Education Program Supervisors in-charge have likewise encouraged school heads during promotional meetings to conduct classroom observation and demonstration teaching particularly for the mother tongue so they could monitor its implementation. District Supervisors also urged Grade 1 and 2 teachers to restructure classrooms by labeling in three languages the facilities inside the classroom, the inclusion of MTB-MLE in their bulletin board and putting some quotes or poems in their native tongue⁷.

Region VII (Central Visayas) – Cebuano

Region VII (Central Visayas) is composed of four provinces and 19 divisions. The language Sinugbuang Binisaya, or Cebuano, is used outside of Region VII, including parts of Mindanao, but there are many different dialects. This study only selected schools using Cebuano as the mother tongue in Region VII. All schools in the region have begun to implement MTB-MLE in 2012.

Learner materials, teachers' guides and big books for Grades 1 and 2 have been produced following a quality assurance process. Teachers' trainings were conducted before the start of classes in 2012 for Grade 1 and 2013 for Grade 2. Teachers were oriented on how to use the teachers' guides and the learners' materials and on how to develop their own materials, such as big books. Teachers were also encouraged to be resourceful and look for materials suited to the lessons/competencies. Some of the materials have been uploaded to the learner resource portal for the region (<http://lrmds.deped.gov.ph/>).



In 2013, the northern part of the region was affected by the typhoon Haiyan, and by an earthquake before that. Therefore school was disrupted, with children out of school for a period and delivery of Grade 2 materials delayed.

Some schools in the region have received some support from USAID/Basa Pilipinas, which supported quality control of the learner materials. Basa conducted training of teachers for improvement of reading skills among the Grade 1 learners.

⁷ Description provided by Dinah Bonao, Regional Coordinator for MTB-MLE.

The region conducted EGRA training with education focal persons, district supervisors, and school heads in June 2013, including practice through simulations. The provinces of Bohol and 15 other divisions conducted EGAs with their Grade 1 pupils—and some Grade 2—from 951 public elementary schools. The regional office reproduced the EGRA materials into hard copies and on CD-ROMs. The results, though not compiled and analyzed formally, contributed to developing interventions such as appointing a reading teacher and reading clinic in each school for remediation.⁸

Region VI (Western Visayas) - Hiligaynon

Region VI consists of six provinces; Aklan, Antique, Negros Occidental, Capiz, Guimaras and Iloilo. Hiligaynon (also known as Ilongo), is spoken in Iloilo, Negros Occidental, Guimaras, Capiz, and Aklan, but other languages also overlap these regions, in particularly Kinaray-a which is spoken in Antique and parts of Iloilo, Capiz and Aklan. The region is considered as one of the biggest regions in the Philippines having different languages as medium of instructions: Most dominant languages used in schools are: Hiligaynon, Kinaray-a, Cebuano, and Akeanon.

The region began implementation of MTB-MLE in all schools, grades K – 3, in school year 2012-2013. Prior to the implementation of MTB-MLE under the K to 12 Basic Education Program, the region piloted use of the Lingua Franca as medium of instruction in the primary levels, and Hiligaynon was used by some schools as medium of instruction. This practice was inspired by the first Iloilo First Experiment 1948-1954 done by Supt. Jose Aguilar.

With the advent of MTB-MLE, DepEd through the Regional Curriculum and Learning Division, Training and Development and the Learning Resource Management Division has conducted series of 10-day regional trainings of teachers on MTB-MLE and continuous materials development activities.

These activities aimed to:

- Develop participants' knowledge and skills on the bridging processes of the different languages such as L1, L2, L3 or mother tongue, Filipino and English for basic literacy;
- Enhance their knowledge and skills on the theoretical and pedagogical foundations of MTB-MLE;
- Enrich their knowledge and skills on the use of the mother tongue in developing the language, and the cognitive, academic and socio-cultural aspects of the learners;
- Develop and contextualize existing materials for appropriateness and relevance to the grade level and culture of the pupils.

Demonstration teaching were also made available by Master Teachers facilitating instructional delivery to delve deeper into each subject area and apply correct and more suitable teaching strategies, styles and techniques. Instructional materials like the Teachers' Guide and Learners' Materials were developed by selected writers of each region. Series of writeshops were called for this purpose. Materials were developed and submitted to the

⁸ Description of activities provided by Nena Meñoza, Regional Coordinator for MTB-MLE. Note that the version of EGRA used in this study was different from the one distributed within the region; however, some items—particularly letter sounds and familiar words—have some overlap.

Instructional Material Council Secretariat (IMCS), a division/unit in the department taking charge of all the instructional materials of the schools. The materials are being reproduced by IMCS and duly distributed to the schools through the different school districts. However, for Grade I, (SY 2012-2013) the soft copy (CD) of the teacher's guide and learner materials were given to the division offices and the corresponding amount for reproduction and distribution.

ARMM – Maguindanaoan

The Autonomous Region of Muslim Mindanao (ARMM) in Mindanao, the southernmost landmass of the Philippines, includes five provinces, Lanao del Sur and Maguindanao on the mainland, and Basilan, Sulu, and Tawi-Tawi on the Sulu archipelago. The language, Maguindanaoan, is spoken primarily in the province of Maguindanao on the mainland and is one of about 8 dialects spoken in the region. Variations in vocabulary (pronunciation and spelling) exist between the north and south even within the Maguindanaoan dialect. Residents of each area can understand each other fairly well, but for children the differences create difficulty.



Implementation of the mother-tongue curriculum addresses 100% of pupils who speak Maguindanaoan in the home and with teachers. In August 2013 the regional office conducted a training of trainers on how to conduct EGRA, and after one week the division conducted an orientation for Grade 1 teachers and school heads for the implementation of EGRA. However, formal assessments were not carried out by the time this EGRA had been completed by USAID. The materials are standardized to use the vocabulary most common to all students.

C. Study design and methodology

About EGRA

EGRA is aligned with essential and teachable reading skills, and as such the results provide clear guidance for changing instructional methods and measuring progress over time. Measurements of how quickly and accurately children can read a text out loud, and how much of it they understand, also align with a scientific and a popular understanding of what it means to be able to read. The key differences between EGRA and other kinds of national diagnostic tests are that:

- EGRA is a *criterion-referenced* measure rather than a *norm-referenced* measure, and as such does not assess any specific curriculum or method. Instead, the same instrument and items can be used to compare results across different curricula, instructional methods, or grades.
- EGRA is *orally administered* by a trained assessor to one child at a time.

- EGRA uses *timed* subtests that provide a measure of fluency (words per minute⁹).

An adaptation process conducted in the Philippines in November 2013, involving the DepED and language specialists from around the country, served to ensure that the instruments used for this exercise:

- were culturally appropriate;
- corresponded to the grade-level material students were expected to be able to read;
- reflected the way the language was being used locally in reading materials for the target grade level;
- reflected the way that reading is typically taught;
- took into consideration the linguistic structure of the language; and
- addressed issues of mother-tongue interference, local dialects, and pronunciation.

The instruments developed were pilot tested with more than 150 students and analyzed using Rasch analysis to ensure that items measured a range of skills and were neither too easy nor too hard for the target population.¹⁰ The final instruments for all four target languages included the subtests summarized in **Table 2**.

Table 2: EGRA instrument subtests in the Philippines

Language/Subtest	Skill	Description The child is asked ...	Grade(s) implemented
Mother Tongue (MT)/ Initial-sound segmentation	Phonemic awareness	... “Tell me the first sound you hear in the word ‘bat’” (Answer is “b”). 10 items. Untimed.	1 and 2
MT/Differentiating initial sounds (Ilokano only)	Phonemic awareness	... “Tell me which syllable begins with a different sound: ba, bo, ka.” (Answer is “ka”). 10 Items. Untimed.	1 and 2
MT/Letter-sound identification	Alphabetic principle—letter-sound correspondence	... “Tell me the sound of this letter”. 100 items, random order, upper and lower case. Timed.	1 and 2
MT/Familiar word reading	Fluency—automatic word recognition	... read a list of 50 common words printed on a page. For example, “stop,” “the,” “was.” Timed.	1 and 2
Filipino/Familiar word reading	Fluency—automatic word recognition	... read a list of 50 common words printed on a page. For example, “stop,” “the,” “was.” Timed.	2 only
MT/Non-word reading	Alphabetic principle—letter sound correspondence and fluency—automatic decoding	... read a list of 50 non-words printed on a page. Words are constructed from actual orthography, but are not real words. For example “jaf,” “tob.” Timed.	1 and 2
MT/Oral reading fluency	Fluency—automatic word reading in context	... read a short story out loud (up to 60 words in length) printed on a page. Timed.	1 and 2
English/Vocabulary	Vocabulary	... Point to “head”, “shoes” (14 items). “Put the pencil on the paper” (6 commands). Untimed.	2 only

⁹ More information on the development of the instruments, subsequent studies including the 2013 Philippines EGRA report, and the growing body of research can be found at www.eddataglobal.org. For a summary of conclusions and recommendations from 2013, see Annex 1.

¹⁰ For more information or data on the Rasch analysis for this study, please contact spouez@rti.org.

Language/Subtest	Skill	Description The child is asked ...	Grade(s) implemented
MT/Reading comprehension	Comprehension	... respond to 5 questions that the assessor asks about the short story.	1 and 2
MT/Listening comprehension	Language comprehension and vocabulary	...listen to a story that the assessor reads out loud, then answer questions about the story.	1 and 2
English and Filipino/Listening comprehension	Language comprehension and vocabulary	...listen to a story that the assessor reads out loud, then answer questions about the story.	2 only

For children in Grade 2, the assessment included a short measurement of basic oral English vocabulary, and two Filipino subtests (listening and familiar word reading) in order to determine whether the MTB-MLE transitional program is being implemented as planned. (Children are expected to be introduced gradually to Filipino and English orally at first, and then to reading skills later so that instruction can proceed in English and Filipino in Grade 4 and beyond. See also **Table 4**.)

Each instrument, administered by trained assessors in one-on-one sessions with individual learners, requires about 20 minutes for each child, including a brief questionnaire for the student to gather basic information about the home and academic environment.

Other study instruments

Teacher survey. The teacher survey was administered to teachers in grades 1 to 3 whose children were selected for the EGRA study. It collected information about basic demographic characteristics, background related specifically to EGRA and early grade learning, and attitudes and practices towards reading in the early grades. This information was complementary to the information provided in the classroom observation tool and helps to understand the overall context in which teachers are working.

Classroom observation tool. In addition to EGRA, the study administered a teacher questionnaire and a classroom observation instrument. The teacher questionnaire asked teachers about their background, training they have received, and strategies they use in the classroom. These questions provide descriptive information for understanding whether or not MTB-MLE is being implemented and whether teachers have specific challenges teaching using the mother tongue. The classroom observation tool uses timed “snapshots” throughout a lesson. Every three minutes, field researchers would record information about the language in use, the content, and the classroom configuration at the moment of the snapshot. This information was then compiled to provide an overall picture of what methods were being used in the classroom. Each teacher was observed for one mother-tongue language arts lesson and one other subject area.

The data for all instruments, including EGRA, were gathered using the Tangerine® tool for mobile data collection. For full methodological details related to data collection—including assessor training, fieldwork procedures, and quality control—please refer to a separate report prepared by TNS, Methodological Report for the 2014 Early Grade Reading Assessment (April, 2014).¹¹

¹¹ For a copy please contact Yazmin.Tolentino@tnsglobal.com.

Notes on Interpreting EGRA Data

This report presents the results of all four languages at once, for convenience. However, caution should be taken when comparing results across languages and contexts, particularly for timed measures that report scores as words per minute. Each language has its own specific linguistic features that affect the expected rate of literacy acquisition; for example, agglutinative languages like Ilokano consist of very long words, where another language might have broken the same concept into several words. Therefore, a reasonable expectation of word reading fluency by grade for one language might look very different for another language. However, for the purpose of comparison, it is possible to compare the proportion of “zero scores” because this is an indicator of children who are not able to demonstrate any reading ability on a given task.

As the sampling design below explains, the sample is based on a list, provided by the DepED, of schools in the region that are using the mother tongue. However, there may also be schools *outside* of the target regions that are using the mother tongue, and there may be schools within the region that were not included on the list. Therefore although the sample size was intended to be large enough to be representative, it is not representative of the target region, and it may not be fully representative of all of the schools that teach using the mother tongue. It is only representative of the population of schools that was documented in the aforementioned list.

Sample design

In each region, the 2014 EGRA study was carried out in 40 schools and reached at least 800 students (400 in Grade 1 and 400 in Grade 2). **Table 3** presents a breakdown of the study sample. The data were collected in February 2014, which corresponded to the end of the academic year, in the second year of MTB-MLE implementation. Schools were sampled on the basis of a list provided by DepED of schools in each region that reported using the target mother tongue as language of instruction. The 40 schools were chosen proportional to enrollment. See **Annex C** for more details.

Table 3: Study sample

Mother tongue	Grade level	Sample size	No. of schools
Ilokano	Grade 1	400	40
	Grade 2	400	
Hiligaynon	Grade 1	400	40
	Grade 2	400	
Cebuano	Grade 1	400	40
	Grade 2	400	
Maguindanao	Grade 1	400	40
	Grade 2	400	
TOTAL		3,200	160

II. Findings

A. Is MTB-MLE being implemented?

Introduction

Before drawing conclusions about how well children are learning to read in the mother tongue, we should know what inputs are expected to lead to that particular output. The MTB-MLE program consists of an ambitious plan to use the mother tongue in Kindergarten through Grade 3 while gradually introducing English and Filipino. Children are expected to have academic fluency in English and Filipino by Grade 4, since these languages become the medium of instruction for the remainder of upper-primary and secondary school. **Table 4** summarizes the expected progression of language skills acquisition in primary school.

Table 4: National curriculum standards related to reading

Grade	Mother tongue	Filipino	English
Kindergarten	<ul style="list-style-type: none">• Oral fluency• Pre-reading activities• Medium of instruction	-	-
Grade 1	<ul style="list-style-type: none">• Oral fluency• Academic vocabulary• Reading and writing• Medium of instruction	<ul style="list-style-type: none">• Oral (listening and speaking) in Q2• Reading (Q4)	<ul style="list-style-type: none">• Oral (listening and speaking) in Q3
Grade 2	<ul style="list-style-type: none">• Oral fluency• Literacy development• Medium of instruction	<ul style="list-style-type: none">• Oral (communicative competence)• Literacy development	<ul style="list-style-type: none">• Oral (communicative competence)• Reading (Q2)
Grade 3	<ul style="list-style-type: none">• Oral fluency• Literacy development• Medium of instruction for most subjects	<ul style="list-style-type: none">• Oral (communicative competence)• Literacy development• Medium of instruction for some subjects (Q1)	<ul style="list-style-type: none">• Oral (communicative competence)• Literacy development• Medium of instruction for some subjects (Q3)

To determine the extent to which MTB-MLE is being appropriately implemented, the 2014 study interviewed teachers and observed lessons in Grades 1 and 2 in each school. This permitted the survey teams to gather information about the content and methods used in mother tongue lessons (reading and language skills development) and whether mother tongue was also used consistently as the medium of instruction in other subject areas as well. Data from the teacher interviews and classroom observations were combined across languages and classified into three categories covering the following aspects of MTB-MLE implementation:

- Teacher training, language ability, and the supports teachers receive at their schools
- The materials available and being used in classrooms
- The instructional practices observed during reading and other subject lessons.

Indices were created to summarize the data related to each of these three areas. **Annex B** provides detailed information regarding which variables were used in each of the three indices and how scores were attributed to a school for Grade 1 and Grade 2. The remainder of this section provides an overview of what each index is attempting to measure and provides a summary of how schools across the Philippines scored on each index.

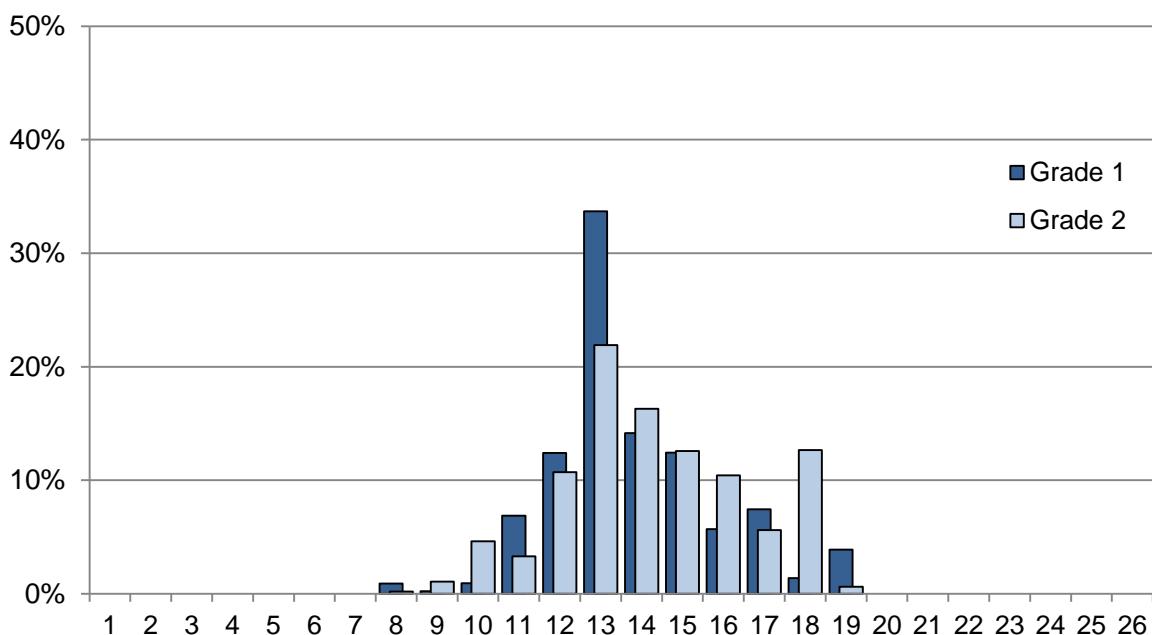
Findings

Index 1: Teachers, Language and Teacher Support. For the MTB-MLE policy to be effectively implemented, teachers must be appropriately oriented to the new curriculum, undergo the training that enables them to provide effective instruction, and master the language of instruction well enough to help children learn to read in it. In addition, good practice in early grade literacy should include regular assessment to monitor student progress and use of assessment results to adapt and plan instruction so as to meet the needs of students. The first index therefore includes the following information:

- Whether a teacher has prior training related to teaching in the early grades, teaching reading, and teaching in mother tongue
- Whether a teacher is a native speaker of the mother-tongue language s/he is being asked to teach in
- If a teacher describes him/herself as very comfortable in teaching in the mother-tongue language
- If a teacher states that s/he is very familiar with the grammar, alphabet sounds, spelling, and pronunciation of the mother-tongue language
- If the principal checks the teachers' lesson plans at least once per week
- If the principal observes the teachers at least once per week.
- If a teacher has used EGRA in his/her classroom
- If a teacher uses appropriate assessments to measure student progress in reading
- If a teacher uses assessment results to plan and adapt teaching activities to suit students' needs
- If a teacher's expectations for when students should learn to read in the mother tongue, Filipino, and English correspond to MTB-MLE curriculum expectations

The above information, drawn from interviews of one Grade 1 and one Grade 2 teacher in each school, was summarized in such a way as to assign a score from 0 to 26 for each grade in each school. **Figure 1**, below, shows the weighted percentages of schools scoring all along that possible range for both Grade 1 and Grade 2 classrooms. These indexes can be used as baseline measurements against which future monitoring can compare level of implementation.

Figure 1: Scores on the MTB-MLE Implementation Index—Teachers



The above graph shows that no school in either grade received more than 19 out of 26 points. Three-quarters of schools had a score from 12 to 16 out of 26. The average scores for Grades 1 and 2 on the teacher-related index were both 14. Analysis of where schools were gaining points and where they were losing points reveals some useful insight into the strengths and weaknesses of MTB-MLE implementation. For example, 75 to 80% of teachers reported not having been trained specifically for early grades or for teaching beginning reading.¹² Only around 30% reported having been trained in how to teach in mother tongue. In contrast, almost all teachers reported that they are native speakers of the language they are teaching; that they are very familiar with the language's grammar, alphabet sounds, spelling, and pronunciation; and that they are very comfortable teaching in it. Encouragingly, 85 to 95% of teachers also reported being observed and supported frequently. Around 80% of teachers reported using oral evaluations to assess reading progress, but very few are doing so with formal assessments such as EGRA or Phil-IRI. However, only 35 to 40% of teachers stated that they use assessment results to inform and adapt their instruction. Most teachers' expectations for literacy acquisition did not necessarily align with the objectives of MTB-MLE. While all teachers stated that students should learn to read in mother tongue in Grade 1, less than half indicated that students should learn to read Filipino in Grade 2 and English in Grade 3 (see also **Table 5** for a summary of responses to select indicators by language and grade).

Index 2: Teaching and Learning Materials. The second important component of MTB-MLE implementation concerns the availability and use of materials in the mother-tongue languages. Information for this index was drawn from the classroom observations and from teacher interviews and is related to the following.

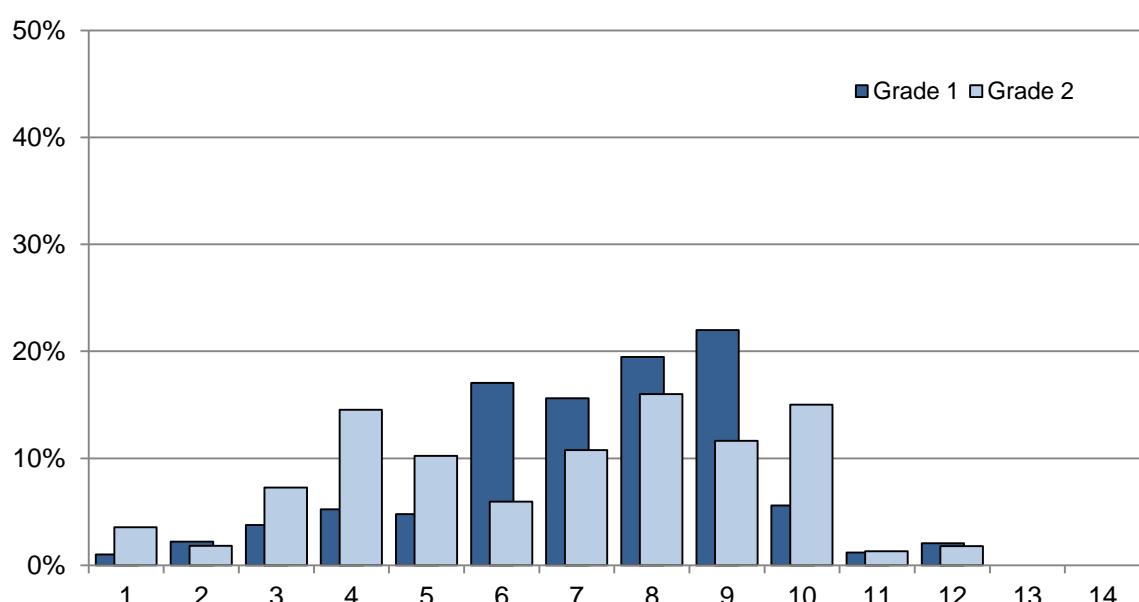
- Whether most students have textbooks in the mother-tongue language

¹² Caution should be used when interpreting this finding. According to DepED, all teachers have received in-service training related to MTB-MLE and reading in the mother tongue. This question asked about pre-service, but some teachers still answered with information about in-service training that had been received. Therefore it may be an inconsistent measure of training received.

- Whether there are many posters in mother-tongue language in the classroom
- Whether students have subject matter books in the mother-tongue language
- Whether teachers report having other materials for mother-tongue language, such as big books, workbooks, textbooks, teachers' guides, visual aids and multimedia materials
- Whether teachers think they have sufficient materials to support instruction.

Information relative to these points was compiled into a 14-point index, with separate scores calculated for each grade in each school. The distribution of those scores is shown in **Figure 2** below.

Figure 2: Scores on the MTB-MLE Implementation Index—Materials



Three-quarters of Grade 1 classrooms had scores of from 6 to 9 out of 14 on the materials index. Grade 2 classrooms had more schools scoring 10 out of 14, but also had a higher percentage of schools that scored below 6 (more than a third compared to just 17 percent of Grade 1 classrooms scoring that low). The average score for both grades on the materials-related index was 6. In about 60% of Grade 1 and Grade 2 classrooms, most students did have books in the mother tongue and about half the interviewed teachers reported having a teachers' guide for mother-tongue reading. In more than 80% of classrooms teachers reported having visual aids to support the mother tongue and posters in the mother tongue were visible in 60 to 70% of classrooms, according to the classroom observation tool. Yet 70 to 75% of teachers stated that they feel that they do not have sufficient resources for teaching reading in their class.

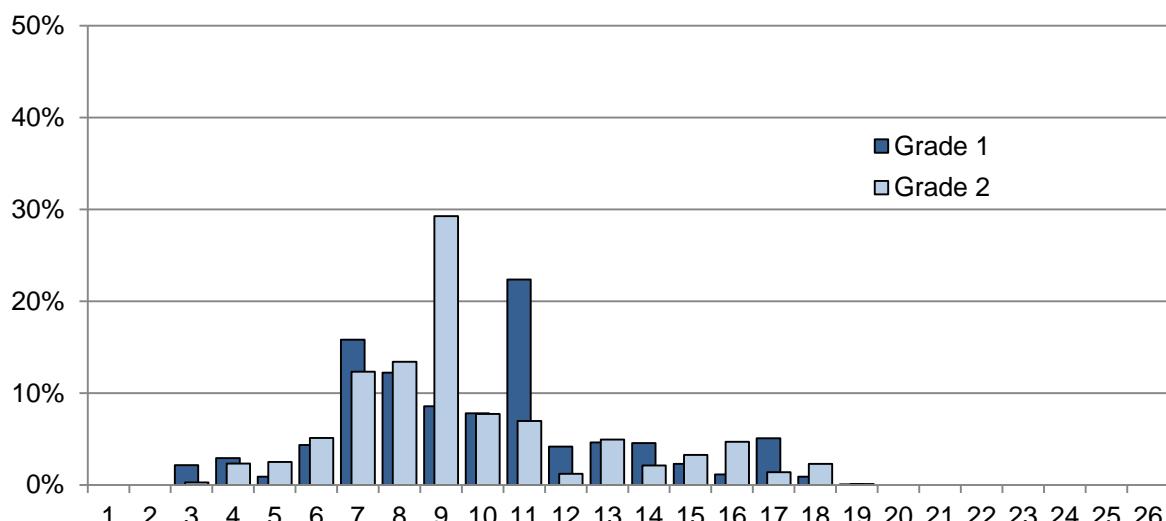
Index 3: Instructional Practices. The third and final aspect of MTB-MLE implementation that was evaluated concerned observed instructional practices. Observations were conducted at each school during a reading lesson and during another subject lesson in Grade 1 and 2. The third index takes into account several aspects of what was observed, including the following:

- Whether the mother-tongue language is in use for 80% or more of the class time

- Whether pupils are actively engaged in the lesson while the mother-tongue language is being used
- Whether the teacher models proper mother-tongue language use for pupils to emulate, praises or rewards pupils for using mother-tongue language correctly, points out and corrects mother-tongue language errors, monitors pupil comprehension, and uses visual aids and simpler forms of the language to explain something the students struggle to understand
- If the teacher shifts to another language only in order to clarify or emphasize a concept first explained in mother tongue or to give procedural directions to students
- If 25% or more of a reading lesson is spent on activities that involve students reading
- If during a reading lesson, students also spend time productively listening, speaking, or writing in the mother-tongue.

The index awards points on a scale going from 0 to 26 depending on the above observed practices. The distribution of school scores for Grades 1 and 2 are shown in **Figure 3**.

Figure 3: Scores on the MTB-MLE Implementation Index—Instruction



No school had higher than 18 or 19 out of 26 on this index, and the average score for both grades was 9. Two-thirds of the schools scored between 7 and 11 on this index in either grade. Close to 90% of classrooms in Grades 1 and 2 were using the mother-tongue during more than 80% of the observed reading lessons. During the lessons in other subject areas, the mother tongue was being used more than 80% of the time in 75% of the classrooms. Pupils were noted as actively engaged in 75 to 85% of the observed reading and other subject area lessons. Where implementation was weakest, however, was in how teachers supported the use of mother-tongue. In only about 20% of observed reading lessons did students spend adequate time on reading activities, and in only 10% of the observed lessons were students engaged regularly in productive speaking, listening, and/or writing activities. Furthermore, less than a third of teachers were observed to be modeling correct use of the mother tongue; only about 30% praising students for correct usage, less than 10% correcting students when they made errors in mother-tongue language use, and 30% monitoring pupils' comprehension during the lesson. The most common way that teachers supported use of the mother tongue

was to use a simpler form of the mother-tongue to explain something the students failed to understand. About 30% of teachers used another language to clarify or emphasize a point during reading lessons, and about 35 to 40% did so during other subject lessons.

Discussion

What the above indices show is that teachers and students were using the mother-tongue in Grades 1 and 2, both during mother-tongue reading lessons and during other subjects. However, it would appear that not all the necessary supports are in place to help teachers provide the most effective instruction possible in mother tongue. Most teachers have not received training in how to teach reading in the mother tongue. Teachers are being supported, but would benefit from training more focused on the instructional practices most conducive to effective acquisition of literacy in the mother tongue. While the majority of schools had mother-tongue language materials, in a significant percentage (40%), not all students had the materials and half of teachers did not have teachers' guides. Teachers were not regularly employing techniques that would most productively reinforce proper use of and reading in mother tongue. In fact, the amount of time on reading and the nature of activities in reading, speaking, listening, and writing were the weakest aspects of MTB-MLE implementation. Effort should be mobilized to provide professional development in how to plan reading lessons to maximize the time children spend reading and how to employ specific techniques to reinforce correct mother-tongue language development.

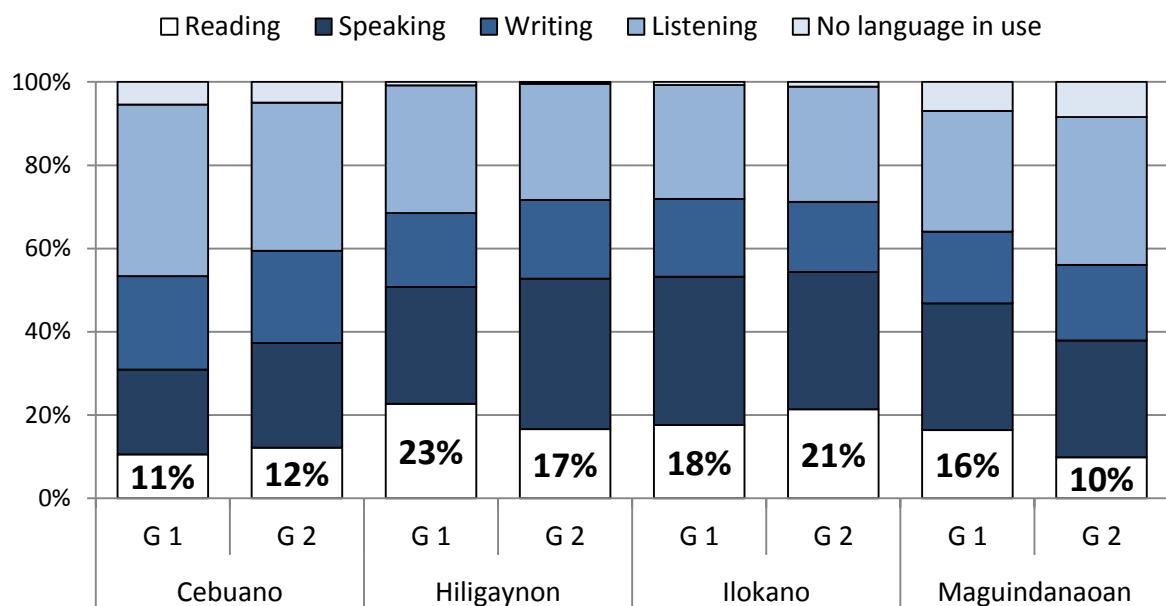
B. What are the predominant teaching methods being used?

If children are not spending time reading during the part of the day dedicated to mother-tongue language development, then what are they doing? The classroom observations collected data at regular intervals during the mother-tongue language lesson allowing us to estimate how much time was being spent on different types of language activities. A balanced approach to literacy instruction promotes use of a mix of methods and activities that reinforce listening, speaking, reading, and writing. However, in the early grades, when children are learning to read in their mother tongue (and therefore oral language acquisition is not a key concern), a focus on reading and writing should take precedence with listening and speaking occurring naturally as part of this instructional focus. Learning to read requires a significant and sustained amount of practice, especially when consolidating foundational skills to build reading fluency.

Figure 4, below, shows that reading was the primary instructional focus for the majority of pupils less than 25% of the time in all languages, in both Grade 1 and Grade 2.¹³ The amount of time spent reading does not appear to increase from Grade 1 to Grade 2 either, although we would expect that as children acquire the skill of reading they would spend more time practicing and building fluency.

¹³ The data tell us more precisely the total number of observations recorded in each category. We can use this as a proxy for the predominant activities and teaching methods. The assessors were trained to record the primary instructional intent of the majority of pupils at the observation moment; for example, if one child was reading aloud (i.e., reading off of the board), the observation should have been counted as “reading,” not “speaking” (which the individual child was doing) or “listening”, (which the majority of other pupils would have been doing—an assumption is made that they were supposed to be passively reading). We recognize that sometimes several things are happening at once, and this could bias the results depending on the assessor’s background and ability to surmise the intent of the teacher.

Figure 4: Pupil language activity during MT lesson



The relatively small proportion of time spent reading can be explained by understanding the teaching methods that teachers were using. Across all languages, the top three teacher actions at the observation moment—accounting for between 56 and 71% of total observations—were:

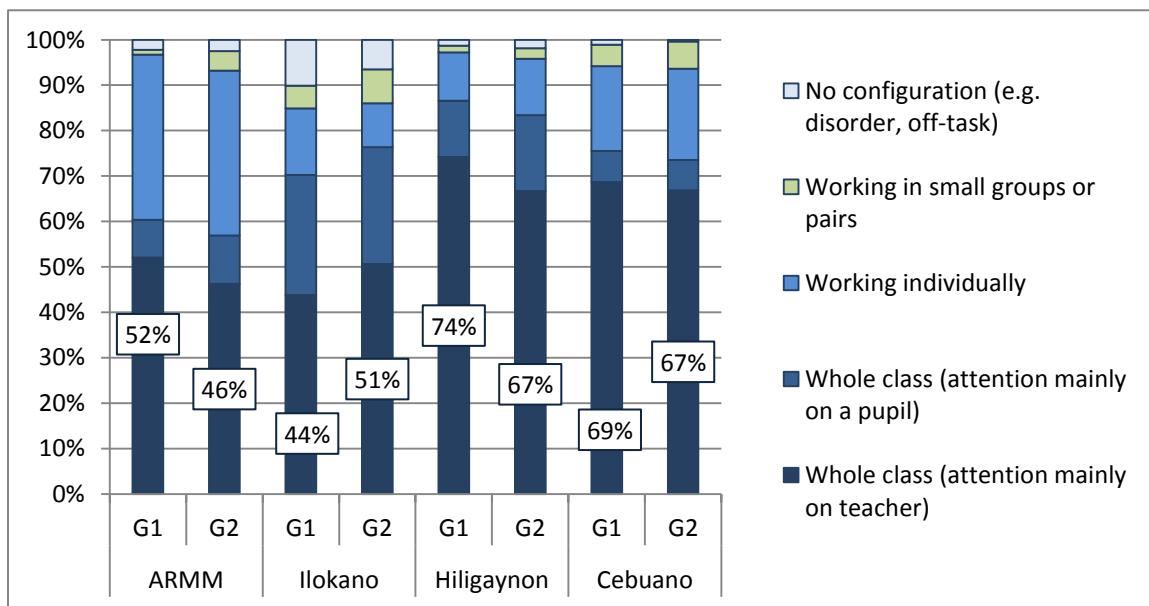
- Asking/answering questions, calling on a child
- Waiting for pupils to respond/complete task
- Explaining/lecturing/teaching subject matter.

Teachers were predominantly using teacher-centered, whole class instruction. Students participate in turn, when asked either to read something out loud or to answer a question. This method results in most time being spent listening to or speaking the language, or passively reading as a group. It is not guaranteed that all children have equal opportunities to practice reading or speaking the language actively and receiving personalized feedback. The observation intervals also recorded what type of instructional configuration was in place at the time, and this confirmed that whole class instruction was by far predominant, with very little group work or individual attention being done. **Figure 5**, below, shows that 50 to 75% of the time the configuration was such that the majority of pupils had their attention on the teacher. In the ARMM region only a large amount of time was spent doing individual work. Across all languages and grades, working in small groups was rare.

While instructional configuration alone is not an indicator of good instruction—whole class instruction can be effective and group work can be ineffective, as well as the inverse—it is more likely that children are getting individualized instruction that meets their needs if the classroom is configured to do group work or individual work. In these situations, the teacher may be circulating and providing personalized instruction. Similarly, if children are working individually that can be productive depending on the tasks and whether if the teacher is circulating and providing feedback. On the other hand, it may also be relatively unproductive if children are merely copying from the board. The frequency of “copying” appears to explain

the high proportion of individual work seen in the sample from Maguindanaoan schools; 12% of observations were coded as “Writing (major activity): Copying (sub-category of activity)”. More in-depth research would need to be done to confirm what type of instruction is typically associated with different types of classroom configuration, and whether this has an effect on reading.

Figure 5: Classroom configuration during MT lesson



Equally as important is the amount of time where the observations recorded were “off task”. Children in Cebuano and Hiligaynon classrooms were off-task approximately 1% of the time, Ilokano classrooms 8% of the time, and Maguindanaoan classrooms 2% of the time. If we judge only by what the teacher was doing at the observation moment, counting the observations that were coded as “off task” or “giving procedural directions or managing behavior”, the percentage of time where little or no productive instruction is taking place during the reading classroom is as follows (see **Table 5**).

Table 5: Percent of classroom observation where no instruction was taking place

Mother Tongue	Grade 1	Grade 2
Cebuano	10%	8%
Ilokano	33%	28%
Hiligaynon	13%	13%
Maguindanaoan	15%	15%

Table 6: Summary of select indicators related to MTB-MLE implementation, by language and grade

Item (Weighted)	Cebuano		Hiligaynon		Ilokano		Maguindanaoan	
	G1	G2	G1	G2	G1	G2	G1	G2
Teacher's MT is the EGRA language (teacher reported)	95%	90%	94%	91%	95%	94%	83%	83%
This is the first year teaching in the MT	26%	68%	22%	90%	18%	89%	51%	75%
Teacher mainly uses the EGRA language this year (student reported)	87%	78%	90%	86%	97%	89%	89%	84%
Teacher mainly used the EGRA language last year (student reported)	85%	82%	95%	90%	91%	92%	85%	84%
Teacher used MT "most of the time" (> than 75%) during the lesson (observed)	98%	97%	100%	94%	95%	89%	71%	82%
Percent of total observations using the MT in other subject areas (observed)	90%	91%	99%	96%	97%	97%	59%	39%
Teacher <u>never</u> uses the MT learners' guide (student reported)	13%	24%	35%	14%	13%	11%	45%	43%
Teacher <u>never</u> uses the MT math learners' guide (student reported)	26%	37%	36%	53%	14%	21%	50%	50%
Student did not miss any days of school last week (student reported)	46%	37%	24%	25%	40%	41%	25%	27%
Child speaks the MT in the home (student reported)	96%	97%	98%	97%	87%	85%	99%	97%
All pupils have MT learner materials (teacher reported)	63%	48%	43%	67%	84%	77%	49%	32%
The majority of pupils were offtask or socializing during the reading lesson (observed)	1%	0%	1%	1%	17%	13%	10%	14%
Teacher has teachers' guide to support MTB-MLE (teacher reported)	49%	50%	80%	74%	28%	38%	46%	23%
Teacher has NO supporting materials for MTB-MLE (teacher reported)	0%	0%	0%	0%	0%	0%	2%	4%
Thinks the MT teaching materials are sufficient (teacher reported)	31%	21%	22%	20%	32%	21%	29%	28%
Principal reviews lesson plans at least once per week (teacher reported)	72%	54%	41%	51%	78%	72%	37%	37%
Percent of teachers who think children should read fluently in the <u>MT in G1</u> (teacher reported)	21%	21%	18%	2%	22%	16%	0%	0%
Percent of teachers who think children should read fluently in <u>Filipino in G2</u> (teacher reported)	62%	53%	71%	26%	50%	58%	31%	25%
Percent of teachers who think children should read fluently in <u>English in G3</u> (teacher reported)	45%	49%	35%	55%	28%	27%	40%	49%
Has received some kind of instructional support from an instructional advisor at least <u>once per month</u> since the start of the school year	2%	4%	0%	5%	0%	4%	22%	26%
Top two methods of getting support for MT teaching, when needed								
- Peer support	61%	68%	85%	96%	82%	71%	46%	41%
- Consults principal	64%	64%	82%	64%	52%	64%	84%	85%

C. How well are children learning how to read in the MT?

Sections A and B, above, describe expectations and findings related to what teachers should be and are doing in classrooms. Although learning how to read is not the only outcome expected of students in these classrooms, it is a critical one, and the one that this study is designed to measure. **Table 4** above outlines the expected progression of reading in target languages from Grades 1 to 3. **Table 2**, earlier, described the different measurements included in EGRA to determine whether or not children are acquiring the skills required by the curriculum. In this section, we will describe whether these measurements (EGRA subtests) are identifying levels of language acquisition in line with curricular expectations.

Scores on pre-reading skills (*listening comprehension, phonemic awareness*)

The first two EGRA subtests that children are given focus on oral language skills that are considered important pre-requisites for learning how to read. **Listening comprehension (LC)** measures the child's basic oral mastery of the language (receptive and productive capacity). The EGRA administrator reads a short story out loud and then asks three questions related to the story. The stories used in each language were 30–50 words long and centered on a familiar concept or theme. The score is calculated as either the number of questions correct out of three, or a percentage (0%, 33%, 67%, or 100% are the four possible scores). **Phonemic awareness (PA)** measures the ability of children to distinguish individual sounds in words—in this case, the beginning sound.¹⁴ There are 10 items to which the child is asked to give a response; if the child responds incorrectly to the first 5 items, the subtest is discontinued.

Figure 6a and b show the results on these two subtests, both of which are scored as a percent of correct responses out of the total items possible. The bars in the chart show the average percent by grade and language, including scores of children who did not provide a single correct response (zero score). For phonemic awareness, the results show that the highest score was near 60%, or 6 of 10 initial sounds identified correctly. For listening comprehension the scores remained under 60% across languages, but this indicates fewer than 2 out of 3 questions answered correctly for a short story read out loud. The results also show that average scores were higher in Grade 2 than Grade 1 in all languages, though the scale of that grade-to-grade improvement varies by language.

Figure 7 goes one step further by presenting the two tasks together, including a marker (x or triangle) for the averages excluding the non-readers. Across the four languages and two grades, anywhere between 10 and 50% of children got a zero score for one of these subtests. What is important to note in both graphs is that zero scores decreased from Grade 1 to Grade 2. As a reminder, the listening comprehension exercise consists of only three questions, so where the mean score is represented as “50%”, this means that on average, children answered 1.5 questions correctly. To state it another way, many children only answered 1 of 3 questions correctly, while many others answered 2 correctly.

¹⁴ In the case of Ilokano, the exercise asked children to isolate which syllable is different from the other. For example, the assessor asks “Listen to the syllables, and tell me which one begins with a different sound: ba, ma, bi.” The child would be expected to answer “ma.”

Figure 6: Average percent correct responses on PA and LC tasks

a) Phonemic awareness

b) Listening comprehension

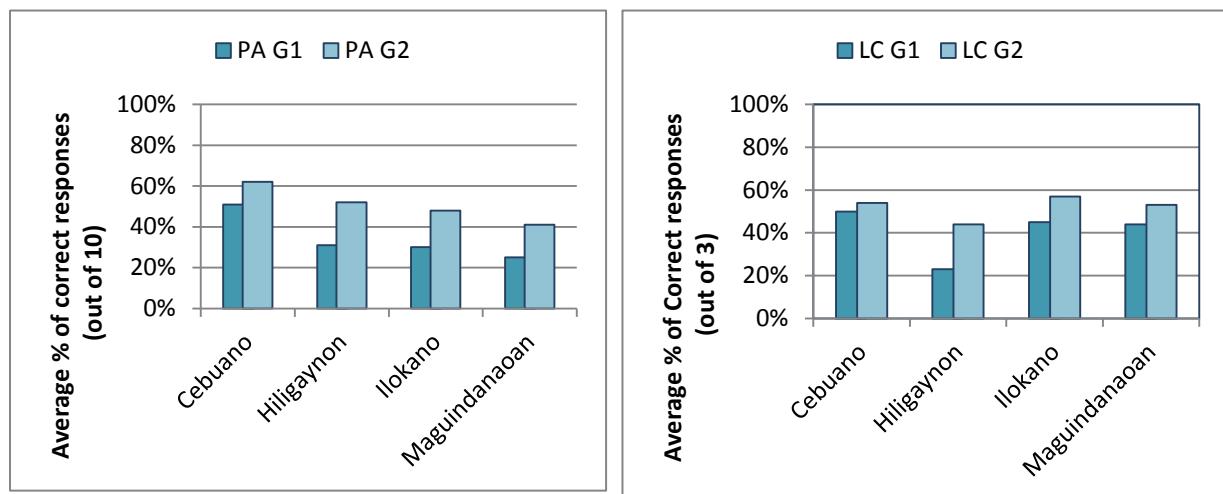
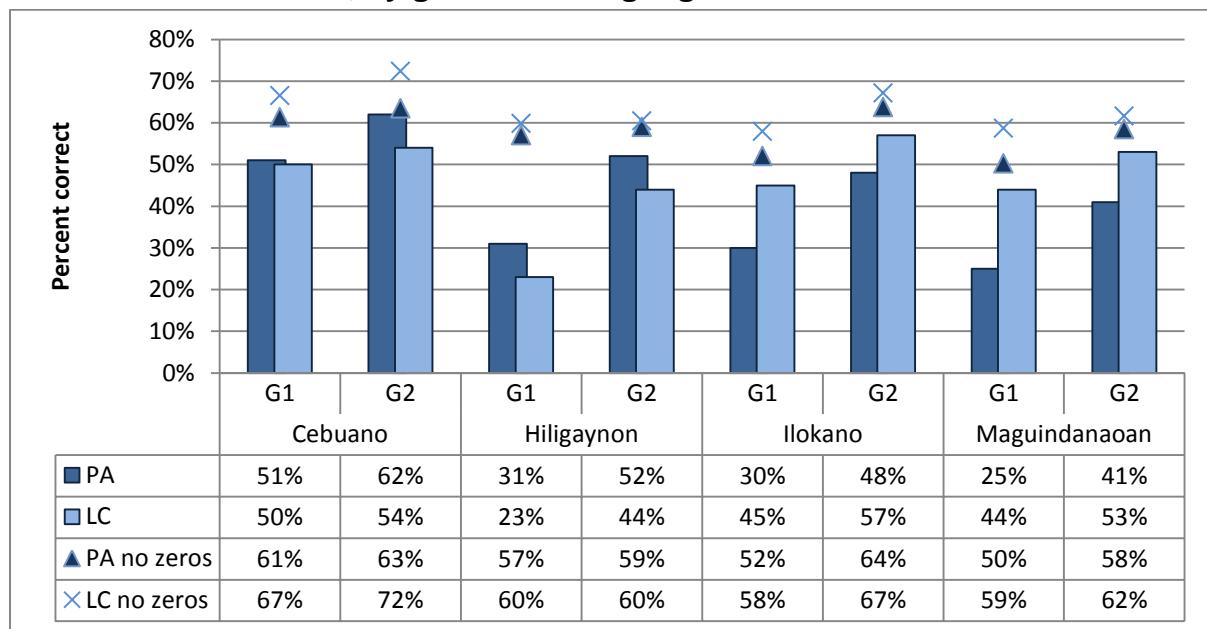


Figure 7: Scores on listening comprehension and phonemic awareness exercises, by grade and language



However, it is important to note that average scores may misrepresent the overall range of ability. As **Figure 7** shows, there were cases in which there was a remarkable gain in average score from one year to the next when all children are considered, but when the average is calculated without the zero scores, there is little to no change in the average (see Hiligaynon, for example, where there was almost no change in average score from Grade 1 to Grade 2 as shown by the symbols X and Δ). Looking at the distribution of responses can help us understand whether this was due to:

- A reduction in the number of children who get zero scores,
- An increase in the average of children who have some ability, or
- Another type of change in scores at the extremes.

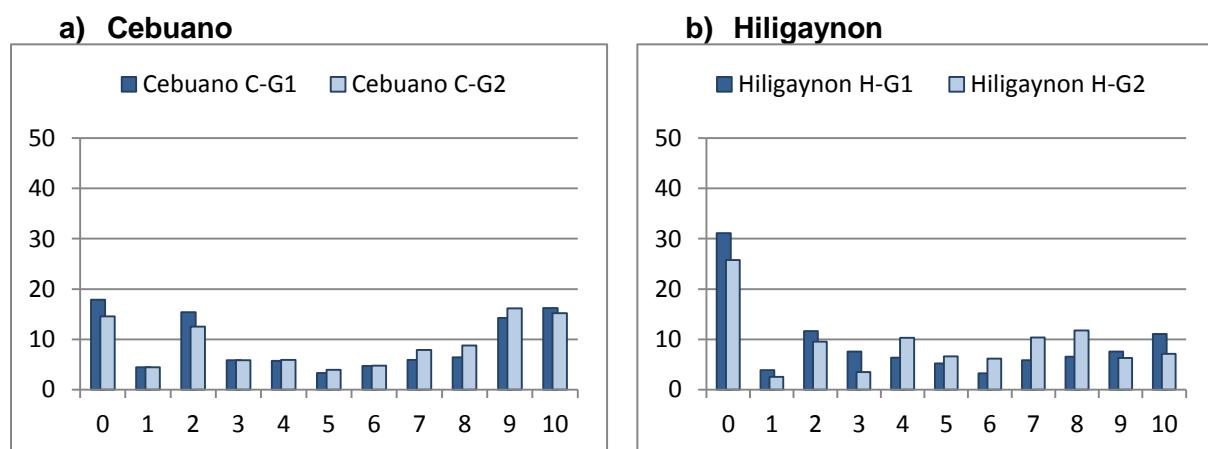
Figure 8a-d, below, show distributions for **phonemic awareness** across the languages, where the horizontal axes are the possible scores (total items correct out of 10) and the vertical axis is the percent of children who achieved that score. We can see some remarkable differences among the language groups. For example, Maguindanaoan, the region with the highest proportion of zero scores in both grades, shows a drop in zero scores of 20 percentage points from Grade 1 to Grade 2, and an increase in the number of children in the highest skill range (8, 9, and 10 items correct). The Grade 2 average score was 16 percentage points higher than the Grade 1 score.

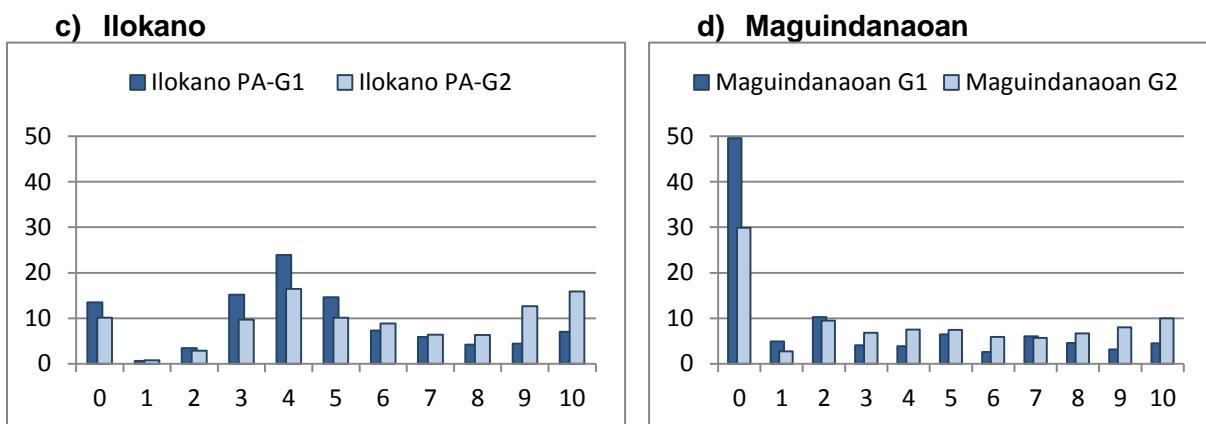
On the other hand, Cebuano had a relatively low proportion of zero scores, a relatively high proportion of children in the highest range (9 and 10), and an almost equal distribution of children in each category in between. Yet there was little change in the middle categories from Grade 1 to Grade 2.

The phenomenon cited above with relation to Hiligaynon (almost no change from Grade 1 to Grade 2 when zero scores are removed) can be explained by looking at the distribution below. There was a lower proportion of zero scores from Grade 1 to Grade 2, but also fewer children in Grade 2 achieving scores at the high end of the range (9 and 10). Therefore the large increase in the average scores (all children included)—from 31 to 52% was actually the result of a higher proportion of children in the ranges from 4 to 8, but with a large proportion of zero scores remaining.

With the exception of the high proportion of zero scores, achievement peaked in the middle range only for Ilokano, Grade 1. However, by Grade 2 many more children were in the higher range of ability (8, 9, and 10) while the peak was still present—but smaller—in the range of 3 to 5 correct responses.

Figure 8: Phonemic awareness: Distribution of scores by grade and language

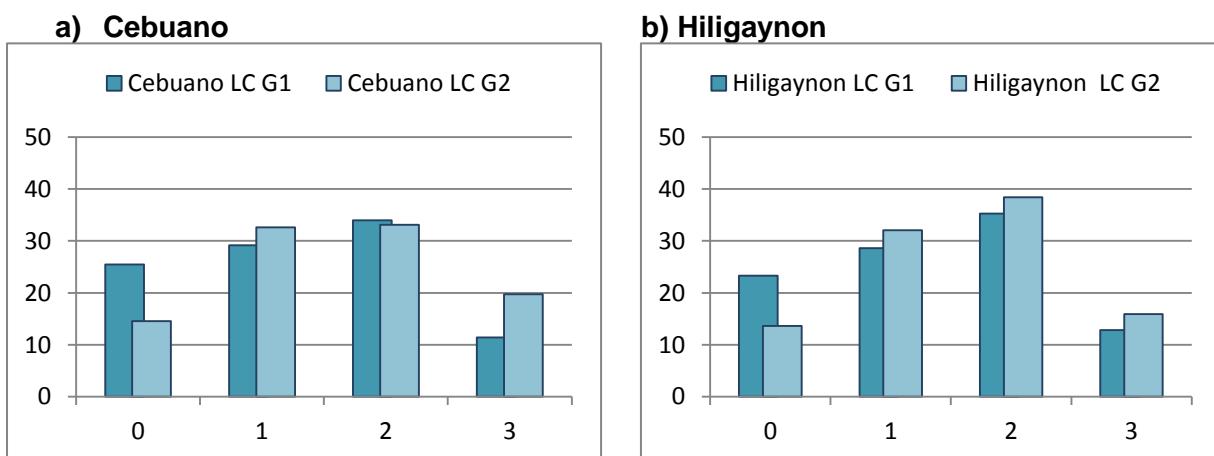




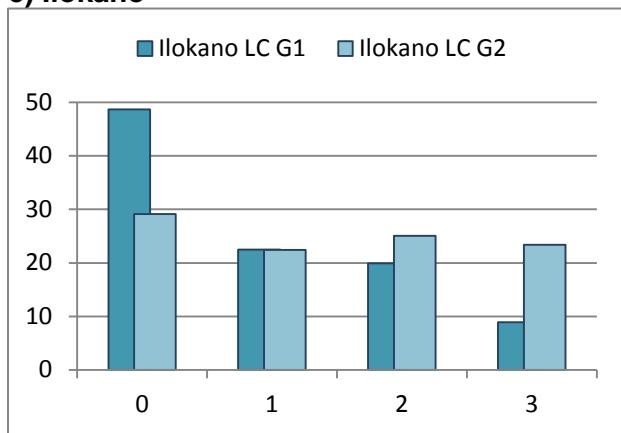
While the details differ from region to region, one thing that is similar is that children across all languages had difficulty with the phonemic awareness subtest. This may be because phonemic awareness is not explicitly taught, especially sounding letters in isolation. In many cases, children gave the first syllable, rather than the first sound on this subtest. For all languages, the phonemic awareness skill did improve from Grade 1 to Grade 2, but one should expect that by the end of Grade 2 a much higher proportion of children would be demonstrating a high range of ability on this basic task.

Figure 9a-d, below, provide the distribution of scores on listening comprehension; the horizontal axis is the number of questions answered correctly out of three that were asked, and the vertical axis is the percentage of children who achieved that score. Where we see large percentages of zero scores (the child answered no questions correctly), this may be an indication that children were not native speakers of the target language or that their pre-school education was in a language other than the target language.

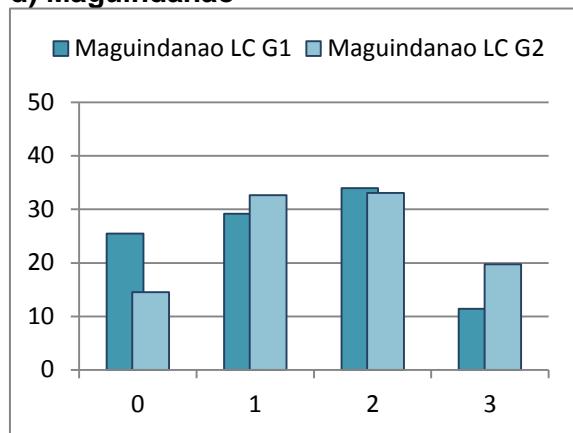
Figure 9: Listening comprehension: Distribution by grade and language



c) Ilokano



d) Maguindanao



Again, the results show important decreases in zero scores from Grade 1 to Grade 2 across all languages, but especially in Ilokano where there was a 20 percentage point drop from Grade 1 to Grade 2 in the percentage of children who answered no questions correctly. Furthermore, the number of Grade 2 children who could answer all three questions on listening comprehension in Ilokano was nearly three times that of Grade 1. Similarly to phonemic awareness, in the case of Hiligaynon, we see the largest overall percentage gain in average score from Grade 1 to Grade 2 when all children are included—21%; however, there is almost no change when the average is calculated without zero scores. We can see that this is explained because the distribution in all categories of children who answered at least one question increases by the same amount—in other words, the number of children who could only answer one question also increased in Grade 2. Only for Cebuano and Ilokano do we see that the overall distribution was shifting in favor of more children with higher scores, and few children with lower scores.

Although there is some improvement across languages from Grade 1 to Grade 2, it is somewhat surprising that there were still so many children in Grade 2 unable to answer any questions, since children should have been listening to the story in their own mother tongue. The stories were all very short, so memory recall should not have been a factor preventing children from remembering. Some possible explanations for these low scores are that:

- Grade 1 students were still struggling with listening comprehension at the end of the year if they did not attend kindergarten in the mother tongue. **Table 6** provides the percentage of children in each language and grade who reported that their (previous grade) teacher mostly spoke the target language last year. In the case of Grade 1 students, we can assume this meant the language of instruction in preschool. The results were between 85% for Maguindanaoan and Cebuano and 95% for Ilokano and Hiligaynon.
- Children were unaccustomed to listening to stories out loud or being asked to answer questions related to these stories. When a score is reflected as “incorrect,” it could be either that the child gave a response that was not correct, or that the child simply stayed silent and didn’t provide any response. Therefore it is possible that a proportion of the children felt inhibited and did not respond for some reason (shyness, or ability to understand the language but difficulty in speaking it).

- Children were not actually speakers of the language and may have been attending a school that taught a different mother tongue. **Table 6** above also provides the proportion of children who reported speaking the mother tongue at home. This is lowest (85%) for children in the Ilokano sample, where zero scores were also the highest in Grade 1

The listening comprehension exercise has some important limitations. First, there are only three questions, so it can be argued that this doesn't provide much opportunity for children to demonstrate their skills. However, to ask more questions a longer story would be required and then the task relies much more on cognition and memory. The task also requires language production (speaking to answer the question). Despite these limitations, the task provides information about the cognitive and academic skills of the children even if it is not a highly reliable measure of mother tongue ability—or at least prompts us to pose relevant questions about classroom instruction. For example, if children are perfectly fluent in the language, but still can't answer question about a short story, then what does this indicate about their ability to be active in class across subject areas? If children aren't prepared to listen actively and answer questions without having either the story or the questions repeated several times, what does that indicate about their preparedness for internalizing explanations and content delivered by the teacher (recall from Figure 4 and the surrounding explanation that a large proportion of class time is spent with the teacher talking and the children listening)?

Importantly, children were improving their skills from one grade to the next. Both phonemic awareness and listening comprehension are skills that are considered fundamental building blocks of learning to read and are indicators of important cognitive skills that are needed for success in school.

Note on comparison of scores: It is important to recall that for the letter-sound identification task, the alphabets differ somewhat between languages and scores should be interpreted/compared with this in mind. The key difference is in the number of graphemes: Hiligaynon and Cebuano have 20; Maguindanaoan has 17; Ilokano has 26 (including borrowed letters like ñ, j, and q). When trying to understand what might constitute an adequate score, it is useful to remember that a rate of 60 correct letters per minute would be one letter per second with no errors; 30 correct letters per minute is 1 letter every 2 seconds, and so forth. Letter identification is a skill that should become automatic, and therefore require less than 1 second to process regardless of language, though the age/grade at which this happens may be different based on the language. This principle is the same for word reading tasks that follow.

Scores on initial reading skills (letter-sound identification, non-word reading, familiar word reading)

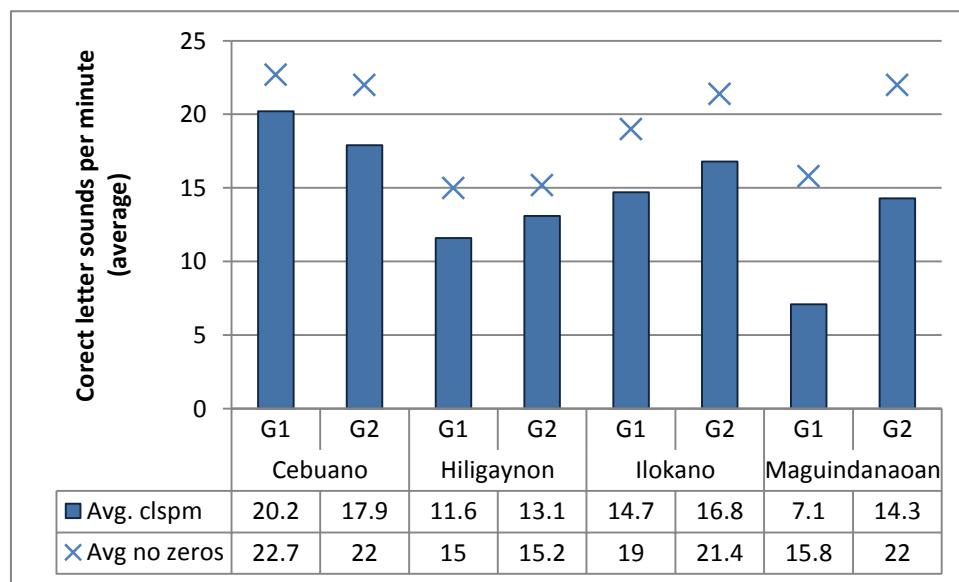
The two previous components of EGRA measure skills that are oral only; in other words, no print is involved. The next three subtests that will be reported in this section are those that require the child to read print from stimuli sheets prepared for the assessment.

Letter-sound identification. The first subtest, letter-sound identification, requires the child to provide the sound of letters of the alphabet, presented in random order and in a mix of upper- and lower-case forms. The task is timed, providing a measure of correct letter sounds

per minute (clspm). Additionally, we can identify the proportion of children with zero scores (those unable to identify a single letter-sound correctly in one minute). **Figure 10** below shows the scores by region on letter-sound identification, both with and without the zero scores included.

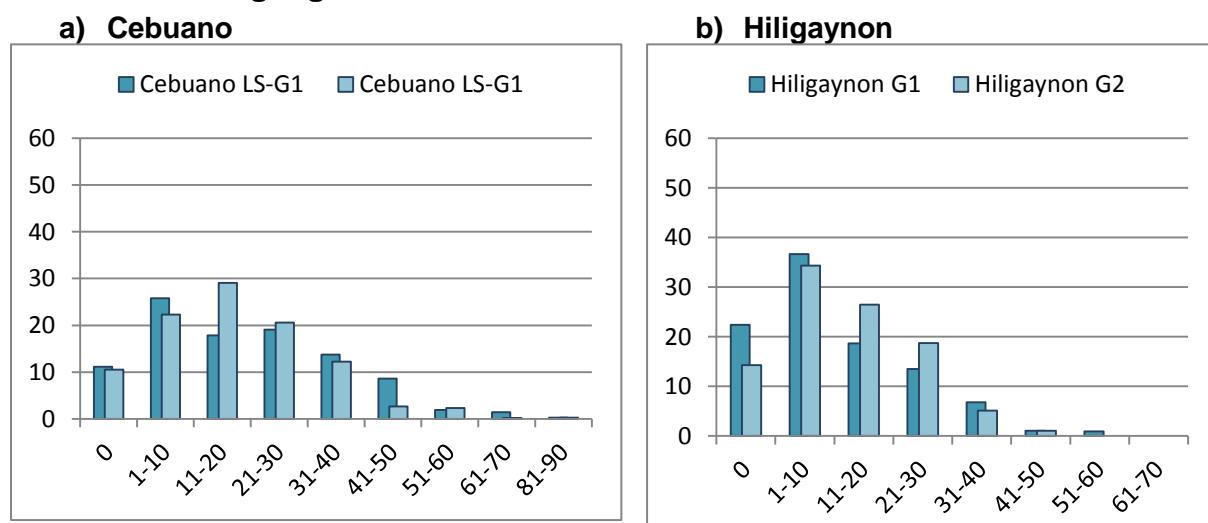
While in most regions fewer than one-fourth of the children had zero scores in either grade, more than half of children learning to read in Maguindanaoan couldn't correctly identify a single letter sound in Grade 1. This drops to 35% in Grade 2, again showing the most dramatic change from Grade 1 to Grade 2, but still the highest proportion of zero scores across the regions.

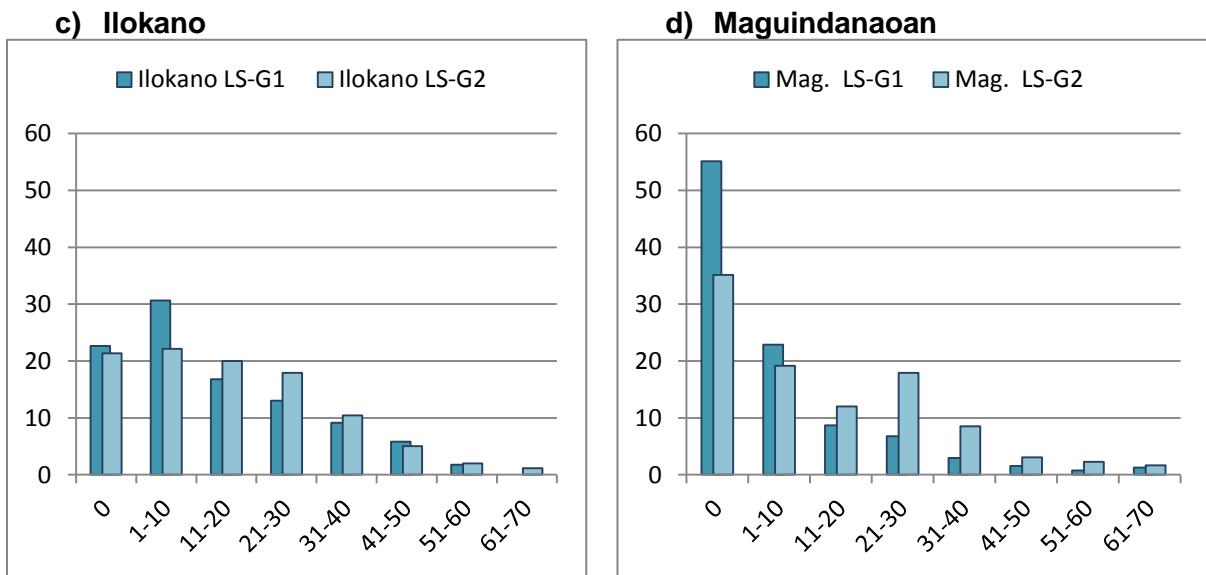
Figure 10: Letter-sound identification scores, by region and grade



Across the regions, children were identifying on average fewer than 20 letter sounds correctly in one minute. While the score improved twofold (from 7 letters to 14) for Maguindanaoan, most other languages showed little change from Grade 1 to Grade 2. Again, looking at the distribution of responses (see **Figure 11a-d**) can provide some additional insight.

Figure 11: Letter-sound identification: Distribution of scores, by grade and language





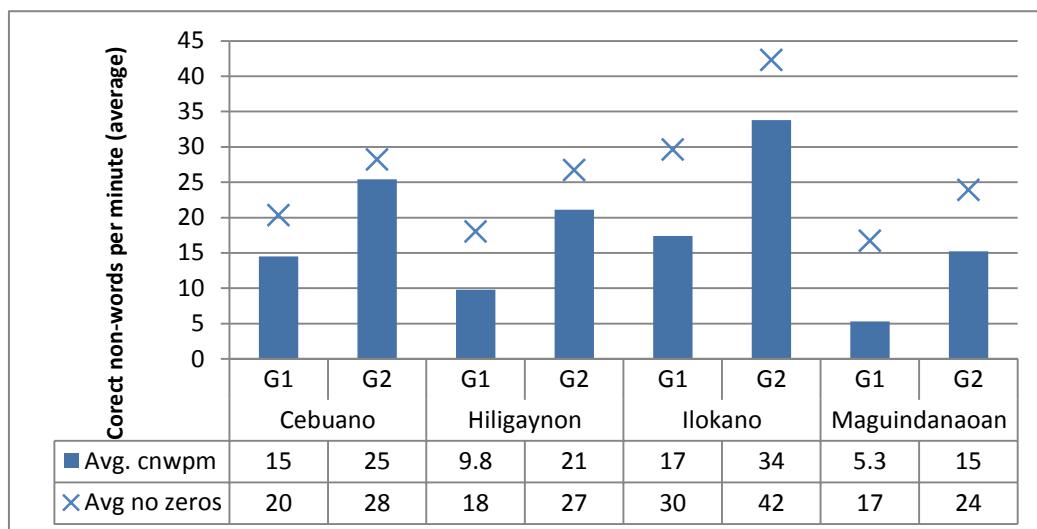
With the exception of Maguindanaoan, where scores were concentrated at the lower end of the scale, the regions mostly showed a peak at around 10 or 20 letters identified correctly in one minute. A key difference between languages is in the scores at the extremes. There were no children learning to read in Hiligaynon that achieved more than 60 clspm, and very few that achieved more than 50. In all other languages, there were at least some children with this ability.

Across regions, letter sound identification showed a reasonably strong correlation to other word reading skills, but the correlation was strongest for Maguindanaoan, where the correlation between the letter-sound identification score and scores on familiar and invented word reading and oral reading fluency were all above .6 (Pearson coefficient). For Hiligaynon and Ilokano the correlation coefficients fell within the .5 to .6 range, and for Cebuano the correlation was the least apparent, but still present, with a coefficient of .4 to .5. One reason letter sound scores may not be correlated strongly with other reading skills is that it is very difficult to get high inter-rater reliability on this task. For many reasons (children who speak softly, background noise, different assessor backgrounds, and native language interference) it can be hard to distinguish some letter sounds, and different assessors have a different tolerance for the precision expected in a child's response (for example, to say /b/ instead of /buh/ for the letter B.). This year assessors were asked to explain the main reason children were getting letters wrong. That information indicated that assessors not hearing student responses correctly was not usually an issue. In 46% of cases (all languages and grades combined), assessors indicated that the child said letter names instead of the sound, and in 32% of the cases they indicated that the child said syllable sounds (/ba/).

Non-word reading. Becoming a reader requires using explicit knowledge of letter sound correspondence to decode a word that one has never seen before. Therefore a non-word reading subtest is included in EGRA to measure whether children have learned how to blend letter sounds into words. All of the words respect a plausible orthography in the given language, even if they are not real words. For example, the word "qat" would not be used as an invented word in English because the letter "q" is never used without combining with "u," or the word "caj" would not be allowable because the "j" is never used as a word ending. Like the letter-sound subtest, this is a timed test, resulting in a measure of correct non-words

per minute, or cnwpm. **Figure 12**, below, summarizes average scores by grade for each language. The bars include all children, while the X represents the average of only those children who showed some ability (excluding zero scores). In the following figures, as a reminder, across all subtests, including non-word reading, Grade 1 and Grade 2 children read the same set of words.

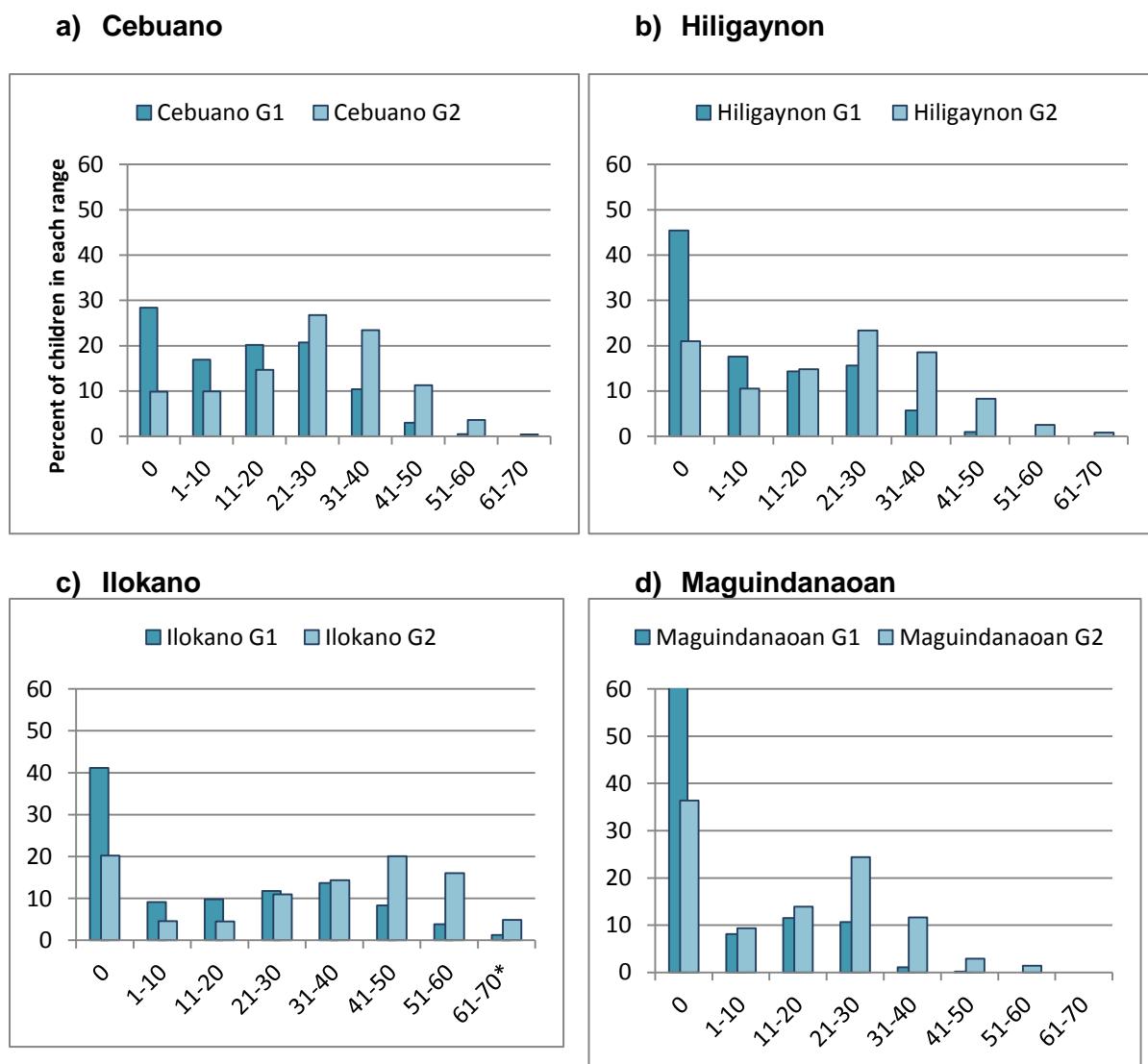
Figure 12: Correct non-words per minute: Average scores by region and grade, with and without zero scores



In this chart we can see the effect of the zero scores, especially in Grade 1. We can also see that the children reading in Ilokano were well ahead of children in other regions by Grade 2; however, the high average masks a large proportion of children with zero scores.

Looking at the distribution of scores provides more detail for analysis, including the proportion of zero scores by grade and region. Zero scores represent the largest proportion of students in Grade 1 for all languages and often remain the largest concentration of children in Grade 2 (**Figures 13a-d**). A notable exception is for Region VII (Cebuano) where in Grade 1, 45% of children read fewer than 10 non-words per minute. By Grade 2 this had dropped to 19%, and the share of zero scores dropped to less than 10% from 28%.

Figure 13: Non-word reading: Distribution of scores, by grade and language



* there are an additional 5% of children in a range higher than 70 cwpm not shown here, for the sake of consistency between graphs

Region I (Ilokano) and ARMM (Maguindanaoan) show a similar trend with the share of zero scores remaining as high as, or higher than, the next largest category of performance (1 to 10). For Ilokano, half of the children in Grade 1 could not read more than 10 non-words per minute correctly. This amount was cut in half by Grade 2, and nearly half of all children (45%) were reading above 40 non-words per minute correctly—three times more than in any other region. For Maguindanaoan the change from Grade 1 to Grade 2 is also notable, with 76% of children unable to read more than 10 non-words per minute in Grade 1, of which 68% read none. By Grade 2, this dropped to 46%, but the children who read no words correctly remained the largest category of performance. Any peak in scores is least pronounced in Region VI (Hiligaynon), with large proportions of children in the lowest categories of performance (fewer than 10 cnwpm). Nonetheless, the proportion of children who could not read more than 10 non-words per minute dropped by half in Grade 2.

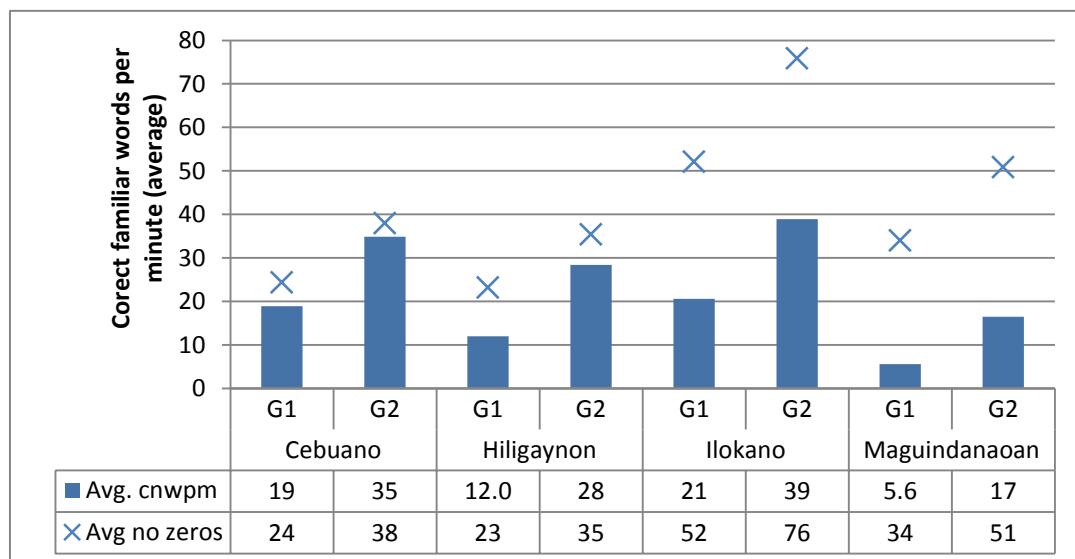
Thus we can see that from Grade 1 to Grade 2 children were acquiring word decoding skills and increasing the fluency (speed and accuracy) with which they were able to process these new words. The challenge remains of ensuring that children with the most difficulty are not

being left behind, by reducing the proportion of low-achievers while at the same time increasing the proportion of children in the highest categories.

Familiar word reading. Automaticity in reading happens when children are exposed over and over again to the same words. The decoding that is required for attacking a new word (and which is demonstrated in the non-word reading subtest) becomes increasingly automatic as words are recognized as whole units with familiar patterns. The **familiar word reading subtest** measures this skill. The words selected for this subtest come from an analysis of Grade 1 and Grade 2-level texts in use in the Philippines. From a list of 150 frequent words established by the teams of language and teaching experts who participated in the instrument adaptation, 50 words were selected, with attention to selecting a variety of word types and words that are familiar across the divisions in the region, regardless of different dialects or variations in the language in use.

Figure 14, below, shows the average scores by region and grade. Markers above each bar show what the average would be if children who scored zero were not included. The average scores were slightly higher than for non-word reading, which is expected. The effect of the zero scores was particularly pronounced for children learning to read in Ilokano and Maguindanaoan. The averages in Grade 1 and Grade 2 were very similar among the three languages Cebuano, Hiligaynon, and Ilokano (slightly less in the case of Hiligaynon), where children in Grade 1 read close to 20 correct familiar words per minute. This nearly doubled in Grade 2, with the largest Grade 2 average of these three regions being Ilokano at 39 cwpm. Although the average score in Grade 2 was triple the average in Grade 1 for children learning to read in Maguindanaoan, the Grade 2 average was still less than the Grade 1 average for two of the other languages.

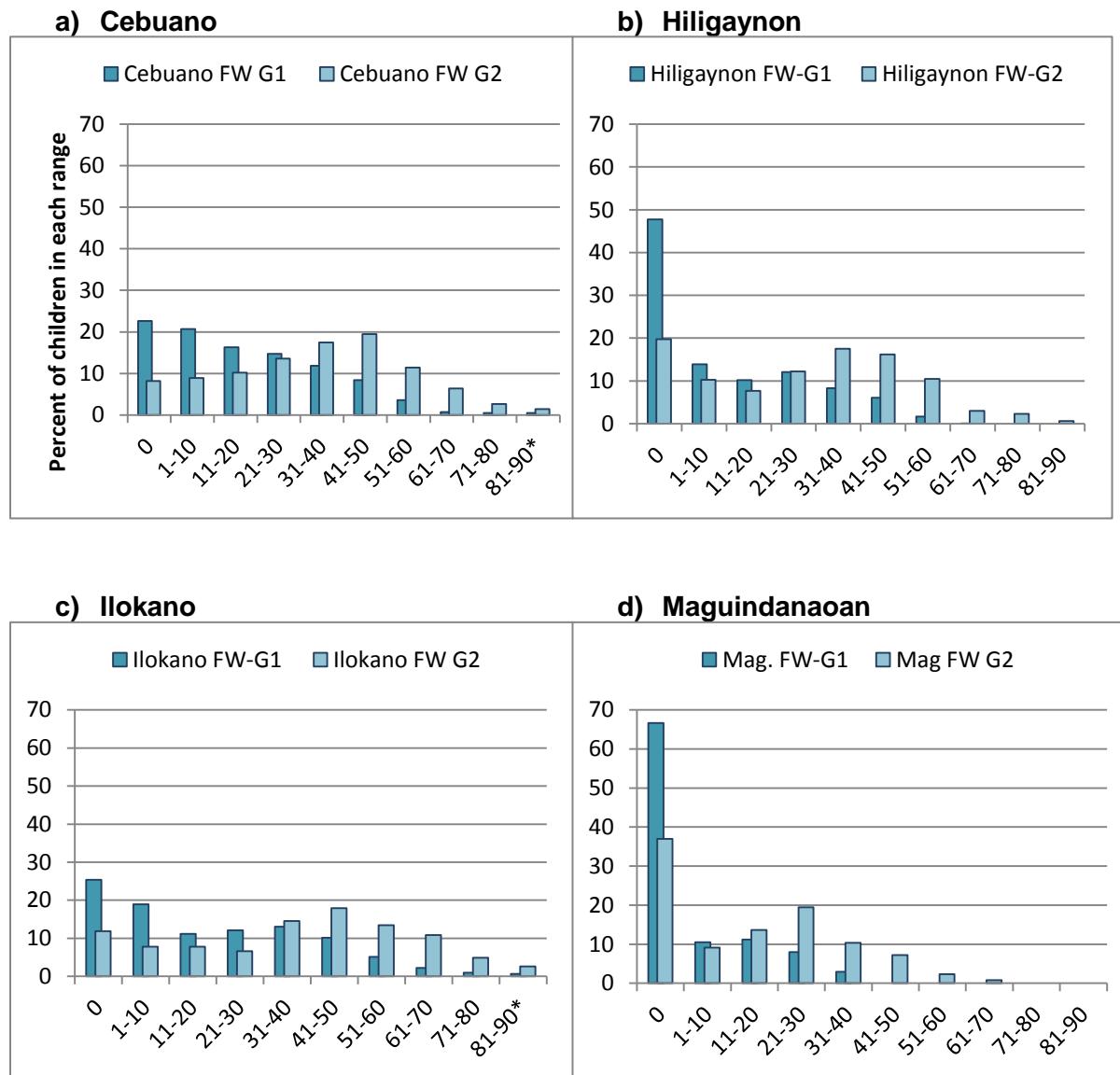
Figure 14: Correct familiar words per minute: Average scores by region and grade, with and without zero scores



Looking at the distribution of scores we can see a trend similar to that of non-word reading—zero scores remained high, at times as high or higher than scores in any other category of performance (**Figures 15a-d**). This is noticeably different for Cebuano, for which zero scores decreased to under 10% in Grade 2 from 23% in Grade 1. For Ilokano and Hiligaynon, although zero scores decreased, there were still more children in this category in Grade 2 than in the three following categories (1–10, 11–20, and 21–30). In Maguindanaoan, the share of

children with zero scores was greater than for any other category of performance in Grade 1 and in Grade 2.

Figure 15: Familiar word reading: Distribution of scores, by grade and language



* Cebuano scores above 90 cwpm left off the chart are less than 1% in each grade. Ilokano scores above 90 cwpm left off the chart are less than 1% in G1 and close to 2% in G2. Some G2 students read between 100 and 110 cwpm.

The importance of these underlying skills becomes apparent when we look at the correlations between different subtests. **Table 6** shows the Pearson correlation between the pairs of subtests listed in the first column. This type of correlation shows whether there is a linear relationship between two things. In the case of EGRA, this answers the question: To what extent are skills in one area such as letter sound identification, related to skills in another area—if a child does well in one task, will s/he also do well in the other? The closer the Pearson correlation is to 1, the stronger that relationship is. Correlations above .9, for example, show a very tight relationship. While recognizing that there is always more than just one thing influencing any of the subtest outcomes, this is an interesting way to look at how underlying skills such as recognizing letter sounds, reading familiar words with automaticity,

and word decoding contribute to the ultimate goals of connected passage reading and, ultimately, comprehension. For example, being able to decode words with ease, as demonstrated by the non-word reading task, is highly correlated with reading fluency in all languages, as is familiar word reading. Furthermore, automatic recognition of a corpus of familiar words is associated with stronger reading comprehension in most languages except Ilokano. The next section will look at the scores across regions on the connected passage (story) reading and comprehension subtests.

Table 7: Correlations between subtests

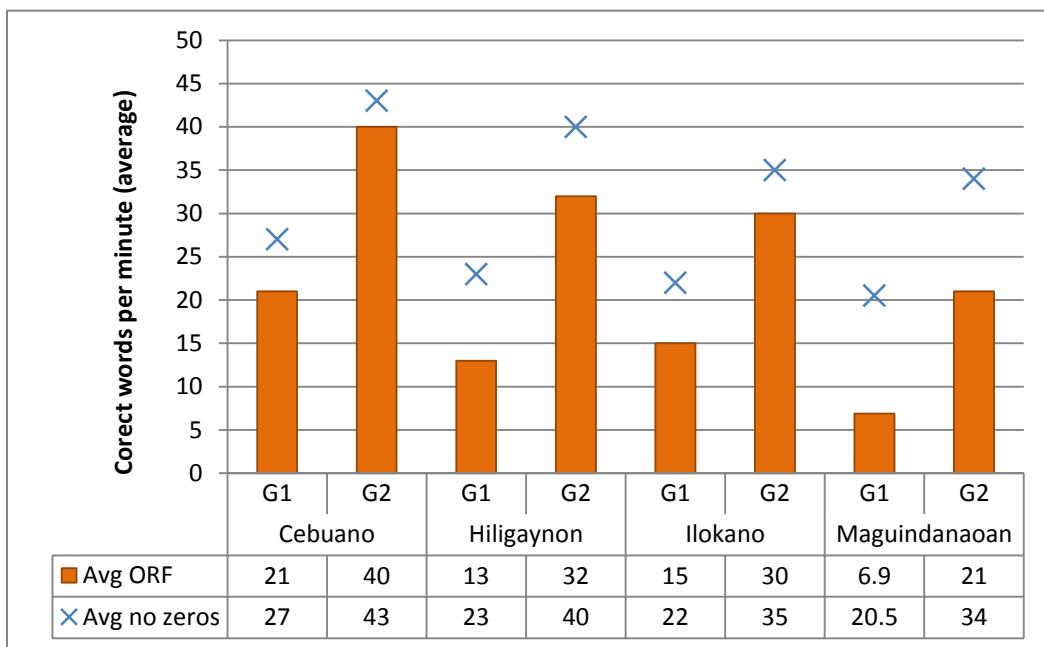
Relationship	Pearson correlation			
	Cebuano	Hiligaynon	Ilokano	Maguindanaoan
Letter sounds to familiar words	.499	.579	.480	.699
Letter sounds to non-words	.472	.576	.412	.690
Letter sounds to oral reading fluency	.467	.569	.434	.679
Familiar words to non-words	.954	.965	.935	.975
Familiar words to oral reading fluency	.970	.976	.936	.970
Familiar words to reading comprehension	.857	.890	.542	.851
non-words to oral reading fluency	.945	.955	.916	.969
non-words to reading comprehension	.825	.857	.501	.835

Reading fluency and comprehension

The ultimate goal of reading in any language is to be able to make meaning out of passages of printed text. Reading words in isolation (out of the context of a sentence) is a useful diagnostic for beginning readers, but eventually we need to know if children are capable of reading words in the context of sentences and paragraphs and understanding the meaning. For this subtest, children read a short story composed of 40–60 words (depending on the language) and are then asked questions about the story. The subtest is timed, so children read only as much as they can in 60 seconds and then are asked to stop and answer questions corresponding only to how much of the passage they completed. The basic measurement is similar to that of the timed tasks above—the rate of fluency expressed in terms of correct words per minute. In order to distinguish from the correct familiar words task and correct invented words task, we have adopted the term “oral reading fluency,” or ORF, as shorthand for the score derived from the calculation of the number of correct words read per minute in reading.

With most languages, we would expect the rate of reading connected text (sentences, paragraphs) to be slightly higher than that of reading words in isolation, since children have context cues that they can use to anticipate subsequent words. **Figure 16**, below, shows that this was the case for all languages except Ilokano. This may be explained by the fact that Ilokano is an agglutinative language (where words are formed by adding prefixes and suffixes that add a single unit of meaning, such as “plural,” or “past tense”). Most of the words in the familiar and non-word subtests for Ilokano were 2 syllables in length, with a few 1- and 3-syllable words as well. The longest familiar word was 6 letters. However, words of 6 letters and up to 12 were common in the Ilokano reading passage.

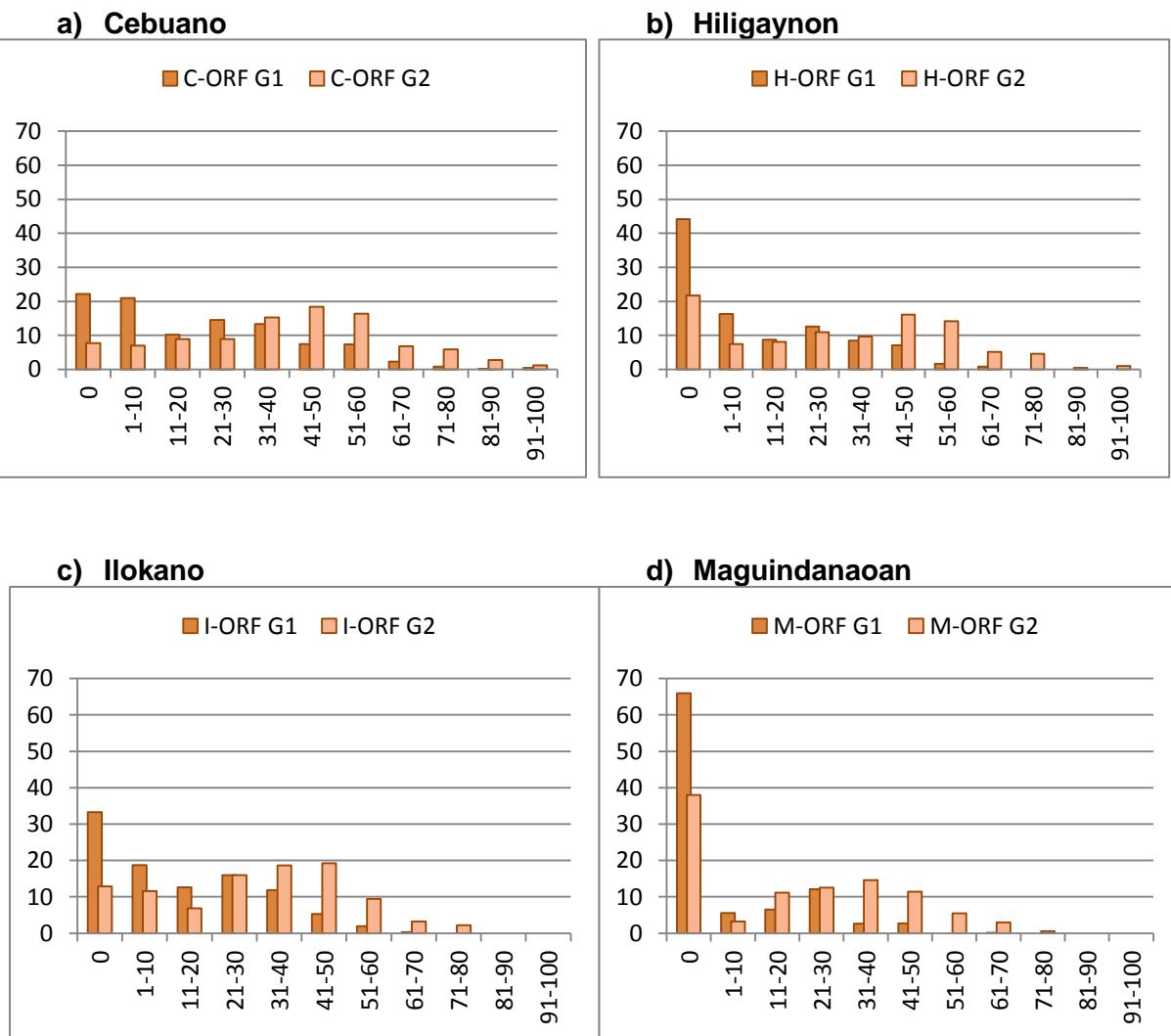
Figure 16: Oral reading fluency (with and without zero scores), by language and grade



The highest average ORF in any region was 40 cwpm (or 43 excluding zero scores), which was the score of the Grade 2 children reading Cebuano. Children reading in Hiligaynon and Ilokano averaged 32 and 30 words per minute in Grade 2, and for all languages the Grade 2 rate of reading was at least double what was achieved in Grade 1. In Maguindanaoan, although a large increase in ORF was seen from Grade 1 to Grade 2, the latter was equivalent to the highest Grade 1 score of the other three languages.

The distribution of scores, as shown in **Figures 17a-d**, points out there were children reading Cebuano and Hiligaynon who were able to achieve upwards of 80 correct words per minute on a reading passage (these two languages both had less than a half percent of students in the range of 101 to 110 and 111 to 120, but these are not shown on the graph for better readability). There was no single ORF above 79 for Ilokano, although some children surpassed this rate of fluency in isolated word reading (see **Figure 13c** and **Figure 14**).

Figure 17: Oral reading fluency: Distribution of scores, by grade and language



Apart from seeing the overall range of ability, the distribution helps see at a glance the extent to which the proportion of children in the lowest categories of ability (zero scores, ORF less than 10) is lower and the proportion of those in the highest categories is greater for Grade 2 compared to Grade 1. All the regions show a similar pattern—reductions in the lowest categories and increases in the middle to higher ranges. This indicates that students along the full range of oral reading proficiency levels were doing better in Grade 2 than in Grade 1.

With the measure correct items per minute on the timed subtests, we cannot know whether children were reading very fast with a lot of errors (so perhaps carelessly), or whether they were reading slowly, but with attention to accuracy. Therefore it can be useful to also look at accuracy, or the proportion of words that were correct out of those attempted.

Table 8: below, shows this indicator for all languages and grades for the three word reading tasks (non-word reading, familiar word reading, and story reading). These figures do not include zero scores.

Table 8: Reading accuracy: Percent correct out of total attempted

	Cebuano		Hiligaynon		Ilokano		Maguindanaoan	
	G1	G2	G1	G2	G1	G2	G1	G2
Non-words	74%	81%	73%	80%	81%	89%	67%	73%
Familiar words	76%	88%	79%	86%	77%	89%	67%	77%
Short story	79%	90%	76%	87%	76%	85%	72%	84%

With a few exceptions, for all languages and tasks accuracy improved by less than 10 percentage points from one grade to the next. This means that Grade 1 children who showed some reading ability were already reading with a great deal of accuracy, and by Grade 2 their cognitive skills could be used for increasing automaticity and therefore overall fluency. We can also see that as reading becomes more embedded in context and familiarity (short story reading as opposed to reading words in isolation), accuracy usually improves. Ilokano was the only exception where accuracy was lower when reading a short story, compared to when reading words in isolation in both Grade 1 and Grade 2. It is likely that this was due to the highly agglutinative nature of the language.

Reading comprehension. It is widely accepted that adequate reading fluency is necessary for comprehension (though not sufficient); this does not mean that fast reading always enables comprehension, but slow reading (hesitant, decoding) will most likely interfere with it.¹⁵ Reading accurately and knowing the vocabulary, the context/setting of the story or text, grammar and sentence structure, and other conventions such as punctuation are also equally important. This early reading assessment does not attempt to measure all possible influences on reading comprehension, but the listening comprehension and familiar words subtests, with the measure of reading fluency from the ORF subtest, provide important clues. **Table 9** shows for each language the minimum number of words required to attempt each comprehension question. While the actual breakdown varies by language, there is usually a question corresponding to every 10 words.

Table 9: Minimum number of words attempted for each comprehension question, by language

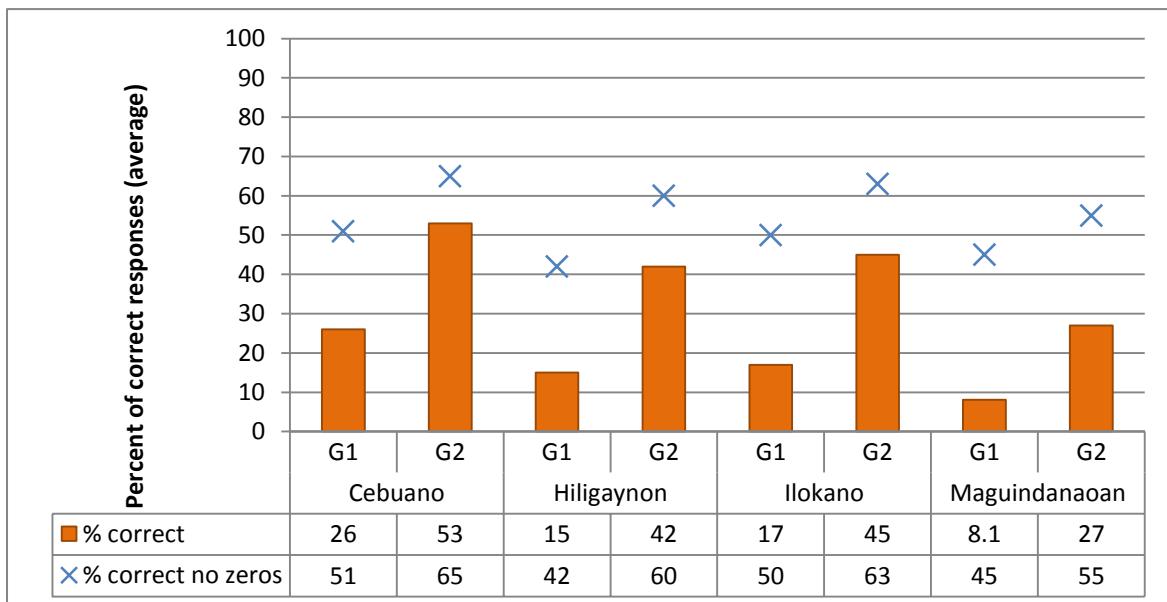
Language	Number of words read in order to attempt each question				
	1	2	3	4	5
Cebuano	10	21	34	43	52
Hiligaynon	7	16	26	34	49
Ilokano	8	11	23	30	45
Maguindanaoan	10	20	26	33	40

Figure 18, below, shows the average score for the reading comprehension task. The score is the percent of questions answered correctly out of the total possible—five in each case. Therefore, a child who only got one question correct would have a score of 20%, and a child

¹⁵ Wagner, Daniel A. (2012). Smaller, Quicker, Cheaper: Better learning assessments for developing countries. Prepared for IIEP under contract from FTI under the IIEP Quality Learning Indicators Project (QLIP).

who answered four questions correctly would have a score of 80%. Keep in mind that children may not have attempted all five questions since the test is designed to discontinue after one minute, regardless of where the child was in the short story. Therefore this measurement gives us an indication of both fluency and comprehension, since children who attempted five questions were reading above a certain range of fluency.

Figure 18: Reading comprehension: Average percent correct by language and grade



From this table we can see that the average of all children's scores did not exceed 70% for any language or grade, even when zero scores are omitted from the calculation. This doesn't mean that high scores were not present, however. As in the other subtests, the distribution of responses gives us a better indicator of what children were able to achieve. To get a truer sense of whether students comprehended what they were reading—at whatever level of fluency—**Figure 19a-d** below show the percentage of students in each of the four languages in Grades 1 and 2 who achieved different levels of comprehension ranging from answering no questions correctly (no matter how many they attempted) to answering all the questions they attempted correctly.

A very particular pattern is evident in each of the languages, with some important distinctions to note. To start, in all the languages there were large proportions of zero correct responses, especially for Grade 1 students, of whom 50% or more did not answer any attempted questions correctly.¹⁶ In Maguindanaoan, 80% of the Grade 1 students had 0% correct, but many fewer Grade 2 students had 0% correct. For Cebuano, almost half the students fully comprehended the text that they read and were able to answer correctly all the questions they attempted. In all the languages we saw a greater share of Grade 2 students able to answer more than half the questions correctly, with between 15 and 50% of students able to answer 100% of questions correctly.

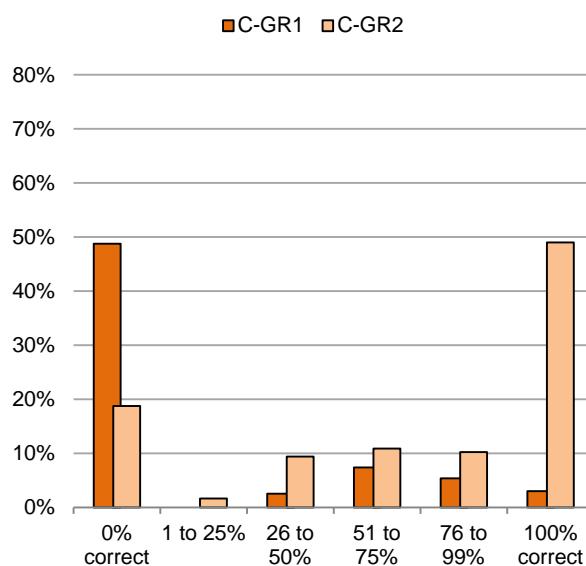
It would appear that as students become more comfortable reading in the language (for example, as they move from first to second grade), they are more fully comprehending what

¹⁶ Bear in mind that these figures already exclude children who attempted no questions because they were unable to read any words in the story.

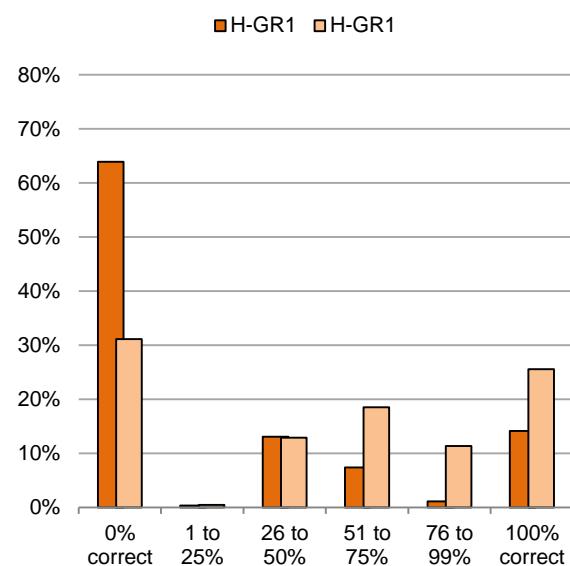
they read. However, there is still cause for concern as even in Grade 2 there were appreciable percentages of students failing to comprehend even part of a short passage—19% in Cebuano; 31% in Hiligaynon; 30% in Ilokano; and 50% in Maguindanaoan.

Figure 19: Distribution of comprehension scores (percent correct out of comprehension questions attempted) by language and grade

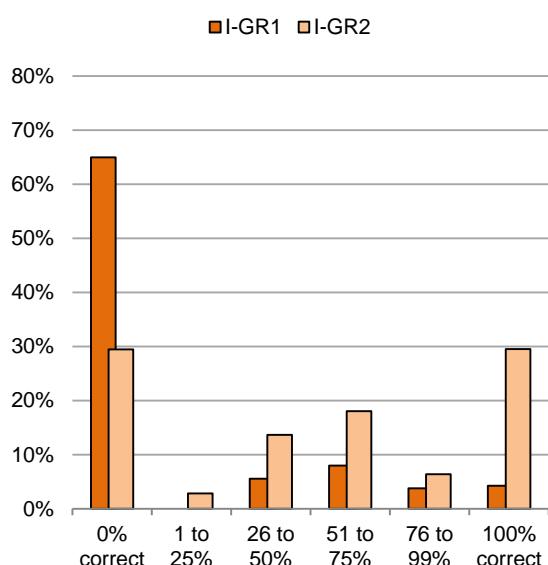
a) Cebuano



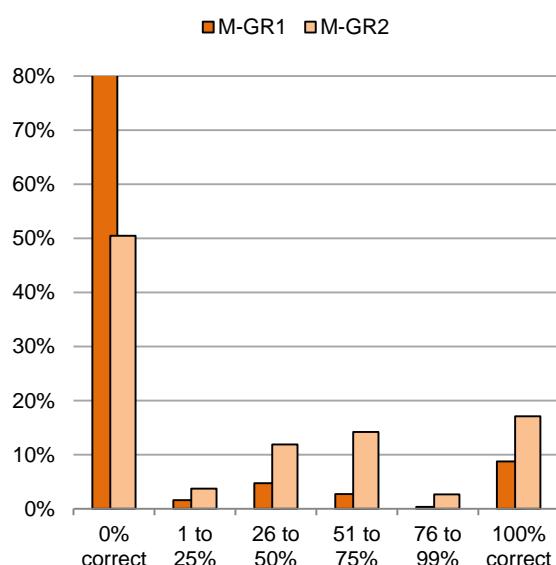
b) Hiligaynon



c) Ilokano



d) Maguindanaoan



Conclusion

Across subtests, there are some similar trends, notably the persistence of a large share of zero scores even at the end of Grade 2. Efforts should be made to ensure that classroom instruction is sufficiently differentiated to take into consideration the range of abilities and needs in the classroom and to ensure that children who are falling behind receive the help they need to catch up.

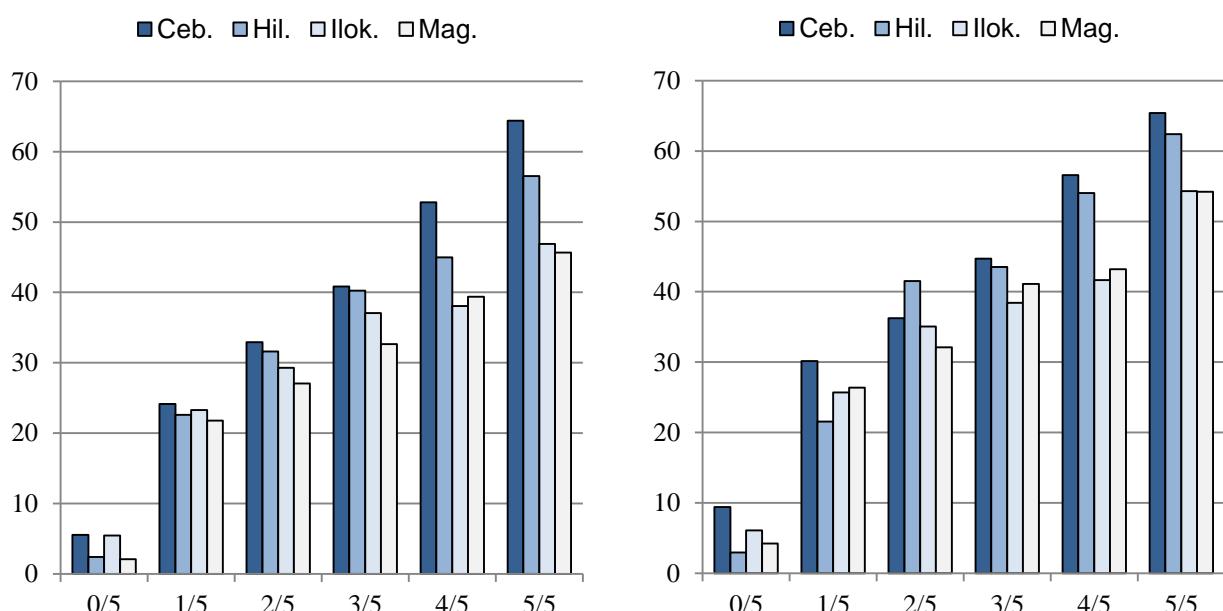
Students are reading accurately, even those who have limited fluency. This implies that the basic skills of sounding out unfamiliar words or sight reading familiar words are developing, but reading must become increasingly automatic, especially for students at the end of Grade 2. More practice oral reading should help address this (and as we will see in Section E below, students whose teachers ask them to read in the mother tongue have a higher likelihood of being in the top 25% of fluent readers for their grade and region).

As DepED considers benchmarks for oral reading fluency in these four mother-tongue languages, the charts presented in **Figure 20** may provide some useful information. The average oral reading fluencies corresponding to different levels of reading comprehension are plotted for each language for Grades 1 and 2. The horizontal axes of these two graphs show the possible comprehension scores, ranging from zero out of five to five out of five. The vertical axis is the number of words read correctly per minute for students in each of the five comprehension categories.

Figure 20: Comprehension and oral reading fluency

a) Grade 1

b) Grade 2



What we see is that students at the highest levels of comprehension across Grades 1 and 2 had oral reading fluencies in the ranges shown for each language in **Table 10** below.

Table 10 ORF and comprehension

	80% comprehension	100% comprehension
Cebuano	52-55 wpm	62-65 wpm
Hiligaynon	45-54 wpm	55-65 wpm
Ilokano	38-41 wpm	45-55 wpm
Maguindanaoan	39-41 wpm	45-53 wpm

D. How well are children acquiring L2 and L3?

Filipino competency

As outlined in **Table 4**, children are expected to begin learning Filipino orally in the second quarter of Grade 1 and begin reading in the fourth quarter. They continue oral competence and literacy development throughout Grade 2, so that by Grade 3 some subjects can be taught in Filipino as the medium of instruction. **Table 5** above provided some detail about teachers' expectations related to the grade in which they believe that students should be able to read in Filipino. Of concern is whether Grade 2 teachers believe that their students should be reading Filipino. **Table 11** below shows how that varies across the regions.

Table 11: G2 teacher expectations for Filipino reading

Mother Tongue	% of G2 teachers stating students should read in Filipino in:		
	Grade 1 or 2	Grade 3	Grade 4 or higher
Cebuano	62%	28%	10%
Ilokano	75%	18%	7%
Hiligaynon	28%	52%	20%
Maguindanaoan	34%	49%	17%

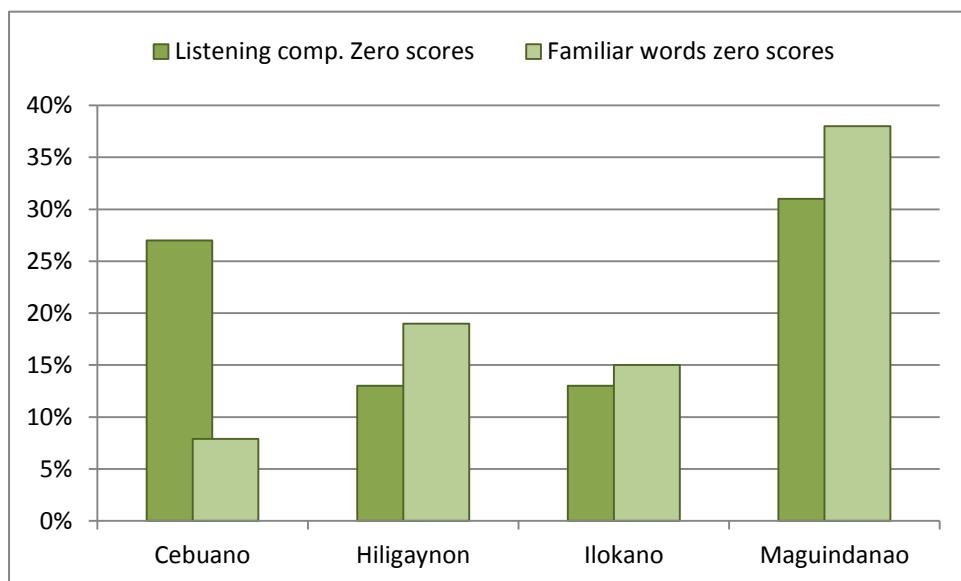
What these data show is that in Cebuano (62%) and Ilokano (75%) classrooms the majority of Grade 2 teachers expected their students to be learning or have already learned to read in Filipino. In Hiligaynon and Maguindanaoan only 28% and 34% expected that. Interestingly, larger proportions of Grade 1 teachers had the expectation that students should read Filipino in Grade 1 or Grade 2: 82% in Cebuano, 94% in Ilokano, 81% in Hiligaynon, and 50% in Maguindanaoan.

Teacher expectations regarding literacy in Filipino were raised because the assessment of reading capacity included evaluations of student performance in Filipino as well as in the mother tongues. This assessment measured two competencies in Filipino at the end of Grade 2. The two tasks were familiar word reading and listening comprehension. The frequent word items that were used were from the same EGRA that was developed for the 2013 EGRA administered to Grade 3 students, but the order was changed. The listening comprehension passage was one that had been piloted but not used in 2013.

Before learning to read familiar words students would be expected to understand the language orally, and this is the sequence in which instruction is introduced. The results of the listening comprehension and familiar word tasks in Filipino, in terms of how many children scored zero on each task, are shown in

Figure 21 below.

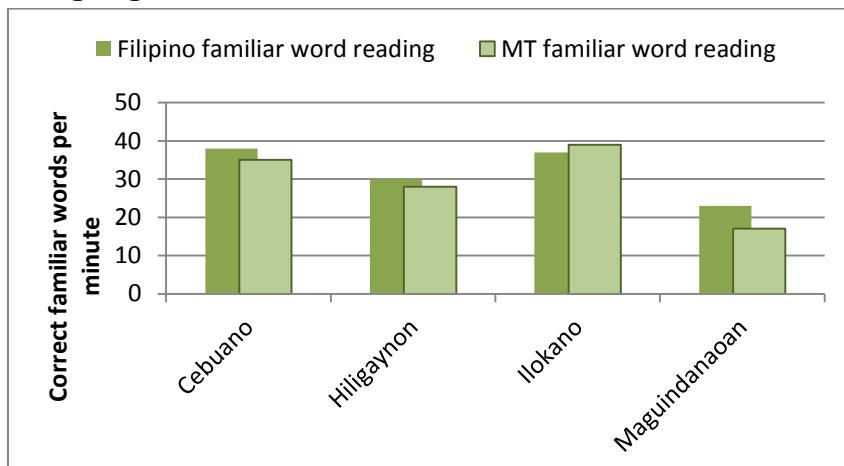
Figure 21: Filipino—Proportion of zero scores



It is surprising to see that there was such a high proportion of zero scores for Filipino listening comprehension in Cebuano schools, but a much lower proportion for familiar word reading. This could indicate that children were learning the words by rote, but not acquiring oral comprehension. Or there may be another reason comprehension skills were low that was not measured by this assessment (for example, inhibition or the unfamiliar context of the story). The highest listening comprehension score was 68% in the schools using Ilokano as the mother tongue followed by Hiligaynon schools with an average of 58% of questions correct. Cebuano and Maguindanaoan classrooms scored 47% and 48% correct, respectively.

For familiar word reading, the rates of correct words per minute are shown in **Figure 22**, below, next to the scores for mother tongue familiar word reading. Despite cases where listening comprehension was low, children were learning to recognize Filipino familiar words at a rate as good, or better, than for the mother tongue.

Figure 22: Filipino familiar word reading: Correct words per minute by language



It is assumed that these Grade 2 students were in Grade 1 in schools that had been teaching using the mother tongue. However, 13% of Grade 2 students in Maguindanao and Cebuano reported that their teacher mostly spoke Filipino last year. It is also possible that many of the children in this cohort also attended preschools using Filipino before MTB-MLE was rolled out, but this indicator was not measured in this study.

It is useful to recall that the average familiar word reading score of grade 3 students nationwide tested in 2013 with the same items was 54 correct words per minute. Therefore, if we take the average intergrade progression in mother tongues as a guide—children acquire approximately 15 correct words per minute between Grade 1 and Grade 2—children are on track for achieving equivalent Filipino capacity in Grade 3 through the MTB-MLE system as they did under the previous system.

While it is encouraging to see that the MTB-MLE program is at least partially achieving the goal of Filipino literacy development in Grade 2, one wonders why one year of instruction would provide the same results for Filipino familiar word reading as two years of instruction for mother-tongue literacy. This may be due to some of the factors above (some additional influence of Filipino in Kindergarten or Grade 1) or it may also demonstrate that once children are able to acquire foundational reading skills such as letter-sound correspondence and phonemic awareness in their mother tongue, they are able to transfer these skills to other languages. In fact, the Pearson correlation tables show a strong positive relationship between performance in the mother tongue and performance in Filipino familiar word reading, as shown in **Table 12:** Correlations between , below.

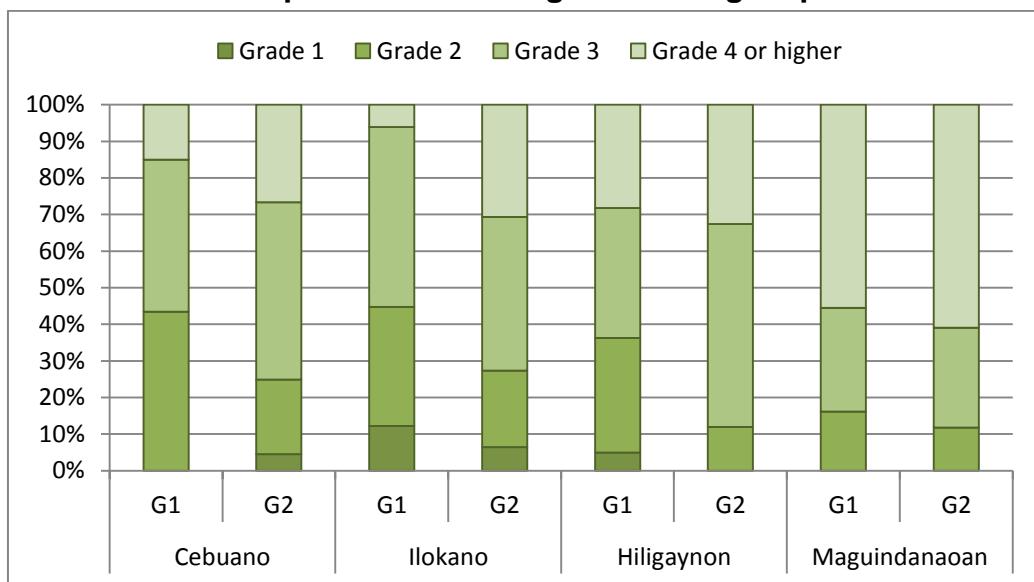
Table 12: Correlations between subtests

Relationship	Pearson correlation			
	Cebuano	Hiligaynon	Ilokano	Maguindanaoan
MT familiar words to Filipino familiar words	.946	.978	.970	.960
MT invented words to Filipino familiar words	.889	.954	.966	.966
MT story reading to Filipino familiar words	.967	.980	.971	.977

English competency

Figure 23, below shows teacher expectations related to learning English. As with Filipino, opinions were very different among teachers as to when children should acquire English reading fluency. The teachers in schools using Maguindanaoan disproportionately believed that children should only acquire fluency in Grade 4 or later. Most other Grade 2 teachers believed that fluency is possible in Grade 3.

Figure 23: Teacher expectations for English reading acquisition

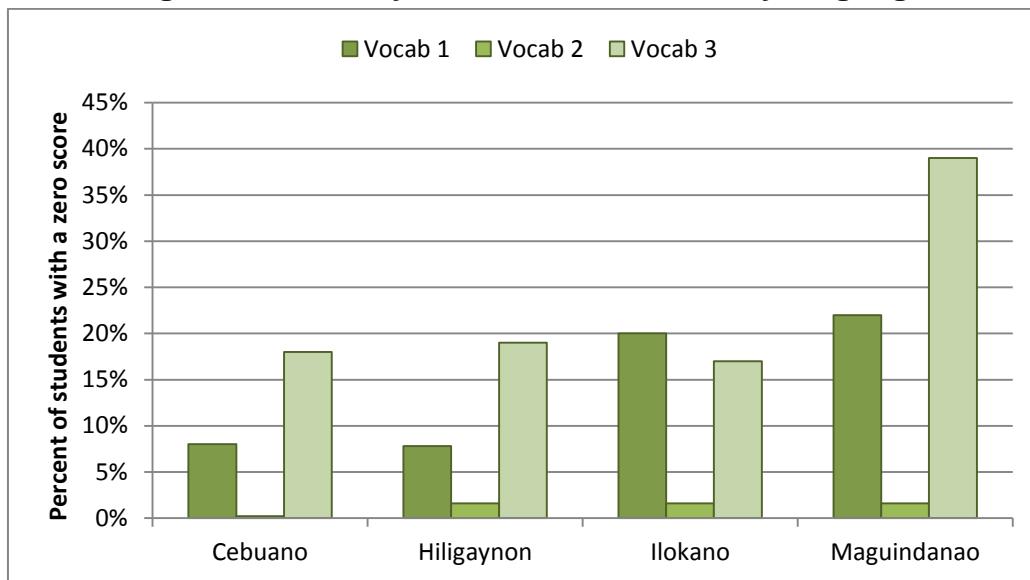


The English diagnostic was an oral vocabulary test that had not yet been used in the Philippines, but has been tested in other countries. Children were asked to point to familiar objects and to execute simple commands. A few of the vocabulary items were changed to be more locally appropriate. According to DepED specialists, the items used are all present in the curriculum, and it can be reasonably expected that children would know these words by the end of Grade 2.

- The first task (Vocab. 1) asked children to identify body parts—arm, mouth, neck, foot, knee, back, shoulder, elbow.
- The second task (Vocab. 2) asked children to identify objects in the room—shoes, chair, pencil, bag, paper, wall.
- The third task asked children to execute commands: Put the pencil...behind you, next to the paper, on the paper, under the paper, in front of you, to the left of the paper.

The assessment results (see **Figure 25**) showing the percentages of zero scores, show that children had the most difficulty with the third task, executing simple commands. They had the least difficulty with the task related to identifying objects in the room, where almost no students scored zero in any region.

Figure 24: English vocabulary: Percent zero scores by language



While there was some variance among the four regions in the proportion of children who were unable to answer any items correctly, performance was remarkably similar for the percent of correct responses, as shown in **Figure 25** below, but with students learning in Maguindanaoan scoring once again lower than the students in the other regions.

Figure 25: English vocabulary: Percent of correct responses by language

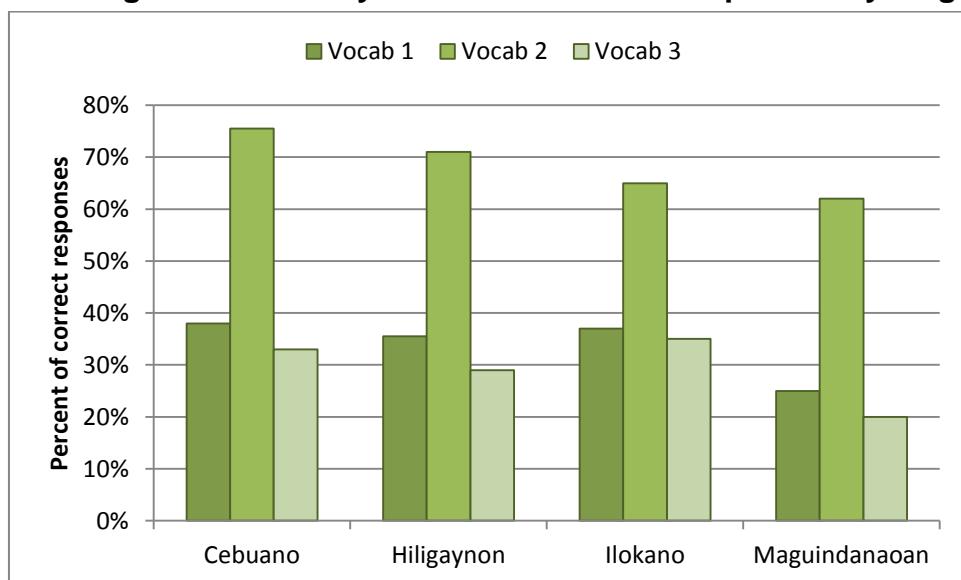


Table 13: Summary of EGRA results: Average by subtest and grade

Item	All	Cebuano		Hiligaynon		Ilokano		Maguindanaoan	
		G1	G2	G1	G2	G1	G2	G1	G2
Pre-reading skills (average percentage of correct responses) _pctscore									
Phonemic awareness (10 questions)		51%	61.5%	31%	52%	30%	48%	25%	41%
Listening comprehension (3 questions)		50%	54%	23%	44%	45%	57%	44%	53%
Basic reading skills (average correct items per minute)_score									
Letter-sound identification (100 letters)	20	18	12	13	15	17	7.1	14	
Non-word reading (50 words)	14.5	25	10	21	17	34	5.3	15	
Familiar word reading (50 words)	19	35	12	28	21	39	5.6	16.5	
Story reading (35-60 words)	21	40	13	32	15	30	6.9	21	
Reading comprehension (average percentage of correct responses)_pctscore									
Percentage correct out of total questions asked	26%	52.5%	15%	41%	17%	45%	8.1%	27%	
English and Filipino skills (Grade 2 only, percent correct)									
English vocabulary 1 (body parts, 8 items)	-	38%	-	35.5%	-	37%	-	25%	
English vocabulary 2 (objects, 6 items)	-	75.5%	-	71%	-	65%	-	62%	
English vocabulary 3 (prepositional commands, 6 items)	-	33%	-	29%	-	35%	-	20%	
Filipino listening comprehension (3 questions)	-	47%	-	58%	-	68%	-	48%	
Filipino familiar word reading	-	66%	-	53%	-	64%	-	57%	

Table 14: Summary of EGRA results: Proportion of zero scores by subtest and grade

Item	All	Cebuano		Hiligaynon		Ilokano		Maguindanaoan	
		G1	G2	G1	G2	G1	G2	G1	G2
Pre-reading skills (average percentage of correct responses)									
- Phonemic awareness percent of zero scores		18%	15%	31%	26%	14%	10%	50%	30%
- Listening comprehension percent of zero scores		23%	15%	23%	14%	49%	29%	26%	15%
Basic reading skills (average correct items per minute)									
- Letter sounds percent of zero scores		11	10.5	22	14	23%	21%	55%	35%
- Familiar words percent of zero scores		23	8.2	48	20	25%	12%	67%	37%
- Non-words percent of zero scores		28	9.8	45	21	41%	20%	68%	36%
- Story reading percent of zero scores		22	7.8	44	22	33%	13%	66%	38%
Reading comprehension (average percentage of correct responses)									
- Reading comprehension percent of zero scores		49%	19%	64%	31%	65%	29.5%	82%	50%
English and Filipino skills (Grade 2 only, percent correct)									
- Vocabulary 1 percent of zero scores	-	8.0%	-	7.8%	-	20%	-	22%	
- Vocabulary 2 percent of zero scores	-	0.2%	-	1.6%	-	1.6%	-	1.6%	
- Vocabulary 3 percent of zero scores	-	18%	-	19%	-	17%	-	39%	
- Listening comprehension percent of zero scores	-	27%	-	13%	-	13%	-	31%	
- Filipino familiar words percent of zero scores	-	7.9%	-	19%	-	15%	-	38%	

E. What factors are associated with reading skills?

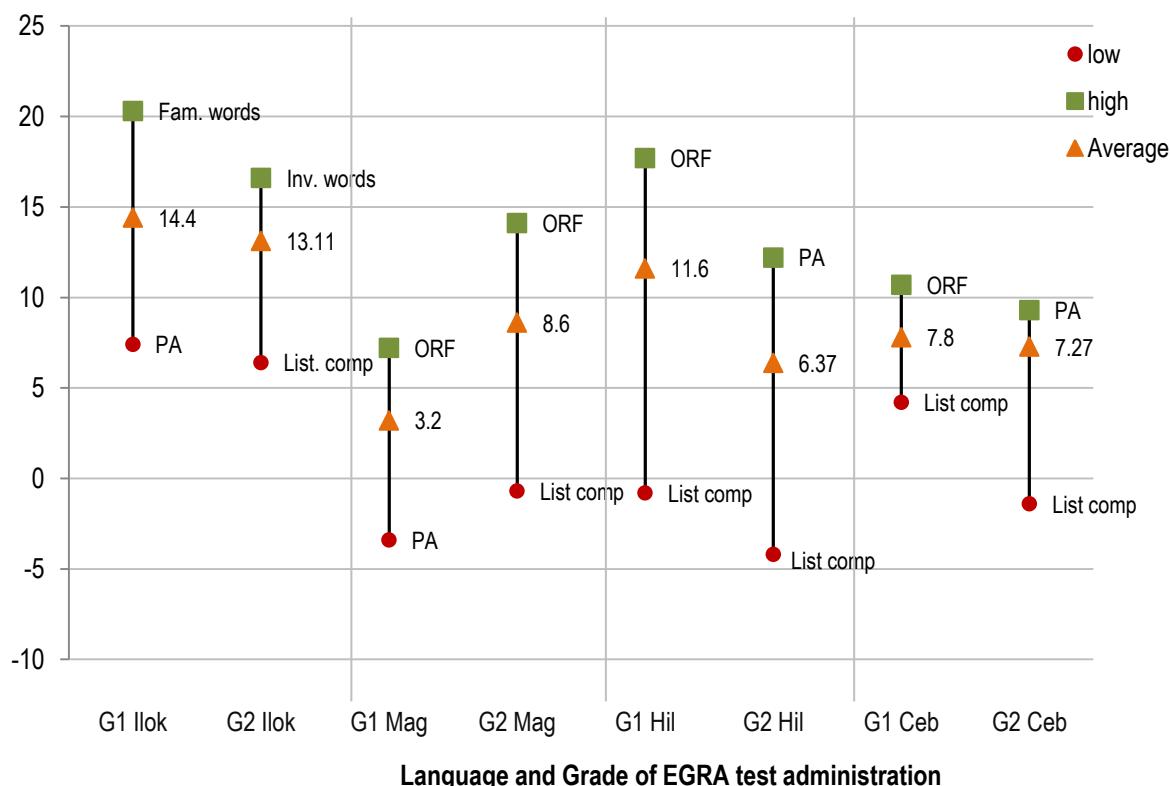
An important conclusion from the 2013 EGRA studies was the existence of significant differences in achievement among the regions of the Philippines, between boys and girls, and across children of different age levels (especially children who are overage for grade). The summary results above clearly show that the inequity among regions still exists, especially for the ARMM region (children learning to read in Maguindanaoan), but to some extent also in schools that are using Hiligaynon as the mother tongue. In this section, we explore home, school, and other demographic factors that may be associated with reading performance.

Note on interpretation of data: It is important to keep in mind that inter-language comparisons should be done with caution. Each language has its own specific characteristics, including linguistic characteristics, but also such things as how well established it is, whether it has a long literary history, whether previous programs existed to use the language as the mother tongue, and more. Languages are also directly associated with geographic regions, and regions have their own characteristics such as geography, urban and rural divides, communications infrastructure, and more, which distinguish them from each other. This research did not measure all of these different possible factors that could explain differences among regions. Furthermore, although statistical analysis was done to establish correlations between variables within a region/language, the sample size is generally too small to get strong statistical power. Therefore the correlations presented in this section are derived from combining all languages into one large sample and then running the variables against output variables that are comparable across languages. Comparable variables are: % of zero scores, and top 25% of achievers.

Gender

Figure 26, below, shows that girls fairly consistently outperformed boys in both grades in all four regions. The graph shows the percentage point differences between the average scores for girls and boys in several different ways. The orange “Δ” shows the average difference in the percentage of correct responses given by girls compared to boys across all subtests (shown by language and grade). So for example, in Grade 1 in Ilokano, girls on average scored 14 percentage points higher than boys, across all EGRA subtests. The other points plotted for Grade 1 Ilokano, show the subtest that had the greatest gender difference—in this case familiar word reading, on which girls outscored boys by about 20 percentage points—and show the subtest that had the lowest gender difference—phonemic awareness, on which girls outperformed boys by about 7 percentage points.

Figure 26: Summary of differences in performance by gender, across languages and subtests (percentage point difference, girls-boys)



* Differences may not be statistically significant

While on average girls outperformed boys across all subtests, there were a few exceptions where boys had average percent correct scores higher than girls (where the bottom points on the graphed lines go below zero). These include phonemic awareness in Grade 1 and listening comprehension in Grade 2 in Maguindanaoan, listening comprehension in Grades 1 and 2 in Hiligaynon, and listening comprehension in Grade 2 in Cebuano.

Other factors associated with reading skills

To examine how different factors in children's homes and schools impact their learning to read, we conducted two kinds of analysis on the data collected through the student interviews. The first analysis looked at the impact of different factors on the average oral reading fluency of students in Grades 1 and 2. The second analysis looked at how these factors impact the likelihood of a student being in the top 25% of performance on oral reading fluency for his/her grade.

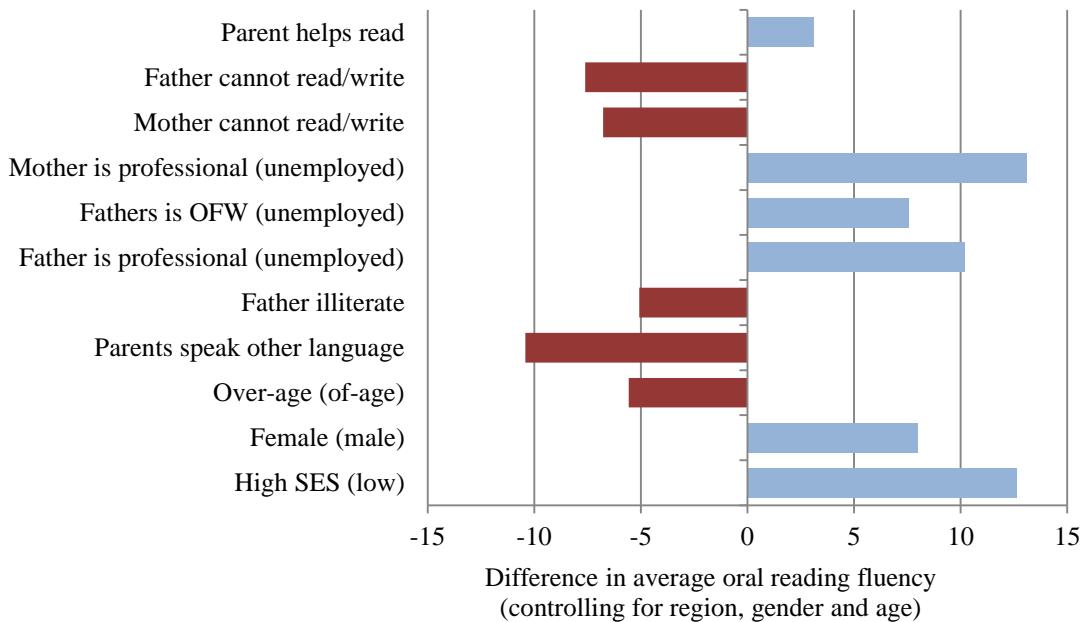
Each student included in the sample was asked a battery of questions about such things as his or her home situation, what goes on in school, and what reading materials s/he may have access to. The responses to these questions were analyzed using multivariate regression to determine how they impact the average reading performance of students. The analysis controls for:

- Region—because there are distinct differences in the average performance across the four regions included in this survey
 - Grade—because Grade 2 students have higher average performance than Grade 1

- Gender—because on average, girls outperform boys on story reading
- Socio-economic status
- Whether students are the correct age for their grade.

Figure 27, below, shows the set of home factors that had statistically significant correlations with oral reading fluency (story reading).

Figure 27: Differences in oral reading associated with home factors



The bars show the amount by which oral reading fluency was above or below the average performance for the comparison group. For example, students whose parents helped them read at home (the first factor at the top of the graph) on average read 2.5 more words per minute than those whose parents did not. The biggest positive impacts were associated with the following factors:

- Students whose mother was a professional read on average 13 more words per minute than those whose mothers were unemployed
- Similarly, students whose father was a professional on average read 10 more words per minute than those whose fathers were unemployed
- Students in the highest socio-economic quartile (the top 25% of households on the relative scale created for this survey) read on average 12.5 more words per minute than those in the lowest quartile.

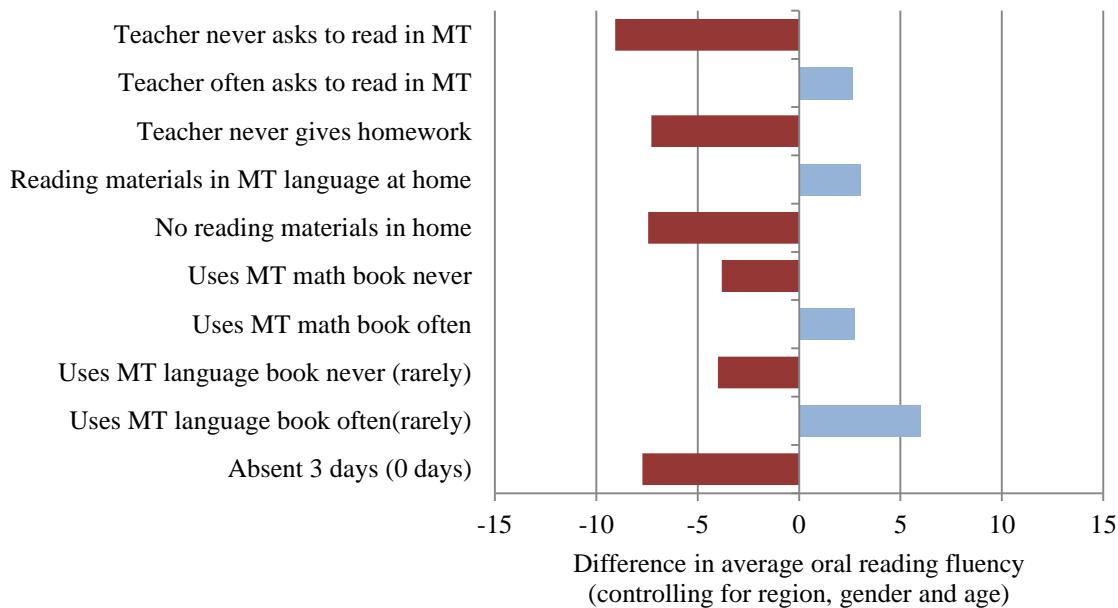
Note that these three characteristics were likely highly correlated themselves. In other words, students in the highest socio-economic quartile were those whose parents were professionals.

Factors that were negatively associated with average oral reading performance included having parents who spoke a language other than the mother tongue being used as the medium of instruction and who could not read or write. The former factor accounts for a reduction of 10.5 words per minute in a student's average oral reading fluency.

School-related factors that showed statistically significant relationships with oral reading fluency included such things as the use of language, assignment of homework, and access to

and use of reading materials. **Figure 28**, below summarizes the differences in average reading performance associated with these kinds of variables.

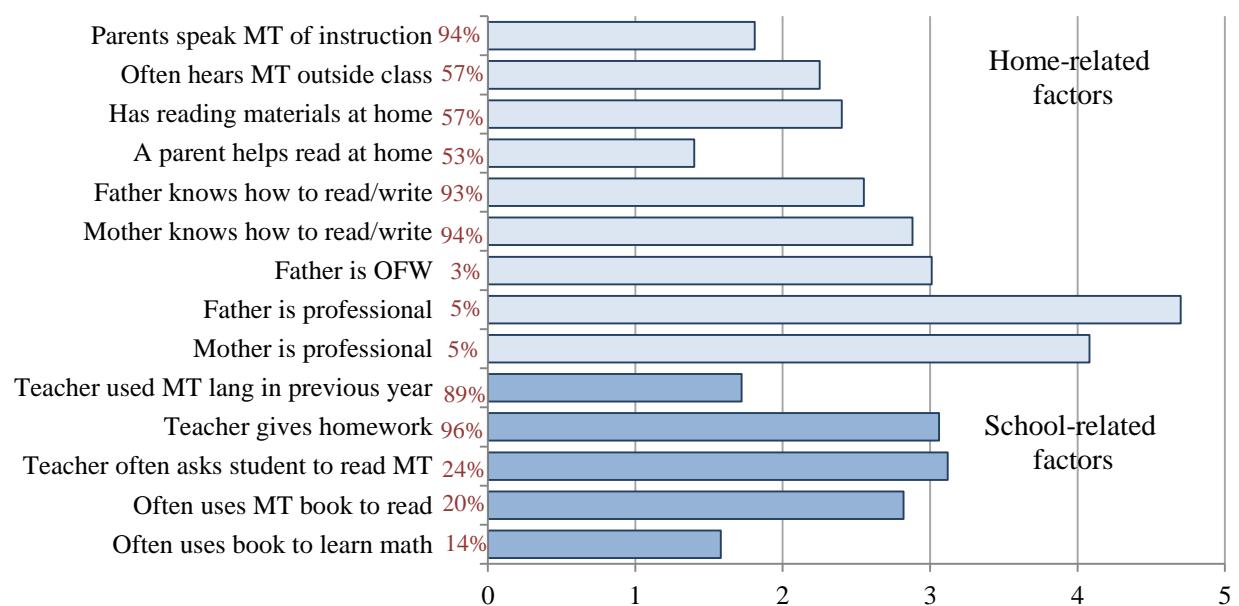
Figure 28: Differences in average oral reading fluency associated with school-related factors



We see that students of teachers who never asked students to read in the mother tongue and never assigned homework had average oral reading fluency 9 and 7 words per minute lower than those who teachers at least occasionally asked students to do those two things. Students without reading materials at home and who never used their mother-tongue books (for reading or math) had lower average oral reading fluency than those who had books and used them. Students who reported being absent three days in the previous week had average oral reading fluency that was about 8 words per minute lower. Students who often used their mother-tongue book to read had average reading fluency 6 words per minute higher than those who only rarely did so.

One concern we had when using regression to analyze the relationships between the above mentioned kind of in-school and at-home factors was that average oral reading fluency in the different languages may have varied solely due to linguistic or orthographic differences between the four languages. To account for that, we looked at EGRA performance within each language and identified the top 25% of students based on oral reading fluency. We then used logistic analysis to determine how home- and school-related factors affected the probability of a student being in the top 25% of readers. The results of such an analysis yielded odds ratios, which indicated the extent to which a factor increased or decreased the likelihood of a student being in the top 25%. **Figure 29** below shows the factors that increased the probability of a student performing in the highest quartile. Many of the factors were the same as those displayed previously.

Figure 29: Odds ratios for variables associated with being in the top 25% of readers



The bars in this graph represent the size of the odds ratio for each factor, and the percentages in red at the base of each bar indicate the proportion of the population of students that exhibit that characteristic. Among the home-related factors, a student whose mother or father was employed as a professional (defined by such occupations as a lawyer, nurse, teacher, police officer, etc.) was 4 and 4.7 times more likely to be in the top 25% of readers than those whose parents were unemployed. Such students represented only 5% of the Grade 1 and 2 students in the survey. Students with reading materials at home (57% of the population) were 2.4 times more likely to be top performers than those who did not have materials at home, and those whose parents spoke the language of instruction were 1.8 times more likely to be top performers than those whose parents did not speak the mother tongue of instruction (representing only 6% of the population in these four regions).

School-related factors that increased the likelihood of being in the top 25% of readers include teachers giving homework (3 times more likely than students whose teachers did not), teachers often asking students to read in the mother tongue (3.1 times more likely than students whose teachers never did), and students often using their mother-tongue language books to read (2.8 times more likely than those who never did). Additionally, classrooms that had “many (more than 8)” posters on the walls, according to the classroom observations, were 4.5 times more likely to be in the top 25% of classroom.

The relationships in both of the analyses discussed above are fairly consistent. There are really no surprises in terms of the out-of-school factors that were associated with reading performance—higher socio-economic status, parents who are literate and who speak the language of instruction, and parents who help read at home. These analyses also confirm the in-school factors one would normally associate with higher reading performance—teachers regularly using the language of instruction, having students read to them in that language, and assigning homework, as well as students not missing school and making use of their mother-tongue textbook for reading practice.

We also tried to associate MTB-MLE implementation with reading outcomes. However, when looking at school level variables such as those used for the MTB-MLE indices, the sample size reduced dramatically from the 3,200 students down to 160 schools. Because of this reduction and because of the limited amount of variation among schools on most of the variables included in the indices, no relationships showed up as statistically significant.

III. Summary and Conclusions

A. Summary of EGRA results by language

In this section we will summarize findings by language, including information about the characteristics of teachers, students, and classrooms in each population. Figures in the tables here, as above, are weighted, and therefore can be considered representative of the full population of schools in the four regions. As a reminder, the population is defined as schools that reported that they were using the target mother tongue languages as the medium of instruction. When a figure is provided without mentioning the grade level, it can be assumed that it is for Grade 1 and Grade 2 combined. Readers may also refer to **Table 6** for a summary of statistics related to MTB-MLE implementation or Section 3-C and D for summary of performance on reading and language tasks.

Cebuano

As described in the background section above, Region VII has been actively implementing MTB-MLE since 2012, though the school year 2013–2014 was disrupted by Typhoon Haiyan. All Grade 1 and Grade 2 teachers have been trained on how to use the teachers' guides and learner materials, and also on conducting EGRA. This region is also receiving support from USAID/Basa Pilipinas.

According to data from this study, most teachers in schools using Cebuano were native speakers and were using this language to teach (though less often in Grade 2 on both measures); two-thirds of teachers considered themselves “very comfortable” teaching in the mother tongue. Less than half felt very comfortable with the mechanics of the language. Eighty-two percent of students said their teacher used mostly the mother tongue for teaching, and observation data also suggested that teachers used Cebuano a large majority of the time, even in other subject areas. However, nearly a quarter of students never used the mother-tongue learners’ guide for language or math (the prevalence of this was higher in Grade 2 than in Grade 1). Furthermore, although all teachers had been trained, half of teachers said they did not have the teachers’ guide to support MTB-MLE.

Just over half (54%) of teachers reported having one or more children with a diagnosed learning disability in their classroom, and 83% had over-age children. The main reason for being over age was repeating a grade, though late start or a break in schooling also came up as a reason about 40% of the time. Very few teachers received regular (monthly or more) instructional support from a pedagogical advisor.¹⁷ In fact, half of Grade 1 teachers said they received support only once in the year. On the other hand, principals were active in checking

¹⁷ The exact question reads “Since the start of the year, how often have you received instructional support from any type of instructional supervisor, for example, district/division/regional supervisors, grade coordinator, MTB-MLE coordinators, etc.?”)

lesson plans once per week or more (64%) or observing mother-tongue lessons once per week or more.

Table 15: Select indicators for Cebuano

Item	G1	G2
Proportion of over-age students (+7 in G1; +8 in G2)	12%	17%
Student repeated a grade	11%	3%
Student was in the same school last year	98%	99%
Teachers native language is Cebuano	95%	90%
Student's home language=Cebuano	96%	97%
Student's home language=Filipino/Tagalog	2%	0%
Student's home language=other local	2%	2%
Mean ORF	21	40
Mean comprehension (% correct out of total)	26%	52.5%

This region stands out as having some of the highest performance averages. In pre-reading skills, children learning to read in Cebuano outperformed all other regions with averages close to 50% for both listening comprehension and phonemic awareness in Grade 1. However, there was very little change for Grade 2, where children scored on average 60% on phonemic awareness and 51% on listening comprehension.

Also for basic reading tasks, this population had some of the highest scores among all of the regions assessed, surpassed only by Ilokano in familiar and non-words. Children in Grade 1 read on average 20 correct letter sounds per minute, 15 correct non-words per minute, and 19 correct familiar words per minute. In Grade 2 the non-word and familiar word reading averages increased to 25 and 35, respectively; however, the average letter-sound score decreased slightly, to 18. This is not to be interpreted as an actual loss of skills from one grade to the next, but rather the effect of a large confidence interval for these averages. On the other hand, it does suggest that there is not a *gain* in letter reading skill—notably automatic recognition of letter sounds. Moreover, in Grade 1 and Grade 2, approximately 11% of children could not identify a single letter correctly. This was the lowest proportion of zero scores among all of the regions, but it is the only region where there was practically no decrease from Grade 1 to Grade 2.

On the other hand, zero scores showed a noticeable decrease from Grade 1 to Grade 2 for non-word and familiar word reading. And more children in Grade 2 read above 30 correct words per minute. On these letter- and word-reading tasks, girls outperformed boys in both grades by an average of 7.5 items per minute in Grade 1 and 6.4 in Grade 2.

Children reading short stories in Cebuano showed the lowest proportions of zero scores in Grade 1 and Grade 2, and a large decrease in zero scores—to less than 10%—in G2. There were also more children scoring above 60 correct words per minute (17%) and higher accuracy (90%) in Grade 2 than elsewhere. This capacity was also reflected in comprehension scores. Almost half of Grade 2 students fully comprehended the text that they read and were able to answer correctly all the questions they attempted. Nonetheless, nearly 20% of students still could not answer any questions correctly in Grade 2, from nearly two-thirds in Grade 1.

Hiligaynon

For most Grade 2 teachers (90%), this study year was their first year of teaching using Hiligaynon, though this was only the case for 22% of Grade 1 teachers. Teachers were using the mother tongue more consistently in this population than in others, according to student reports of language use “this year” and “last year”; 95% of Grade 1 students reported that their teacher used mainly Hiligaynon as the language of instruction last year, compared to 85 to 91% in other regions. In fact, this was the only region in which 100% of classroom observation snapshots of reading lessons showed the mother tongue in use (and 99% in other subject lessons). Similar to other regions, however, there was a slightly lower percentage of teachers whose native language was Hiligaynon in Grade 2 than in Grade 1.

Many children did not have the mother-tongue learner materials, and surprisingly, 35% of students reported that the teacher never used the learners’ guide in Grade 1 for the mother tongue or math lesson, though this dropped to 14% in Grade 2 for mother tongue. The only other region with such a high figure in this category was ARMM (Maguindanaoan language) where 45% of Grade 1 students reported never using the book. More than half of students in Grade 2 reported that their teacher never used the mother-tongue math learners’ guide. On the other hand, more Grade 1 and Grade 2 teachers in this region have the teachers’ guide.

Hiligaynon was also similar to Maguindanaoan in the student self-reported absenteeism rate. Only one-quarter of students said they didn’t miss any days of school the week before; this translates into 75% of students who were absent one or more days. Although there are limitations to the accuracy of this information and whether it represents absenteeism throughout the year, there was a strong positive correlation across the entire sample between this particular indicator (student reported no absences in the last week) and reading scores.

Very few Grade 1 or Grade 2 teachers believed that children should be able to read the mother tongue fluently by the end of Grade 1. It is particularly interesting that almost no Grade 2 teachers (2%) believed that children can read in Grade 1, though at least a few Grade 1 teachers (18%) believed this. On the other hand nearly 71% of teachers thought that it is feasible for children to learn how to read Filipino in Grade 2.

Table 16: Select indicators for Hiligaynon

Item	G1	G2
Proportion of over-age students (+7 in G1; +8 in G2)	12%	22%
Student repeated a grade	8%	5%
Student was in the same school last year	99%	96%
Teacher’s native language is Hiligaynon	94%	91%
Student’s home language=Hiligaynon	98%	97%
Student’s home language=Filipino/Tagalog	1%	1%
Student’s home language=other local	1%	2%
Mean ORF	13	32
Mean comprehension (% correct out of attempted)	26%	53%

Performance in this region is characterized by moderate performance and gains between Grade 1 and Grade 2, and generally a smaller standard deviation than the other regions (i.e., fewer children at the extremes of demonstrated performance). However, that there were high

proportions of zero scores on most subtests. Scores on listening comprehension were the lowest of all four languages, with Grade 2 students at approximately the same level as Grade 1 students in the other languages. This may reflect the linguistic diversity of the region, where Hiligaynon and Kinaray-a languages are both found within the same communities, and it is not always clear which schools are using Hiligaynon as the medium of instruction (data collection ran into several instances of schools that were mis-classified).

The average scores on pre-reading skills for students learning to read in Hiligaynon were between 23% (listening comprehension) and 31% (phonemic awareness), but with a much larger average gain between Grade 1 and Grade 2 than any other language on both tasks. There were not large differences between boys and girls, though girls showed a tendency to perform better in phonemic awareness while boys did better on listening comprehension. This difference was much more pronounced by Grade 2, with nearly 13 percentage points more boys scoring zero on phonemic awareness than girls.

On the initial reading tasks children read, on average, 12 correct letter sounds per minute, 10 correct non-words per minute, and 12 familiar words per minute. These increased to 13, 21, and 28 in Grade 2, with notable decreases in the proportion of children with zero scores. While there were almost no children who could read more than 60 correct non-words per minute in Grade 1, there were 12% in this category by the end of Grade 2. Girls were outperforming boys by an average of 7 items per minute in Grade 2 and 5 in Grade 1.

The average ORF was 13 correct words per minute in Grade 1 and 32 in Grade 2, which is very similar to performance in Ilokano and just behind Cebuano. Comprehension was also very similar to that in Ilokano, with about 60% of students in Grade 1 who could not answer any questions and 30% in Grade 2. This was most likely linked to the proportion of children who couldn't read any words of the story. Of the children who could read all or part of the story, about one-quarter answered all questions correctly in Grade 2 and 15% in Grade 1. The Grade 1 comprehension in Hiligaynon was actually higher than in any other languages.

Ilokano

In schools in Region 1 that reported using Ilokano as the mother tongue, teachers were mostly Ilokano native speakers (95%), but a much lower proportion of students reported speaking the language at home (86%). About half (47%) of teachers felt very comfortable teaching using Ilokano as the medium of instruction, though only 21% felt very comfortable with the mechanics of the Ilokano language (grammar, orthography, etc.) The rollout of the mother-tongue program was reflected by 13% of teachers in Grade 1 who said this was their first year teaching using the mother tongue (32% in Grade 2). Additionally, some of the indicators related to mother-tongue usage and materials showed a lower rate of implementation in Grade 2; for example, observers noted that teachers used mother tongue most of the time in 95% of observations in Grade 1, but only 89% in Grade 2. While 14% of Grade 1 students reported that the teacher never used the mother-tongue math materials, 21% of children in Grade 2 said the same. Of teachers, 77% reported that all children had the learner materials in Grade 2, but 84% had materials in Grade 1. On the other hand, more Grade 2 teachers said they had the teachers' guide than Grade 1. Fewer than one-third of teachers in either grade thought that the mother-tongue teaching materials were sufficient, and this was lower for Grade 2 (21%) than Grade 1 (32%).

This population had the lowest prevalence of over-age children (as self-reported). According to teachers, about 77% of classrooms sampled had over-age children, but this was lower than the average across regions (83%). The main reason for being over-age for grade, according to these teachers, was repeating a grade, though there were also reports of children who had stopped school for a period because of financial problems. In fact, this region also showed the lowest prevalence of repeating a grade, with only 2% of children reporting that they were in the same grade this year as last year. In the Ilokano sample 26% of teachers reported having one or more children in the class with a diagnosed learning disorder. Nearly 80% of teachers reported that they had only received instructional support either once in the year or once every 2–3 months. Almost none were receiving instructional support monthly in either Grade 1 or Grade 2. Instead, they were relying heavily on peer support and school-based support from the principal.

Table 17: Select indicators for Ilokano

Item	G1 (n=399)	G2 (n=400)
Proportion of over-age students (+7 in G1; +8 in G2)	5%	11%
Student repeated a grade	2%	0%
Student was in the same school last year	99%	98%
Teacher's native language is Ilokano	99%	98%
Student's home language=Ilokano	87%	85%
Student's home language=Filipino/Tagalog	11%	11%
Student's home language=other local (Pangasinan or Kankanaey)	2%	3%
Mean ORF	15	30
Mean comprehension (% correct out of attempted)	25%	54%

Children learning to read in Ilokano demonstrated relatively high levels of reading skill compared to their peers learning to read in other languages, with the exception of the task of listening comprehension, in which the region had the highest proportion of children unable to answer one listening comprehension question correctly in Grades 1 (50%) and 2 (30%). Ilokano also showed some of the highest averages and individual scores across the regions on the three initial reading subtests (letter sounds, non-words, and familiar words), but also some of the largest ranges of ability in the distribution of scores. For example, for the non-word reading subtest, where Grade 2 students had the highest average correct words per minute score (34), nearly 40% of children in Grade 1 and 20% in Grade 2 could not read a single word, while at the same time 25% of children (Grade 2) were reading more than 50 correct words per minute. In no other language could children read more than 70 correct non-words per minute, yet the average number of familiar words read in Ilokano was over 70 when children with zero scores were left out of the analysis. Fluency and accuracy both decreased, however, when children read short story passages—the average ORF was 30 and accuracy was 4 to 5 percentage points fewer than for non-word reading. This may be explained by the agglutinative nature of the language, which causes words to become much longer in context because of markers that signal tense, possession, or other context cues.

We can see that there is a great deal of potential for children to master early reading skills in Ilokano, but many children are still being left behind—in many cases, the boys. There was an average difference in score of 11.2 items per minute on these three subtests, with the highest difference in familiar word reading, in which girls read 16 more correct words per minute than boys in Grade 1. The trend was similar in Grade 2, with only a slightly smaller gap (10.4 items per minute average).

The appearance of a portion of children whose home language was Pangasinan and Kankanaey is illustrative of one of the challenges of MTB-MLE implementation in general, and sampling for this study, in particular. As mentioned above, the sample was drawn from a list provided by DepED of schools that were using Ilokano as the language of instruction. In some cases this list came from a regional language mapping exercise, or alternatively from records of where mother-tongue instructional materials had been shipped. It is not surprising that there were some children who speak Pangasinan at home—this was also reported in 2013—since the sample was drawn from schools in Regions 1 and 2 and it is known that this is an area where both languages are used. In a given school, half the children may come from an Ilokano-speaking home and the other half from a Pangasinan-speaking home. However, this year one school was visited in which all the children sampled spoke Kankanaey at home, yet the school had declared the language of instruction to be Ilokano, and the students reported that the teachers were using Ilokano and they had the Ilokano learner materials.

Compared to Grade 1 children assessed in 2013 using the same EGRA test (with only the reading passage adjusted, but verified to be of equivalent difficulty through pilot testing), overall average scores were not significantly different.

Maguindanaoan

The data collected suggest that most children spoke Maguindanaoan at home, but many teachers were not native speakers, and they were using other languages in the classroom with a higher frequency than teachers in other regions sampled in this study. This was particularly the case in other subject areas (not the mother-tongue reading lesson), according to classroom observations. In at least half of the classrooms the children did not have mother-tongue learner materials, and the teacher was not using mother-tongue or math learner materials during lessons. Similarly, fewer than half of Grade 1 teachers had the teachers' guide, and only 23% of Grade 2 teachers had it. Classroom observations indicated more time spent off-task than in other regions, and a prevalence of copying as a method.

For many Grade 1 and Grade 2 teachers (51% and 75%), this was the first year of teaching in the mother tongue. More than half of teachers (57%) felt very comfortable teaching in the mother tongue, but fewer than half, on average (42%), felt very comfortable with the mechanics of the language (orthography, grammar rules, etc.); 40% of teachers reported that they had one or more students with a learning disability.

According to teachers, instructional support was not consistent—some teachers received support monthly (25%), others only every 2–3 months (30%), or even once per year (23%). A majority of teachers believed that children should be able to read English fluently by Grade 3, but Filipino only by Grade 4 or higher. Opinions were very split among teachers who thought that children should be able to read the mother tongue in Grade 2 (32%), Grade 3 (42%), or Grade 4 or higher (25%). However, there was a consensus that runs contrary to the MTB-

MLE curriculum, as almost no teachers felt that a child should be able to read in Maguindanaoan by the end of Grade 1.

Table 18 summarizes some key indicators related to the sample and outcomes, followed by a description of the characteristics of student performance across subtests.

Table 18: Summary of key indicators for Maguindanaoan

Item	G1	G2
Proportion of over-age students	36%	35%
Student repeated a grade	21%	2%
Student was in the same school last year	99%	99%
Teacher's native language is Maguindanaoan	83%	83%
Student's home language=Maguindanaoan	99%	97%
Student's home language=Filipino/Tagalog	0%	1%
Student's home language=other local language	0%	.2%
Mean ORF	6.9	21
Mean comprehension (% correct out of possible)	8.1%	27%

Children learning to read Maguindanaoan scored equally as well as children in other language groups in listening comprehension—on average, 44% of questions were answered correctly in Grade 1 and 53% in Grade 2. However, on other subtests performance in this region was characterized by large proportions of zero scores and the lowest score averages. In several cases the proportion of children with zero scores was higher than for any other category of performance for that language even in Grade 1, and the Grade 2 averages were equivalent to or less than the Grade 1 averages for other languages. By Grade 2, the average ORF score was 21 words per minute, although some children were achieving more than 40 correct words per minute (20.5% in Grade 2). Comprehension scores were on par with scores of other languages in Grade 1, but this is one subtest for which the increase from Grade 1 to Grade 2 was less pronounced than it was in other languages.

Classrooms teaching in Maguindanaoan showed the largest gains from Grade 1 to Grade 2 on most subtests, at times reducing by half the proportion of students with zero scores or tripling the average score. There were not large differences between boys and girls, with the exception of phonemic awareness, for which 13 percentage points more boys had zero scores than girls in Grade 2.

Interestingly, the correlations between subtests were the highest for this language; for example, the letter-sound identification score and scores on familiar and non-word reading and oral reading fluency were all above .6. This was the only language that showed a strong correlation between letter sounds and reading skills. This may be explained by characteristics of the language, by better consistency in scoring by the assessors in this region, or because of the overall low proportion of children who can read at all—if they show some ability, it is usually across subtests or not at all.

Clearly an important finding from the study in the ARMM region is that there was a disproportionate amount of children with zero scores across subtests and persisting in Grade 2. Furthermore, while recognizing that each language must set its own benchmarks, it would appear that in relation to other languages in this study, children learning to read in

Maguindanaoan were taking two years to acquire the level that could be expected of one year of instruction.

The data from this study do not provide all the information needed to know what, specifically, is the most influential factor in the context. It may be low teacher expectations, the lack of instructional materials or training; it could also be a lack of instructional time including frequent absences and children out of school or repeating grades due to conflict (and thus a high proportion of over-age children in the early grades). This is not the first study to have identified low levels of achievement in the region and the needs in this context. But this study did show that the lower performance of students in this part of the country seen at all levels of education begins in the earliest grades of elementary school.

B. Limitations

At the conclusion of the 2013 survey, several limitations of that work were discussed and strategies for improving the reliability of the data were put forth. Chief among the methodological concerns that surfaced in 2013 was the use of non-educators in most instances as assessors of students. This year teams still combined personnel from TNS who were experienced at conducting surveys and collecting data with personnel from DepED regional offices who were accustomed to working with students. However, the implementation of the student assessments was the responsibility of the DepED personnel, while TNS staff conducted the classroom observations and teacher interviews and served as field coordinators.

While this accommodation responded to the concern that the children being assessed be put at ease by someone who is experienced at interacting with young students, concern about the reliability of an individually administered oral test with Grade 1 and 2 aged children is still valid. A degree of uncertainty about the reliability of the results is introduced whenever we test children with a methodology and technology to which they are not accustomed (although this limitation was probably reduced in the 2014 assessment when compared with the 2013 survey by the more systematic inclusion of DepED staff as assessors).

Another limitation of this kind of survey arises from variation among assessors in how they administer the tests, observations, and questionnaires. A significant source of potential variability in test results is assessor consistency in conducting the test, and more importantly, assessor consistency in interpretation of students' oral responses to test items. To diminish this limitation, the assessor training for this study included ample practice and also emphasized consistency in coding responses. Practice was done in small groups, with pairs of assessors marking responses to one student. This allowed assessors to compare their responses, observe behaviors, and provide feedback to each other. Emphasis was placed on inter-rater reliability (IRR), or ensuring that everyone would judge a given response the same way. Peers would compare responses and discuss items for which they disagreed on the correct answer.

Several formal and informal IRR exercises were conducted with the groups to determine how consistently assessors scored responses. In all cases, assessors listened to the same stimuli and marked responses. During informal IRR, trainees discussed which items were marked incorrect and why (or why not). During formal IRR, the assessors used tablets to record a series of subtests without discussing the results immediately. The results were tabulated and then discussed, referring back to the items for which assessors were not in agreement. These

activities helped to get a global picture of IRR and to identify test items that had a high rate of inconsistency and assessors who were often in disagreement with the others. These assessors were provided individual feedback and additional coaching.

IRR was consistently measured at over 90% for the assessors in each of the four regions. This reflects the percentage of assessors who were both in agreement with the “gold standard” response as well as in agreement with each other—thereby ensuring more than 90% reliability that the coded responses match the correct coding and that consistency is ensured across assessors. Even with over 90% IRR, there is a possibility that some student responses were either incorrectly or inconsistently coded, thus introducing a small percentage of error into the findings—something that is impossible to completely eliminate in this kind of survey work.

Other limitations arose when field data collection teams encountered problems when visiting schools. The most frequent issue encountered was the need to replace some schools in the sample. This was either because schools that had originally been included on the list for sample selection were not actually teaching in the mother-tongue language that was the target of the assessment in a given region or because schools were located in inaccessible areas. Under these circumstances, two types of replacements were made. When possible, randomly pre-selected replacement schools were used. If those replacement schools were also not useable for the same reasons, then we reselected replacement schools. The total number of schools that had to be replaced included 6 in Ilokano, 9 in Hiligaynon, 1 in Cebuano, and 9 in Maguindanaoan. We are confident in our methods for replacing schools and ensuring the randomness of those that were included. The integrity of the sample was protected, but some degree of uncertainty is introduced whenever last minute changes are needed.

Another aspect of the field data collection that limited the data we were able to collect was the method of classroom observations. In one instance when a tablet device malfunctioned in the Ilokano region, six classroom observations had to be completed using pen and paper rather than the tablets. Those data were never entered and transmitted to RTI for inclusion in the analysis, reducing the classroom observation sample by about 4%.

The use of tablets on the whole offered a much more reliable technique for collecting, transmitting, and compiling field data. RTI statisticians were able to regularly monitor data as they came in and on a few occasions were able to catch errors and communicate with the field supervisors so that teams could return to a school and recapture data when necessary. However, the tablets did not function 100% as desired. A few tablets caused some problems for field teams because the software was not as responsive as it should have been, interfering with data capture (a major concern, especially during timed assessments) or because the tablet did not hold a charge as long as it should have. Replacement tablets were available throughout the country, so the teams were able to make sure that the EGRA assessors used the best-performing tablets (giving first priority to the integrity of the student-level data).

In only two or three instances were DepED team members unable to fulfill their roles as assessors because of scheduling conflicts. On those days, the teams used substitute assessors who had been trained and were on call for just such an occasion. Using a different assessor for one or two days during a field team’s work probably introduced some uncertainty into the data collection for those two or three cases where this occurred.

Another limitation of the survey this year arose from the fact that previous versions of the instruments have been circulated and in some regions pilot EGRA studies have already been

conducted. Although the items used for this version of the assessment were chosen based on a larger bank of items and unique stories established during an adaptation workshop prior to fieldwork, some students may have been more familiar with the assessment process than others. There is always a possibility that the actual EGRA was disseminated prior to fieldwork.

The questionnaires attempted to quantify exactly how many teachers were familiar with the assessment by asking them if they had ever heard of EGRA and if they had implemented EGRA in the classroom before. It emerged during this process that the term “EGRA” means different things to different people. For example, when asked “What is your experience with EGRA?”, answers included things like “it is a strategy for reading more easily” or “it is the program for using the mother tongue.” Similarly, we found that in one question teachers may have said they never heard of EGRA, but then in another question they answered that they have used EGRA in the classroom. This may have also been due to a miscommunication in the way the question was phrased. The result is that we were not able to get an accurate picture of whether prior experience with EGRA may have affected the student scores.

Another similar miscommunication occurred with the question related to training programs. First of all, during the instrument adaptation process RTI was asked to remove a question about in-service training related to teaching in the mother tongue and instructed to only ask about pre-service. The rationale was that DepED already knew that all teachers have been trained through the national program. However, many teachers ended up responding to the question about pre-service with accounts of training from the DepED K–12 initiative, Basa Filipinas, or other sources. Again this may have been due to a miscommunication in the way the question was asked. The result is that we missed the opportunity to verify whether all teachers had in fact participated in the national training programs, and how many of them had received additional training on top of this.

The sampling procedure required selecting only one classroom per grade per school in order to match teachers and student outcomes. A Kish grid was used for the random selection of classrooms.¹⁸ In some large schools with many different classrooms for the same grade, students may have been grouped by ability. We excluded special needs student (children with physical or learning disabilities that were more serious than those of “slow learners”) classrooms from the classroom selection, but left other homogenous types of classrooms in the sample on the assumption that classroom sizes (and therefore probability of student selection) were still equivalent to heterogeneous classrooms in the same school. Furthermore, classrooms were not always clearly labeled as advanced or slower; sometimes classrooms were numbered, and sometimes they were given a name. Where classrooms were numbered, sometimes the higher numbers held the slower children and sometimes the lower numbers did. Therefore the decision was made to sample randomly and expect that the proportions would end up representative. Teachers were asked to describe what type of classroom theirs was. **Table 19** shows the final teacher-reported breakdown of sampled classrooms by type.

¹⁸ See methodological report, TNS (2014).

Table 19: Breakdown of classroom sample by type of classroom

Column1	Cebuano	Ilokano	Hiligaynon	Maguindanaoan
Heterogeneous (kids of different ability levels)	93%	94%	98%	93%
Homogenous, science aptitude	1%	3%	1%	4%
Homogenous, for fast/gifted learners	7%	3%	1%	0%

C. Conclusions

The most prevalent sign of MTB-MLE being implemented in the four regions included in this study was the consistent use of the regional mother tongues as the languages of instruction in both reading and other lesson periods. All teachers observed were using the mother tongue for the vast majority of their instructional activities and interactions with students. Students were also observed to be engaged in the activities being conducted in mother tongue. Also, Filipino and English are being taught as expected, and children are acquiring beginning skills according to the progression expected by the curriculum. However teachers were not exhibiting the full set of skills needed to appropriately support literacy acquisition in the mother tongue during their reading lessons. Many teachers indicated that they had not been sufficiently trained for mother-tongue instruction, and more than 70% stated that they did not have sufficient materials and resources for teaching reading. Majorities of teachers reported not having important materials like the mother-tongue teachers' guides in three of the four regions. A particular concern is whether the activity of reading practice is taking up a sufficient amount of time during reading lessons, or if teachers are disproportionately speaking and asking children to read or recite in chorus. There were also very different expectations of when children should be acquiring reading skills in different languages (or possibly just different expectations of what "fluently" means).

For the vast majority of students in the regions included in this study, the language of instruction matched their home language. However there were students who did not speak the mother tongue of instruction at home, especially in Region 1 where as many as 12% of the students were in a school that was using a language of instruction that was not their mother tongue. Children's self reports of language use in the home may be somewhat unreliable, and assessors noted in some circumstances children were unfamiliar with the terms 'language' and so additional prompting needed to be done to make sure the children were answering accurately. This is an important finding in itself—how are children being prepared for instruction in multiple languages, and how are teachers bridging between the first and second (and third) languages if children are not exposed to the meta-language associated with code switching in the classroom, the community and the home? What are children learning about being multilingual?

Another somewhat accidental finding is that teachers have been hearing about EGRA through a variety of different sources including K–12 workshops, Basa Pilipinas, and different regional initiatives. In their responses it became clear that there were different understandings about what EGRA is and what it is supposed to do. In some cases the term "EGRA" is confused with the notion of teaching reading in the mother tongue rather than as a specific type of reading assessment tool.

As this was still early in the process of rolling out a complex reform, it is understandable that the resources, supports, and practices for MTB-MLE were not yet all in place. Notably, some important aspects of MTB-MLE implementation show strong relationships with student performance in reading. For example, the 20% of students who often used their mother-tongue book for reading were 2.8 times as likely as those who never used it to be in the top 25% of reading performance. Similarly, the 24% of students whose teachers often asked them to read in the mother tongue were more than three times as likely to be in the top 25%. This indicates that if MTB-MLE implementation conditions could be better assured across a much broader range of schools, then it is very likely that reading performance will improve.

The uneven implementation that these data show helps explain the types of outcomes presented in this report. According to the MTB-MLE expectations, children should be learning to read in their mother tongues by the end of Grade 1. A large proportion of children learning in Maguindanaoan were scoring zero on all the EGRA subtests, except listening comprehension in Maguindanaoan. This indicates that many children were not acquiring even the building blocks of literacy in that language during Grade 1. Ilokano and Hiligaynon also had high proportions of children in Grade 1 with zero scores on several subtests.

For all the languages, reading performance in Grade 2 was much better than in Grade 1, indicating that as children and teachers spent more time working in these mother tongues, the teaching and the learning of reading skills was improving. However, since reading accuracy was already very good in Grade 1, more time should be spent working on reading fluency in Grade 2, and we could expect reading fluency to be much higher. This would be the case if more time were spent practicing reading—silently or aloud. However, the classroom observation data suggest that Grade 2 classrooms still spent a disproportionate amount of time on teacher-centered listening and speaking activities. By late in Grade 2, students on average were reading more than 30 words per minute of connected text in three of the four languages, with Maguindanaoan being the exception. In Cebuano, Hiligaynon, and Ilokano, 51%, 44%, and 33% of Grade 2 students were reading 40 or more words per minute. As MTB-MLE implementation improves and instruction is more consistently supported, it is likely that these levels of performance should begin to be seen at the end of Grade 1 instead of at the end of Grade 2.

D. Recommendations

The conclusions, as summarized above already point to several clear recommendations, for example:

- Ensuring the presence of sufficient materials (including basic learner materials and teachers guides) in classrooms;
- Ensuring time on task during the reading lesson, especially time spent practicing reading;
- Raising teacher expectations about what children can do grade by grade—and in this case, *must do*, because the curriculum and students' progression to fully multilingual learners by Grade 4 depends on it;
- Continue efforts to make sure that as much as possible children and teachers are matched to the school that uses their mother tongue as medium of instruction;

Additionally, the results were discussed in Manila with Department of Education stakeholders, including regional representatives who participated in data collection. In small groups, stakeholders of each language discussed findings and implications for their region and made specific recommendations. Although the findings were specific to each region, they also raised many common themes, which have also been found to be critical components of reading improvement programs worldwide. These are now commonly referred to as “the 5 T’s” (see, for example, <http://blog.usaid.gov/2012/11/education-week-2012-reading-improves-with-the-5-ts/>): time, texts, teaching, tongue and testing.

Time. Reading is a skill that is acquired through practice. Children need to spend an adequate amount of time on learning to read and practicing the skill with a variety of materials that are matched to their level of ability. Time in school is just one factor, which is why the home environment also makes a great deal of difference. In terms of time spent at school, one needs to consider not only how time is used, or lost, during the reading lesson (as this study was able to point out to some extent), but also throughout the school day, the week and the year. Other factors that could be associated with reading ability in the Philippines, but not measured by this study, include time lost due to conflict and natural disasters when schools were closed and time lost when teachers are absent for personal or professional reasons (such as training). This study, and recommendations made by groups during the workshop, point to a need to increase instructional time on reading. This can be achieved by making the reading lesson more productive (especially by reducing time spent off-task, or on indirect skills such as listening and speaking), or by adding other opportunities to read during the day. This can also be achieved by increasing the time allocated to reading instruction in the curriculum, especially in the early grades. If more emphasis is placed on learning to read in the first half of grade 1 then it is likely that this year’s Grade 2 scores could be achieved in Grade 1. While at first it may be at the ‘expense’ of other subject areas, in the long term, it will make instruction and acquisition of other subject matter easier and more efficient. This recommendation seems especially pertinent for the ARMM region, where children are barely achieving Grade 1 level reading scores at the end of Grade 2.

Some specific recommendations made by the regions that focus on time on task include: increasing time on storytelling and story reading, providing more follow up home activities and feedback.

Texts. It follows that if time on reading is going to increase, then there needs to be a sufficient amount of materials to read, at a variety of levels of difficulty. This study pointed out a lack of available reading materials for children and for teachers, although significant efforts have been made by DepEd to develop such materials to support the curriculum. An issue seems to be cost effective and timely printing and distribution of these materials in the regions. In some cases, schools and even teachers are expected to print the materials locally, which appears result in insufficient quantity and quality of reproduction. On the other hand, teachers do have visual aids, many of which may have been produced locally, and some regions like Cebuano have also been successful in producing and distributing big books. Such efforts should continue (with obvious attention to quality control); local NGOs, communities, and the private sector should be encouraged to contribute to materials production, purchase and dissemination. An important issue that needs to be addressed with regards to the production of materials is that in many regions there is still debate and inconsistency in the accepted and used orthographies of the languages, despite the issuance of official orthographies by DepEd (via KWF).

Some specific recommendations made by the regions include: providing quality-assured reading materials, increasing the use of word walls as a strategy; improving the teacher's guides and learner materials; developing localized supplementary materials.

Teaching. The teacher is of course a critical component for children in learning how to read. While there needs to be sufficient time spent on reading, and with appropriate reading materials, the teaching methods that are used also need to be appropriate. In this context, that means teaching strategies that take into consideration the characteristics of students and the context (for example, different methods for boys and girls; methods that are proven effective for bilingual learners; methods that consider the home environment, etc.) Specific teaching strategies may need to be developed to address the specificities of each language as well, in particularly the agglutinative nature of Ilokano. Teaching needs to explicitly focus on the components of reading, such as phonemic awareness, vocabulary building and fluency. All languages should develop and use frequent word lists that are linked to the most common words in the language. This study suggests, based on the analysis of classroom observation data, that classrooms are still using a primarily teacher-centered, whole-class approach to teaching reading that may not allow for the differentiation and individualized attention that students need. Furthermore, the reading classes are disproportionately focusing on speaking/listening activities rather than individual reading activities.

All regions recommended that teachers needed to be trained on techniques for teaching reading, and they should be capacitated to use varied teaching methods. While it is clear that teacher training is important, the type of training and methods used need to be carefully considered. In ARMM, for example, participants lamented the fact that teachers have been trained over and over by international projects, yet results never change. The model of one-off, workshop-based teacher training, and even cascade-type trainings is not proving effective. More school-based, continuous professional development needs to be considered, so that teachers rare receiving strategies and feedback on implementing those strategies when and where they can apply them immediately. The teacher questionnaires from this study revealed that teachers rely heavily already on school-based support from the director or from peers, so the question is whether or not the school director or a school-based instructional leader has been appropriately trained to provide specific support to the teaching of reading. Investments in this area may prove more effective than removing teachers from the classroom for more training.

Tongue. The Philippines is already demonstrating commitment to the proven fact that children learn to read best when they are reading in the language they know best. As mentioned previously, given the linguistic diversity of the Philippines and shifting language boundaries (See McEachern, 2013), there are still many children who are not in schools where their mother tongue is used as the medium of instruction and therefore have to learn three new languages in three years. For these children, if it is impossible for them to be in school in that uses their mother tongue, then specific bridging strategies are needed or other support mechanisms.

Testing. Understanding what children know is important in order to provide them with remediation before it is too late. Testing refers here to the need to be continuously assessing children's mastery task by task throughout the year, rather than waiting until end-of-grade assessments to find out whether children have acquired the content of the curriculum. The Philippines is also clearly showing a commitment to diagnostic testing to help identify

weaknesses in the system and to monitor effectiveness of programs. DepEd has also promoted the use of classroom-based diagnostics such as Phil-IRI and more recently, EGRA. However, both of these types of assessments have shown their limitations for improving reading ability—perhaps because Phil-IRI is a classroom diagnostic that is being used at the system level, and EGRA is a system-level diagnostic that is being promoted as a classroom-based tool. Elements of EGRA can be adapted for classroom continuous assessment, and it can be vastly simplified. For example, items can be standardized to a limited extent based on the scope of what is in the textbooks, but teachers can also select these items themselves based on what is being taught in the classroom day-to-day rather than waiting for official standardized tools to be released; tests can be untimed (for example, testing mastery of letter sounds can be done as sounds or introduced, or periodically using only a randomized list of letters. Teachers can estimate reading fluency simply by listening to the cadence of a child reading any passage out loud). Moreover, classroom assessments administered by teachers should not be aggregated and used as regional or national diagnostics. If accountability is needed, school heads or regional supervisors can periodically conduct monitoring visits and can use EGRA with a random sample of children, but beyond that the resources used to conduct frequent, large scale classroom-based data collection (and its analysis) might be better spent at this point in improving instruction. Reducing the complexity of classroom-based and teacher-administered reading assessments will encourage this to happen more frequently and should help teachers adapt instruction to meet identified needs and to measure their own children's progress throughout the year. This will have an impact on other sample-based, standardized assessments that are done periodically to measure the impact of reforms.

This study has also been able to provide the Department of Education with data that can be used to establish benchmarks. Such benchmarks can be used to set standards of achievement by grade, and also to measure progress based on the number of children meeting those standards. Table 20, below, shows the range of oral reading fluency associated with successful comprehension, and the proportion of children who are currently reading within that range. This can be a starting point for determining reading fluency benchmarks.

Table 20: Reading fluency and comprehension as benchmarks

	Grade 1	% in range	Grade 2	% in range	
Cebuano	80% comp	42 – 56	11%	46 – 64	27%
	100% comp			55 – 76	19%
Hiligaynon	80% comp	38 – 51	10%	44 – 62	25%
	100% comp			53 – 69	15%
Ilokano	80% comp	34 – 44	10%	35 – 45	25%
	100% comp			45 – 58	20%
Maguindanaoan	80% comp	32 – 57	4%	35 – 56	20%
	100% comp			45 – 62	13%

Annex A: Summary of results from 2013 EGRA study

One of the key takeaways from this study is the existence of important inequities in achievement between boys and girls, and across regions. **Figure A-1** and **A-2** below present again the same data that were presented previously, but in a format that shows the differences among regional populations. The challenges in these lowest-performing regions are well-known—instability, poverty, geography, lack of access to early childhood education, etc.—and many programs are in place to try to address these factors. However, understanding which factors are having the most influence and what specifically can be done most effectively and efficiently to address the achievement gaps will be critical going forward.

Figure A-1: Comparison of Filipino scores across regions

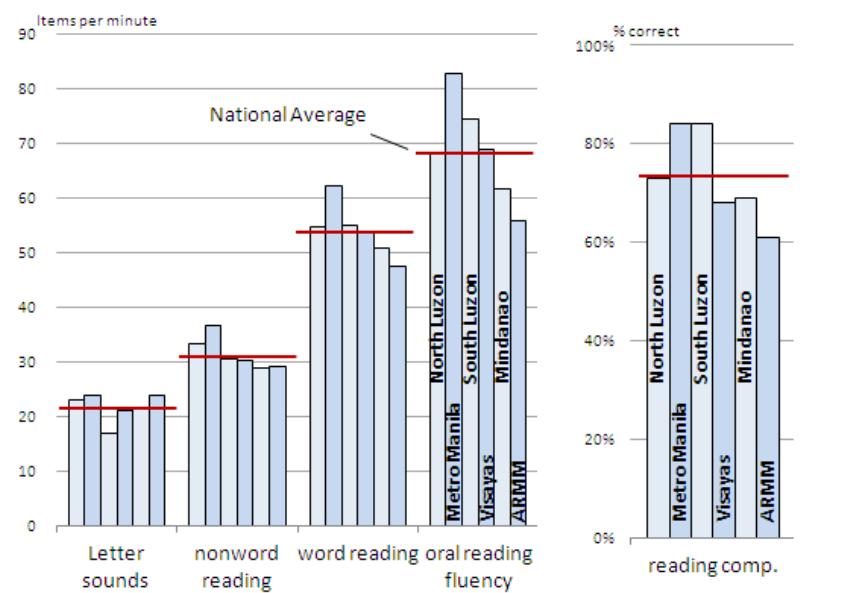
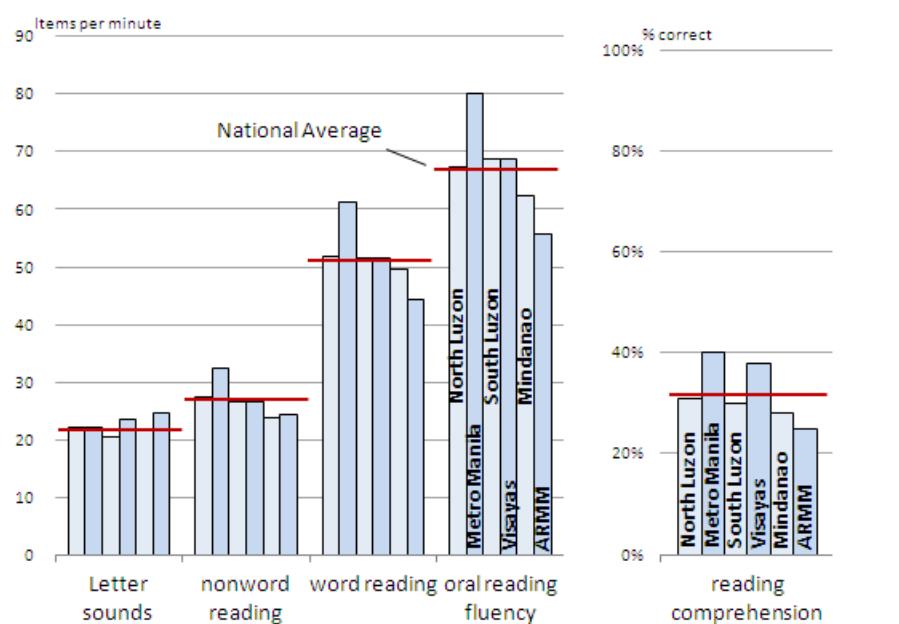


Figure A-2: Comparison of English scores across regions



Additionally, the gap in achievement between boys and girls is remarkable, and specific strategies are needed to better engage boys in reading. However, equally important is ensuring that all children start Kinder at the appropriate age and continue their schooling consistently, grade by grade, with a low rate of absenteeism. Combined with common-sense teaching and learning inputs such as textbook availability and activities that promote reading both in school and out of school, policies to promote improvements in these areas would address many of the factors that were revealed to be associated with reading scores in this study.

While some factors cannot be controlled by policy makers, it is important to weigh the magnitude of the effect against the percentage of the population that is affected by the factor when making policy decisions. For example, although living more than an hour away from school was associated with lowest levels of ORF in English and Filipino in grade 3, policies to build schools in rural areas would affect only 2% of the population, and may not, therefore, be as cost-effective as interventions that target changes in larger populations. On the other hand, one third of the children in grade 3 classrooms sampled were over-age, indicating a wide prevalence of children who start late or policies that encourage holding children back to repeat a grade. It may be worthwhile to look into ways in which classrooms can be made more homogenous (including avoiding multigrade classrooms) in order to better address the needs of over-age students.

The reasons why children were over-age were not uncovered by this study, but the children likely started late or were required to repeat a grade or grades. However, the correlation between age in grade and reading fluency supports findings in other international literature and other EGRA studies. Children who are required to repeat a grade by definition are experiencing difficulty in school, possibly because of other characteristics such as poverty, high absenteeism, etc. Therefore, they likely were poor readers before being asked to repeat a grade. What is important to note, however, is that repeating a grade does not usually help to raise the level of achievement of the child. See, for example, research from Senegal by Glick and Sahn,¹⁹ which recommended “alternative measures to improve the skills of lagging children” than repeating (p. 93); or the literature review by CfBT Education Trust on grade repetition in sub-Saharan Africa, which concluded that “The majority of the studies undertaken to explore the effects of repetition on learning suggest that the practice does more harm than good (p. 5).”²⁰

The issue of age is just one factor that creates classrooms that are very diverse, both linguistically and demographically. This presents a significant challenge for teachers trying to match their instruction to learners’ needs when those needs are necessarily so different. It has been found to be the case in other countries that teachers typically teach to those at the top of the class—teachers may always call on children they know have the ability to answer and as long as some children have understood and mastered the lesson, they will continue to move on, leaving others behind. Finding time and resources to intervene with remedial instruction for those who need it is a worldwide challenge, but it is one that the Philippines will need to address, given the large percentages of children who are being left behind, according to this

¹⁹ Glick, P., & Sahn, D. (2010). Early academic performance, grade repetition, and school attainment in Senegal: A panel data analysis. *The World Bank Economic Review*, 24(1): 93–120. Washington, DC.: World Bank.

²⁰ Ndaruhtse, S. (2008). *Grade repetition in primary schools in sub-Saharan Africa: An evidence base for change*. Reading, Berkshire, UK: CfBT Education Trust. Retrieved from http://www.cfbt.com/evidenceforeducation/pdf/Grade%20Repetition_FINAL_8FEB08.pdf

study. In fact, a 1999 study published in the World Bank Economic Review found that out of four experiments to improve education quality in the Philippines, the one that was found to be the most promising was to provide teachers with learning materials, “which helped them to pace lessons according to students’ differing abilities (p. 493).”²¹

A key question at the grade 3 level is whether, in either language, children are ready to read to learn in the more cognitively complex forms of the language used in subject areas like science and mathematics, and whether they are prepared to do so through independent study as well as classroom instruction. The results of this study raise important doubts about whether the reading skills of grade 3 students are strong enough to “read to learn” in grade 4 and beyond. What is clear is that multilingual education goals of the curriculum are very ambitious and will require a great deal of highly contextualized language input, in a language-rich environment, supported by trained teachers throughout the primary grades.

A one-day event was held in Manila to present the results of this study to DepEd staff, including regional representatives who participated in data collection, local academic experts in language, curriculum and instruction, including those who helped develop the instruments, and other stakeholders such as international organizations and donor institutions. Some of the recommendations made by this group after reviewing and discussing the results include:

- ensuring that pre-service training addresses specific strategies for reading instruction in this multilingual context and that the results are shared and discussed at the pre-service level
- ensuring that the study is used to inform instruction, by building capacity of teachers to understand and use the data
- that the benchmarks be revisited in light of the results, taking into consideration evidence-based expectations for the children and the linguistic characteristics of the language (i.e., average word length.)

The actual use of EGRA was discussed and debated and it should be determined whether it will continue to be used as a national, sample-based diagnostic and measurement of the effectiveness of MTB-MLE, or if it will be modified for use as a classroom-based formative assessment. It has been communicated that in its current form, EGRA is a sample-based diagnostic; for it to be used as a classroom-based assessment significant modifications would be required. In its present form, EGRA can continue to be used for system level monitoring of the impact of implementing MTB-MLE – either in terms of students learning to read in mother tongues or in terms of their transitioning their literacy skills to Filipino and English.

Further analysis of these data. The data collected by this study represent hundreds of individual variables for each language and thus many possible levels of analysis. Further analysis could be done in areas such as:

- More detailed, word-level analysis of dictation data.
- More cross-language analysis per child—i.e., did children who did well in Filipino also do well in English? Did children present the same error types across subtests?
- More detailed analysis the relationships between home and school environment, and particularly the language environment, and results. This is especially pertinent

²¹ Tan, J., Lane, L., & Lassibille, G. (1999). Student outcomes in Philippine elementary schools : An evaluation of four experiments. *The World Bank Economic Review*, 13(3): 493–508. Washington, DC.: World Bank.

for the Ilokano data in which more detailed questions about language spoken at home and in the community were asked of children.

Additionally, whereas this report presents a great deal of data and attempts to provide explanations and interpretations based on that data, the validity of this interpretation depends on participation of local experts—linguists, teachers, DepEd officials—who know more about how teachers are trained, what materials are in circulation, and what actually happens in classrooms. Therefore teachers and curriculum specialists should be encouraged to review the results with a critical eye on how teaching practices, linguistic characteristics, and classroom environments may explain some of the results. Some of this dialogue began at a high-level presentation of the EGRA results that took place in Manila on July 25, 2013. Issues raised by participants at that time have been integrated into this report, but secondary analysis and interpretation of the results is encouraged as well.

Potential follow-on studies. Interpreting these results, and monitoring the implementation of MTB-MLE, will also benefit from more in-depth studies that measure the actual language use in the classroom. See, for example, studies done in Africa using classroom observations of instruction and language use.²² These studies provide detailed information about how much time in the day was being spent—across curricular subjects—by teachers and students using different languages, as well as what types of activities teachers and children were engaged in (i.e., reading silently, reading out loud, etc.). These results can also be associated with reading outcomes measured by EGRA and help determine how variations in instructional practice can influence reading outcomes. Given the overall low levels of performance measured in the Ilokano portion of this survey, and the large variations seen in the Ilokano reading data, such further study would prove especially useful. However, in any evaluation of reading performance in mother tongues, one must first determine whether MTB education was actually being implemented in the schools in the way it was intended by the program. Similarly, such a study could also look into teacher competencies in the language of instruction. As discussed briefly above, national studies of teacher competency in English have shown low levels of proficiency; additionally, across the Philippines, different linguistic backgrounds affect how English is spoken and understood.

Theories of language acquisition, which inform different models of bilingual education, stress the importance of “comprehensible input”,²³ which means not only that the speaker needs to be speaking clear and correct language, but also that whatever is said needs to be appropriate to the level of language that children are capable of understanding, with gradual additions to increase vocabulary and comprehension. This “input +1” model is similar to the concept of leveled readers for children learning to read: Providing texts that are too difficult for a child will result only in frustration and failure, but providing decodable texts reinforces the concepts the child has already learned and creates confidence and ability to attack increasingly difficult texts. To summarize, teachers’ competency needs to be measured not only by their own language speaking and reading competency but also by the extent to which

²² Studies were conducted by RTI in Mali, Senegal, Kenya, and Uganda with funding from the William and Flora Hewlett Foundation. See the two Piper reports cited earlier, as well as Varly, P. (2010). *The monitoring of learning outcomes in Mali: Language of instruction and teachers' methods in Mali grade 2 curriculum classrooms.* and Varly, P. (2010). *The monitoring of learning outcomes in sub-Saharan Africa: Senegal languages of instruction and teachers' methods in senegal grade 3 classrooms.* Both prepared for the William and Flora Hewlett Foundation under the program Monitoring of Learning Outcomes in Sub-Saharan Africa. Research Triangle Park, North Carolina: RTI International. Retrieved from <https://www.eddataglobal.org/>

²³ Krashen, S. (1982). *Principles and Practice in Second Language Acquisition.* New York: Pergamon Press.

they know how to adjust their language use to match the level of the children in the classroom and provide comprehensible input that results in “comprehended intake.”

The results presented in this report are the first of their kind on this scale for the Philippines, and represent a baseline against which progress can be measured in the future. It will also be interesting for the Ilokano group to determine intergrade progression in mother-tongue reading by assessing grade 2 learners as well as grade 1 learners. This could be done, for example, in January 2014 using the same instrument in both grade 1 and grade 2, then in 2015 in grade 2 and grade 3, again using the same instrument. In the future, it is recommended to limit this externally administered, sample-based EGRA to grade 2 students who are a bit more mature and may be less intimidated by the procedure.

Similarly, a study completed in grade 3 after the MTB-MLE implementation had been under way for several years would help detect whether children were able to acquire English and Filipino competency on the foundation of their mother-tongue literacy as fast as, or faster than, in the previous structured bilingual immersion method. Again, before definitive conclusions are drawn about the effectiveness based only on student outcome measures, more detailed studies need to be carried out to determine whether mother-tongue literacy was actually taught and acquired by the children. Thus, for example, in 2015 the children who were tested in Ilokano will be completing grade 3. At this time, a measurement of Ilokano, English, and Filipino literacy could be conducted with each child to determine proficiency across languages.

Note that for repeated use of the EGRA instrument, the items will have to be changed and equivalency studies conducted to ensure that the measurements are the same, yet the items cannot be known by teachers in advance such that they can artificially prepare students for success on the assessment. More details will be provided on this and other considerations for comparability and scale-up in a separate EGRA toolkit that is under development for DepEd Philippines.

Finally, the results concerning of over-aged students and the gender gap deserve further attention. The issue of over-aged students should be looked into in more detail to identify the reasons why overage children are present and why they are not succeeding—are they children who started late, repeated a grade, and if so was this due to some kind of learning disability or other factors that could be addressed through policy actions. Studies to find out more about the origins of, and how to address, the gender and age gap would also be of interest. Why are there so many overage children in the early grades? What is preventing boys from achieving as high as girls? Does the gender gap widen or disappear with age? How is it reflected in retention and completion rates?

This proposed list of follow-on studies is not exhaustive, and there are certainly many other qualitative and quantitative studies that could add important contextual information with which to improve the interpretation of these findings and to improve implementation of programs to address them.

Annex B: MTB-MLE Index

Procedure: Each school is assigned a "score" based on the variables mentioned here, which are derived from questionnaires and classroom observation data.

Component 1: Teacher preparation, readiness and support for MTB-MLE

Question (variable name)	Value (label)	Index contribution
During your pre-service training, did you receive any specific training on how to teach reading in the early grades? (t_pre_reading_specify_1)	if =1 (yes)	1
During your pre-service training, did you receive any specific training on how to teach beginning reading? (t_pre_reading_specify_2)	if =1 (yes)	1
During your pre-service training, did you receive any specific training on how to teach in the mother tongue? (t_pre_reading_specify_3)	if =1 (yes)	1
What is your native language? [the language of assessment?] (t_NATIVELANG_EGRA)	if =1 (yes)	1
How comfortable do you feel teaching using [the EGRA language] as a medium of instruction (t_7)	if =1 (somewhat)	1
	if =2 (Very)	2
Are you familiar with grammar rules, alphabet sounds, spelling, and pronunciation in [the EGRA language] and how to teach it? (t_8)	if =1 (somewhat)	1
	if =2 (Very)	2
Have you ever used EGRA with your students in this classroom? (t_12)	if =1 (yes)	1
How often does the principal check your mother tongue/reading lesson plans or log? (t_20_FREQ_CAT2)	if 1x/week or more	1
	if less frequently than 1x/week	0
How frequently does your principal perform an observation of your Mother Tongue lesson? (t_22_CAT3)	if never	0
	If less frequently than 1x/week	1
	if 1x/week or more	2
How many times since the start of the school year have you received instructional support from any type of instructional supervisor? (t_23_CAT3)	if never	0
	If less frequently than 1x/week	1
	if 1x/week or more	2
How do you measure your pupils' progress in reading (any languages)?: Frequent oral evaluations (other than EGRA/Phil-IRI) (t_27_2)	if =1 (yes)	1
How do you measure your pupils' progress in reading (any languages)?: Portfolios and other projects" (t_27_3)	if =1 (yes)	1
How do you measure your pupils' progress in reading (any languages)?: Homework (t_27_4)	if =1 (yes)	1
How do you measure your pupils' progress in reading (any languages)?: EGRA (t_27_8)	if =1 (yes)	1

Question (variable name)	Value (label)	Index contribution
How do you measure your pupils' progress in reading (any languages)?: Phil-IRI (t_27_10)	if =1 (yes)	1
How do you use the results of those assessment activities in your teaching? : To design teaching materials. (t_28_2)	if =1 (Yes)	1
How do you use the results of those assessment activities in your teaching? : To plan teaching activities. (t_28_3)	if =1 (Yes)	1
How do you use the results of those assessment activities in your teaching? : To adapt teaching to better suit students' needs. (t_28_4)	if =1 (Yes)	1
At what grade level do you expect children to be reading [the EGRA language] fluently? (t_31)	if =1 (Grade 1)	1
At what grade level do you expect children to be reading English fluently? (t_32)	if =3 (Grade 3)	1
At what grade level do you expect children to be reading Filipino fluently? (t_33)	if =2 (Grade 2)	1
Total index score possible for this component		14

Component 2: Mother Tongue Materials

Question (variable name)	Value (label)	Index contribution
Do children have books [learner materials, workbook, reader, textbook, other] corresponding to the subject area? (Observed). (CO_BOOKS_SCORE)	if =1 (some)	1
	if =2 (All)	2
If they have books, are they in the mother tongue [EGRA language]? (CO_BOOKS_LANG_SCORE)	if =1 (yes)	1
Are there posters and reading materials on the wall in the EGRA language? (CO_POSTERS_LANG_SCORE)	if =1 (yes)	1
How many posters are visible around the room? (CO_POSTERS_SCORE)	If none	0
	If some	1
	If many	2
What other resources do you have for supporting the MTB-MLE program? : Big books (t_14_1)	if =1 (yes)	1
... : Textbooks (t_14_2)	if =1 (yes)	1
... : Workbooks (t_14_3)	if =1 (yes)	1
... : Posters (t_14_4)	if =1 (yes)	1
... : Multimedia (t_14_5)	if =1 (yes)	1
... : Teachers guide (t_14_6)	if =1 (yes)	1
... : Visual aids (t_14_8)	if =1 (yes)	1
Do you think the resources you have are sufficient? (t_15)	if =1 (yes)	1
Total index score possible for this component		14

Component 3: Instruction

Question (variable name)	Value (label)	Index contribution
% of total classroom observations where MT was in use (COR_LANG_%)	>=80%	1
In the reading lesson, pupil participation using the EGRA language was: (COR_GENERAL_LESSON_3)	if =2 (Pupils participate when called on to do so and some pupils volunteer)	1
	if =3 (Pupils participate actively (including showing a willingness to ask and answer questions, make guesses.)	2
During the reading lesson, how did the teacher support the pupils' understanding and use of the EGRA language?: ...: By modeling the proper EGRA language use for pupil to repeat or emulate (COR_GENERAL_LESSON_4_2)	if =1 (yes)	1
...: By praising or rewarding pupils for using the EGRA language appropriately (COR_GENERAL_LESSON_4_4)	if =1 (yes)	1
...: By pointing out or correcting EGRA language errors and misunderstandings (COR_GENERAL_LESSON_4_5)	if =1 (yes)	1
...: By monitoring comprehension (e.g. 'Do you understand?' 'Need more explanation?') (COR_GENERAL_LESSON_4_6)	if =1 (yes)	1
...: Using visual aids like pictures, charts (COR_GENERAL_LESSON_4_7)	if =1 (yes)	1
...: By using simpler form of EGRA language to explain more difficult language. (COR_GENERAL_LESSON_4_8)	if =1 (yes)	1
During the reading lesson, when teacher used a different language (code-switching), it was for what reason? ... : To clarify or emphasize a concept explained first in the EGRA language (COR_GENERAL_LESSON_6_1)	if =1 (yes)	1
...: To give procedural directions (COR_GENERAL_LESSON_6_3)	if =1 (yes)	1
Classroom where the EGRA language lesson content was "reading" on 25% or more of the observations (COR_LC_%reading (new var))	>=25%	1
If more than 25% of the observations coded as "speaking" are coded in any of the "on task" subtasks COR_LC_SUB_%speak	>=25%	1
If more than 25% of the observations coded as "listening" were for the subtasks listening to instructions about sounds, letters and syllables or listening to instructions about words. COR_LC_SUB_%listen	>=25%	1
If more than 25% of the observations codes as writing were for the subtasks dictation, writing word or sentence length answers to questions, or creative writing. COR_LC_SUB_%writing	>=25%	1
COO*_LANG_% (new var)	>=80%	1
Pupil use of EGRA language during other lesson COO*_GENERAL_LESSON_3	if =2	1
	if =3	2
How teacher corrects use of EGRA language: COO*_GENERAL_LESSON_4_2	if =1	1
COO*_GENERAL_LESSON_4_4	if =1	1
COO*_GENERAL_LESSON_4_5	if =1	1
COO*_GENERAL_LESSON_4_6	if =1	1
COO*_GENERAL_LESSON_4_7	if =1	1
COO*_GENERAL_LESSON_4_8	if =1	1
COO*_GENERAL_LESSON_6_1	if =1	1
COO*_GENERAL_LESSON_6_3	if =1	1
Total index score possible for this component		26

* Same as above but for other subject area lesson

Annex C: Sampling methodology

The principal goal of the 2014 EGRA conducted in the Philippines in 2014 was to assess the reading ability of students in 1st and 2nd grade who attended government schools known to instruct the four specific mother tongue languages in the four target regions. The Regions and languages are: Region I – Ilokano, Region VI – Hiligaynon, Region VII - Sinugbuonong Binisaya, ARMM – Magindanaon. However, the BEIS currently does not collect information on schools' language of instructions. Therefore the final list of the schools located in the 4 Regions instructing in the 4 specific languages was compiled by regional DepEd staff and may not be 100% accurate or complete.

Exclusions. The 2013 BIES list of public primary schools was used to extract a then random sample of schools known to have a grade 1 and grade 2 enrollment, were located in the 4 Regions, and were known to instruct in the specified mother tongue language. Of the 46,603 schools in the national list, 8,474 schools were not excluded and thus were considered our final 'universe' of schools from which 160 schools were sampled at random. A list of the excluded schools and reasons for the exclusion may be found in the Table C-1.

Table C-1. Number of Excluded schools from the 2013 BIES school list and reasons why schools were excluded.

Reason for Exclusion	N schools
Total Number of Schools in the BIES list.	46,603
School was not located in one of the four Regions.	-33,376
School is not a primary school	-399
School contained zero enrollment in Grade1 and/or Grade2	-1,559
School was used for the Practice Study	-8
School was used for the Pilot Study	-47
Region ARMM school did not Instruct in Maguindanaon	-1,786
Region VI school does not instruct in Hiligaynon	449
Region I school did not instruct in Ilokano.	-505
Not Excluded	8,474

* Note: It is believed that 97% of the population in Region VII speak Sinugbuonong Binisaya so no additional exclusions were made within this Region.

Sample Methodology and Sample Counts. After excluding schools as mentioned above, schools were stratified by the 4 Regions and 40 schools from each Region were randomly sampled proportional to the combined Grade1+Grade2 enrollment. Within each sampled school, one Grade1 and one Grade 2 classroom were sampled with equal probability. Within each of the sampled classrooms, about five or six girls and five or six boys were randomly sampled with equal probability. Table C-2 provides the 3 stage sample details. Table C-3 provides the final counts of the schools, classrooms, and students sampled for each Region.

Table C-2: Sample methodology for the 2014-Four Region EGRA-SSME.

Stage Number	Item Sampled	Stratified by	Probability of Selection
Stage 1	Schools (160) (40-in each Region)	Region (4)	Proportional to Grade 1 + Grade 2 Enrollment
Stage 2	Classrooms (320) 160 Grade1, 160 Grade2	Grade (2) [Grade1, Grade2]	Equal
Stage 3	Student(800) 200 grade1-Boys 200 grade1 girls, 200 grade2 boys, 200 grade 2 girls)	Gender (2) [Male, Female]	Equal

Table C-3: The four Regions' final sample counts of schools, classrooms (by grade) and students (by grade, gender).

Region	Schools	Classrooms		Students					
		G1	G2	G1	G1 Male	G1 Female	G2	G2 Male	G2 Female
Region I	40	40	40	400	202	198	399	200	199
Region VI	40	40	40	402	205	197	400	200	200
Region VII	40	40	40	472	237	235	434	216	218
ARMM	40	40	40	402	194	208	404	203	201
Total	160	160	160	1,676	838	838	1,637	819	818

Assessments Conducted and Final Counts. This EGRA consisted of four instruments: a teacher interview, classroom observation (for the mother tongue reading class), a classroom observation (for a non-reading class), a teacher interview, and a student EGRA assessment and a student questionnaire. Once the assessors arrived at the sampled school and randomly selected a grade 1 and grade 2 classroom, they conducted a teacher interview for each teacher associated with the selected classroom. They also sat in and observed the classroom's reading class, as well as another class (math, social studies, or a MAPEH [Music, Art, Physical Education, and Health]). Table C-4 provides the completed classroom level questionnaires captured and used in this analysis.

Table S4: Counts by Region-Grade of the completed assessments/questionnaires at the classroom level.

Region	Completed Teacher Interview		Completed Reading Classroom Observation		Completed Other Classroom Observation		
			Grade1	Grade2	Grade1	Grade2	
	RegionI	39	40	35	33	37	37
RegionVI	39	38	35	23	40	40	
RegionVII	39	39	23	25	40	40	

ARMM	39	40	39	40	40	40
Total	157	157	132	121	157	157

For each of the sample students, the assessment team collected a student questionnaire as well as a mother tongue EGRA for both Grade 1 and Grade 2 students. For the grade 2 students only, the assessment team collected a mini-EGRA in English and Filipino. Table S5 displays the counts of the completed student level assessments/questionnaires.

Table C-5: Counts by Region-Grade of the completed assessments/questionnaires at the student level.

Region	Completed Student Interview		Completed Mother Tongue EGRA		Completed English EGRA	Completed Filipino EGRA
	Grade1	Grade2	Grade1	Grade2	Grade2 - only	Grade2 - only
RegionI	400	399	400	399	399	399
RegionVI	402	400	402	400	400	400
RegionVII	378	371	432	432	402	371
ARMM	400	399	400	401	400	401
Total	1,580	1,569	1,634	1,632	1,601	1,571