USAID
FROM THE AMERICAN PEOPLE

USAID/PHILIPPINES BASA PILIPINAS PROGRAM

FINAL OUTCOME EVALUATION REPORT: 2013-2016

JULY 2016

This publication was produced for review by the United States Agency for International Development. It was prepared by Education Development Center, Inc. (EDC).

USAID/PHILIPPINES BASA PILIPINAS PROGRAM

FINAL OUTCOME EVALUATION REPORT: 2013- 2016

Program Title:	USAID/Philippines Basa Pilipinas Program
Sponsoring USAID Office:	USAID/Philippines - Office of Education
Contract Number:	AID-492-C-13-00004
Contractor:	Education Development Center, Inc. (EDC)
Date of Publication:	July 29, 2016
Author:	Education Development Center, Inc. (EDC)

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

EXECUTIVE SUMMARY

Basa Pilipinas is a four-year early grade reading project that supports the Philippine's Department of Education's national reading program. Basa is helping the Department of Education (DepEd) implement transformative literacy practices in selected divisions of Regions 1 and 7 by providing teacher and student materials, training teachers and school heads, and providing post-training support for Grade 1, 2 and 3 teachers.



The project evaluation is designed to measure changes in Grade 2 student achievement associated with the Basa intervention, focusing on the effects of two years of the intervention, and to measure growth in reading performance from Grade 2 to Grade 3 in Filipino and English. The evaluation study also measures improvement in teachers' instructional practices in reading over the course of the Basa intervention. To explore factors that are associated with better early grade reading performance of Grade 2 students in Filipino and English, this report also includes the results of an additional exploratory study, Best Practices Study (BPS).

The study follows a cross-sectional design to examine changes in student achievement in reading. In school year 2013/14, prior to the full implementation of the project, the Basa project team collected data on Grade 2 student achievement in literacy, teachers' classroom practices, teacher beliefs on literacy instruction, and data on school environment, to serve as a comparison cohort for the Basa intervention. Data were collected from 40 schools in Cebu and La Union. In the two subsequent school years, after the rollout of the full Basa intervention, data collection was expanded to 80 intervention schools in Bohol, Cebu, Ilocos Norte, Ilocos Sur, and La Union. In school year 2014/15, data was collected on student achievement in Filipino in Grade 2, and in school year 2015/16 data was collected on student achievement in Grades 2 and 3 in Filipino and English.

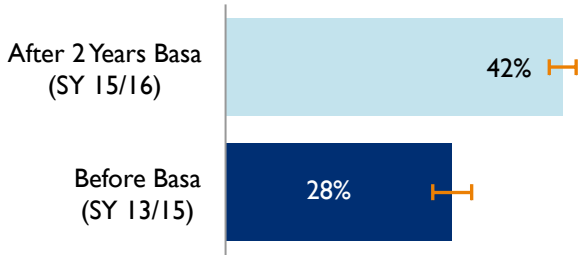
This report focuses on a comparison of student achievement in Grade 2 from baseline in school year 2013/14 to endline in school year 2015/16. In addition, we examine reading performance in Grades 2 and 3 in school year 2015/2016. To measure improvement in teacher practice, the study follows a longitudinal design in which the same Grade 2 teachers were observed annually. Additional data was collected from 10 schools, four top performing and six bottom performing, to examine factors that may contribute to student achievement in literacy. Key findings are outlined below.

GRADE 2 STUDENT LITERACY FINDINGS

After two years of Basa intervention, Grade 2 students are performing significantly better in Filipino than students performed prior to the intervention on four out of eight of the EGRA subtests, specifically, Initial Sound Identification, Letter Sounds, Reading Comprehension and Dictation. Of particular interest, analysis showed significant improvement in learner performance in two key EGRA measures – Filipino Oral Reading Fluency and Reading Comprehension. For Reading Comprehension, the average percent correct for students after two years of Basa was 42% versus 28% before Basa. An analysis of the effect size, which is used to measure the magnitude of change, was also conducted. The largest effect sizes were in Dictation (.52) and in Reading Comprehension

(.43). There were also moderate effects in the percent correct for Initial Sound Identification (.42), Letter Sounds (.32) as well as the Prosody score (.33).

Improvements in Grade 2 Filipino Reading Comprehension over the Basa Intervention



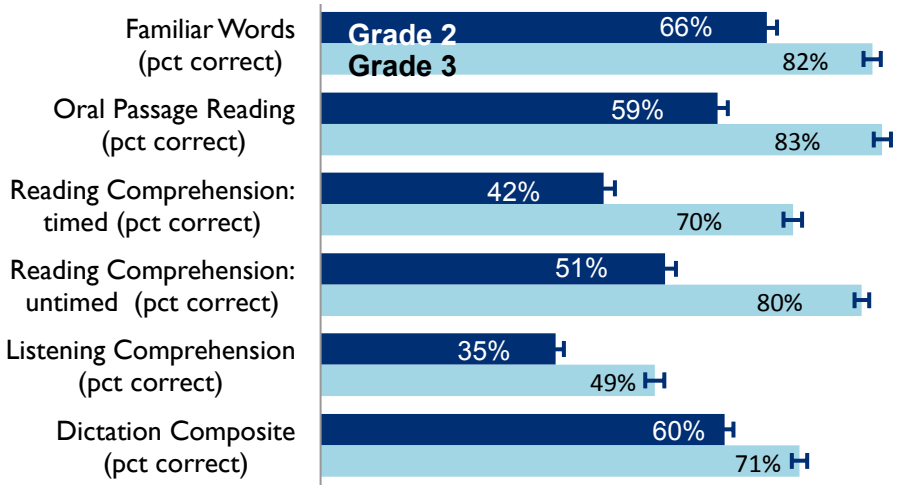
In English literacy, Grade 2 students perform best on initial sound identification and oral reading subtests. The English assessment was not administered to students in school year 2013/14, before the Basa intervention, as a result, the report provides results for intervention cohort students only in school years 2014/15 and 2015/16. Overall, performance in English remained relatively similar in school years 2014/15 and 2015/16. The subtests with the lowest scores were English Reading Comprehension, Listening Comprehension and Dictation. In both years, about half of second graders scored zero on English Reading Comprehension.

At Grade 2, girls, on average, continue to out-perform boys in all Filipino and English subtests in school year 2015/16. The difference in reading performance between boys and girls is statistically significant at the $p<.001$ level for nearly all Filipino and English subtests. Girls also have showed larger improvements than boys from 2013/14 to 2015/16. One exception to this trend is Filipino Reading Comprehension; results show that boys showed larger gains than girls. In fact, boys have nearly caught up to girls in Filipino Reading Comprehension after two years of Basa intervention.

FINDINGS FOR LEARNING TRAJECTORY IN READING FROM GRADE 2 TO GRADE 3

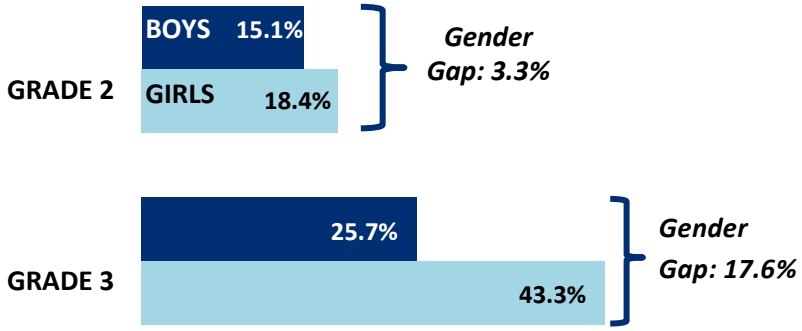
Overall, performance in Grade 3 on all subtests in Filipino was substantially better than in Grade 2. For each subtest, Grade 3 learners scored between 11 to 28 percentage points higher than Grade 2 students. Average fluency in Grade 3 was 57.8 words correct per minute, whereas in Grade 2 it was 39.4 words correct per minute. Reading comprehension in Grade 3 averaged 70% correct, whereas the Grade 2 average was 42% correct. There was also a significant reduction in zero scores on all subtests; the largest reduction was in reading comprehension, where zero scores fell from 21% in Grade 2 to 4% in Grade 3. These results indicate that most students are reading at a sufficient level to transition to Filipino as the primary language of instruction in most subjects in Grade 4.

Comparisons of Grades 2 and 3 Average Filipino EGRA Subtest Results for SY 2015/16



Grade 3 pupils also demonstrated higher scores in English than second graders, scoring between 8 to 19 percentage points higher on all subtests. Third graders performed best on Familiar Word Reading, where they averaged 76% correct, and on measures of Oral Reading, where they read an average of 76% of words in a passage correctly at a rate of 60.5 words per minute. However, they did not perform as well in Reading Comprehension, where the average percent correct was only 35%. As seen in Filipino, there were statistically significant reductions in zero scores on English EGRA subtests from Grade 2 to Grade 3. Large reductions in zero scores were seen on the Reading Comprehension and Listening Comprehension subtests, in which zero scores dropped by 11% and 12% respectively, but zero scores were still high on these subtests. This suggests that while there is significant progress in English reading skills from Grade 2 to Grade 3, most students are not prepared to transition to learning math and science in English in Grade 4.

Widening Gender Gap in English Reading Comprehension from Grade 2 to Grade 3



In measures of Filipino reading, the gender gap remained relatively consistent from Grade 2 to 3 with girls outperforming boys by about the same margin in both grades. However, the gender gap increased in some subtests in English. For example, in English reading comprehension, second grade girls scored only 3.3% higher, on average, than boys, while in third grade the gender gap increased to 17.6%.

CHANGES IN TEACHER PRACTICE FINDINGS

Observed second grade teachers' practices in a Filipino reading class showed improvements in both classroom structure and language literacy instruction over the course of two years of Basa intervention. The largest gains were seen in classroom structure in ensuring accessible classroom materials, providing learners the opportunity to reflect on reading and writing, and effective management of reading and writing instruction. Overall, 62% of teachers scored a rating of "strong" on classroom materials, and 30% scored a "strong" on effective management of reading and writing instruction. Teachers scored lowest in the area of effective grouping strategies, a less familiar practice.

In classroom practices specific to literacy instruction, teachers scored lower overall than on classroom structure but showed significant improvement in the areas of providing opportunities for oral language development, developing reading fluency and developing comprehension. The growth in providing opportunities for developing reading fluency is particularly striking, as we saw almost no evidence of this practice prior to the Basa intervention. Oral language development is also a key skill for teachers, particularly in a multi-lingual context where bridging opportunities from one language to another needs to be intentionally planned by teachers.

Higher scores in measures of teacher practice were significantly positively correlated with average Grade 2 Filipino EGRA scores, however, the relationship was relatively weak (r=.22). Providing opportunities for reflection, oral language development, developing vocabulary and comprehension had the strongest correlations with average Filipino EGRA scores.

FINDINGS ON FACTORS THAT MAY CONTRIBUTE TO STUDENT ACHIEVEMENT IN LITERACY

Results from the BPS study, showed that the school environment in bottom and top performing schools was very similar, in that, schools had comparable class sizes, numbers of teachers employed, and attendance of principals and teachers. However, teachers in bottom performing schools reported higher levels of student absenteeism and roughly a half hour less of class time, on average, for students in Grades 1, 2 and 3. Top performing schools tracked student attendance more closely than bottom performing schools. Principals in top performing schools also appeared to be more engaged, reporting checking teachers' lesson plans and observing classrooms more frequently than bottom performing schools.

Teachers' method for delivering lesson content in bottom performing schools was **overwhelmingly teacher-directed** and not focused on active student engagement. Conversely, teachers in top performing schools, had a more **balanced lesson delivery**, in which roughly two-thirds of the class time was spent instructing the whole class, and the other third was devoted to independent or small group work.

Teachers in top performing schools also demonstrated better teaching practices. They were substantially better at ensuring a positive learning environment, providing opportunities for developing fluency, comprehension and meaningful reading, as well as effective grouping strategies. For example, half of teachers in top performing schools were rated "strong" or "exemplary" in conducting effective grouping, while the majority of teachers in bottom performing schools were rated "inadequate" in this practice. Teachers' method for delivering lesson content in bottom performing schools was overwhelmingly teacher-directed and not focused on active student engagement. Conversely, teachers in top performing schools had a more balanced lesson delivery, in which roughly two-thirds of the class time was spent instructing the whole class, and the other third was devoted to independent or small group work. A small, but statistically significant relationship was found between the percent of time teachers devote to independent or group work and higher Filipino reading fluency scores (r=.26).

All schools reported having parent-teacher associations that met regularly during the school year. However, principals in top performing schools reported meeting with a larger percentage of parents or guardians during the school year. They were also slightly more satisfied with the level of parental involvement than those in bottom performing schools. There was a moderate positive relationship between principal satisfaction with parental involvement and higher Filipino reading scores (r=.32).

CONCLUSIONS AND RECOMMENDATIONS

By the end of Grade 2, after one full year of reading instruction in Filipino, students are demonstrating beginning reading skills in the language. The Basa approach to literacy instruction is effective, as demonstrated by the finding that students who benefitted from the Basa intervention for two years show larger gains on key measures of literacy such as oral reading fluency and comprehension than those who did not benefit from Basa. Earlier exposure to language instruction is also likely to improve performance in the early grades, as attendance in kindergarten was positively associated with higher performance in Grade 2.

By the end of Grade 3, results show that students are fairly proficient Filipino readers with the majority attaining both the fluency and comprehension benchmarks set by DepEd. This suggests that most students are ready to transition to Filipino as the primary language of instruction in most

subjects in Grade 4. However, by the end of Grade 3 there is still a significant proportion of students who do not demonstrate that they understand what they read in English. This indicates that more must be done to help prepare students to learn math and science in English in Grade 4.

It is also a concern that, although reading performance of both girls and boys in Grade 2 has improved after two years of Basa intervention, findings suggest that girls are showing larger improvements as a result of the intervention compared to boys, with the exception of Filipino reading comprehension, where boys have nearly caught up to girls. This is an important issue given that a persistent gender gap has continued in which girls outperform boys in Filipino and English. Of equal concern is that the gender gap in Filipino and English between boys and girls does not close, or even narrow, by Grade 3. In fact, it widens in some subtests. There is an urgent need for additional research in order to find strategies that can increase boys' achievement in reading in the early grades in order to better prepare them for academic success.



While teachers demonstrate good progress in applying effective classroom management and literacy instructional practices, there is room for improvement particularly in the more unfamiliar and/or complex teaching strategies. Evidence from the best practice study of low and high performing schools suggests that continuing to improve teacher practices in areas such as effective grouping is likely to have a positive effect on student achievement in reading. This indicates that teachers will need continuing professional development to reinforce and deepen their understanding of effective literacy instruction in order to see continued progress in student achievement. In particular, training should focus on managing differentiated instruction and reinforcing the reading-writing connection.

In addition to more training and support to apply specific literacy instruction strategies in class, reinforcing supervision in general is likely to have a positive effect on teacher performance. Findings from the BPS study found that schools with principals who report checking lesson plans and observing classrooms frequently tend to perform better than those who report less frequent checking. There was no difference among top and bottom performing schools in how often they convened LAC sessions. However, we did find positive associations between student achievement in reading and principals who reported attending LAC training and providing instructional support to teachers on teaching reading.

Not surprisingly, results from the analysis of best practices also indicate that students in schools where there is more time on task, due to low absenteeism and more class time, have higher learning gains. Our observations show that class time in both top and bottom performing schools is spent almost entirely on instruction. This finding suggests that additional time dedicated to literacy instruction may be required in order to increase learning gains beyond what can be achieved by improving instructional practice during regular class time. Students can benefit from additional instruction during remedial reading time, or other structured opportunities to hone their reading skills. In addition, providing more opportunities for students to access a variety of reading materials may help improve achievement in reading. Our study showed that schools where learners visited the school library more frequently tended to have higher Filipino comprehension scores.



Parental literacy was also found to have a significant positive association with student performance. Top performing schools reported meeting with a larger percentage of parents during the school year, and principals reported higher levels of satisfaction with parents' involvement in their children's school work. These findings suggest that regular participation of all parents in their children's learning may help improve student performance.

CONTENTS

EXECUTIVE SUMMARY 2
CONTENTS 8
List of Figures 10
List of Tables 12
ACRONYMS 14
INTRODUCTION 1
Structure of Report 2
EVALUATION DESIGN 4
Methodology 4
Sample 6
1. STUDENT CONTEXT INTERVIEW FINDINGS 8
School and Teacher Environment 8
Socio-Economic Factors 9
Home Environment 10
2. SCHOOL AND LEARNING ENVIRONMENT FINDINGS 12
Classroom Environment 12
School Leadership 13
Learning Environment 14
Accessibility of Teaching and Learning Materials 14
Teacher Perceptions of Teaching and Learning Materials 15
Teaching Practices 16
3. CHANGES IN GRADE 2 EGRA RESULTS OVER THE LIFE OF THE BASA PROJECT 18
Summary of Filipino and English EGRA Findings 18
Overall Grade 2 Filipino EGRA Results 18
Overall Grade 2 English EGRA Results 20
Zero Scores 22
Results by Sex 23
Benchmarks 24
Grade 2 Filipino EGRA Subtest Results 26
Initial Sound Identification, Letter Sounds, Familiar and Nonsense Words 26
Oral Reading Fluency and Reading Comprehension in Filipino 28
Listening Comprehension and Dictation 33
4. FILIPINO & ENGLISH EGRA FINDINGS – COMPARISON OF GRADE 2 & GRADE 3 36
Summary Comparisons of Grade 2 and Grade 3 EGRA Results 36
Overall Grade 2 and Grade 3 Filipino EGRA Results 36
Overall Grade 2 and Grade 3 English EGRA Results 37
Benchmarks 42
Grades 2 and 3 Filipino Subtest Results 43
Familiar Words 43
Oral Reading Fluency and Reading Comprehension 44
Listening Comprehension and Dictation 48
Grades 2 and 3 English Subtest Results 50
Familiar Words 50

Oral Reading Fluency and Reading Comprehension 51
Listening Comprehension and Dictation 56
5. IMPACT OF CONTEXTUAL FACTORS ON ACHIEVEMENT 59
Learner Characteristics and EGRA Results 59
Impact of Home Language 60
Impact of School Leadership 63
Impact of Classroom Environment 63
6. CHANGES IN TEACHER PRACTICE 65
Classroom Structure Results 65
Language and Literacy Instruction Results 66
7. BEST PRACTICES STUDY (BPS) 68
Description of the Study 68
Study Methodology 68
Study Findings 69
Barangay Socioeconomic Profile 69
Principal and Teacher Demographics 70
School Environment 70
School Management 71
Reading Environment 73
Teacher Beliefs and Motivation 74
Teaching Practices 75
Time on Task in Reading 77
Parental and Community Involvement 79
ANNEXES 1
ANNEX 1. METHODOLOGY 2
Design 2
Sample 3
Data Collection Tools 5
Outcome Evaluation Tools 5
Best Practice Study (BPS) Evaluation Tools 13
Data Collection 15
Scope Data Collection 15
EGRA Data Collection 16
Best Practices Study 18
Data Analysis 18
Limitations 18
ANNEX 2. EGRA SUBTESTS 20
ANNEX 3. SUMMARY EGRA RESULTS 22
Grade 2 EGRA Results 22
Detailed Grade 2 Filipino EGRA Results 22
Detailed Grade 2 English Results 26
Grade 3 EGRA Results 29
Detailed Grade 3 Filipino EGRA Results 29
Detailed Grade 3 English EGRA Results 32
ANNEX 4. SUMMARY SCOPE RESULTS 34

LIST OF FIGURES

Figure 1: Basa Theory of Change 1
Figure 2. Basa Pilipinas Geographic Scope 3
Figure 3. Timeline of Basa Evaluation Activities Across Cohort 6
Figure 4. Kindergarten Attendance Before Grade 1 (n=3,255) 8
Figure 5. Household Assets (n=3,255) 9
Figure 6. Parental Literacy and Help with Reading at Home (n=3,247) 11
Figure 7. How often do you miss school? (n=3,254) 11
Figure 8. Observed Condition of School and Classroom Infrastructure in Sampled Schools (n=231) 12
Figure 9. Average Principal's Perceptions and Attitudes of Literacy Instruction in Sampled Schools (n=96) 14
Figure 10. How do you usually use the Leveled Readers in your classroom? (n=231) 16
Figure 11. Can Learners Bring Home Books from the Classroom? 17
Figure 12. Average Percent Correct for Grade 2 Filipino EGRA Subtests, by Cohort 20
Figure 13. Average Percent Correct on Grade 2 English EGRA Subtests for Cohort 3 (SY 2015/16) 21
Figure 14. Percent of Tested Grade 2 Students Scoring Zero on Filipino EGRA Subtests 22
Figure 15. Percent of Tested Students in SY 2015/16 Scoring Zero on Grade 2 English EGRA Subtests 23
Figure 16. Average Percent Correct on Filipino and English Grade 2 EGRA Subtests, by Sex 24
Figure 17. Percent of Grade 2 Students Meeting Filipino Fluency and Comprehension Benchmarks. 25
Figure 18. Percent of Grade 2 Students Meeting Filipino ORF and Comprehension Benchmarks in SY 2015/16, by Sex 25
Figure 19. Filipino Initial Sound Identification – Percent Correct, by Cohort 26
Figure 20. Filipino Letter Sound Subtest Percent Correct, by Cohort 27
Figure 21. Filipino Familiar Word Identification Subtest Results, by Cohort 27
Figure 22. Filipino Nonsense Words Subtest Results, by Cohort 28
Figure 23. Filipino Oral Passage Reading Results, Percent of Words Read Correctly, by Cohort 29
Figure 24. Filipino Oral Passage Reading Average Percent Correct, by Cohort and Sex 29
Figure 25. Filipino Oral Reading Fluency (wcpm), by Group 30
Figure 26. Filipino Fluency Measure (wcpm), by Sex and Cohort 30
Figure 27. Filipino Oral Passage Reading - Prosody Score, by Cohort 31
Figure 28. Grade 2 Reading Comprehension Results, Filipino EGRA, by Cohort 32
Figure 29. Average Percent Correct on Reading Comprehension, by Sex and Cohort 32
Figure 30. Measuring Reading Comprehension 33
Figure 31. Comparison of Cohort 3 Comprehension Results Using Timed and Untimed Reading 33
Figure 32. Filipino Listening Comprehension Questions – Number Correct, by Cohort 34
Figure 33. Grade 2 Filipino Dictation Results (Percent Correct), by Cohort 35
Figure 34. Filipino Dictation Breakdown, by Group 35
Figure 35. Average Filipino EGRA Subtest Results for Cohort 3 (SY 15/16), by Grade 37
Figure 36. English EGRA Subtest Results for Cohort 3 (SY 2015/16), by Grade 39
Figure 37. Percent of Tested Grade 2 and Grade 3 Students Scoring Zero on Filipino EGRA Subtests 40
Figure 38. Percent of Tested Grade 2 and Grade 3 Students Scoring Zero on English EGRA Subtests 40
Figure 39. Average Percent Correct on Filipino and English EGRA Subtasks (SY 2015/16), by Grade 41

Figure 40. Gender Gap in English Reading Comprehension Percent Correct, by Grade (Cohort 3 – SY 2015/16) 41
Figure 41. Percent of Cohort 3 (SY 15/16) Learners Meeting Filipino Fluency and Reading Comprehension Benchmarks, by Grade 42
Figure 42. Percent of Cohort 3 (SY 15/16) Learners Meeting Combined Filipino Fluency and Reading Comprehension Benchmarks, by Grade 42
Figure 43. Percent of Cohort 3 (SY15/16) Learners Reading at least 60 wcpm in Filipino, by Grade and Sex 43
Figure 44. Filipino Familiar Word Identification Subtest Results, by Grade 44
Figure 45. Filipino Oral Passage Reading Results, Percent of Words Read Correctly, by Grade 44
Figure 46. Filipino Oral Passage Reading Average Percent Correct, by Grade and Sex 45
Figure 47. Filipino Oral Passage Reading Subtest Results for Cohort 3 (SY 15/16) by Grade 45
Figure 48. Filipino Fluency Measure (wcpm) for Cohort 3 (SY 15/16), by Sex and Grade 46
Figure 49. Filipino Oral Passage Reading - Prosody Score, by Grade 46
Figure 50. Reading Comprehension Results, Cohort 3 (SY 15/16) Filipino EGRA, by Grade 47
Figure 51. Average Percent Correct on Filipino Reading Comprehension (SY 15/16), by Sex and Grade 47
Figure 52. Comparison of Grade 3 Comprehension Results Using Timed and Untimed Reading (SY 15/16) 48
Figure 53. Filipino Listening Comprehension Questions – Number Correct, by Grade (SY 15/16) 49
Figure 54. Filipino Dictation Results (SY 2015/16), by Grade 50
Figure 55. Filipino Dictation Breakdown (Out of Possible 16 Points) (SY 15/16), by Grade 50
Figure 56. English Familiar Word Identification Subtest Results, by Grade 51
Figure 57. Average Percent of English Familiar Words Identified (SY 15/16), by Grade and Sex 51
Figure 58. English Oral Passage Reading Results, Percent of Words Read Correctly, by Grade 52
Figure 59. English Oral Passage Reading Average Percent Correct (SY 15/16), by Grade and Sex 52
Figure 60. English Oral Passage Reading Subtest Results for Cohort 3 (SY 15/16) by Grade 53
Figure 61. English Fluency Measure (wcpm) for Cohort 3 (SY 15/16), by Sex and Grade 53
Figure 62. English Oral Passage Reading - Prosody Score (SY 15/16), by Grade 54
Figure 63. Reading Comprehension Results, Cohort 3 (SY 15/16) English EGRA, by Grade 55
Figure 64. Average Percent Correct on English Reading Comprehension (SY 15/16), by Sex and Grade 55
Figure 65. Comparison of Grade 3 Comprehension Results Using Timed and Untimed Reading (SY 15/16) 56
Figure 66. English Listening Comprehension Questions – Number Correct, by Grade (SY 15/16) 56
Figure 67. English Dictation Results (SY 2015/16), by Grade 57
Figure 68. English Dictation Breakdown (SY 15/16) (Out of possible 17), by Grade 57
Figure 69. Grade 2 EGRA Results for Cohort 3 (SY14/15) Disaggregated by Language Spoken at Home 61
Figure 70. Grade 3 EGRA Results for Cohort 3 (SY14/15) Disaggregated by Language Spoken at Home 62
Figure 71. Changes in Average SCOPE Scores from SY 2013/14 to SY 2015/16 (n=28) 65
Figure 72. Change in Average Classroom Structure Scores for SCOPE Observation from SY13/14 to SY15/16 (n=28) 66
Figure 73. Change in Average Language and Literacy Instruction Scores for SCOPE Observation from SY13/14 to SY15/16 (n=28) 67
Figure 74. Average Percent Correct on Filipino and English EGRA Subtests for Cohort 2 (SY 14/15) Learners in Sampled Schools (n=79) 68

Figure 75. Percent of K-3 Students in 4P Program in School, by School Performance....70
Figure 76. Frequency Principals Apply School Management Training to Work (n=9) 70
Figure 77. How do you know whether your students are progressing academically? (n=9, multiple response) 73
Figure 78. Percent of Schools That Have Received Appropriate Number of Learning Materials at Beginning of School Year (n=9)....73
Figure 79. Teacher Reported Percent of Readers in the Classroom (n=10) 74
Figure 80. Do you find it easy to teach reading? (n=10)....74
Figure 81. Teacher Beliefs on Teaching Reading and Math (n=10) 75
Figure 82. Average SCOPE Score for Top Performing and Bottom Performing Schools (n=10)....75
Figure 83. Average Scores for Select *SCOPE* Domains for Top and Bottom Performing Schools (n=10) 76
Figure 84. Average Percent of Reading Lesson Teacher Spends on Core Reading Skills During Filipino Reading Lesson (n=10) 77
Figure 85. Percentage of Instructional Time Spent Teaching Specific Reading and Comprehension Skills (n=10)....79
Figure 86. On average, what percent of the pupils' parents and guardians did you meet with during the school year? (n=9)....79
Figure 87. Principals' Satisfaction with PTA and Parental Involvement (n=9)....79
Figure A-1. SCOPE-Literacy Dimensions and Indicators....6
Figure A-2. Timeline of Basa Evaluation Activities Across Cohort....15
Figure A-3. School/Teacher are supportive of data collection (n=116)....16
Figure A-4. Assessments were conducted with Few interruptions (n=116) 17
Figure A-5: Percent of Students Who Understood the Language the Assessor Used (n=116)....17
Figure A- 6. Average SCOPE Scores from SY 2014/15 to SY 2015/16 (n=27)....34

LIST OF TABLES

Table 1. Introduction of Filipino and English by Grade Level and Quarter....4
Table 2. EGRA Assessment Schedule....5
Table 3. Student Sample....6
Table 4. Principal and Teacher Sample, Cohort 3 (SY 15/16)....7
Table 5. Most Common Reasons Grade 2 and Grade 3 Students Like and Dislike School....9
Table 6. What language do you speak at home? (n=3,255)....10
Table 7. Where do your parents work? 10
Table 8. Percent of Grade 2 and Grade 3 Classrooms with Teaching and Learning Materials (n=231) 15
Table 9. Filipino Grade 2 EGRA Results, by Subtest and Cohort....19
Table 10. Grade 2 English EGRA Results, Intervention Cohort Sample....21
Table 11. Average Percent Correct on Grade 2 Filipino EGRA, by Sex....24
Table 12. Average Grade 3 and Grade 2 Filipino EGRA Subtest Results for Cohort 3 (SY 15/16)....37
Table 13. English EGRA Subtest Results for Cohort 3 (SY 15/16) by Grade....38
Table 14. Correlations between Student Context Interview Results and Cohort 3 (SY15/16) EGRA Results....59
Table 15. Correlations between School Leadership Practices and EGRA Results, Cohort 3 (SY15/16) 63
Table 16. Correlations of Classroom Infrastructure and Environment with Filipino and English EGRA Results for Cohort 3 (SY 15/16)....64

Table 17. Correlations of Teacher Practices with Filipino and English EGRA Results, Cohort 3 (SY 15/16) 64
Table 18. Best Practices Study Sample 69
Table 19. Use of an Official Form (n=9) 72
Table A-1. Student Sample by Cohort and Grade 4
Table A-2. Cohort 3 (SY 2015/16) Evaluation Sample, by Province 5
Table A-3. SCOPE Literacy Reliability, SY 2015/16 7
Table A-4. SCOPE Classroom Structure Items Reliability, SY 2015/16 7
Table A-5. SCOPE Language and Literacy Instruction Reliability, SY 2015/16 8
Table A-6. Grade 2 EGRA Reliability, Cohort 3 (SY 2015/16) 10
Table A-7. Grade 3 EGRA Reliability, Cohort 3 (SY 2015/16) by Grade 10
Table A-8. Correlations of Grade 2 Filipino EGRA Subtest Cohort 3 (SY 2015/16) 12
Table A-9. Correlations of Grade 2 English EGRA Subtests for Cohort 3 (SY 2015/16) 12
Table A-10. Correlations of Grade 3 Filipino EGRA Subtest Cohort 3 (SY 2015/16) 12
Table A-11. Correlations of Grade 3 English EGRA Subtests for Cohort 3 (SY 2015/16) 13
Table A-12. Assessment Schedule 16
Table A-13. Effect Size (ES) Difference between Intervention (All Provinces) and Comparison Cohort (Cebu/La Union) Means at Baseline 19
Table A-14. Grade 2 Filipino EGRA Subtests 20
Table A-15. Grade 2 English EGRA Subtests 20
Table A-16. Grade 3 Filipino EGRA Subtests 21
Table A-17. Grade 3 English EGRA Subtests 21
Table A-18. Overall Grade 2 Filipino EGRA Results, by Subtest and Cohort 22
Table A-19. Grade 2 Filipino EGRA Subtests – Percent of Learners with Zero Scores 23
Table A-20. Overall Grade 2 Filipino EGRA Results, by Sex and Cohort 23
Table A-21. Grade 2 EGRA Filipino Subtests Summary, by Province and Cohort 24
Table A-22. Overall Grade 2 English EGRA Results, by Cohort 26
Table A-23. Grade 2 English EGRA Subtests – Percent of Learners with Zero Scores 27
Table A-24. Grade 2 English EGRA Results, by Sex and Cohort 27
Table A-25. Grade 2 English EGRA Subtests Summary, by Province and Cohort 28
Table A-26. Grade 3 Filipino EGRA Results for Cohort 3 (SY15/16) 29
Table A-27. Grade 3 Filipino EGRA Results – Percent of Learners with Zero Scores 29
Table A-28. Grade 3 Filipino EGRA Results by for Cohort 3 (SY 15/16) by Sex 30
Table A-29. Grade 3 Filipino EGRA Subtests Summary for Cohort 3 (SY 15/160, by Province 31
Table A-30. Grade 3 English EGRA Results for Cohort 3 (SY 2015/16) 32
Table A-31. Grade 3 English EGRA Results for Cohort 3 (SY 2015/16) – Percent of Learners with Zero Scores 32
Table A-32. Grade 3 English EGRA Results for Cohort 3 (SY 15/16), by Sex 32
Table A-33. Grade 3 English EGRA Subtests Summary for Cohort 3 (SY 15/160, by Province 33
Table A-34. Average SCOPE Scores from SY 2013/14, SY 2014/15 and SY 2015/16 (n=27) for SCOPE Domains 34

ACRONYMS

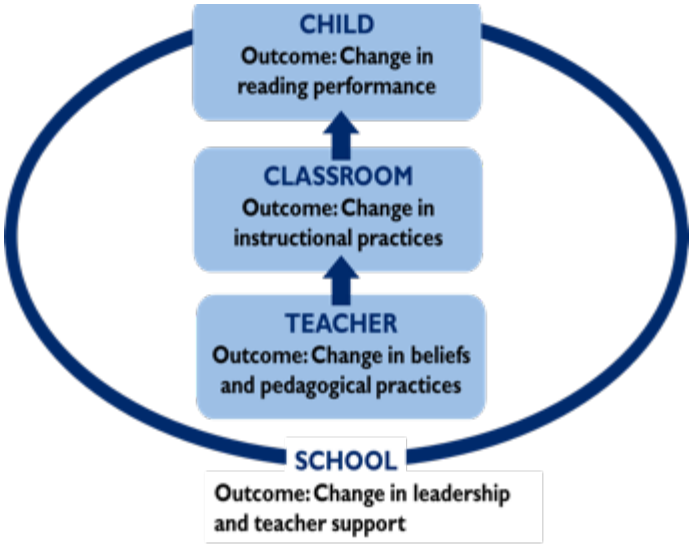
4P	Pantawid Pamilyang Pilipino Program
Basa	Basa Pilipinas
BBF	Brothers Brother Foundation
BIPI	Beliefs and Instructional Practices Inventory (Survey)
BPS	Best Practices Study
DepEd	Department of Education
DNK	Do not know
EGRA	Early Grade Reading Assessment
ICC	Inter-class correlation
IRR	Inter-rater reliability
K-3	Kindergarten through Grade 3
LAC	Learning Action Cell
M&E	Monitoring and Evaluation
MTB-MLE	Mother Tongue-based Multilingual Education
ORF	Oral Reading Fluency
PTA	Parent Teacher Association
RTI	Research Triangle Institute
SCOPE	Standard Classroom Observation Protocol for Literacy
SY	School Year
TG	Teacher Guide
TLM	Teaching and Learning Material
wcpm	Words correct per minute

INTRODUCTION

The Basa Pilipinas (Basa) Project is a four-year project in support of the Philippine Government's literacy component of the K to 12 curriculum, and is implemented in close coordination with the Department of Education (DepEd) and other key education stakeholders in selected divisions and regions nationwide. The project is aligned within the framework of USAID's Global Education Strategy, USAID-Philippines' emerging Country Development and Cooperation Strategy, and the Philippine Government's priorities for basic education. Basa began its transformative work in 2013 to improve the reading skills for one million children in the early grades in English, Filipino and selected Mother Tongues (MT). To this end, Basa has worked on improving reading instruction, reading delivery systems, and access to quality reading materials. In addition, Basa works closely with DepEd to support and strengthen the literacy component of its K–12 Integrated Language Arts Curriculum for Grades 1 – 3.

Basa's core approach and theory of change aims to improve early grade reading performance in the Philippines through: 1) improving teacher supervision and in-service support for teaching reading of DepEd school heads and administrators; 2) training early grade teachers in effective teaching techniques; and, 3) producing innovative and appropriate teaching and learning materials (such as Teacher's Guides, Read Alouds, and Leveled Readers).

Figure 1: Basa Theory of Change



In 2012, DepEd adopted a K-12 Curriculum to be phased in over a six-year span. The new Grade 1 curriculum was implemented nationwide in 2012/13, prior to the start of the Basa project. Grade 2 was implemented in 2013/14, and Grade 3 was rolled out nationwide in 2014/15. Basa's training focuses on enhancing the literacy component of the curriculum, with a particular emphasis on training teachers in techniques for bridging across languages. As per the DepEd policy of Mother Tongue-based Multilingual Education (MTB-MLE), Mother Tongue is the language of instruction in school for Grades 1 to 3 while Filipino and English are introduced gradually. The Basa project works in Regions 1 and 7, where the Mother Tongues of instruction are Ilokano and Cebuano, respectively.

In 2013, Basa provided initial training in enhanced literacy instruction techniques to schools in Cebu (Region 7) and La Union (Region 1). In 2014, the project expanded to include the divisions of Bohol and Mandaue City in Region 7 as well as Ilocos Norte and Ilocos Sur in Region 1. All schools in these divisions were provided with a complete set of Grades 1 and 2 literacy teaching and learning materials, including Teacher Guides (TGs), Read Alouds and Leveled Readers, as well as training for school heads and teachers. In 2015, the project also added the city divisions of Tagilaran and San Fernando. The project also developed and provided teaching and learning materials and training for Grade 3 teachers in all Basa divisions.

STRUCTURE OF REPORT

As part of its scheduled monitoring and evaluation (M&E) activities,[1] Basa is conducting outcome evaluations to measure changes at the school level with Grade 2 and 3 teachers and students, as well as principals, participating in the Basa intervention. The results of the evaluation activities will be used to not only inform the project technical and management teams (program management) but also, to measure outcomes (change in learner and teacher performance as well as principal viewpoints).

The purpose of this report is to provide data on improvements in reading performance of Grade 2 and Grade 3 learners in Basa targeted regions, as well as improvements in early grade reading instruction as a result of the Basa Pilipinas project. To do this, we report information from the outcome evaluation study that includes Early Grade Reading Assessments (EGRAs) in Filipino and English, classroom observations, teacher, principal and student interviews, as well as results from, an additional exploratory study, a Best Practice Study (BPS), that explores best practices in school leadership and early grade reading instruction in top and bottom performing schools in the Basa program. The report focuses on these key research questions:

1. After two years of the Basa intervention, do students, both male and female, demonstrate improved reading and comprehension skills in Filipino at the end of Grade 2?
2. Do students, both male and female, demonstrate improvement in reading and comprehension skills in Filipino and English from Grade 2 to Grade 3?
3. Have teachers' instructional practices in reading (in the Filipino language) improved in target schools?
4. What are the best practices/factors in early grade reading instruction that are associated with better early grade reading performance of Grade 2 students in Filipino and English?

This report presents data that address each of the Research Questions and draws conclusions and recommendations based on those data. The report structure is aligned with the evaluation research questions. The report starts with a brief overview of the evaluation design and methodology. A more detailed description of the methodology, data collection and tools can be found in Annex 1.

Chapters 1 and 2 provide an overview of the context of early grade reading in study schools in terms of the classroom environment, school leadership, learning environment as well as findings from the student interview. Following that, in Chapter 3, Grade 2 EGRA results in Filipino and English are presented. Reading performance of Grade 2 learners before the Basa intervention are compared to results after two years of project implementation in order to measure changes in Grade 2 student achievement associated with the Basa intervention. Chapter 4 examines the growth in reading performance from Grade 2 to Grade 3 in Filipino and English in sample schools. The next chapter, explores the impact of contextual factors such as learner characteristics, home language, and the classroom environment on reading performance in Filipino and English. Chapter 6 details results on improvements in teachers' instructional practices in reading (in the Filipino language) in target schools. Lastly, results from a Best Practices study are presented in Chapter 7, which explores various factors impacting learner achievement that may contribute to schools falling in the top or bottom performing schools in terms of Filipino and English EGRA results.

[1] For the evaluation design and details see the project's M&E (Performance) Plan with Contract Monitoring Plan, April 20th 2013.

Data reported in this document were collected in Bohol, Cebu, Ilocos Norte/Sur and La Union—the focal divisions of the Basa intervention in 2014/15 and 2015/16. The outcome evaluation study was conducted to measure changes in Grade 2 student achievement after two years of the Basa intervention and compare student achievement from Grade 2 to Grade 3. In order to capture changes in literacy outcomes for learners over the course of the Basa intervention, data were collected at three points of time:

- *Cohort 1*: Before the Basa intervention (SY 13/14);
- *Cohort 2*: after one year of Basa intervention (SY 14/15); and
- *Cohort 3*: after two years of Basa intervention (SY 15/16).

Figure 2. Basa Pilipinas Geographic Scope



Cohort 1 second grade students (SY 13/14) were tested only in Filipino at the end of the school year. Cohort 2 (SY 14/15) (second grade) and Cohort 3 (SY 15/16) (second and third grade) students were tested in both Filipino and English at the end of the school year. Data on the quality of classroom instruction were collected using EDC-developed *Standards-Based Classroom Observation Protocol for Educations – Literacy* (SCOPE-Literacy). Additionally, for Cohort 3 (SY 15/16) principals and Grade 2 and Grade 3 teachers were interviewed using a *Classroom Monitoring Checklist* tool, developed by Education Development Center (EDC).

This report presents findings from the student assessment for all 3 cohorts (see Annex 3 for detailed EGRA results for all cohorts), however, given that this is a final outcome evaluation report, the majority of the analysis focuses on gains from Cohort 1, before Basa, to Cohort 3, two years of Basa intervention. Additionally, teacher observations and principal and teacher interviews for Cohort (SY 15/16) are also reported. All survey and observation data were collected in English.

When interpreting the results of student assessment, it is important to recognize that Grade 2 students had different amount of formal instruction in each language. Filipino and English are introduced as subjects during the second and third quarters of Grade 1, with a focus on oral language development. Each language lesson is allotted 30 minutes daily. During Grade 2, students begin reading and writing in Filipino in the first quarter and in English in the third quarter. Each language lesson is allotted 50 minutes daily. In Grade 3, pupils are instructed in reading in all three languages although, Mother Tongue remains the main language of instruction for all other subjects. In Grade 4, Mother Tongue is phased out as both a subject and language of instruction and students are taught all subjects in Filipino, except math and science which are taught in English.

Therefore, by the end of the school year, second graders will have had seven quarters of instruction in the development of oral language in Filipino and four quarters in reading and writing Filipino. They will have received less instruction in English; six quarters in developing oral language skills in English and just two quarters of instruction in reading and writing English.

Table 1. Introduction of Filipino and English by Grade Level and Quarter

	Grade 1				Grade 2				Grade 3			
	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
Mother Tongue	As language and medium of instruction for all other subjects											
Filipino Language		oral			reading/writing							
English Language			oral				reading/writing					

EVALUATION DESIGN

The purpose of this evaluation study is to: 1) measure changes in Grade 2 student achievement associated with the Basa intervention; 2) measure growth in reading performance from Grade 2 to Grade 3 in Filipino and English; 3) measure improvement in teachers' instructional practices in reading over the course of the Basa intervention; and 4) to examine various factors that are associated with higher early grade reading performance in Filipino and English in study schools. Specifically, the evaluation study was designed to answer the following questions:

1. After two years of the Basa intervention, do students, both male and female, demonstrate improved reading and comprehension skills in Filipino at the end of Grade 2?
2. Do students, both male and female, demonstrate improvement in reading and comprehension skills in Filipino and English from Grade 2 to Grade 3?
3. Have teachers' instructional practices in reading (in the Filipino language) improved in target schools?
4. What are the best practices in early grade reading instruction that are associated with better early grade reading performance of Grade 2 students in Filipino and English?

METHODOLOGY

To answer the *first and second questions* on student progress, the evaluation followed a quasi-experimental, cross-sectional design. Data was collected at three points of time:

- *Cohort 1*: Before the Basa intervention (SY 13/14);
- *Cohort 2*: after one year of Basa intervention (SY 14/15); and
- *Cohort 3*: after two years of Basa intervention (SY 15/16).

In 2013/14 before the full Basa intervention began, a comparison cohort of Grade 2 students in a sample of schools in Cebu and La Union was assessed in reading (Filipino) at the end of the school year. The subsequent year (SY 2014/15), after one year of the Basa intervention, to provide a more complete picture of the Basa outcomes, during school year 2014/15 the evaluation was expanded to additional schools in Cebu and La Union, as well as to schools in Ilocos Norte, Ilocos Sur and Bohol. During SY 2015/16, after two years of Basa intervention, a random sample of Grade 2 and Grade 3 students were assessed in the same sample of schools assessed during SY 2013/14 and SY 2014/15.



Given that assessment results between Cohort 1 (SY 13/14) and Cohort 2 (SY 14/15) were compared in the 2014/2015 evaluation report, to answer the first evaluation question, this report will focus on assessing improvement in Filipino reading skills of Grade 2 students after two years of Basa intervention, comparing results from Cohort 1 (SY 13/14), before the Basa intervention, and Cohort 3, after two years of Basa. To answer the second evaluation question, Cohort 3 Filipino and English results are compared across grade to assess growth in reading performance between Grade 2 and 3.

Table 2. EGRA Assessment Schedule

	Grade	Assessment	Feb 2014	Feb 2015	Feb 2016
Cohort 1 (SY 13/14)	Grade 2	Filipino EGRA	X		
Cohort 2 (SY 14/15)	Grade 2	Filipino & English EGRA		X	
Cohort 3 (SY 15/16)	Grade 2 & Grade 3	Filipino & English EGRA			X

To answer evaluation *question #3*, regarding changes in instructional practices among teachers trained by Basa, *SCOPE* data was collected. Forty teachers in Cebu and La Union were tracked and observed each year (SY 13/14, SY 14/15, and SY 15/16., to the extent possible. For those teachers who could not be tracked, due to attrition, replacements were selected. For the purpose of analysis, this report provides findings from November/
December 2013 and December 2015 for only longitudinally matched teachers (28 teachers).

Lastly, to answer the *final evaluation question*, in February and March 2016 a Best Practices Study (BPS) was implemented to examine factors potentially impacting learner achievement in Grade 2 Filipino and English reading performance. The BPS was conducted in four top performing and six bottom performing schools identified using literacy assessment results in 2015 by Basa. The study aims to explore various factors that may contribute to schools falling in the top or bottom performing schools in terms of Filipino and English EGRA results.

The figure below shows the timeline of evaluation activities for data presented in this report.

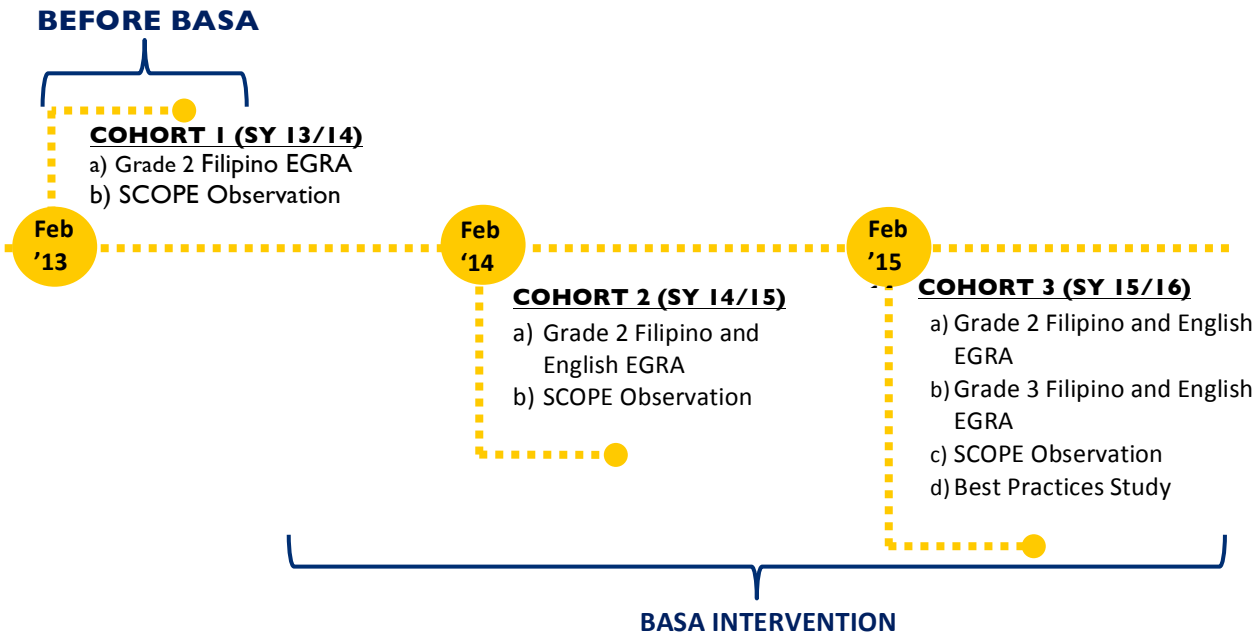


Figure 3. Timeline of Basa Evaluation Activities Across Cohort

SAMPLE

The student sample for Cohort 1 (SY 2013/14) was drawn from Cebu and La Union. In SY 2014/15 and SY 2015/16, the sample was drawn from Bohol, Cebu, Ilocos Norte/Sur, and La Union. The table below details the sample used in this report.

Table 3. Student Sample

	Divisions	# of schools	# of Grade 2 students	# of Grade 3 students
Cohort 1 (SY 2013/14)	Cebu & La Union	40	469	--
Cohort 2 (SY 2014/15)	Bohol, Cebu, Ilocos Norte, Ilocos Sur & La Union	80	1,216	--
Cohort 3 (SY 2015/16)	Bohol, Cebu, Ilocos Norte, Ilocos Sur & La Union	118	1,658	1,597

Sex and Age of Learners. The sample was designed to select an identical number of boys and girls in each grade, in each school. The final distribution by sex was nearly perfect across cohort and grade.

The age range of students was consistent across the three cohorts of Grade 2 students. Although all students were selected from the second grade, students ranged in age from 4 to 11 years in Cohort 3 (SY 2015/16), 6 to 10 years in Cohort 2 (SY 2014/15) and 5 to 12 in Cohort 1 (2013/14). On average, mean ages of students across cohort samples were nearly the same with a median age of 7. For Grade 3 students in Cohort 3 (SY 2015/16) learners' ages ranged from 7 to 11 years old, with a median age of 8.

To enable the computation of estimates of literacy skills among students in *all* schools affected by the Basa intervention, design weights were applied to the analyses of EGRA data. Design weights were applied to compensate for differences in provincial sampling and to ensure an appropriate representation of learners in all provinces in the sample. Consequently, actual *n*'s are only reported in this section and Chapters 1 and 2 that focus on background school, teacher and learner-level context findings; in subsequent sections *n*'s will not be reported and weighted data will be used.

In SY 2015/16, in nearly every school sampled, the principal, and one Grade 2 and Grade 3 teacher were also interviewed. In some schools, principals and/or teachers were unavailable for interviewing. The table below shows the final sample of surveyed principal and teachers. Results from the principal and teacher interviews can be found in the School and Learning Environment Findings section.

Table 4. Principal and Teacher Sample, Cohort 3 (SY 15/16)

	Divisions	# of schools	# of Principals surveyed	# of Grade 2 teachers surveyed	# of Grade 3 teachers surveyed
Cohort 3 (SY 2015/16)	Bohol, Cebu, Ilocos Norte, Ilocos Sur & La Union	118	96	114	117

Annex 1 includes a detailed description of methods and data collection tools, and Annex 3 includes detailed assessment results.

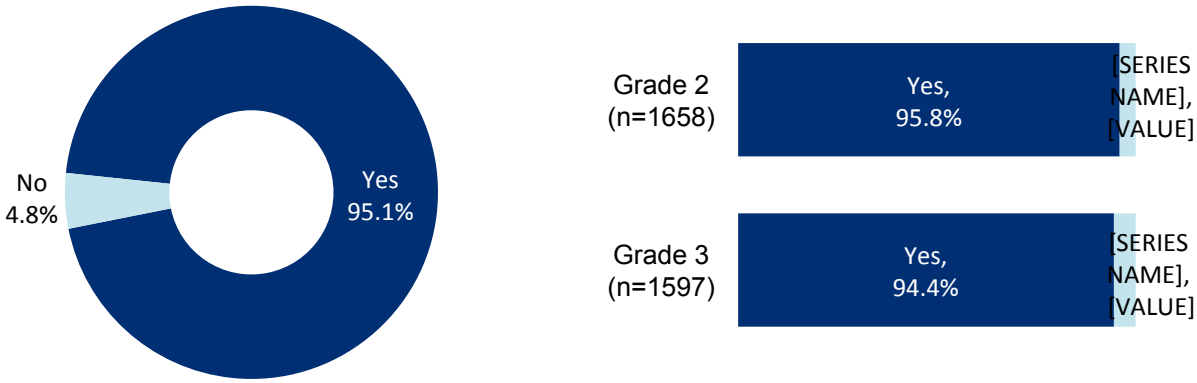
1. STUDENT CONTEXT INTERVIEW FINDINGS

It is widely recognized in the field of education that contextual factors, such as supportive home environments, adequate nutrition, and early exposure to literacy, play prominent roles in helping children succeed academically. Additionally, school factors such as teachers assigning homework or teachers reading to children have also been found to be associated with improved performance. To assess these contextual factors, students were asked a series of questions about their home environment, student/teacher practices and their socioeconomic status. Below are results from the student context interview for Grade 2 and Grade 3 learners from Cohort 3 (SY 2015/16). Most of the data collected for Grade 2 and Grade 3 students were very similar. In the instances where the results differed, they were reported separately. In the instances where the results were similar, they were presented in aggregate.

SCHOOL AND TEACHER ENVIRONMENT

Nearly all students reported that they've been attending their school since the beginning of the school year. Approximately 95% of students overall reported attending kindergarten before starting Grade 1. The large majority of Grade 2 and Grade 3 students reported attending kindergarten (95.1%).

Figure 4. Kindergarten Attendance Before Grade 1 (n=3,255)



The majority (88.2%) of Grade 2 and 3 students reported being allowed to take books home from school., which is largely consistent with teacher interview data. About 87.1% of students reported taking books home from school. In addition, nearly nine out of ten (88.5%) Grade 2 and Grade 3 students reported being able to choose the story books that they read at school. Results were largely consistent among boys and girls and across provinces.



Sampled Grade 2 and Grade 3 students were asked about what they liked the most about school. Common responses among Grade 2 and Grade 3 students were: books in the classroom and reading and writing. Learners were also asked about what they did not like about school. More than 40% of Grade 2 and Grade 3

students replied that there was nothing that they did not like about school. Of those surveyed learners who reported having a dislike about school, learners frequently mentioned that they did not like disruptive, unruly classmates, fighting, and bullying/teasing by their fellow classmates. Playing at school was cited as a reason to like school while playtime at school was mentioned as a dislike; it is unclear why the apparent contradiction but may indicate that bullying or quarreling can occur at playtime. Table 5 below shows the most frequent responses from Grade 2 and 3 learners.

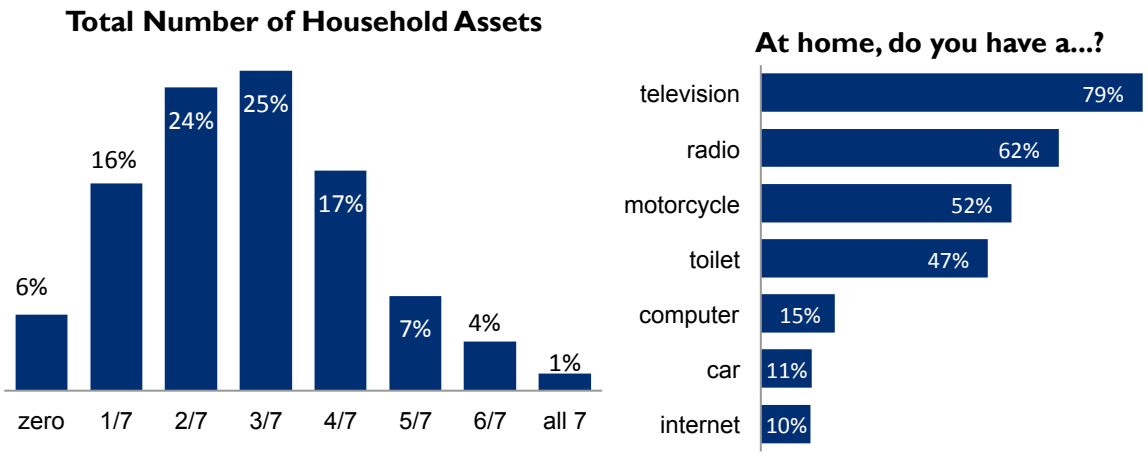
Table 5. Most Common Reasons Grade 2 and Grade 3 Students Like and Dislike School

What do you like most about school?	What do you NOT like about school?
• Books in the classroom • Reading and writing • Learning and studying • Teacher and Classmates • Playing	• Disruptive, unruly classmates • Quarrelling amongst classmates • Bullying/teasing by fellow classmates • Playtime at school

SOCIO-ECONOMIC FACTORS

Reports of common higher priced household items are commonly used as a proxy for household income as well as overall socio-economic status. On average, students said that their families had between two or three household possessions out of the seven possessions listed in the survey. A television, a radio, an indoor toilet and a motorcycle were the most commonly reported household possessions, with almost half or more sampled students in both Grades 2 and 3 reporting having these items in their household.

Figure 5. Household Assets (n=3,255)



Provincial comparisons of socio-economic factors show some variation across provinces. Students in Ilocos Norte, Ilocos Sur, and La Union reported having the most household items with the majority of students reporting having between three to four of the seven household possessions while the pupils in Cebu and Bohol reported between two and three.

Students were also asked whether they had something to eat before they came to school on the day of the reading assessment. Almost all (98.4%) Grade 2 and 3 students reported having had something to eat the day of the survey.

HOME ENVIRONMENT



Students reported using a variety of different languages at home, which was largely consistent among Grade 2 and Grade 3 sampled learners.[2] For each grade, almost half of sampled learners responded that they speak Bisaya/Cebuano at home; 49.0% in Grade 2 and 49.5% in Grade 3. Ilokano was also reported to be widely spoken at home by 44.7% of the Grade 2 students and 45.6% of the Grade 3 students. English was not as widely used at home with 5.4% of Grade 2 and 3.3% of Grade 3 learners reporting speaking it at home. Filipino/Tagalong was reported as being spoken at home by roughly 1 in 5 students in both Grade 2 and 3. Note that Filipino, which is originally based on the Tagalog dialect, is the national language of the Philippines and is introduced gradually beginning in the second quarter of Grade 1, with students' Mother Tongues forming the basis of instruction up through Grade 3. Reading and writing in English is introduced as a subject in the third quarter of Grade 2. Later, in Grade 4, Mother Tongue is phased out as both a subject and language of instruction and students are taught all subjects in Filipino, except math and science which are taught in English.

Table 6. What language do you speak at home? (n=3,255)

Language	Grade 2	Grade 3
Filipino/Tagalog	22.9%	21.0%
English	5.4%	3.3%
Bisaya/Cebuano	49.0%	49.5%
Ilokano	44.7%	45.6%
Other languages	0.9%	0.5%

Of those students who responded, nearly half (47.2%) of surveyed students reported that their mothers were engaged in formal employment, of which the majority (72.4%) of working mothers make a living in the informal economy. Roughly nine out of ten fathers were reportedly employed, largely in the informal economy; only 9.2% were reported as unemployed. Results were consistent across grade. These results must be interpreted with caution since it is likely that children are not always aware of the occupation of their parents.

Table 7. Where do your parents work? [3]

Parental Occupation	Mother (n=3,120)	Father (n=2,993)
Overseas Foreign Worker	6.4%	5.3%
Professional	6.7%	6.2%
Informal/Menial/Self	34.1%	79.3%

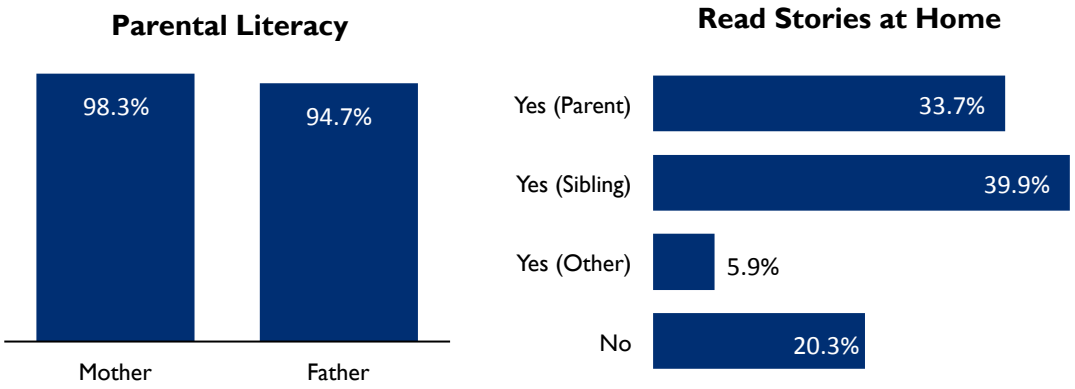
[2] Note, given that learners may speak more than one language at home, respondents were allowed to report multiple responses. As such, the table below does not add up to 100%.
[3] Students who responded "Do not Know" or "No response" are excluded from the percentages presented in these paragraphs. Percentages only include students who responded to the question.

Unemployed	52.8%	9.2%

Parental involvement is a key predictor in early literacy success as well as future academic achievement of children. As such, the student context interview also aimed to find whether students receive any help with reading at home. The large majority of students in both Grades 2 and 3 reported that both their parents were literate though slightly more mothers than fathers were reported as literate.

The majority of students said they receive help at home with reading, either from a parent or from a sibling. Just over 20% of surveyed students said they do not receive help at home with reading. Students also largely (73.2%) reported having books at home.

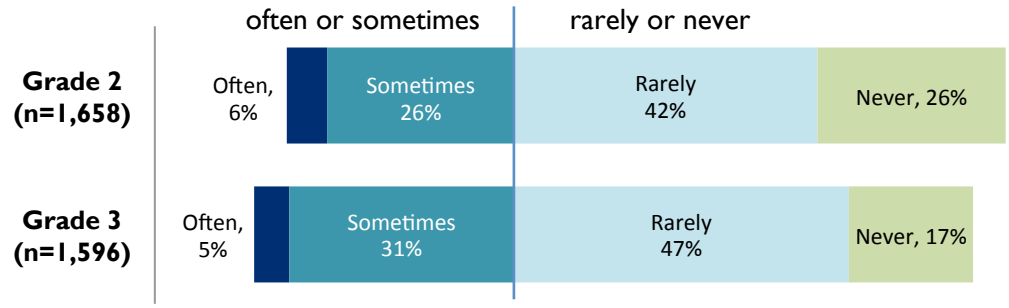
Figure 6. Parental Literacy and Help with Reading at Home (n=3,247)



The large majority of students (85.5%) reported having someone at home check their assignments/ homework. There were no statistically significant differences between grades or sex. Approximately half of learners (48.1%) reported watching TV in English.

About two-thirds (66.2%) of the students in both Grade 2 and 3 reported missing school either rarely or never. There were some differences between the two grades, with more students in Grade 3 reporting to miss school often or sometimes (36.2%) compared to students in Grade 2 (31.6%), and this difference is statistically significant at $p<.05$.

Figure 7. How often do you miss school? (n=3,254)



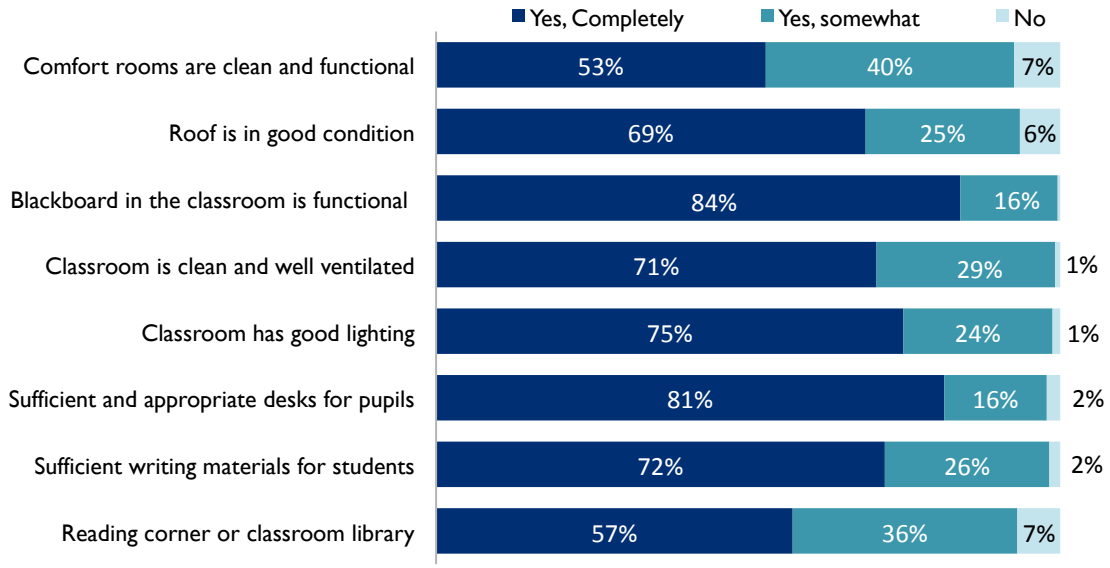
2. SCHOOL AND LEARNING ENVIRONMENT FINDINGS

The school environment and management is critical to understanding the teaching and learning that is taking place in the school. Concurrently with the learner assessment, during SY 2015/16, the data collection team gathered data on the school context and environment, and grade-level resources and practices related to the Basa intervention. In all, 96 principals and 231 Grade 2 and Grade 3 teachers were surveyed in 118 sample schools. Data was collected to provide an overall picture of the school and learning environment; particularly data was collected on: 1) classroom infrastructure and environment, 2) school management and support to teachers, and 3) the learning environment. This information is used as covariates in EGRA assessment data analysis.

CLASSROOM ENVIRONMENT

Data were collected on the school and classroom infrastructure of 118 sampled schools in SY 2015/16. For nearly every school, one Grade 2 and one Grade 3 classroom, whose pupils were sampled to take the EGRA, was observed and rated by assessors. The assessors observed the classroom infrastructure and environment for eight conditions and used a three-point scale ("Yes, completely," "Yes, somewhat," and "No") to rate the conditions. As seen in Figure 8 below, observations of sampled schools showed varying conditions in the school infrastructure and learning environment. The majority of schools were scored by assessors as having either "completely" functional blackboards, clean, ventilated and well-lit classrooms, and sufficient and appropriate desks and writing materials (notebooks, paper) for learners. Slightly more than half of schools had "completely" adequate comfort rooms for boys and girls, reading corners, and classroom libraries. Very few observed schools had inadequate school and classroom infrastructure conditions.

Figure 8. Observed Condition of School and Classroom Infrastructure in Sampled Schools (n=231)



SCHOOL LEADERSHIP

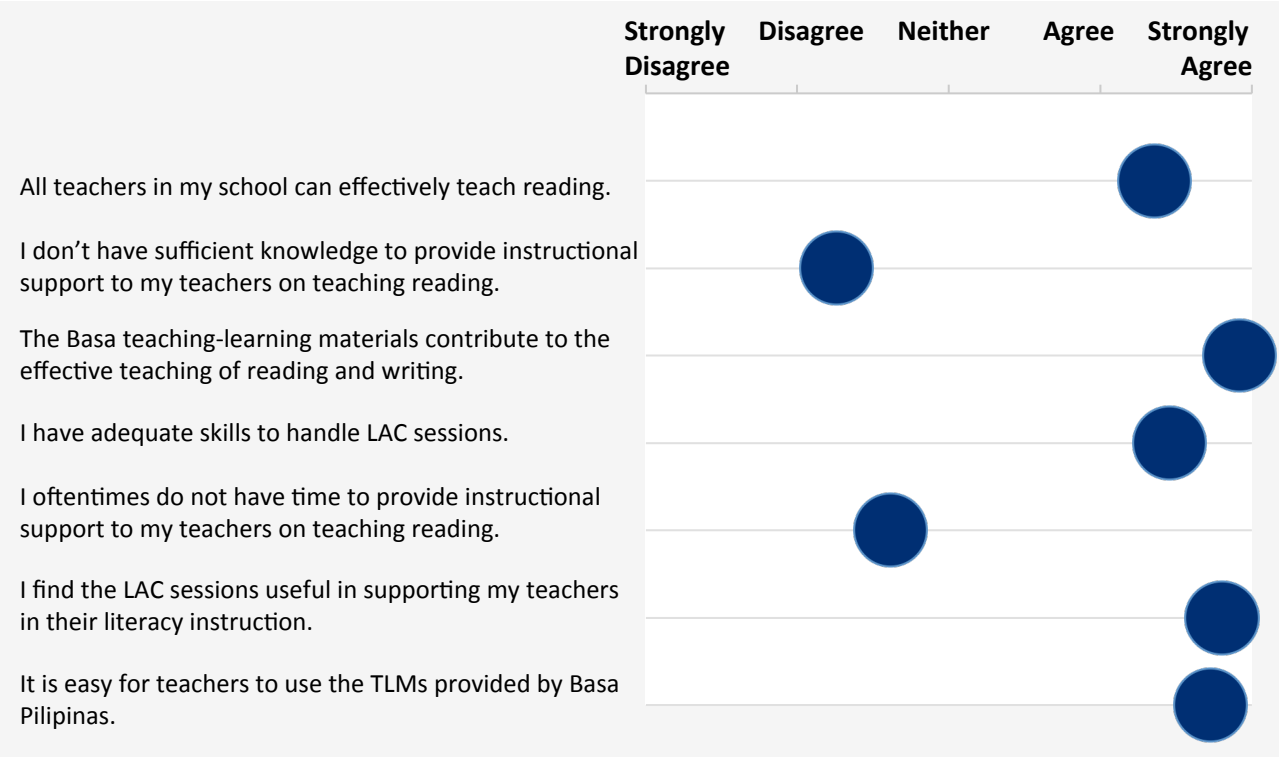
Sampled principals were asked about Basa trainings or orientations that they attended. The most commonly attended trainings were the LAC Training and the School Head Orientation in which 78.9% and 65.3% of sampled principals reported attending each respective training. Very few principals reported attending Teacher Training (13.7%) or Training of Trainers (TOT) training (5.3%). Of the five principals who reported attending a Basa TOT training, three principals reported serving as a facilitator, co-facilitator or lead instructor.



Principals in sample schools were surveyed about their perceptions and attitudes of literacy instruction in their schools. Principals were read seven statements and asked to give their opinion using a five-point scale: Strongly agree, agree, neither agree or disagree, disagree, and strongly disagree. Overall, principals largely had positive perceptions of literacy instruction at their schools. The large majority (93%) of principals reported that they felt that all teachers in their school could effectively teach reading and that it was easy for teachers to use the TLMs provided by Basa Pilipinas (99%). In fact, all principals agreed or strongly agreed that Basa TLMs contributed to the effective teaching of reading and writing in their schools.

Principals largely (72%) felt that they had sufficient knowledge to provide instructional support to their teachers on teaching reading. However, having sufficient amount of time to provide support to teachers on teaching reading, was a concern for a third of principals, however, roughly six out of ten principals felt they were able to devote a sufficient amount of time to support teachers in teaching reading. Principals were also asked about Learning Action Cells (LACs) at their schools. Nearly all (96%) sampled principals indicated that they felt that they had adequate skills to handle LAC sessions. Principals were also very satisfied with LACs, in which all but one sampled principal reported that they either strongly agreed or agreed that the LAC sessions were useful in supporting teachers in their literacy instruction.

Figure 9. Average Principal's Perceptions and Attitudes of Literacy Instruction in Sampled Schools (n=96)



The majority of principals reported that they did not have any Basa-related problems or concerns. For those principals who did raise concerns, a common concern that was raised focused on the need for additional training. Several principals requested additional training on bridging Mother Tongue to Filipino and English lessons to improve reading, as well as refresher trainings for existing and new teachers and the school head. Additionally, some principals requested to receive more teaching and learning materials, either due to the fact that the TLMS were delayed or there was not a sufficient amount for all learners.

LEARNING ENVIRONMENT

ACCESSIBILITY OF TEACHING AND LEARNING MATERIALS

Data collectors observed classrooms for teaching and learning materials. In nearly every observed Grade 2 and Grade 3 classroom, assessors reported observing Basa-revised Teachers Guides (TGs) as well as Basa Leveled Readers in the classroom. All but one observed Grade 2 classroom, had Basa-provided Read Aloud books in the classroom.

In terms of supplemental reading materials in the classroom, roughly three-quarters of classrooms had teacher-bought or produced supplementary materials (not textbooks). Roughly two-thirds (66.1%) of Grade 3 and slightly more than half (56.6%) of observed Grade 2 classrooms had supplementary reading materials provided by Basa, such as, books donated by BBF. Additionally, nearly two-thirds of observed classrooms had DepEd provided supplementary materials. Assessors also noted several classrooms had additional supplementary reading materials which were largely donated from NGOs, the LGU, alumni or parents, private schools, the local school board or international donations. Supplementary materials were largely accessible to pupils on their own, without the need for the teacher to distribute them, in the large majority (94.1%) of classrooms.

Table 8. Percent of Grade 2 and Grade 3 Classrooms with Teaching and Learning Materials (n=231)

Material	Grade 2 (n=113)	Grade 3* (n=118)
Basa-revised Teacher Guides	93.8%	98.3%
Basa Leveled Readers	99.1%	99.2%
Basa-provided Read Aloud Books	99.1%	--
Basa-provided Supplementary Reading Materials (e.g. BBF books)	56.6%	66.1%
Teacher-bought or Produced Supplementary Reading Materials (not textbooks)	77.0%	76.3%
DepEd-provided Supplementary Reading Materials (not textbooks)	60.2%	63.6%

*Grade 3 learners were not provided Read Aloud Books by Basa.

TEACHER PERCEPTIONS OF TEACHING AND LEARNING MATERIALS

Teachers were also asked if they were able to easily follow the revised Teacher Guide (TG) provided by Basa. The majority (76.6%) of sampled teachers felt that the revised TG was in fact easy to follow. Common reasons teachers cited include:

- The Teacher Guide is well sequenced and outlined;
- The Teacher Guide is comprehensive, detailed and complete;
- The Teacher Guide has clear instructions and is easy to follow;
- Activities, stories and learner materials are all provided in the Teacher Guide;
- The Teacher Guide is easy to execute in the classroom.

> *"Pupils can follow the lesson better. The strategy used [in the Basa-revised Teacher Guide] is effective in teaching pupils."*
>
> —Grade 3 Teacher in La Union province

Although the majority of teachers felt that the revised TG was easy to follow, roughly a quarter of teachers reported feeling that the TG was difficult to follow. Teachers commonly cited that they felt that the TG could, at times, be complicated and confusing, especially among teachers who were new or had not been trained by Basa. Other teachers mentioned that they felt that there were too many objectives and that lessons could be too long, and as such they were unable to achieve all objectives in the time allotted. A few teachers indicated that at times the activities were difficult for their students to follow.

Teachers were also asked whether they had any issues or concerns about any of the learning materials (revised TG, Read-aloud or Big Books, Leveled Reader) or if they had any other issues. The most common comments/suggestions from surveyed teachers include:

- Shorten the number of objectives and activities in Teacher Guide lessons;
- Additional read-aloud/Big Books, and Leveled Readers are needed in the classroom;
- Some stories and reading texts are too long for their learners;
- Some stories and reading texts are too difficult for their pupils to understand;
- More colorful Leveled Readers would be more attractive to pupils;
- Additional trainings are needed for teachers, especially new teachers who have not been trained by Basa Pilipinas.

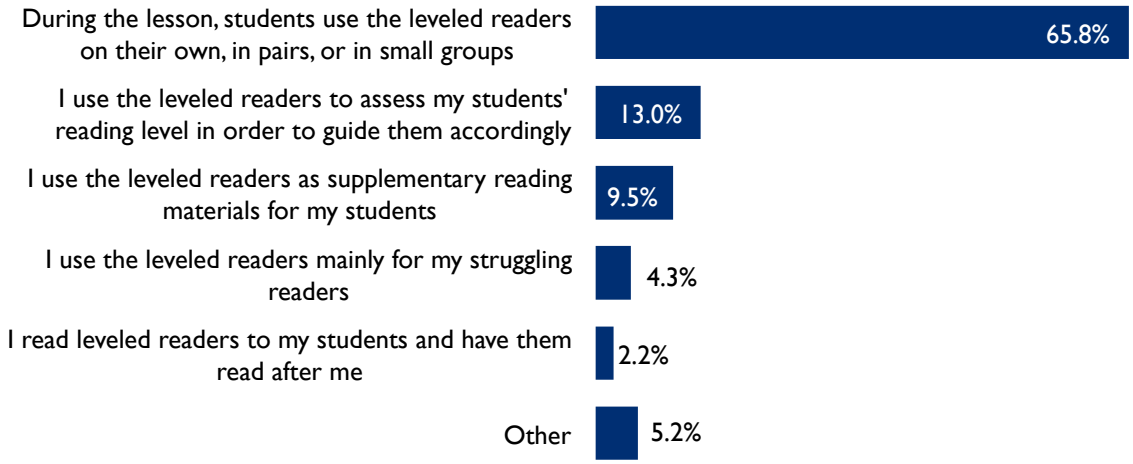
TEACHING PRACTICES

Nearly all sampled teachers (96.1%) reported attending a Basa training. No teachers had attended the online course.

Overall, sampled teachers largely reported using the revised Teacher Guide (TG) provided by Basa when developing their lessons. Roughly a quarter (24.1%) of teachers said that they only use the revised Teacher Guide; two-thirds (61%) said that they mainly use the revised TG provided by Basa, and use the old DepEd TG as reference for supplementary or remedial activities. Given that in Grades 1 and 2, the project references the DepEd guides and Learner Materials, these findings are not contrary to Basa implementation. Only 11.3% of teachers reported "mainly or always" using the DepEd TG for developing lessons.

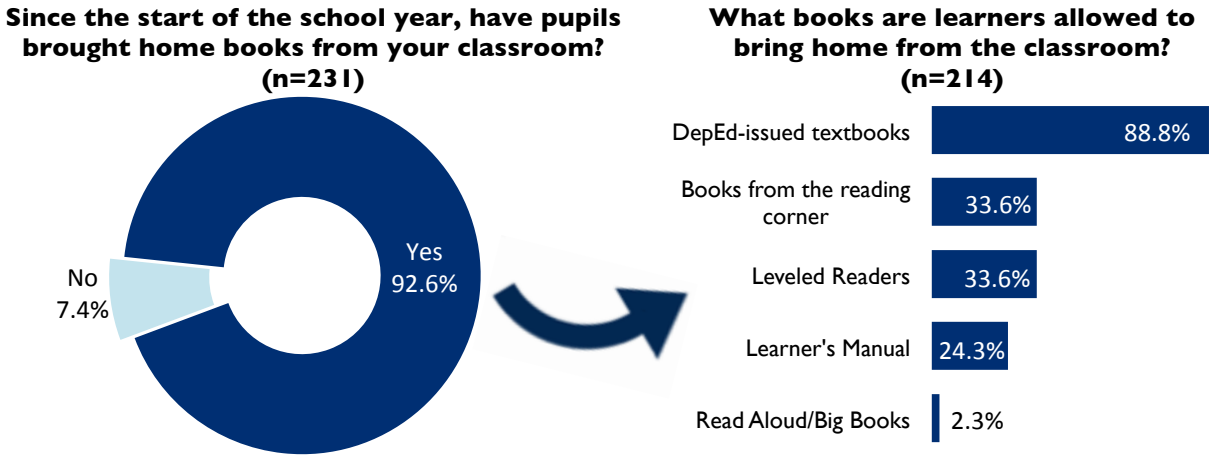
In terms of Leveled Readers, many teachers use Leveled Readers during reading lessons, in which teachers have students use Leveled Readers on their own, in pairs or in small groups. Some teachers reported using Leveled Readers to assess the reading level of their students or as supplementary reading materials. Very few teachers indicated that read Leveled Readers aloud to their students.

Figure 10. How do you usually use the Leveled Readers in your classroom? (n=231)



Overall, the vast majority (92.6%) of Grade 2 and Grade 3 teachers reported that since the beginning of the school year students in their classroom have brought books home from the classroom. The most common type of book, learners were allowed to bring home, was DepEd-issued textbooks, in which eight out of ten teachers reported that students were allowed to bring these books home. About a third of teachers stated that learners could bring home books from the reading corner or leveled readers; roughly a quarter of teachers allowed learners to bring home Learner's Manuals. Allowing students to bring home Read Aloud/Big Books was rare, in which only five teachers reported that they allowed students in their classroom to bring those home. A few teachers also mentioned that they allowed readers to bring home photocopies of Leveled Readers, story books, supplementary reading materials, as well as workbooks.

Figure 11. Can Learners Bring Home Books from the Classroom?



Most teachers tracked books that learners borrow from the class using a log book. The majority (80.4%) of teachers indicated that they personally filled out the log book, noting all books borrowed by learners in in order to keep track of classroom reading resources. A small number of teachers (8.4%) had students sign the books in and out themselves using a log book. A few teachers had less organized approaches to tracking classroom reading materials. For instance, some teachers indicated that they tracked classroom reading materials by memorizing the names and books that learners borrowed; other teaches simply counted the number of books at the end of the school year to see if any were missing. Only nine surveyed teachers said that they did not track or record books check out by learners.

3. CHANGES IN GRADE 2 EGRA RESULTS OVER THE LIFE OF THE BASA PROJECT

To assess changes in reading performance of Grade 2 learners over the course of the Basa intervention, EGRA in Filipino was administered to a random sample of Grade 2 learners before the Basa intervention in SY 2013/14 and then again in SY 2014/15 and SY 2015/16. The Grade 2 English EGRA was administered in SY 2014/15 and again in SY 2015/16. This section provides a summary of Grade 2 Filipino and English EGRA findings as well as detailed sub-test analysis. For the Filipino EGRA, although data is available for Cohort 2 (SY 2014/15), analyses will focus on Cohort 1 (SY 2013/14) and Cohort 3 (SY 2015/16) in order to assess changes in early grade reading performance in Filipino after two years of Basa intervention. Detailed Grade 2 Filipino results can be found in Annex 3. It should be noted that Cohort 1 (SY 2013/14) students were not tested in English. As a result, comparisons could not be made between Cohorts 1 and 3. Analysis can only be made for Cohorts 2 and 3 for the English EGRA.



SUMMARY OF FILIPINO AND ENGLISH EGRA FINDINGS

OVERALL GRADE 2 FILIPINO EGRA RESULTS

Analysis of Grade 2 Filipino EGRA data showed that by the end of Grade 2, learners largely demonstrate initial sound recognition and some word reading, however, have not yet progressed to reading with fluency and comprehension. Overall, results showed that students tended to perform best on the initial sound identification, reading familiar words, and oral reading passage subtests. The subtests with the lowest scores were letter sounds and listening comprehension for the Filipino EGRA. The table below provides the average subtest results for Cohorts 1 and 3 of Grade 2 learners.

Table 9. Filipino Grade 2 EGRA Results, by Subtest and Cohort

Filipino EGRA Subtests	Cohort 1 (SY 2013/14)	Cohort 3 (SY 2015/16)	Gain (Cohort 1 to Cohort 3)	Effect size (Cohort 1, Cohort 3)
Initial Sound Identification (pct correct)	57.6%	72.1%	14.5% (±3.6%)	0.42
Letter Sounds (pct correct)	19.1%	24.1%	5.0% (±1.5%)	0.32
Letter Correct (per min)	19.2	24.3	5.1 (±1.5)	0.31
Familiar Words(pct correct)	68.4%	65.9%	-2.5% (±3.3)	-0.08
Familiar Words Correct (per min)	37.5	35.3	-2.2 (±2.0)	-0.11
Nonsense Words (pct correct)	46.2%	46.8%	0.6% (±2.6%)	0.02
Nonsense Words Correct (per min)	23.2	23.6	0.40 (±1.4)	0.02
Oral Passage Reading (pct correct)	56.1%	58.7%	2.6% (±3.0%)	0.09
Words Correct in a Text (per min)	37.0	39.4	2.4 (±2.1)	0.11
Prosody score	2.1	2.4	0.3 (±0.1)	0.33
Reading Comprehension: timed (pct correct)	28.1%	41.8%	13.7% (±2.9%)	0.43
Reading Comprehension: untimed (pct correct)	-	50.9%	-	-
Listening Comprehension (pct correct)	49.1%	34.7%	-14.4% (±3.7%)	-0.49
Dictation Composite (pct correct)	45.6%	59.7%	14.1% (±2.8%)	0.52

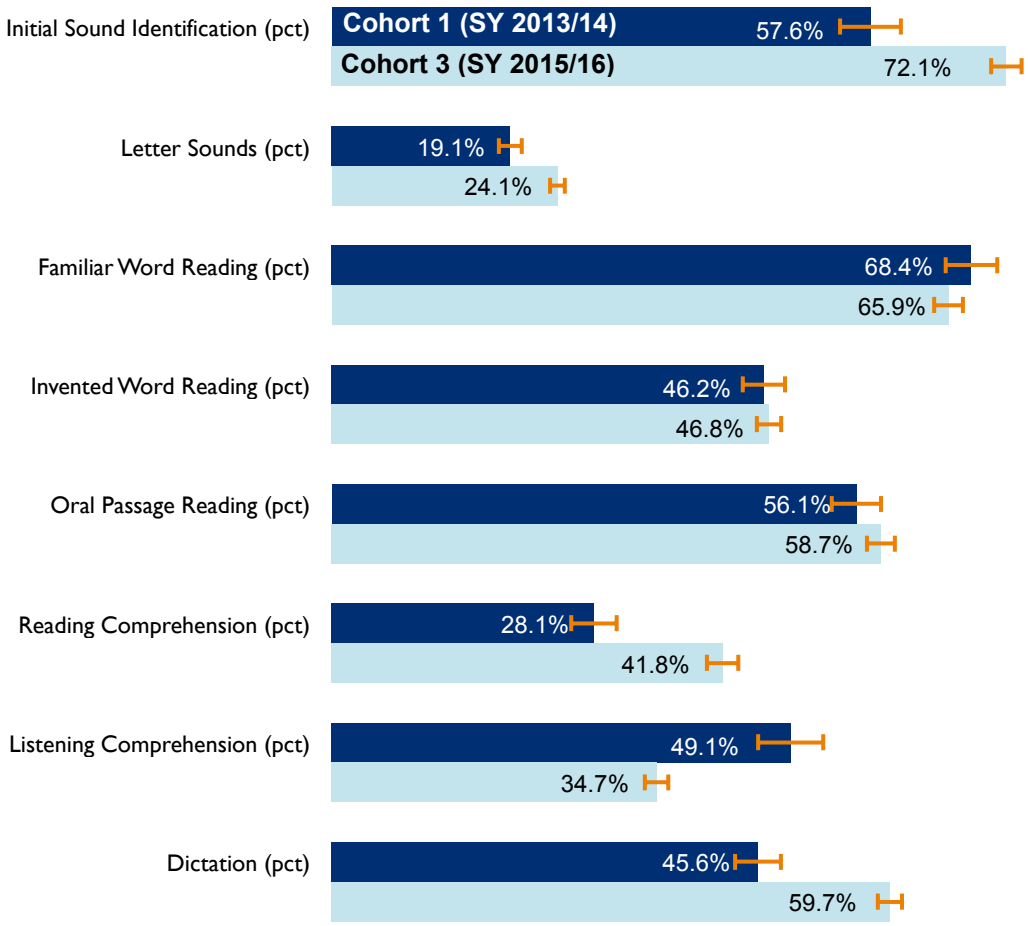
Filipino EGRA results from Cohort 1 (SY 2013/14) and Cohort 3 (SY 2015/16) showed that after two years of Basa intervention, Grade 2 students are performing significantly ($p<.001$) better than students performed prior to the intervention in four out of eight EGRA subtests, with the exception of Familiar Word Reading, Nonsense Word Reading, Oral Passage Reading and Listening Comprehension. Learners in Cohort 1 performed significantly better than Cohort 3 in only one subtest – Listening Comprehension; however, comparisons across cohort should be cautioned given that learners in Cohort 3 were asked to answer five comprehension questions compared to only three for Cohort 1, which could account for differences in learner performance on this subtest between the cohorts. Of particular interest, analysis showed significant improvement in learner performance in two key EGRA measures – Oral Reading Fluency and Reading Comprehension— in which Cohort 3 (SY 2015/16) results showed that Grade 2 students, after two years of Basa intervention, performed significantly better ($p<.05$) than students prior to Basa (SY 2013/14).

Effect size[4] was calculated between Filipino EGRA scores before the Basa intervention (SY 2013/14) and after two years of Basa intervention (SY 2015/16). Six out of thirteen measures showed an effect size of .3 or higher, with the dictation subtest showing the largest effect size of d=0.52. The effect size calculations are reported in Table 9.

[4] Effect size is a statistical measure that is used to estimate the magnitude of difference between two measures. Effect sizes are largely resistant to sample size influence, and thus provide a truer measure of the magnitude of effect. Effect size was computed by dividing the differences between the means of the two groups by the pooled standard deviation. Effect sizes are interpreted as follows, according to Cohen (1998): "small, d = .2," "medium, d = .5," and "large, d = .8". (reference: Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Earlbaum Associates.)

Figure 12 shows average EGRA subtest scores for Grade 2 students before the Basa intervention (SY 2013/14) compared to scores after two years of the Basa intervention (SY 2015/16). As seen in the figure, Grade 2 learners showed the largest improvements in Initial Sound Identification, Reading Comprehension and Filipino dictation, in which students in Cohort 3 (SY 15/16) scored between 13.8 to 14.5 percentage points higher than Grade 2 learners in Cohort 1 (SY 13/14) on these subtests.

Figure 12. Average Percent Correct for Grade 2 Filipino EGRA Subtests, by Cohort



Detailed results of Grade 2 Filipino EGRA by cohort can be found in the next section, *Grade 2 Filipino EGRA Subtest Results*, as well as in the Annex 3.

OVERALL GRADE 2 ENGLISH EGRA RESULTS

An English EGRA assessment was conducted with a sample of students in Basa intervention schools during 2014/15, the first year the project provided the full intervention, as well as during SY 2015/16, after two years of Basa intervention. The English assessment was not administered to Cohort 1 (SY 2013/14) students. As a result, the report provides results for intervention cohort students only (Cohorts 2 and 3), results before the Basa intervention (Cohort 1) are not available.

Analysis of the English EGRA data for Cohort 2 (SY 2014/15) and Cohort 3 (SY 2015/16) showed higher achievement in some EGRA subtests than others. Overall, Grade 2 students tended to perform best on initial sound identification and oral reading passage subtests. The subtests with the lowest scores were reading comprehension, listening comprehension and dictation. The table below shows the results of the English EGRA.

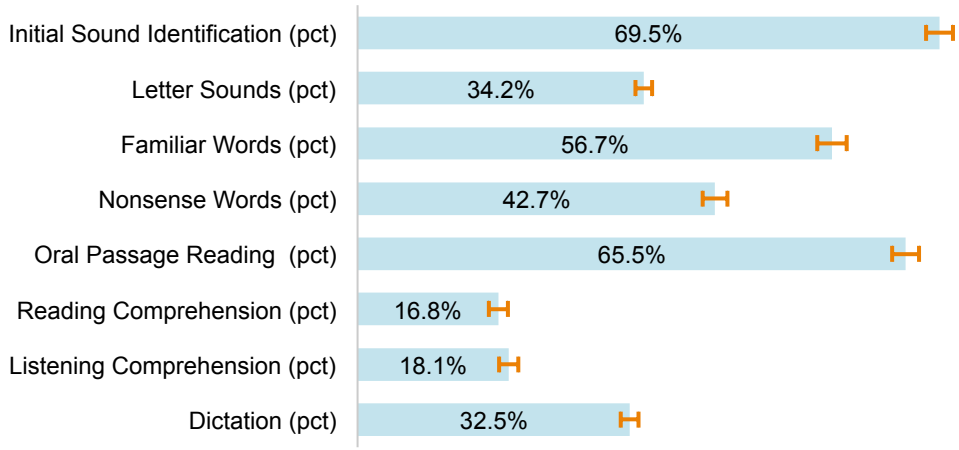
Table 10. Grade 2 English EGRA Results, Intervention Cohort Sample

English EGRA Subtests	Cohort 1 (SY 2013/14)5	Cohort 2 (SY 2014/15)	Cohort 3 (SY 2015/16)	Change Cohort 2 to Cohort 3	Effect Size
Initial Sound Identification (pct correct)	--	73.1%	69.5%	-3.6% (±2.4)	-0.11
Letter Sounds (pct correct)	--	39.1%	34.2%	-4.9% (±1.5)	-0.25
Letter Correct (per min)	--	39.1	34.2	-4.8 (±1.5)	-0.24
Familiar Words (pct correct)	--	63.0%	56.7%	-6.3%(±2.7)	-0.17
Familiar Words Correct(per min)	--	39.5	34.0	-5.5 (±2.0)	-0.21
Nonsense Words (pct correct)	--	49.2%	42.7%	-6.5% (±2.2)	-0.21
Nonsense Words Correct (per min)	--	25.9	21.9	-3.9 (±1.2)	-0.23
Oral Passage Reading (pct correct)	--	72.4%	65.5%	-6.9% (±2.4)	-0.22
Words Correct in a Text (per min)	--	58.1	40.4	-17.8 (±2.0)	-0.65
Prosody score	--	2.8	2.3	-0.5 (±0.1)	-0.51
Reading Comprehension: timed (pct correct)	--	15.9%	16.8%	0.9% (±1.7)	0.04
Reading Comprehension: untimed (pct correct)	--	27.5%	27.9%	0.5% (±2.1)	0.02
Listening Comprehension (pct correct)	--	17.6%	18.1%	0.5% (±1.8)	0.02
Dictation Composite (pct correct)	--	36.2%	32.5%	-3.7% (±1.6)	-0.17

Comparisons of student scores from SY 2014/15 to SY 2015/16, show that for the English EGRA, Grade 2 students in SY 2015/16 performing slightly worse than students in SY 2014/15 in all subtests except reading and listening comprehension. However, as seen in the Table 10 above, effect size difference calculations between SY 2014/15 and SY 2015/16 are relatively small (d=0.3 or smaller) for nearly all subtests, with the exception of Fluency and prosody. These small effect sizes suggest that the difference in English EGRA scores between Cohort 2 and Cohort 3 are relatively small and have largely remained consistent over the two years of the Basa intervention.

Figure 13 below shows visually, the Grade 2 English EGRA results for Cohort 3 students during the SY 2015/16.

Figure 13. Average Percent Correct on Grade 2 English EGRA Subtests for Cohort 3 (SY 2015/16)



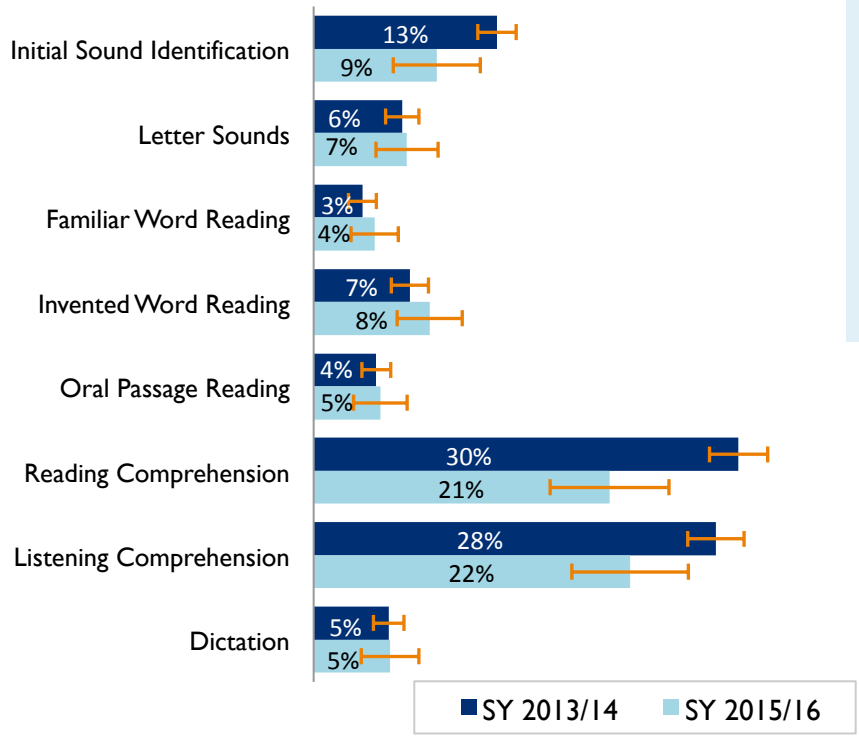
[5] English EGRA results were not collected for Cohort 1 (SY 2013/14).

Detailed English EGRA subtest for Cohort 2 and 3 can be found in the Annex 3. Additionally, English EGRA subtest results for Cohort 3 can be found discussed in more detail in section, *Filipino and English EGRA Findings – Trajectory from Grade 2 to Grade 3*.

ZERO SCORES

The analysis of Grade 2 Filipino assessment results found statistically significant ($p<.01$) reductions in zero scores after two years of Basa intervention in three out of eight Filipino EGRA subtests— Initial Sound Identification, Reading Comprehension and Listening Comprehension. For the Letter Sounds, Familiar Words, Invented Words, Oral Passage Reading, and Dictation subtasks, zero scores were largely unchanged. Figure 14 shows the percent of tested Grade 2 students scoring zero on Filipino EGRA subtests.

Figure 14. Percent of Tested Grade 2 Students Scoring Zero on Filipino EGRA Subtests

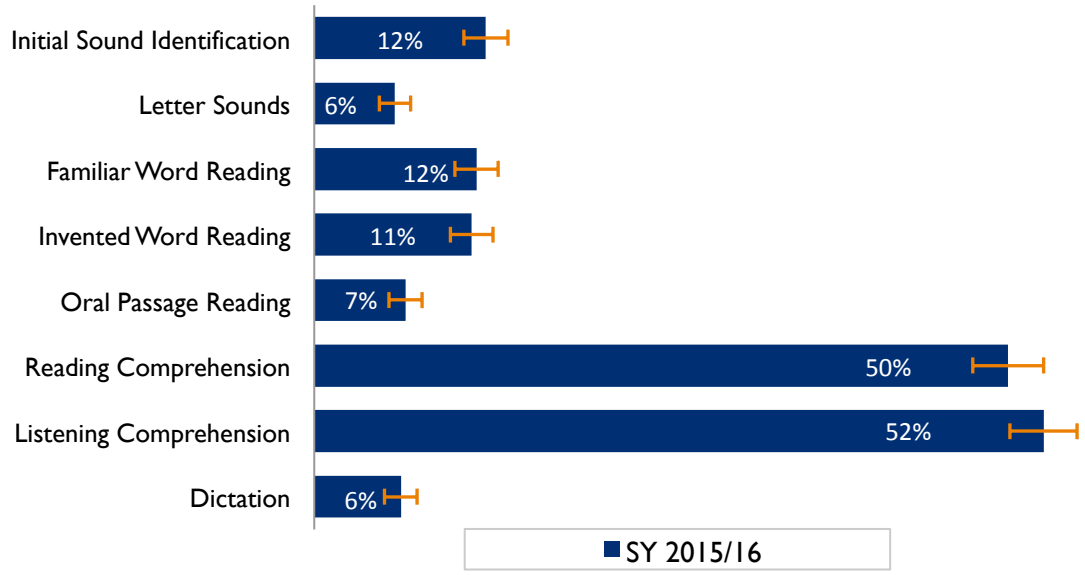


Zero scores on Grade 2 Filipino EGRA subtests ***decreased in 3 out of 8 subtasks****, with the largest reduction in zero scores in reading comprehension. For the Letter Sounds, Familiar Words, Invented Words, Oral Passage Reading, and Dictation subtasks, zero scores were largely unchanged.*

The details of the statistical analyses are found in the Annex 3, which shows the proportion of students with zero scores on each subtest for each study cohort.

Analysis of zero scores on the Grade 2 English EGRA for Cohort 3 (SY 2015/16) showed relatively low percentages of zero scores on English EGRA subtests with the exception of listening and reading comprehension. As seen in the figure below, roughly half of tested Grade 2 students were unable to answer a single reading or listening comprehension correctly.

Figure 15. Percent of Tested Students in SY 2015/16 Scoring Zero on Grade 2 English EGRA Subtests



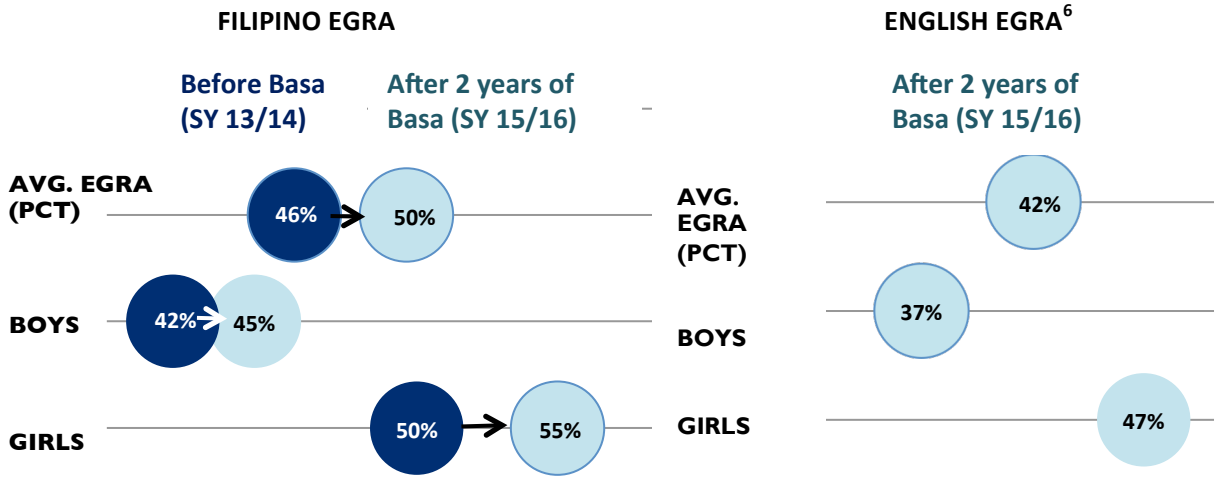
RESULTS BY SEX



Data analysis found that girls, on average, continue to demonstrate far better EGRA results than boys on both the Filipino and English EGRA. The figure below shows the average percent correct across subtests for both Filipino and English by sex. As seen in the figure, in SY 2015/16, after two years of Basa intervention, girls are outperforming boys by ten percentage points on both the Grade 2 Filipino and English EGRA.

Detailed analysis of EGRA subtest results for Cohort 3 (SY 15/16) by sex showed a similar trend. Girls, on average, outperform boys by about two to seventeen percentage points (see Table A-21 for sex-disaggregated Filipino EGRA results, and Table A-25 for sex-disaggregated English EGRA results in Annex 3). The difference in reading performance between boys and girls is statistically significant at the $p<.001$ level for nearly all Filipino and English subtests.

Figure 16. Average Percent Correct on Filipino and English Grade 2 EGRA Subtests, by Sex



Given that results show that girls continue to outperform boys in SY 2015/16, an important facet to explore is whether the difference in girls' and boys' EGRA performance is getting smaller over time or in fact widening. Analysis of Filipino EGRA subtest results show that the gender gap in Filipino has worsened slightly from SY 2013/14 to SY 2015/16. Girls are improving at a particularly faster rate than boys on the Initial Sound Identification and Dictation subtests and as a result, boys are falling even further behind girls on these subtests. One exception to this trend is Reading Comprehension; results show that boys show larger gains than girls in Filipino Reading Comprehension. In fact, boys have nearly caught up to girls in reading comprehension after two years of Basa intervention. Table 11 shows the average effect size for boys and girls on the Grade 2 Filipino EGRA.

Table 11. Average Percent Correct on Grade 2 Filipino EGRA, by Sex

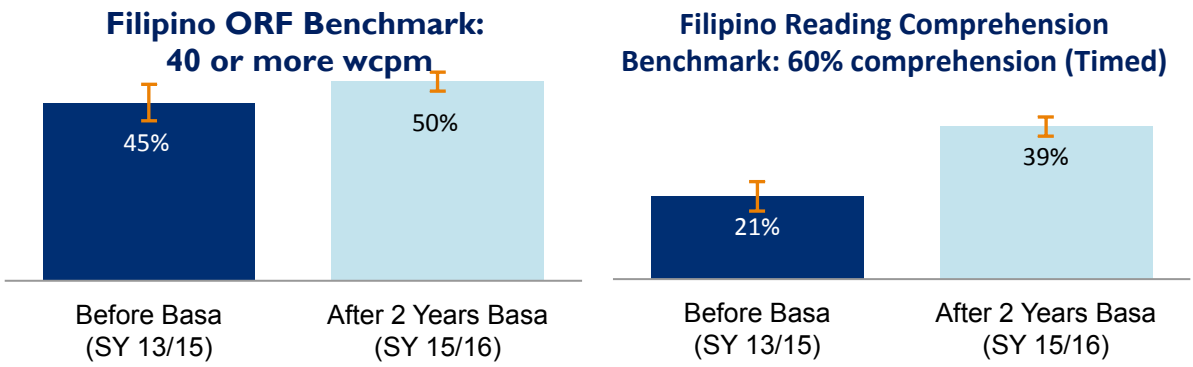
Sex	Cohort 1 (SY 13/14)	Cohort 2 (SY 15/16)	Effect Size of Difference in SY 13/14 and SY 15/16
Boys	42.2%	50.4%	0.12
Girls	44.9%	55.1%	0.23

BENCHMARKS

Basa has worked with DepEd to set Filipino fluency and comprehension benchmarks for Grade 2—40 words correct per minute and 60% reading comprehension. The results of the assessments show that, after two years of Basa, roughly 50% (±2.4%) of Grade 2 students are meeting the Filipino oral reading fluency benchmark and 39% (±2.5%) are meeting the reading comprehension benchmark (Figure 17). On average, Cohort 3 (SY 2015/16) Grade 2 students read 39 words correct per minute, compared to 37 words correct per minute for Cohort 1 (SY 2013/14) students. Twenty-two percent of second graders in the Cohort 3 have already reached the benchmark of 60 wcpm, as compared to 16% of students in Cohort 1 before the Basa intervention began.

[6] Students in the Cohort 1 (SY 13/14) were not tested in English EGRA. Results are only shown for Cohort 3 (SY 15/16).

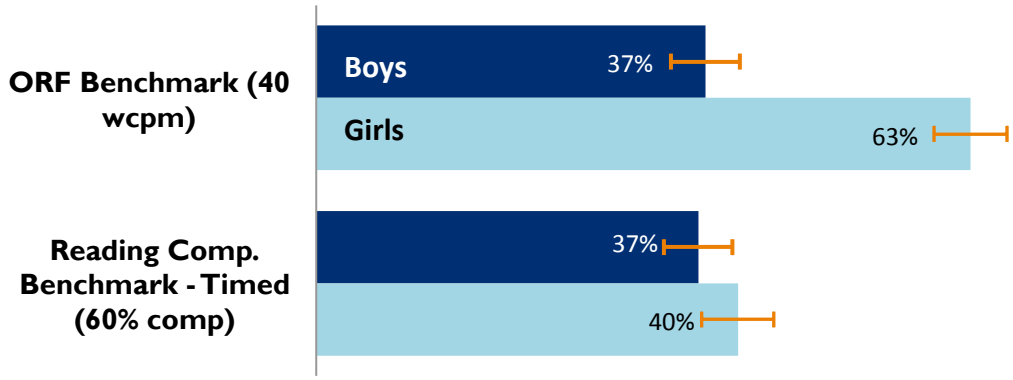
Figure 17. Percent of Grade 2 Students Meeting Filipino Fluency and Comprehension Benchmarks



Additional analysis showed that by the end of Grade 2, 34% of learners met both the Filipino fluency benchmark (40 or more wcpm) and the untimed reading comprehension benchmark of 60% reading comprehension.

As can be seen in Figure 18, in SY 2015/16, analysis by sex showed that nearly twice as many girls meet the Grade 2 Filipino ORF benchmark at the end of the second grade than boys (63% of girls versus 37% of boys); this was statistically significant at the p<.001 level. For the Reading Comprehension benchmarks, results were largely consistent between boys and girls, in which 37% of boys and 40% of girls met the reading comprehension benchmark of 60% comprehension.

Figure 18. Percent of Grade 2 Students Meeting Filipino ORF and Comprehension Benchmarks in SY 2015/16, by Sex



To help better understand the patterns of student achievement by subtest, the next sections of the report present results for each subtest, as well as disaggregations by sex.

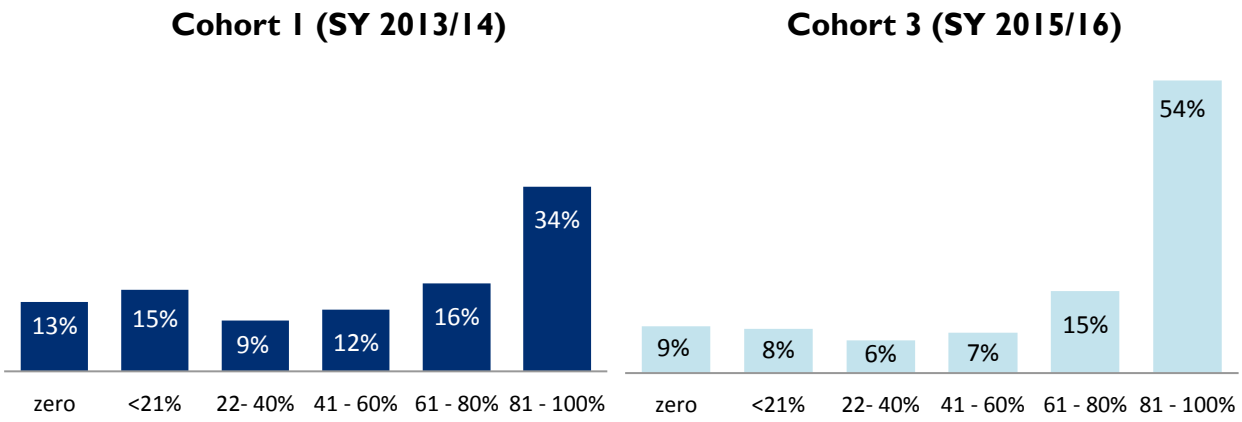


GRADE 2 FILIPINO EGRA SUBTEST RESULTS

INITIAL SOUND IDENTIFICATION, LETTER SOUNDS, FAMILIAR AND NONSENSE WORDS

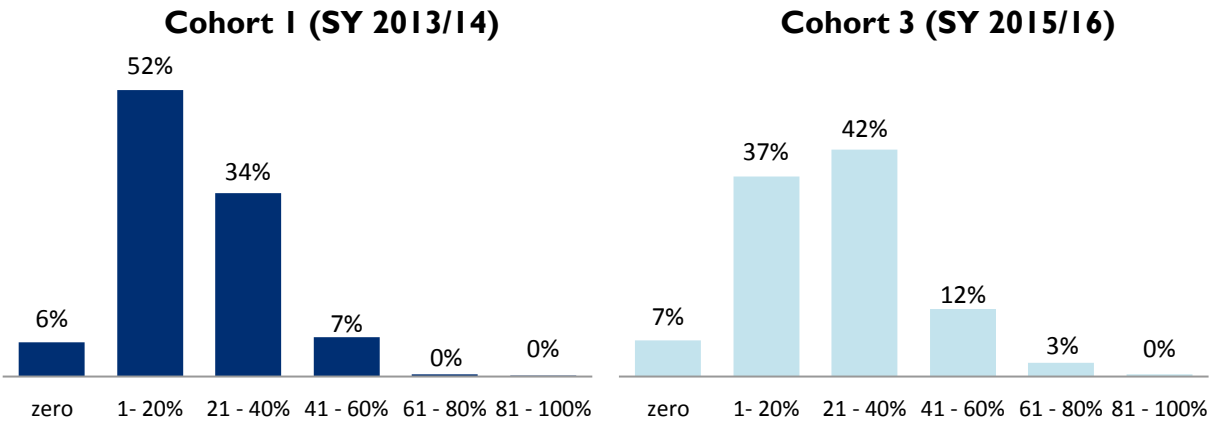
The **initial sound identification** subtest measures phonemic awareness of students. Assessors pronounced ten words and asked learners to listen and identify the correct letter sound. Responses ranged between zero to ten sounds correct, with a mean of 7.2 initial sounds for Cohort 3 (SY 15/16) students, and 5.8 initial sounds for Cohort 1 (SY 13/14) students. The graph below shows differences in the distribution of student's initial sound identification across cohort. For both cohorts, the Filipino EGRA initial sound identification results skews to the left with more than two-thirds of students in Cohort 3 scoring 60% or higher and roughly half of students in Cohort 1. Only 9% of students in Cohort 3 and 13% of students in Cohort 1 scored zero. Comparisons by sex showed that girls and boys in the comparison cohort (Cohort 1) had very similar distributions. In the intervention cohort (Cohort 3), a more distinguished difference between boys and girls was seen with 59.2% of girls scoring between 80 and 100%, while less than half (49.5%) of boys scored similarly.

Figure 19. Filipino Initial Sound Identification – Percent Correct, by Cohort



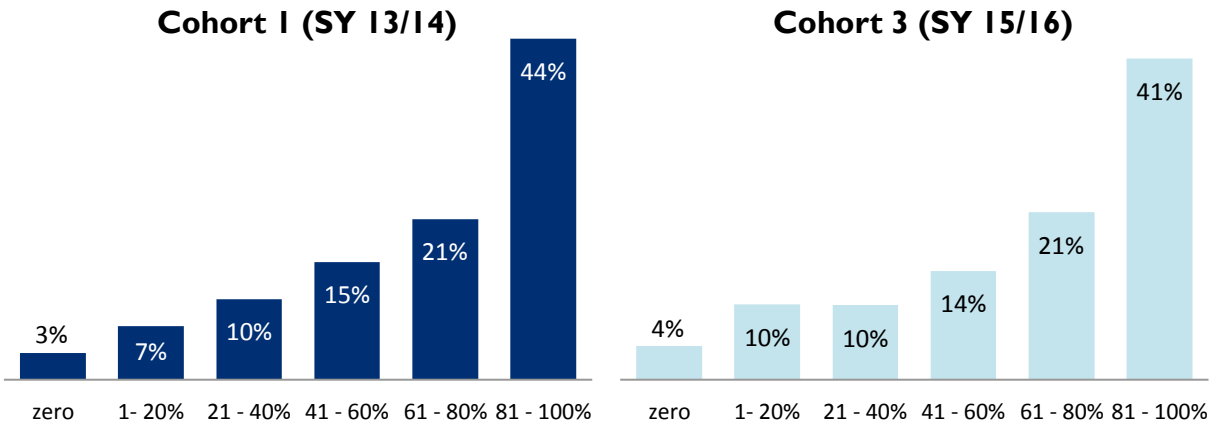
On the **letter sounds subtest** students were given a written list of 100 letters in random order and were asked to articulate the sound of each letter. Students were timed on the responses. The number of sounds pronounced correctly was divided by the seconds it took to answer and then multiplied by 60 seconds to find the correct letter sounds per minute. Number of letters correctly pronounced by students ranged from zero to 84 letters (mean 19.1) in Cohort 1 (SY 2013/14) and from zero to 100 letters (mean 24.1) in Cohort 3 (SY 2015/16). As the graphs below show, the majority of the students in both groups correctly named fewer than 40% of the letter sounds. However, as seen below in the graph, a slight shift is visible, in which slightly more Grade 2 learners in Cohort 3 are able to identify letter sounds correctly than learners in Cohort 1 before the Basa intervention began. The results suggest that students may not be used to being asked to name letter sounds in an assessment context. Analysis by sex showed that girls performed significantly better than boys on the letter sounds subtest ($p<.05$) for both Cohorts.

Figure 20. Filipino Letter Sound Subtest Percent Correct, by Cohort



The **familiar word identification** subtest assesses the next step in reading skill difficulty, in which learners need to utilize their mastery of letter-sound correspondence to decode words. On this subtest, students are given a list of 50 words that children are expected to be able to read at their grade level. Learners are given 60 seconds and are asked to read as many familiar words out loud as they can. Comparison of results by cohort showed that on average, students in Cohort 1 performed slightly better on this subtest than learners in Cohort 3, however, the difference was not statistically significant. Responses ranged from zero to 50 familiar words read correctly, with a mean of 33.0 (65.9%) for Cohort 3 (SY 15/16) students and a mean of 34.2 (68.4%) for the comparison cohort (Cohort 1 (SY 13/14)) students. The graph below shows that the distribution of Grade 2 Filipino familiar word scores for both Cohorts is skewed to the left, with more than half of tested students reading between 61 and 100% of the words correctly. Analysis by sex showed that girls in both cohorts significantly outperform boys ($p<.001$).

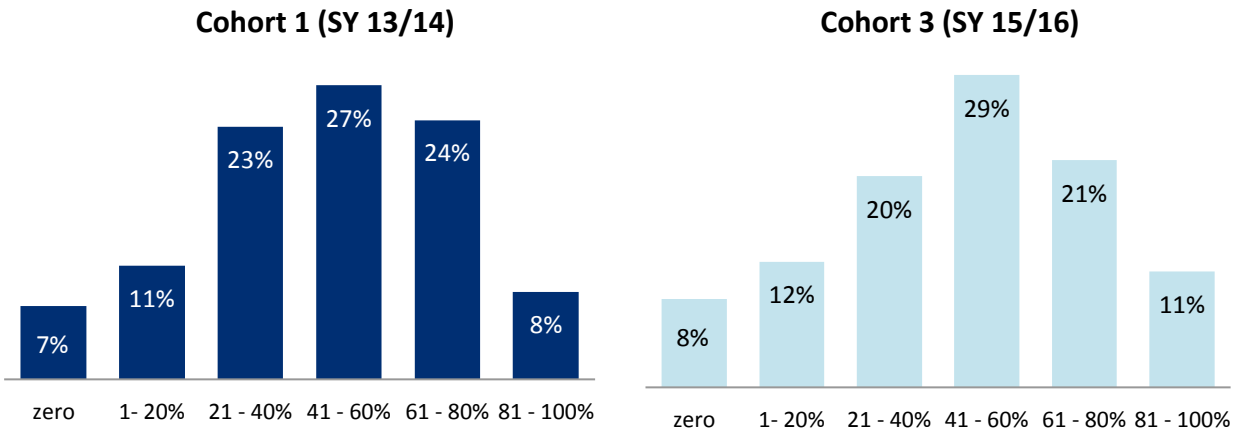
Figure 21. Filipino Familiar Word Identification Subtest Results, by Cohort



On the **simple nonsense word (non-word) decoding,** similar to the familiar word reading subtest, learners are given a list of 50 unfamiliar, invented words to read aloud in order to measure their ability to decode words rather than recognize them. Therefore, it is expected that this test is more difficult than oral passage reading or familiar word reading subtests. Students were timed on their responses. The number of correct responses ranged from zero to 50 nonsense words, with a mean of 23.4 (46.8%) for students in Cohort 3 (SY 15/16) and a mean of 23.1 (46.2%) for students in Cohort 1, prior to the Basa intervention. The distribution of Filipino nonsense words scores is close

to normal for both groups, with nearly a half of students reading between 20 and 60% of nonsense words correctly. Girls generally outperformed boys in both cohorts (p<.001), in which in both cohorts, on average girls scored roughly ten percentage points higher than boys on the nonsense words subtest. The distribution of boys' scores on this subtest is fairly normal for both tests. Conversely, for both cohorts, for girls, the distribution of scores is skewed to the left, showing roughly 40% of girls reading 60% or more of nonsense words correctly compared to only a quarter of boys.

Figure 22. Filipino Nonsense Words Subtest Results, by Cohort



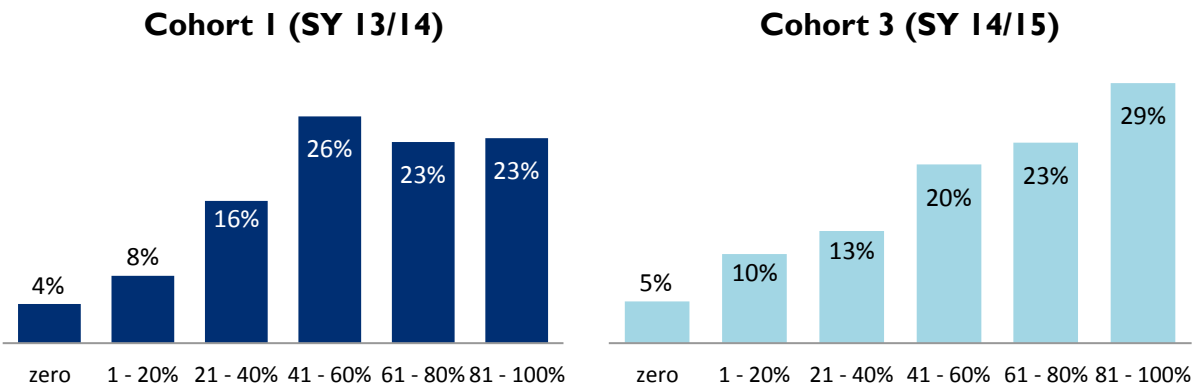
ORAL READING FLUENCY AND READING COMPREHENSION IN FILIPINO

To measure oral reading fluency, the ability to read quickly and accurately with proper expression[7], and reading comprehension, Grade 2 learners were asked to read aloud a grade-level text in Filipino and to answer five reading comprehension questions. On the Filipino passage reading and comprehension subtest, students were scored on the number of words they read correctly in the passage (total possible 64 words), reading comprehension (total possible 5), and prosody (total possible 4).

Oral Passage Reading Percent Correct. The number of words read correctly in Filipino by intervention students ranged from zero to 64 with a mean of 35.9 words (56.1%) for Cohort 1 students and a mean of 37.6 words 58.7%) for Cohort 3 students. The results presented in the graph below show that for Cohort 1 (SY 13/14) and Cohort 3 (SY 15/16) results were largely skewed to the left, with the majority of students reading more than 40% of the text. Although, average number of words read correctly was relatively similar among both cohorts, as seen in the graph below, slightly more Cohort 3 students are able to read 81-100% of the text than Cohort 1 students before the Basa intervention.

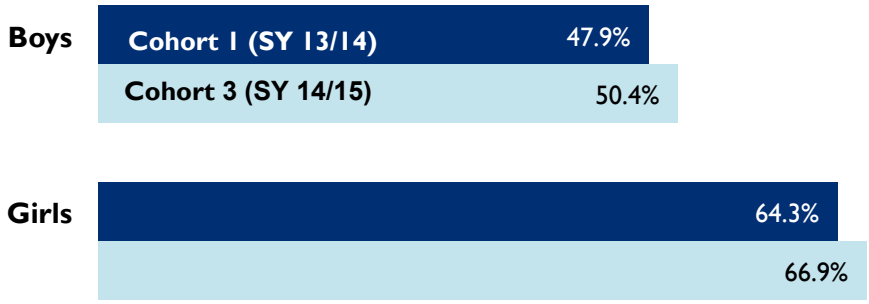
[7] Early Grade Reading Assessment Toolkit, Second Edition. 2015. p 25.

Figure 23. Filipino Oral Passage Reading Results, Percent of Words Read Correctly, by Cohort



Similar to other subtests, girls significantly (p<.001) outperformed boys in the oral reading subtest. As seen below, before the Basa intervention, girls in Cohort 1, read, on average, 64.3% of the text compared to 48.0% for boys. Girls continue to outperform boys in Cohort 3, in which girls were able to read, on average, 66.9% of the text compared to boys who read 50.4% of the text. The average gender gap in oral passage reading has remained largely the same from Cohort 1 to Cohort 3. In fact, as seen below, improvements in oral passage reading were roughly the same between boys and girls, in which on average both boys and girls in Cohort 3 (SY 15/16) were able read 2.5% more of the passage correctly than learners before the Basa intervention began (Cohort 1- SY 13/14). As a result, the gender gap has remained largely consistent over time.

Figure 24. Filipino Oral Passage Reading Average Percent Correct, by Cohort and Sex

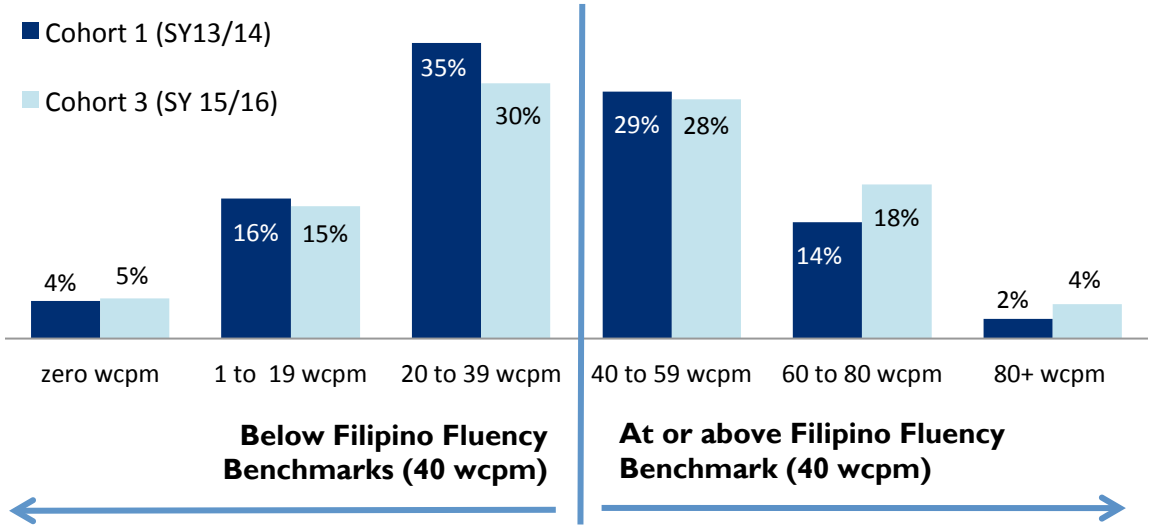


Fluency (Words correctly read per minute). Fluency is often described as a bridge between word recognition and comprehension, in which readers must eventually advance decoding skills to the point that is automatic and as a result, readers can focus their attention on what the text means.[8] Students were timed on reading the text, with the limit of 60 seconds. The number of words read correctly was divided by the seconds it took to read and then multiplied by 60 seconds to find the number of correct words per minute, which is the standard fluency measure used to measure USAID reading interventions. After two years of Basa intervention, Grade 2 learners were able to read significantly faster (p<.05) than learners before the Basa intervention. Cohort 3 (SY 15/16) students, on average, read with the speed of 39.4 words correct per minute, compared to

[8] Armbruster, B. B., Lehr, F., & Osborn, J. (2003). Put reading first: The research building blocks of reading instruction. Washington, DC: Center for the Improvement of Early Reading Achievement (CIERA).

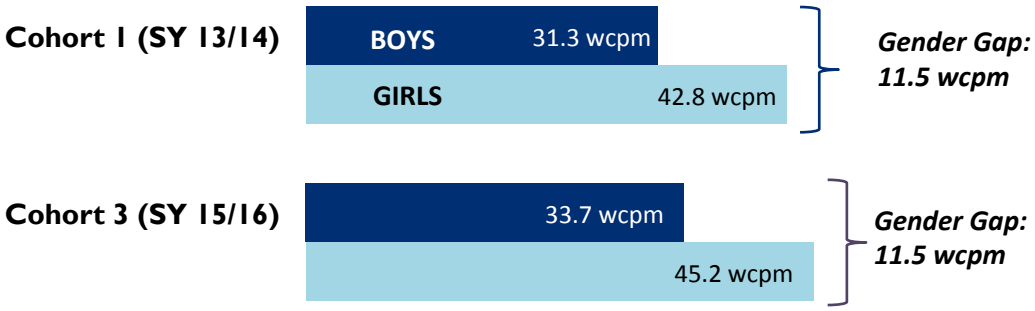
37.0 words correct per minute for students in Cohort 1 (SY 13/14). The graphs below show the distribution of fluency scores by student cohort. As seen below, 50% of students in Cohort 3 were reading at the Grade 2 proposed proficiency level of 40 wcpm, compared to 45% for students in Cohort 1 before the Basa intervention began.

Figure 25. Filipino Oral Reading Fluency (wcpm), by Group



There were significant differences (p<.001) in fluency rates between boys and girls for Cohort 1 students, in which, on average, girls read at a faster rate than boys. This significant (p<.001) difference in fluency rates by sex continued as well for Cohort 3. Analysis of the gender gap showed that, on average, from Cohort 1 to Cohort 3, the gender gap between boys and girls in fluency has remained unchanged.

Figure 26. Filipino Fluency Measure (wcpm), by Sex and Cohort

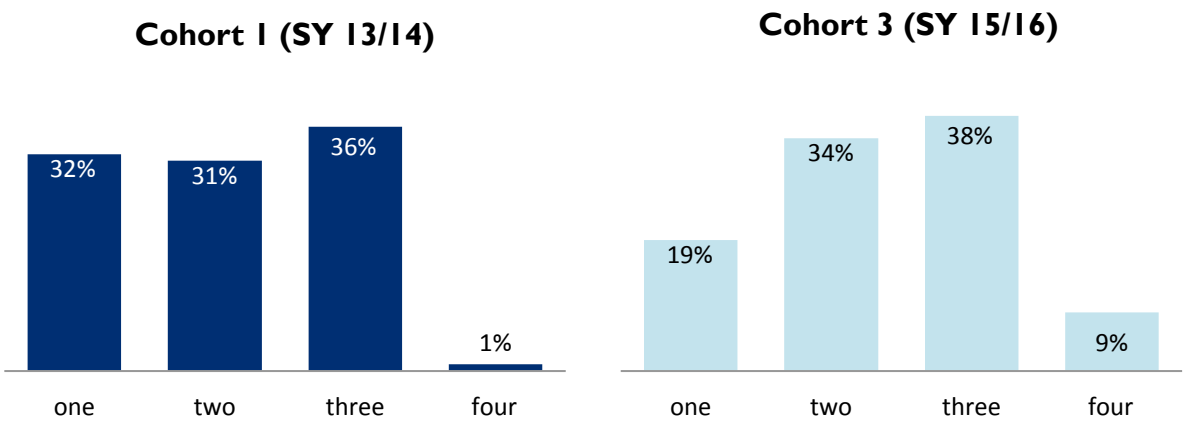


The **prosody scores** measured the degree to which students were able to read with intonation and expression. When students have difficulty reading, their prosody will often be slow and laborious whereas those who have ease in reading can focus on adding expression and tone to their reading. Prosody is just as linked to comprehension as automaticity because reading with good prosody resembles speech, which is embedded with underlying clues about the meaning of a text (Rasinki, 2011). Prosody measures were divided into four categories and scored accordingly:

1 = word-by-word, slow, laborious;
2 = small chunks, awkward;
3 = fluent, but does not mark punctuation; incorrect phrase groups, no expression;
4 = fluent, with expression to mark punctuation and/or direct speech.

Overall, students in Cohort 3 showed significantly (p<.001) higher Filipino prosody scores than Cohort 1 students. Students reading the passage ranged from one to four, with a mean of 2.4 for Cohort 3 students, and 2.1 for Cohort 1 students. The graphs below show the distribution in students' prosody for both cohorts. As seen below, improvements can be seen in prosody from Cohort 1 to Cohort 3. The distribution shows for Cohort 3 that nearly half (47%) of students read fluently (scored 3 or 4); only 19% read laboriously. For Cohort 1, about a third of students read the passage fluently; while nearly a third read laboriously.

Figure 27. Filipino Oral Passage Reading - Prosody Score, by Cohort

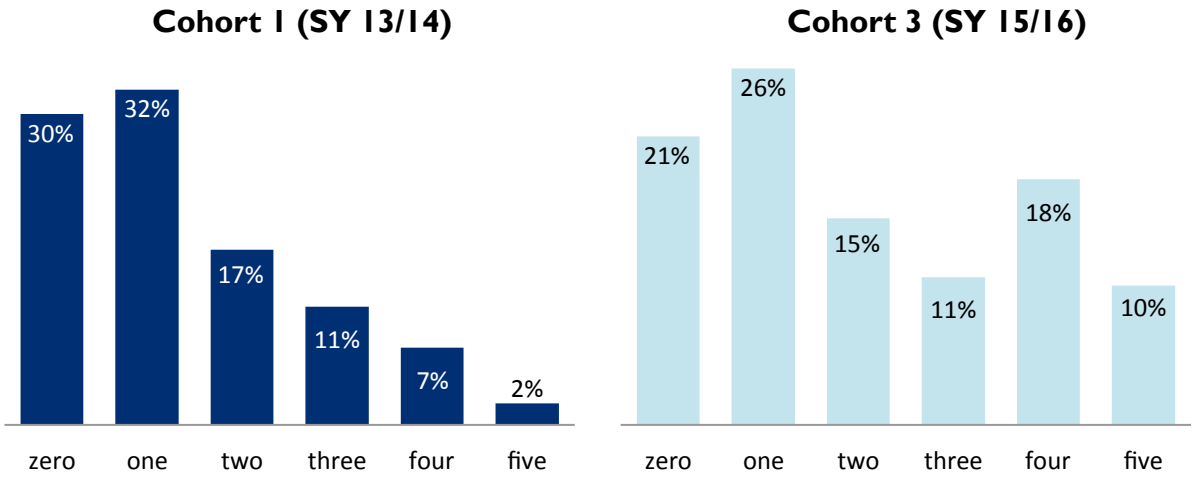


Disaggregation by sex showed that girls from both groups read the Filipino passage with greater prosody than boys, which is consistent with the overall pattern of girls demonstrating higher reading proficiency.

Reading Comprehension. Comprehension is the ultimate goal of reading. To measure comprehension, students were asked five questions about the text after reading a grade-level Filipino passage. Learners were not allowed to look back at the text to help them answer questions. From Cohort 1 (SY 13/14) to Cohort 3 (SY 15/16), students showed a significant (p<.001) increase in the number of reading comprehension questions they could answer correctly. For Cohort 1, the total number of reading comprehension correct ranged from zero to five, with a mean of 1.4 (28.1%) while Cohort 3 students answered on average 2.1 questions correctly (41.8%).

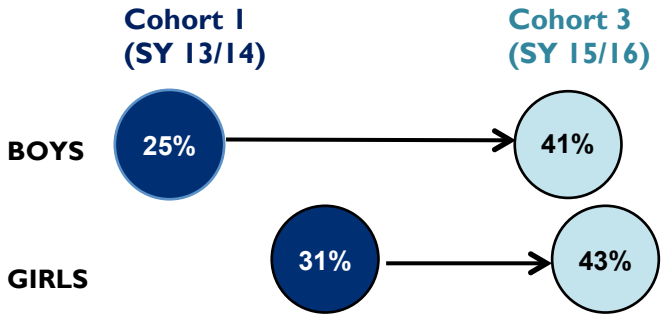
The figures below show the distribution of reading comprehension results by cohort. Students in Cohort 3 were able to answer more reading comprehension questions correctly than students before the Basa intervention in Cohort 1. More than a quarter (28%) of Cohort 3 students answered four or more comprehension questions correctly compared to 9% for Cohort 1 students. Additionally, the percent of Grade 2 learners who were unable to answer a single reading comprehension question correctly decreased, with roughly one fifth (21%) of Cohort 3 students were unable to answer a single reading comprehension question compared to 30% of Cohort 1 students.

Figure 28. Grade 2 Reading Comprehension Results, Filipino EGRA, by Cohort



Comparisons by sex show for Cohort 1, girls significantly ($p<.05$) outperformed boys in reading comprehension at the end of a school year. However, Cohort 3 results showed that reading comprehension results by sex were largely consistent between boys and girls, in which boys had largely caught up to girls in reading comprehension. Detailed results by sex are found in Annex 3.

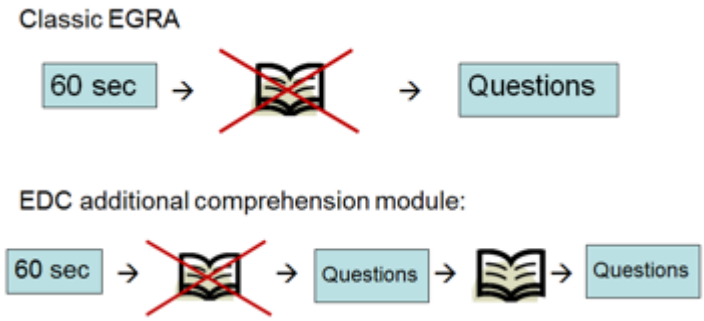
Figure 29. Average Percent Correct on Reading Comprehension, by Sex and Cohort



An examination of the EGRA testing process during SY 14/15 suggested to the Basa evaluation team that low comprehension results might be due to test bias rather than to actual lack of comprehension abilities among students. To test this hypothesis, for Cohort 3, the evaluation team administered the Filipino reading comprehension questions in two rounds. The first round the administration followed the standard EGRA administration procedures. The second round immediately followed the first round. During the second round, assessors gave the text back to the students and allowed them to finish reading the passage[9] (if they haven't done so already), and then asked them comprehension questions without taking the text away from the students. The figure below compares the two rounds:

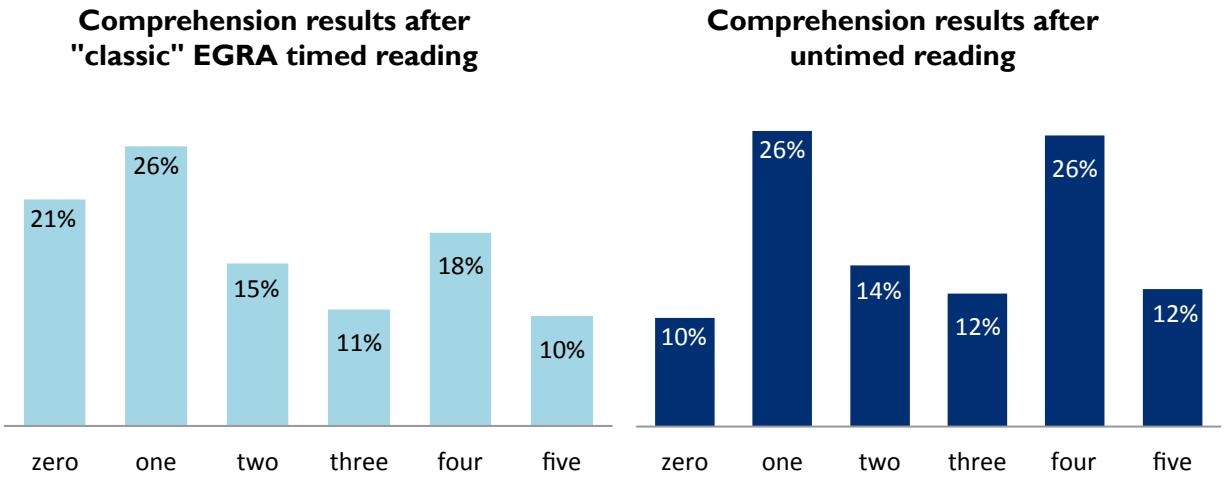
[9] Students who answered four or five comprehension questions during the first round were not asked to read the passage again nor answer reading comprehension questions again given that they demonstrated reading comprehension ability during the first around of administration.

Figure 30. Measuring Reading Comprehension



The results of the comparison between the two models of testing students' comprehension skills are presented in Figure 30. As seen in the figure, the zero scores dropped from 21% to 10%, and the percent of students who were able to meet comprehension proficiency benchmark of 80% increased by ten percentage points (from 28% to 38%).

Figure 31. Comparison of Cohort 3 Comprehension Results Using Timed and Untimed Reading[10]



These findings suggest that reading comprehension results from "classic" EGRA assessments must be interpreted with caution since scores might be as much impacted by the process of administration as by the abilities of the tested students.

LISTENING COMPREHENSION AND DICTATION

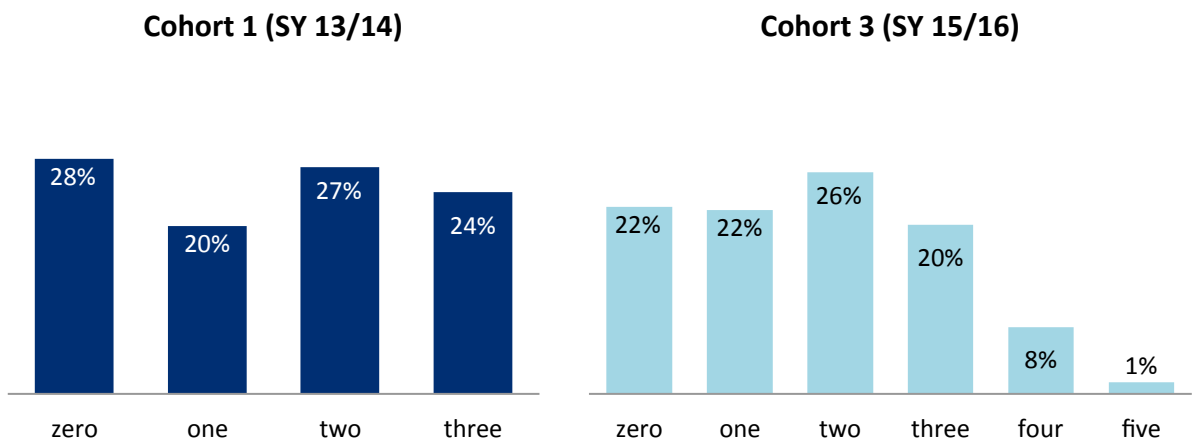
On the **listening comprehension subtest**, students were read a passage and asked comprehension questions. Three comprehension questions were asked after the Filipino passage for Cohort 1 and five questions were asked for Cohort 3. On average, for Cohort 1, students answered 49.1% of listening comprehension correctly compared to 34.7% for Cohort 3. However, it should be noted that since Cohort 3 students were asked a different number of comprehension questions than Cohort 1 caution should be made when comparing results of listening comprehension across cohort.

[10] In order to compare results from the first (timed) round and the second (untimed) round, untimed comprehension results include data for students who, during the second round, were allowed to read the passage (untimed) and answer the comprehension questions, as well as, results for students who answered four or more comprehension questions correctly during the first round, and as a result did not participate in the second, untimed, round.



The figure below shows distributions of the number of listening comprehension questions students were able to answer correctly for Cohorts 1 and 3. For Cohort 1, student scores were relatively evenly distributed. More than a quarter of students were unable to answer a single listening comprehension question; roughly half of students answered two or three questions. For Cohort 3, results showed a decrease in zero scores in listening comprehension compared to learners in Cohort 1, with 22% of students with zero scores compared to 28% for Cohort 1. Very few students (9%) in Cohort 3 were able to answer four or five listening comprehension questions; the majority answered between one and three.

Figure 32. Filipino Listening Comprehension Questions – Number Correct, by Cohort



Listening comprehension results by sex showed that boys and girls in Cohort 1 performed similarly on this subtest. However, as seen in other subtests, in Cohort 3, girls performed better than boys; the difference between boys and girls was statistically significant ($p<.001$).

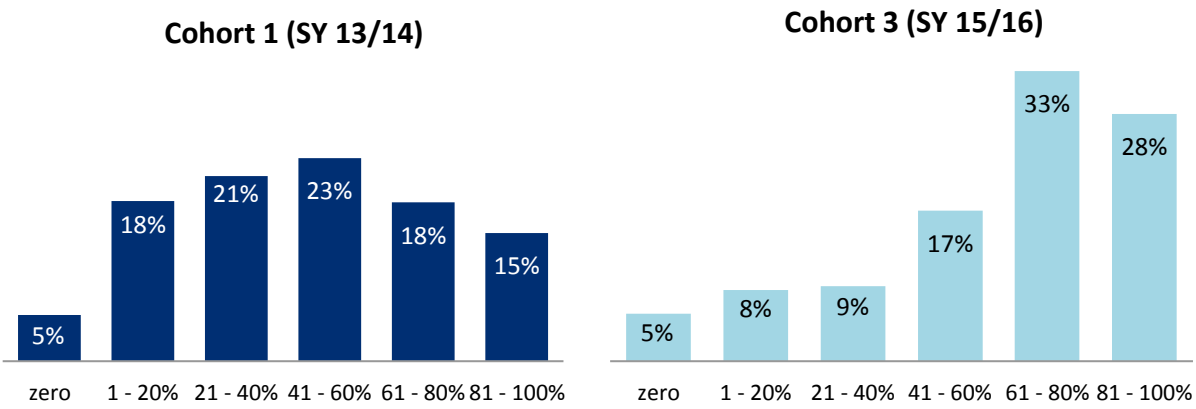
The **dictation subtest** is used to measure both oral comprehension and writing skills of learners. Writing skills are strongly connected to reading skills in that it takes phonological awareness skills to translate sounds into individual graphemes and the written word. On the dictation subtest (total possible 16 correct answers for the Filipino test), Grade 2 students were read a passage once, given a pencil and paper, and read a passage a second time with pauses, then read the entire passage a third time. Students wrote the words on the paper. Dictation scores were broken up into two subtests:

- Number of words spelled correctly (total possible 12)
- Other items relating to conventions of text in writing including, spacing, text direction, capital letter, and using a period at the end of a sentence (total possible 4)

Overall, distribution results show a significant ($p<.001$) improvement in Grade 2 Filipino dictation scores from Cohort 1 (SY 13/14) to Cohort 3 (SY 15/16). Number of correct answers for the dictation subtest ranged from zero to 16, with a mean of 7.3 (45.6%) for Cohort 1 (SY 13/14) students and a mean of 9.6 (59.7%) for Cohort 3 (SY 15/16) students. Figure 33 below shows the distribution of dictation results for Cohort 1 and Cohort 3. Roughly two-thirds of Grade 2 students

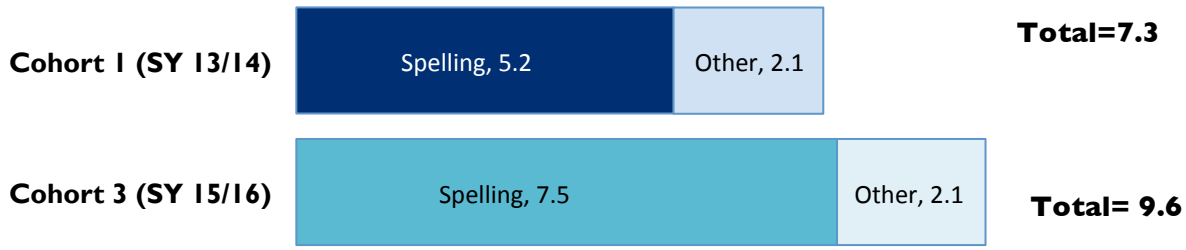
in Cohort 1 scored 60% or less on the dictation subtest compared to 39% of Cohort 3 students. In fact, as seen in the figure below, nearly two-thirds of Cohort 3 students scored over 60%; roughly a quarter (28%) scored between 80-100%.

Figure 33. Grade 2 Filipino Dictation Results (Percent Correct), by Cohort



Breaking down the dictation composite, scores for spelling ranged from zero to 12, with a mean of 7.5 for Cohort 3 students and 5.2 for Cohort 2 students. With regard to the conventions of text, results were similar across Cohort, in which students in both cohorts scored on average 2.1 out of a possible 4. Most students used spacing and the direction of the text correctly. However, only one third of students capitalized correctly in both groups, and less than 20% of students used a period at the end of the sentence.

Figure 34. Filipino Dictation Breakdown, by Group



Disaggregating by sex, girls significantly outperform boys in dictation. This is not surprising given that girls have outperformed boys in all other subtests.

4. FILIPINO & ENGLISH EGRA FINDINGS – COMPARISON OF GRADE 2 & GRADE 3



Grade 3 is an important year for learners in the Philippines. During Grades 1 through 3, Mother Tongue is the main language of instruction in the classroom, with Filipino and English introduced as subjects during these grades. By Grade 4, however, Mother Tongue is phased out as both a subject and language of instruction and students are taught all subjects in Filipino, except math and science which are taught in English. As such, measuring Filipino and English reading outcomes at Grade 3 is crucial to understanding the preparedness of learners to begin instruction in these languages as they proceed to Grade 4.

To measure learners' literacy performance in Filipino and English at Grade 3 and to assess the trajectory of reading skill acquisition from Grade 2 and Grade 3, during SY 2015/16 the EGRA study included an assessment of Grade 3 student performance. The following section highlights the findings from this assessment.

SUMMARY COMPARISONS OF GRADE 2 AND GRADE 3 EGRA RESULTS

OVERALL GRADE 2 AND GRADE 3 FILIPINO EGRA RESULTS

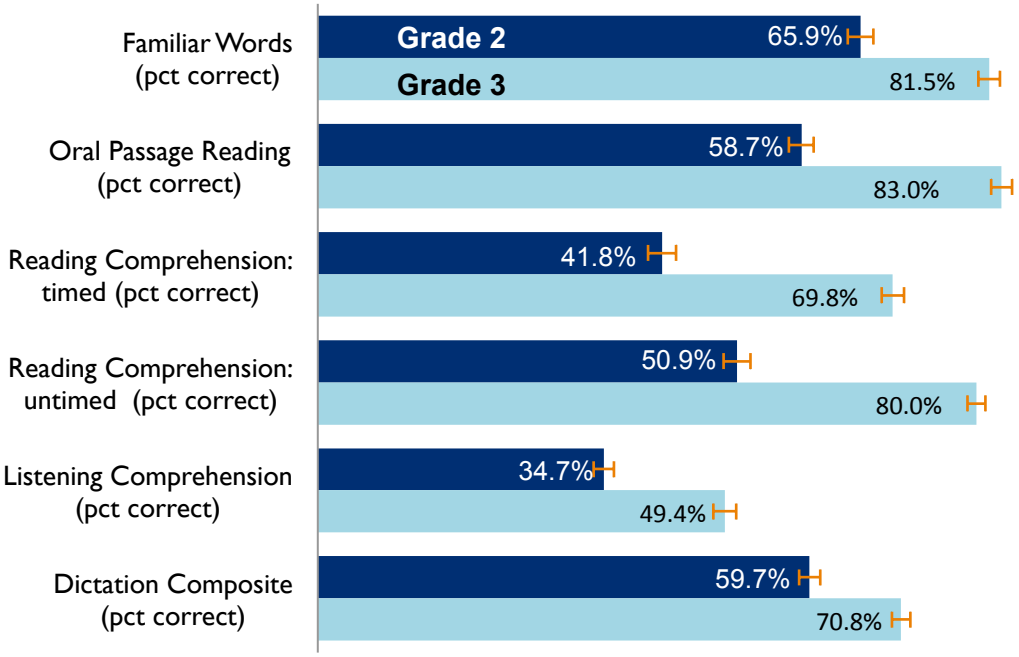
Overall, Filipino EGRA results suggest that on average, by the end of Grade 3, learners demonstrate relatively strong reading skills in Filipino. As seen in the table below, analysis showed that by the end of Grade 3, learners possess relatively strong Filipino word recognition skills. In fact, assessment results showed that learners in Grade 3 performed the strongest on the Familiar Word Reading and Oral Passage Reading subtests. As seen previously for Grade 2 students, the subtest with the lowest scores was Listening Comprehension. Untimed reading comprehension results suggest that learners are demonstrating relatively high reading comprehension which suggests that learners have started to progress to linking the words they read or hear to understanding the meaning of these words. In all, Grade 3 Filipino results suggest that learners are largely prepared for instruction in Filipino in the subsequent grade, Grade 4, and are on track to transition from "learning to read" in Filipino to "reading to learn."

Table 12. Average Grade 3 and Grade 2 Filipino EGRA Subtest Results for Cohort 3 (SY 15/16)

Filipino EGRA Subtests[11]	Grade 2	Grade 3
Familiar Words (pct correct)	65.9% (±1.5%)	81.5% (±1.3%)
Familiar Words Correct (per min)	35.3 (±0.9)	47.8 (±1.0)
Oral Passage Reading (pct correct)	58.7% (±1.5%)	83.0% (±1.3%)
Words Correct in a Text (per min)	39.4 (±0.0)	57.8 (±0.0)
Prosody Score	2.4 (±0.0)	2.9 (±0.0)
Reading Comprehension: timed (pct correct)	41.8% (±1.7%)	69.8% (±1.4%)
Reading Comprehension: untimed (pct correct)	50.9% (±1.6%)	80.0% (±1.1%)
Listening Comprehension (pct correct)	34.7% (±1.2%)	49.4% (±1.4%)
Dictation Composite (pct correct)	59.7% (±1.3%)	70.8% (±1.1%)

Comparisons of Grade 2 and Grade 3 reading performance in Filipino showed that Grade 3 students demonstrate substantially better reading skills in Filipino than their Grade 2 counterparts. For each subtest, Grade 3 learners scored between 11 to 28 percentage points higher than Grade 2 students. These findings suggest that substantial improvements in Filipino reading skills occur between Grade 2 and Grade 3. The figure below shows the Filipino EGRA scores for Cohort 3 (SY 2015/16) students for both Grade 2 and Grade 3.

Figure 35. Average Filipino EGRA Subtest Results for Cohort 3 (SY 15/16), by Grade



OVERALL GRADE 2 AND GRADE 3 ENGLISH EGRA RESULTS

[11] For Grade 2 and Grade 3, different Filipino EGRA tests were designed and developed to measure grade-level reading performance in Filipino. Resultantly, the tests for Grade 2 and Grade 3 are different. Only the dictation subtest was the same for both tests. Given the different tests, gains and effect size are not calculated or reported.

Grade 3 English EGRA results show that, similar to Filipino, by the end of Grade 3, learners demonstrated relatively strong English word recognition skills in which, assessment results showed that learners performed the strongest on the Familiar Word Reading and Oral Passage Reading subtests. Learners struggled particularly with Listening and Reading Comprehension in English, which suggests that although learners have developed skills in word recognition in English, the majority of learners have not progressed to linking the words they read or hear to understanding the meaning of these words, which is needed for English comprehension. In all, Grade 3 English results suggest that learners may not be prepared for instruction in English in the subsequent grade, Grade 4, given that learners continue to struggle with reading and listening comprehension in English at the end of Grade 3.



Table 13. English EGRA Subtest Results for Cohort 3 (SY 15/16) by Grade

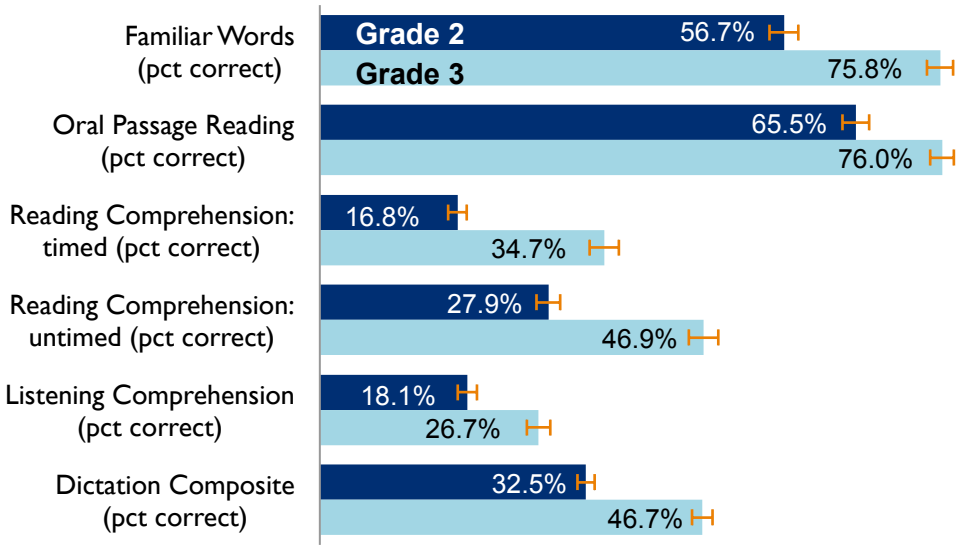
English EGRA Subtests[12]	Grade 2	Grade 3	Difference	Effect Size
Familiar Words (pct correct)	56.7% (±1.8%)	75.8% (±1.6%)	19.1% (±2.4)	0.55
Familiar Words Correct (per min)	34.0 (±1.3)	51.7 (±1.4)	17.7 (±1.7)	0.64
Oral Passage Reading (pct correct)	65.5% (±1.6%)	76.0% (±1.4%)	--	--
Words Correct in a Text (per min)	40.4 (±1.1)	60.6 (±1.6)	--	--
Prosody score	2.3 (±0.0)	2.7 (±0.0)	--	--
Reading Comprehension: timed (pct correct)	16.8% (±1.1%)	34.7% (±1.8%)	--	--
Reading Comprehension: untimed (pct correct)	27.9% (±1.5%)	46.9% (±1.8%)	--	--
Listening Comprehension (pct correct)	18.1% (±1.2%)	26.7% (±1.4%)	8.6% (±1.8)	0.33
Dictation Composite (pct correct)	32.5% (±1.1%)	46.7% (±1.3%)	14.2% (±1.6)	0.59

The figure below shows the English EGRA results for Grade 2 and Grade 3 learners. As seen below, in SY 2015/16, Grade 3 students demonstrate stronger reading skills in English than their Grade 2 counterparts. Across the English subtests, Grade 3 students scored between eight to nineteen percentage points higher on the English EGRA subtests than Grade 2 students, which suggests that learners continue to improve in English reading skills from Grade 2 to Grade 3. In fact, analysis

[12] The English EGRA for Grade 2 and Grade 3 were the same for three out of five subtests. The Oral Passage Reading and Reading Comprehension subtests were different for each grade and were designed to test learners on grade-level text and comprehension. As such, gains are only shown for subtests that were the same for both Grade 2 and Grade 3.

showed that Grade 3 learners performed significantly better than Grade 2 learners on the Familiar Word Reading and Dictation subtests, in which results showed a moderate effect size difference.[13]

Figure 36. English EGRA Subtest Results for Cohort 3 (SY 2015/16), by Grade



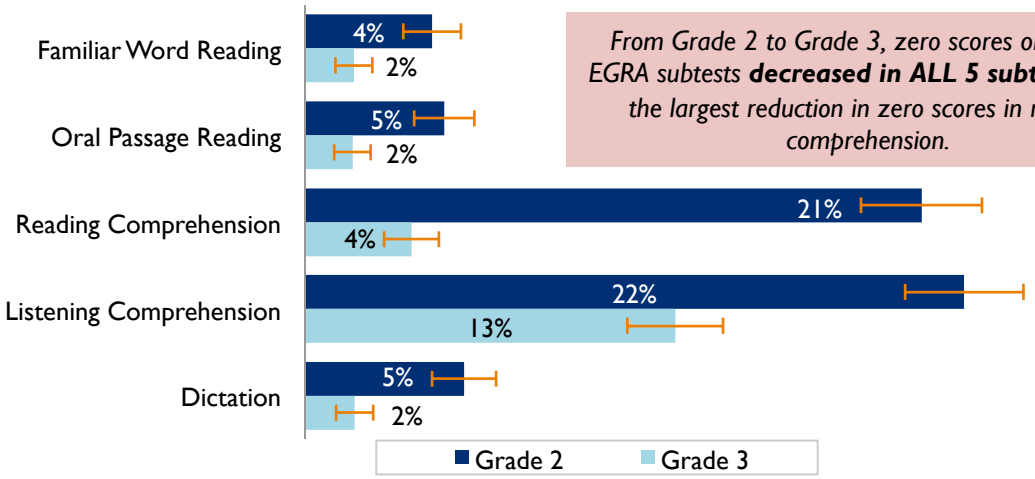
ZERO SCORES

Analysis of zero scores showed that as seen in Grade 2, a larger percentage of Grade 3 students had zero scores on the English EGRA subtests compared to Filipino subtests.

The analysis of Cohort 3 (SY 2015/16) Filipino assessment results found that Grade 3 learners had significantly ($p<.001$) less zero scores than Grade 2 learners on Filipino EGRA subtests. As seen in the figure below, from Grade 2 to Grade 3, large reductions in zero scores were seen in Reading Comprehension and Listening Comprehension. For Reading Comprehension, roughly one in five Grade 2 learners were unable to answer a single reading comprehension question correctly; however, for Grade 3, this percentage had decreased substantially to only 4% of Grade 3 learners. In terms of Listening Comprehension, results also showed an impressive reduction in zero scores from Grade 2 to Grade 3, in which roughly a fifth of Grade 2 learners were unable to answer a single listening comprehension question correctly compared to 13% of Grade 3 learners. Figure 37 shows the percent of tested Grade 2 and Grade 3 students scoring zero on Filipino EGRA subtests.

[13] Effect size is a statistical measure that is used to estimate the magnitude of difference between two measures. Effect sizes are interpreted as follows, according to Cohen (1998): "small, d = .2," "medium, d = .5," and "large, d = .8". Effect size was only calculated for English EGRA subtests that were the same on both the Grade 2 and Grade 3 English EGRA tests – Familiar Word Reading, Listening Comprehension and Dictation.

Figure 37. Percent of Tested Grade 2 and Grade 3 Students Scoring Zero on Filipino EGRA Subtests

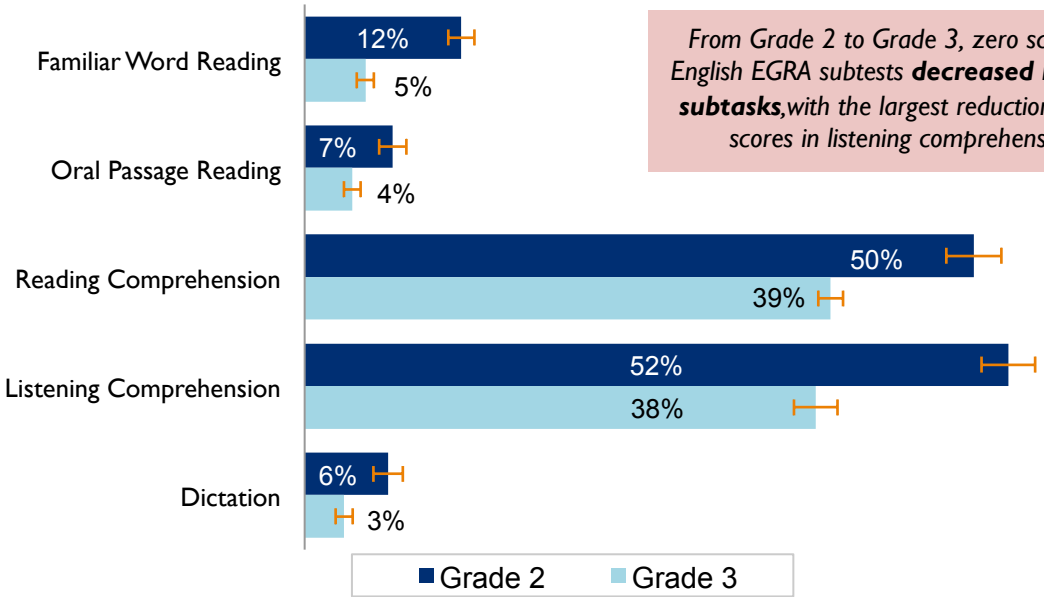


*From Grade 2 to Grade 3, zero scores on Filipino EGRA subtests **decreased in ALL 5 subtasks,** with the largest reduction in zero scores in reading comprehension.*

Analysis of English EGRA results from Cohort 3 (SY 2015/16) also showed statistically significant ($p<.001$) reductions in zero scores from Grade 2 to Grade 3. The figure below shows the zero scores for each English subtest for both Grade 2 and Grade 3 learners. As seen in the figure, large reductions in zero scores were seen in reading comprehension and listening comprehension from Grade 2 to Grade 3. Zero scores decreased by 11 percentage points for Reading Comprehension and 14 percentage points for Listening Comprehension.



Figure 38. Percent of Tested Grade 2 and Grade 3 Students Scoring Zero on English EGRA Subtests

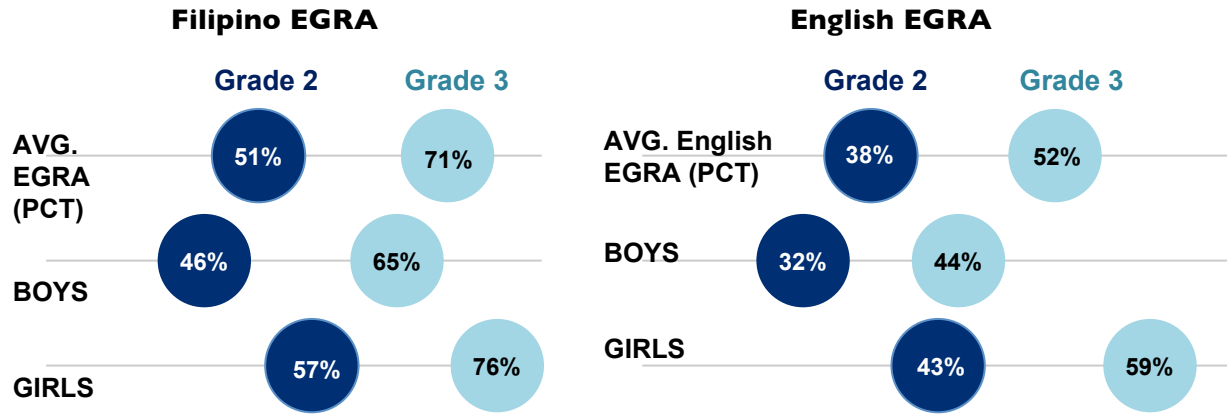


*From Grade 2 to Grade 3, zero scores on English EGRA subtests **decreased in ALL 5 subtasks**,with the largest reduction in zero scores in listening comprehension.*

RESULTS BY SEX

Overall, Cohort 3 (SY 2015/16) EGRA results show that girls continue to outperform boys in Filipino and English in Grade 3. Figure 39 below shows the average EGRA scores for Grade 2 and Grade 3 learners for both the Filipino and English EGRAs. As seen in the figure, average EGRA results suggest that the gender gap in EGRA performance appears to remain largely consistent for Filipino from Grade 2 to Grade 3. However, for the English EGRA, the gender gap appears to be widening.

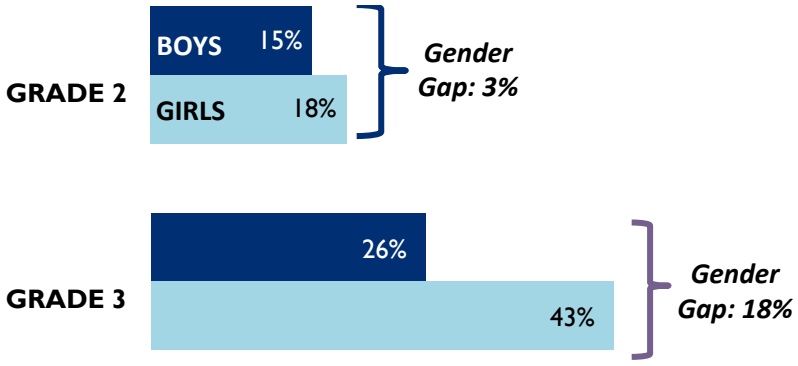
Figure 39. Average Percent Correct[14] on Filipino and English EGRA Subtasks (SY 2015/16), by Grade



Further analysis of Filipino EGRA results by subtest showed that the gender gap remained largely consistent from Grade 2 to Grade 3, in that, the gender gap neither worsened nor improved substantially from Grade 2 to Grade 3. Detailed analysis of Filipino EGRA results by sex and grade can be found in Annex 3.

Conversely, analysis of English EGRA results for Cohort 3 (SY 15/16) by subtest showed that the gender gap increased between boys and girls in three of the five English subtest – Reading Comprehension, Listening Comprehension and Dictation—

Figure 40. Gender Gap in English Reading Comprehension Percent Correct, by Grade (Cohort 3 – SY 2015/16)



which suggests that the gender gap in English may be increasing from Grade 2 to Grade 3. Of particular interest, is the substantial widening of the gender gap in English Reading Comprehension, in which, on average, Grade 2 boys scored only 3.2% worse than girls in English reading comprehension; however, by Grade 3, boys, on average, scored 17.6% worse than girls. Figure 40 shows the substantial widening of the gender gap in English Reading Comprehension from Grade 2 to Grade 3.

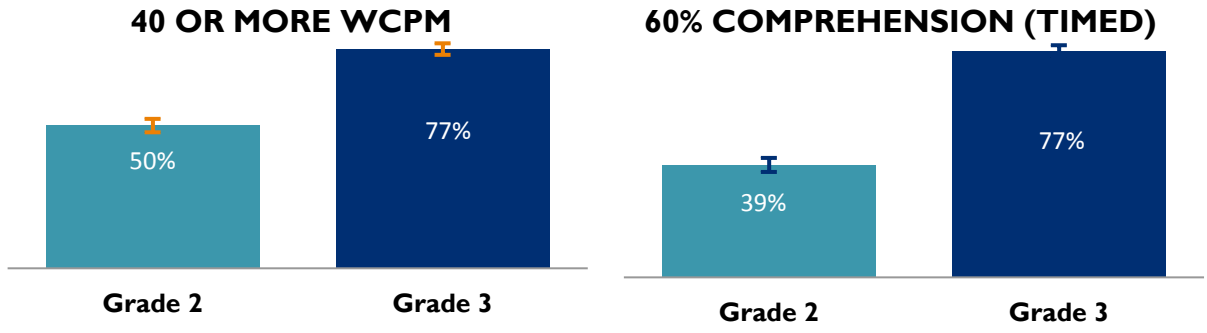
[14] Calculated by averaging five EGRA subtasks that were administered to both Grade 2 and Grade 3 students (Familiar Words, Oral Passage Reading, Reading Comprehension, Listening Comprehension, and Dictation)

BENCHMARKS

Basa has worked with DepEd to finalize Filipino fluency and comprehension benchmarks for Grades 2 and 3 —fluency benchmark of 40 words correct per minute and 60% reading comprehension. The results of the assessments show that, by the end of the school year, half of Grade 2 students were reading at the Basa-proposed proficiency standard of 40 wcpm (Figure 41). However, by Grade 3, more than three-quarters (77%) of learners are meeting the Filipino fluency benchmark of 40 words correct per minute. On average, at the end of the school year, Grade 2 students read 39 words correct per minute, compared to 58 words correct per minute for Grade 3 students, which demonstrates a substantial improvement in fluency in Filipino from Grade 2 to Grade 3. Very few Grade 2 students (4%) were able to read 80 words correct per minute or more; however, by Grade 3, roughly a fifth (20%) of Grade 3 learners had Filipino fluency rates of 80 wcpm or more.

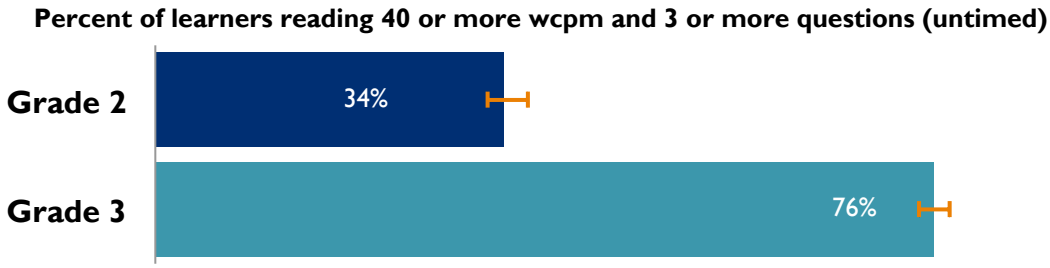
Analysis of Filipino reading comprehension results showed that, similar to fluency, the majority of Grade 2 learners were not able to meet the Filipino reading comprehension benchmark of 60%, in which roughly a third of Grade 2 learners were able to answer 60% of reading comprehension questions correctly. By Grade 3, as seen in the figure below, the percent of learners meeting the reading comprehension benchmark nearly doubled with more than two thirds (77%) of Grade 3 learners who were able to read 60% or more Filipino reading comprehension questions correctly.

Figure 41. Percent of Cohort 3 (SY 15/16) Learners Meeting Filipino Fluency and Reading Comprehension Benchmarks, by Grade



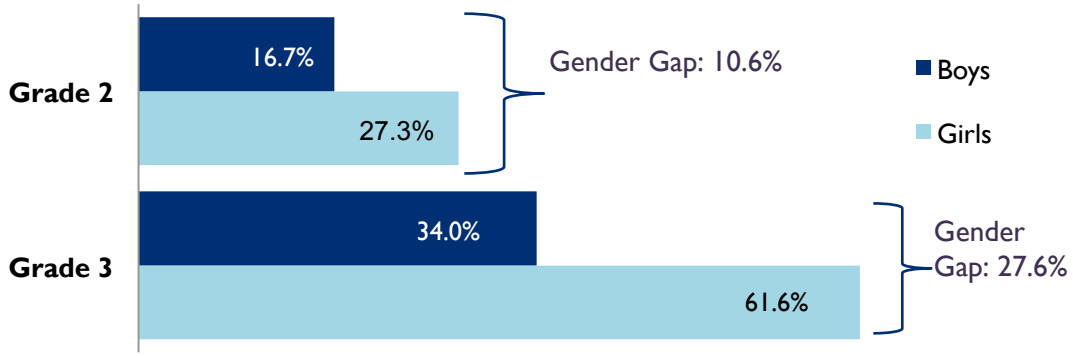
Additional analysis showed substantial improvements from Grade 2 to Grade 3 in the percent of learners who met both the Filipino fluency benchmark of 40 words correct per minute and the reading comprehension benchmark of three or more comprehension questions correctly (untimed). As seen in the figure below, in Grade 2, roughly a third (34%) of students met the combined fluency and comprehension benchmark (untimed); however, by Grade 3, roughly three quarters of learners were able to.

Figure 42. Percent of Cohort 3 (SY 15/16) Learners Meeting Combined Filipino Fluency and Reading Comprehension Benchmarks, by Grade



For both grades, more girls than boys were meeting the Filipino Grade 3 fluency benchmark of 60 words correct per minute. As seen in the figure below, in Grade 3, 61.6% of girls were able to meet Filipino fluency benchmarks, while only a third (34%) of boys were able to. These findings suggest that girls are not only demonstrating higher fluency rates than boys but are also improving in fluency faster than boys. Results show that the gender gap in Filipino oral reading fluency is increasing substantially from Grade 2 to Grade 3 in which the gender gap has more than doubled from 10.6% to 27.6%.

Figure 43. Percent of Cohort 3 (SY15/16) Learners Reading at least 60 wcpm in Filipino, by Grade and Sex



Filipino reading comprehension benchmarks by sex show less pronounced differences between boys and girls. By Grade 3, roughly three-quarters of boys (73.7%) met the Filipino reading comprehension benchmark of 60% reading comprehension compared to 81.0% of girls. This difference was statistically significant at the p<.001 level.

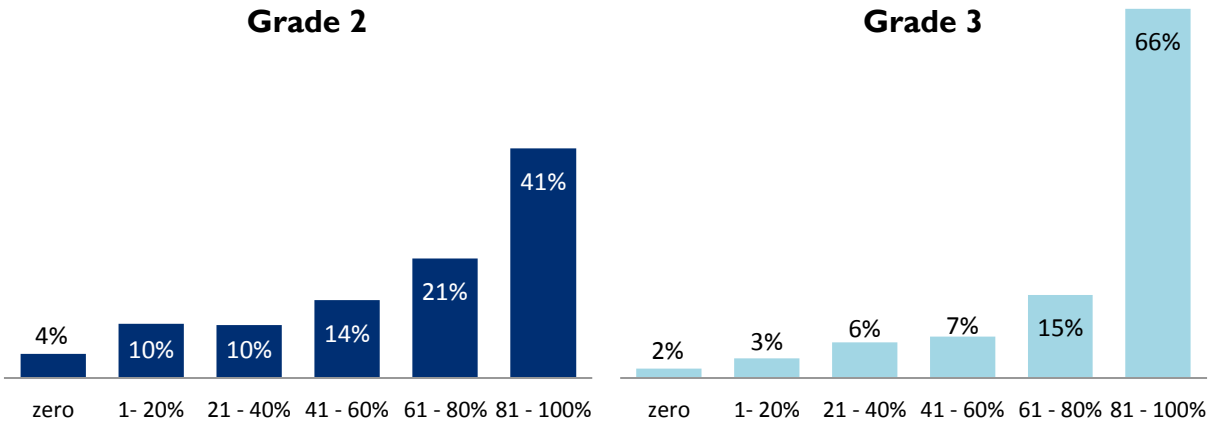
GRADES 2 AND 3 FILIPINO SUBTEST RESULTS

The following section provides detailed results for Filipino EGRA subtests for Grade 3 students compared to Grade 2 students in Cohort 3 (SY 2015/16).

FAMILIAR WORDS

The **familiar word identification** subtest assesses learners' mastery of letter-sound correspondence needed to decode words. On this subtest, students are given a list of 50 words that children are expected to be able to read at their grade level. Learners are given 60 seconds and are asked to read as many familiar words out loud as they can. Responses ranged from zero to 50 familiar words read correctly, with a mean of 33.0 (65.9%) for Grade 2 students and a mean of 40.7 (81.5%) for Grade 3 students. The graph below shows that the distribution of Grade 2 and Grade 3 Filipino familiar word scores are both skewed to the left. As seen in the figure, by Grade 3, nearly all learners (81.0%) are able to read between 61 and 100% of the words correctly compared to roughly two-thirds of learners at Grade 2. Analysis by sex showed that girls in both grades significantly outperform boys (p<.001).

Figure 44. Filipino Familiar Word Identification Subtest Results, by Grade

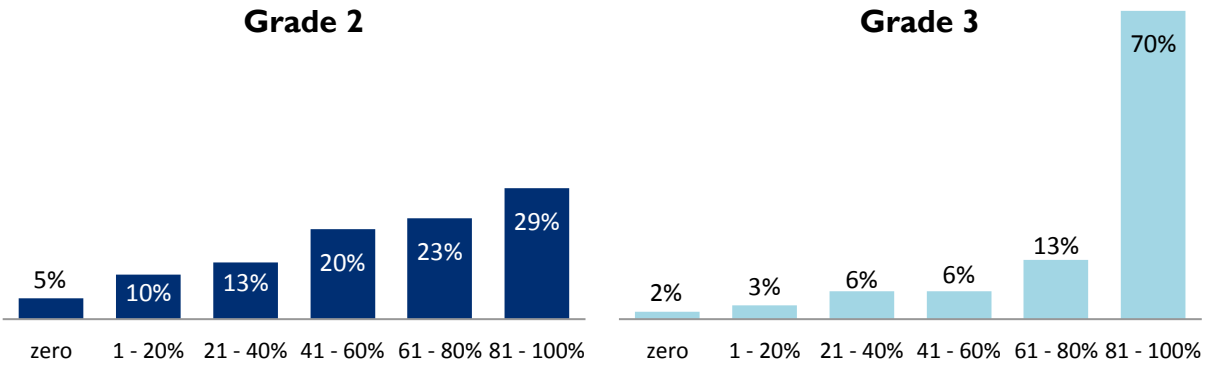


ORAL READING FLUENCY AND READING COMPREHENSION

To measure oral reading fluency, the ability to read quickly and accurately with proper expression, and reading comprehension, Grade 2 and Grade 3 learners were asked to read aloud a grade-level text in Filipino and to answer reading comprehension questions. On the Filipino passage reading and comprehension subtest, students were scored on the number of words they read correctly in the passage (total possible 64 words for Grade 2 and 55 words for Grade 3), reading comprehension (total possible 5), and prosody (total possible 4).

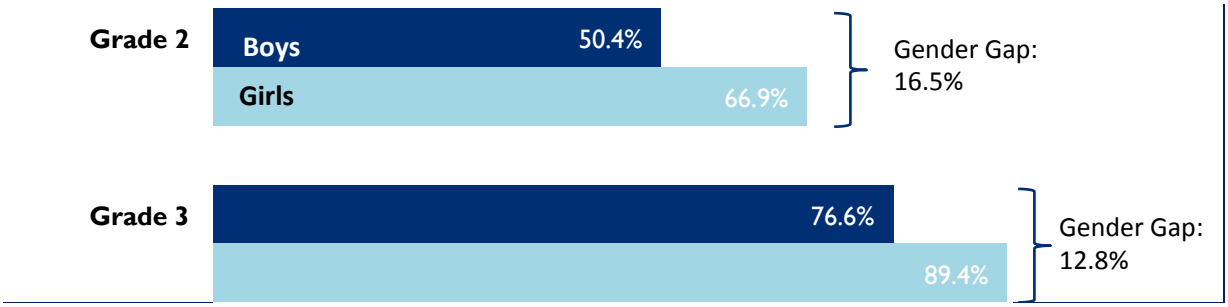
Oral Passage Reading Percent Correct. Oral Passage Reading scores in Filipino ranged from zero to 100% with a mean of 58.7% for Grade 2 students and a mean 83.0% for Grade 3 students. The results presented in the graph below show that by Grade 3 the large majority of learners are able to read the majority of the grade-level oral passage reading text correctly. Comparisons across grade show that there is a large improvement in oral passage reading from Grade 2 to Grade 3 in which only 29% of Grade 2 students were able to read 81% or more of the grade-level text compared to 70% at Grade 3.

Figure 45. Filipino Oral Passage Reading Results, Percent of Words Read Correctly, by Grade



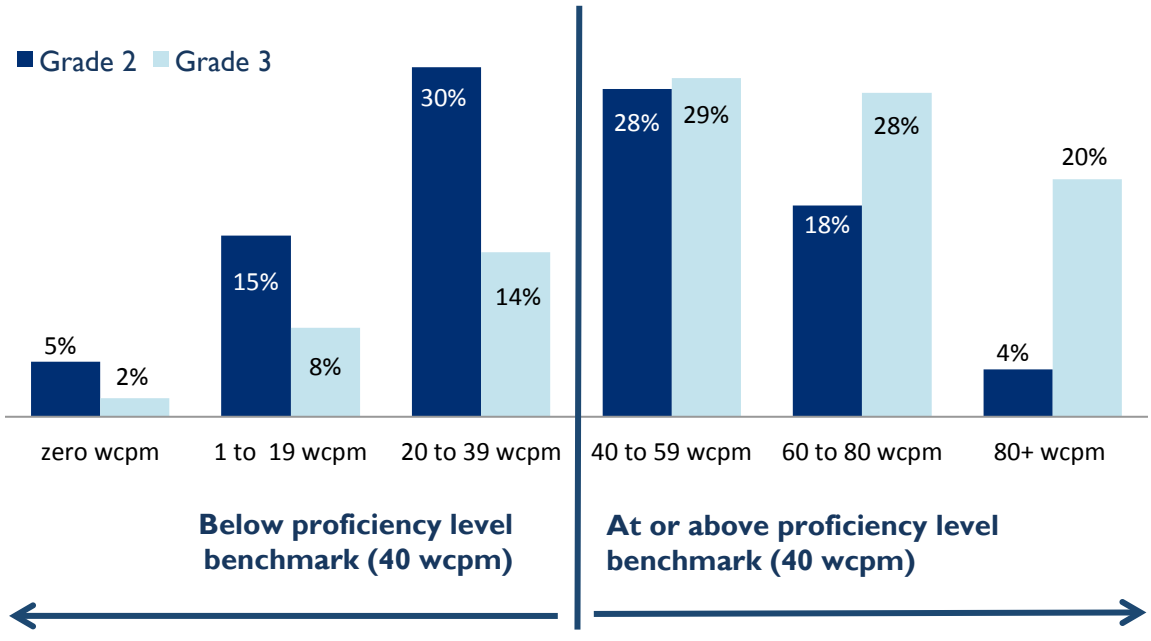
Similar to the familiar words subtests, girls in both Grades 2 and 3 significantly ($p<.001$) outperformed boys in the oral reading subtest. However, as seen below, the difference in boys and girls oral passage reading scores is shrinking slightly from Grade 2 to Grade 3. Girls in Grade 2, read, on average, 16.5% more of the text compared to boys; in Grade 3, girls, on average, read 12.8% more of the text than boys. This shows that the gender gap in oral passage reading is getting slightly smaller as students advance from Grade 2 to Grade 3.

Figure 46. Filipino Oral Passage Reading Average Percent Correct, by Grade and Sex



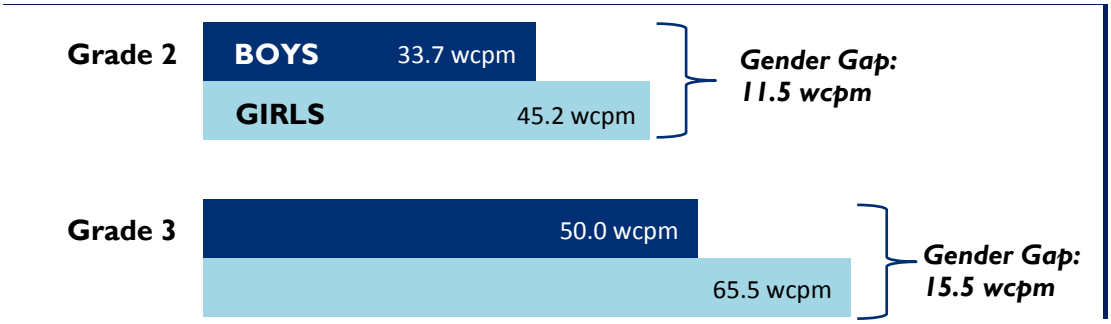
Fluency (Words correctly read per minute). Students were timed on reading the text, with the limit of 60 seconds. The number of words read correctly was divided by the seconds it took to read and then multiplied by 60 seconds to find the number of correct words per minute, which is the standard fluency measure used to measure USAID reading interventions. Grade 3 students on average read 57.8 words correct per minute, compared to 39.4 words correct per minute for Grade 2 students. The graph below shows the distribution of fluency scores by grade. As seen below, 77% of students in Grade 3 were reading at the proposed Filipino proficiency level of 40 wcpm, compared to 50% for students in Grade 2.

Figure 47. Filipino Oral Passage Reading Subtest Results for Cohort 3 (SY 15/16) by Grade



There were significant differences (p<.001) in fluency rates between boys and girls for both Grade 2 and Grade 3 students, in which, on average, girls read at a faster rate than boys. Analysis suggests that girls are not only demonstrating higher fluency rates than boys but are also improving in fluency faster than boys. As seen in the figure below, the girls in Grade 3 can read on average 20.3 more words per minute than girls in Grade 2, while Grade 3 boys can read 16.3 more words per minute on average than boys in Grade 2. These findings suggest that the gender gap in Filipino oral reading fluency is widening slightly from Grade 2 to Grade 3.

Figure 48. Filipino Fluency Measure (wcpm) for Cohort 3 (SY 15/16), by Sex and Grade

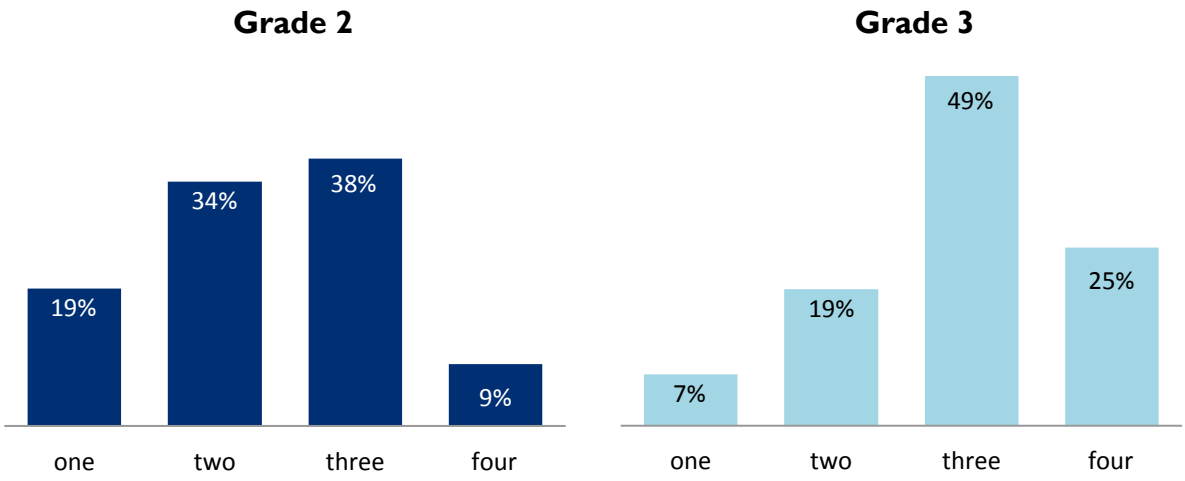


The **prosody scores** measured the degree to which students were able to read with intonation and expression. Prosody measures were divided into four categories and scored accordingly:

1 = word-by-word, slow, laborious;
2 = small chunks, awkward;
3 = fluent, but does not mark punctuation; incorrect phrase groups, no expression;
4 = fluent, with expression to mark punctuation and/or direct speech.

Students reading the Filipino passage ranged from one to four in prosody, with a mean of 2.9 for Grade 3 students and 2.4 for Grade 2 students. The graphs below show the distribution in students' prosody for both cohorts. As seen below, improvements can be seen in prosody from Grade 2 to Grade 3. The distribution shows for Grade 3 that nearly three-quarters (74%) of students read fluently (scored 3 or 4), compared to Grade 2, in which only 47% read fluently. By Grade 3, very few students (7%) read laboriously (score of 1), which is a substantial reduction from Grade 2, where roughly a fifth of students read laboriously in Filipino.

Figure 49. Filipino Oral Passage Reading - Prosody Score, by Grade

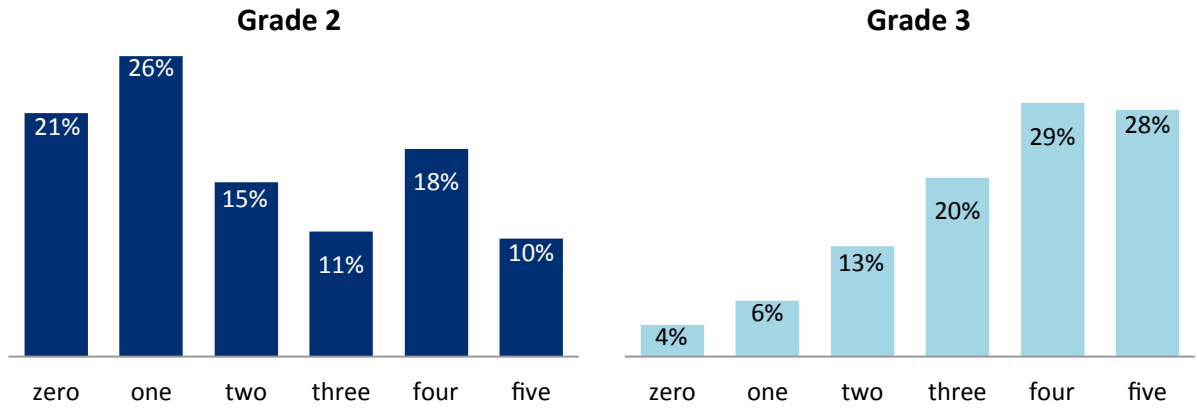


Disaggregation by sex showed that girls from both groups read the Filipino passage with greater prosody than boys, which is consistent with the overall pattern of girls demonstrating higher reading proficiency.

Reading Comprehension. Comprehension is the ultimate goal of reading. To measure comprehension, students were asked five questions about the text after reading a grade-level passage. Learners were not allowed to look back at the text to help them answer questions. Analysis shows learners in Grade 3 are able to answer more reading comprehension questions on average. For Grade 2, the total number of reading comprehension correct ranged from zero to five, with a mean of 2.1 questions correct (41.8%). Grade 3 learners on average answered 3.5 (69.8%) reading comprehension questions correctly.

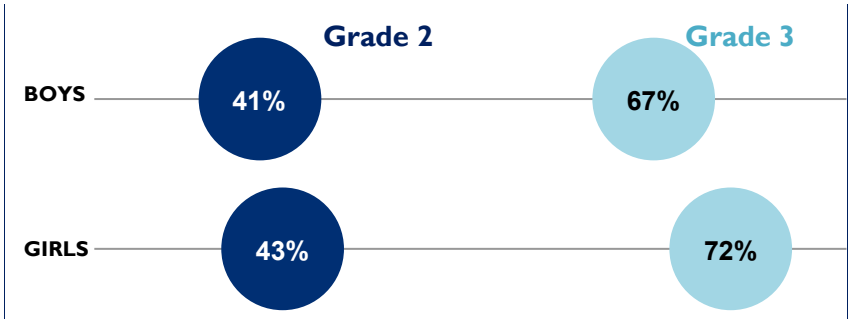
The figures below show the distribution of reading comprehension results by grade. Students in Grade 3 were able to answer more reading comprehension questions correctly than students in Grade 2. More than half (57%) of Grade 3 students answered four or more comprehension questions correctly compared to a quarter (28%) of students in Grade 2. The percent of learners who were unable to answer a single reading comprehension question correctly is substantially smaller for Grade 3 students, with only 4% of learners who were unable to answer a single reading comprehension question compared to one fifth (21%) of Grade 2 students.

Figure 50. Reading Comprehension Results, Cohort 3 (SY 15/16) Filipino EGRA, by Grade



Comparisons by sex show that in Grade 2, boys and girls performed similarly in Filipino reading comprehension. However, in Grade 3, results showed that girls significantly (p<.001) outperformed boys in reading comprehension at the end of a school year. Detailed results by sex are found in Annex 3.

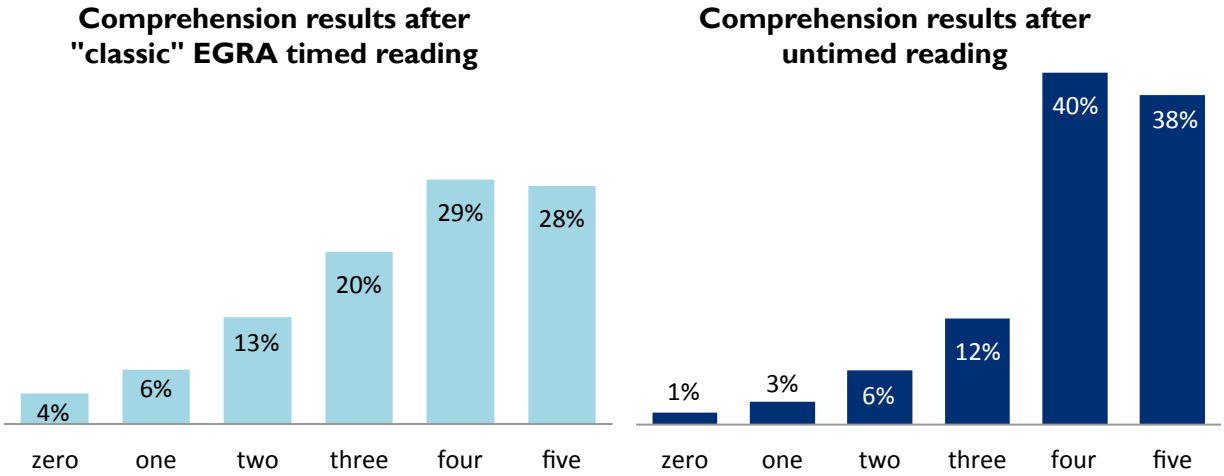
Figure 51. Average Percent Correct on Filipino Reading Comprehension (SY 15/16), by Sex and Grade



Filipino reading comprehension questions were administered in two rounds. The first round the administration followed the standard EGRA administration procedures. The second round immediately followed the first round. During the second round, assessors gave the text back to the students and allowed them to finish reading the passage[15] (if they haven't done so already), and then asked them comprehension questions without taking the text away from the students.

The results of the comparison between the two models of testing students' comprehension skills are presented in Figure 52. As seen in the figure, the percent of students who were able to meet comprehension proficiency benchmark of 60% increased by twenty percentage points (from 57% to 78%).

Figure 52. Comparison of Grade 3 Comprehension Results Using Timed and Untimed Reading (SY 15/16)[16]



As seen in Grade 2 results, these findings suggest that reading comprehension results from "classic" EGRA assessments must be interpreted with caution since scores might be as much impacted by the process of administration as by the abilities of the tested students.

LISTENING COMPREHENSION AND DICTATION

On the **listening comprehension subtest**, Grade 2 and Grade 3 students were read a passage and asked five comprehension questions. On average, for Grade 3, students answered 49.4% of listening comprehension correctly compared to 34.7% for Grade 2. The figure below shows distributions of the number of listening comprehension questions Grade 2 and Grade 3 students were able to answer correctly. In Grade 2, distributions

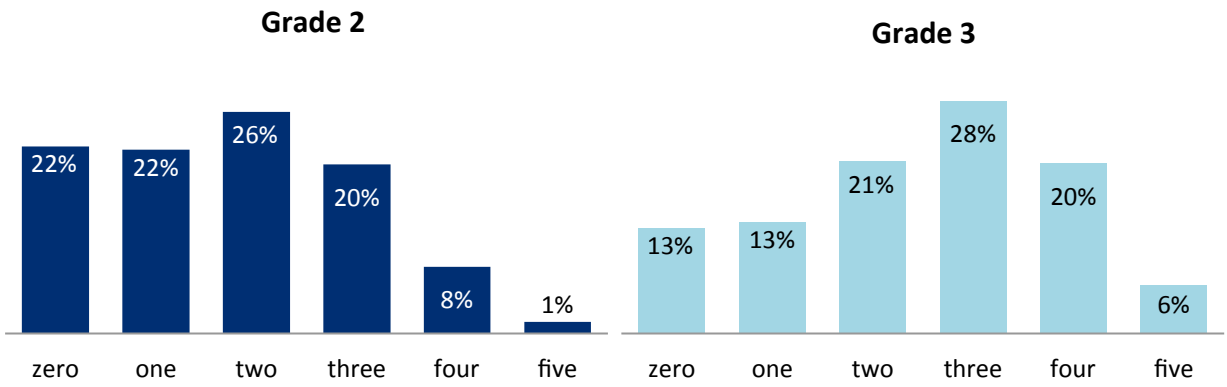


[15] Students who answered four or five comprehension questions during the first round were not asked to read the passage again nor answer reading comprehension questions again given that they demonstrated reading comprehension during the first around of administration.

[16] In order to compare results from the first (timed) round and the second (untimed) round, untimed comprehension results include data for students who, during the second round, were allowed to read the passage (untimed) and answer the comprehension questions, as well as, results for students who answered four or more comprehension questions correctly during the first round, and as a result did not participate in the second, untimed, round.

of learners' listening comprehension scores were skewed to the right, with the majority (70%) of learners answering 2 or less comprehension questions correctly. Very few students (9%) were able to answer four or five listening comprehension questions. Conversely, in Grade 3, listening comprehension scores were skewed to the left, with more than half of learners who were able to answer three or more questions correctly.

Figure 53. Filipino Listening Comprehension Questions – Number Correct, by Grade (SY 15/16)



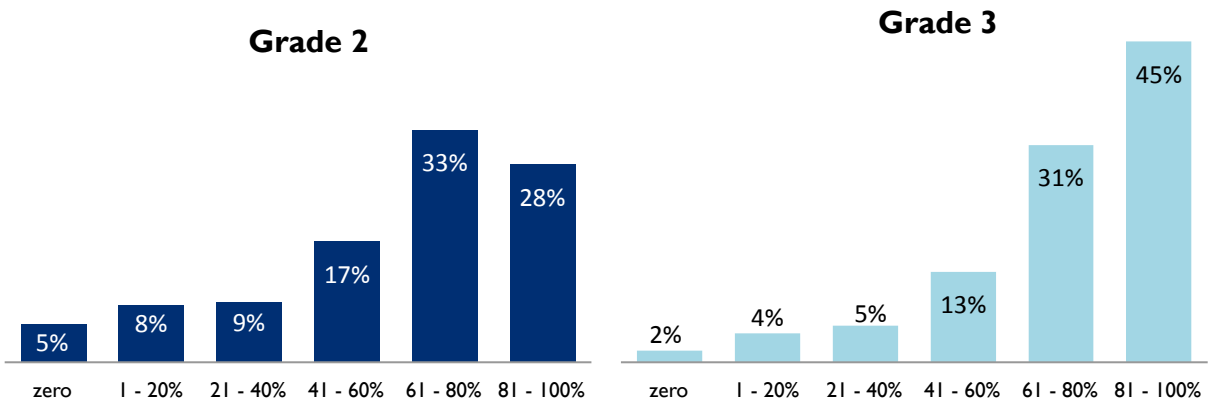
Listening comprehension results by sex showed that, as seen in other subtests, girls in performed better than boys in both Grade 2 and Grade 3; the difference between boys and girls was statistically significant ($p<.001$).

The **dictation subtest** is used to measure both oral comprehension and writing skills of learners. Writing skills are strongly connected to reading skills in that it takes phonological awareness skills to translate sounds into individual graphemes and the written word. The dictation subtest was the same for both Grades 2 and 3. On the dictation subtest (total possible 16 correct answers for the Filipino test), students were read a passage once, given a pencil and paper, and read the passage a second time with pauses, then read the entire passage a third time. Students wrote the words on the paper. Dictation scores were broken up into two subtests:

- Number of words spelled correctly (total possible 12)
- Other items relating to conventions of text in writing including, spacing, text direction, capital letter, and using a period at the end of a sentence (total possible 4)

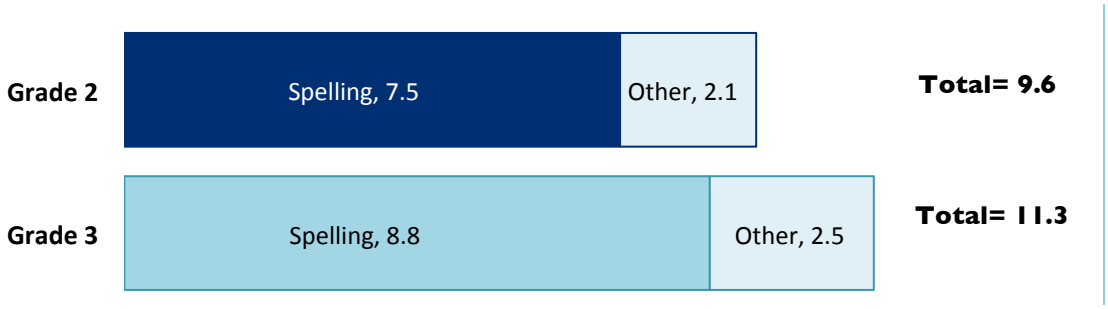
Number of correct answers for the dictation subtest ranged from zero to 16, with a mean of 9.6 (59.7%) for Grade 2 students and a mean of 11.3 (70.8%) for Grade 3 students. Figure 54 below shows the distribution of dictation results for Grade 2 and Grade 3. Overall, distribution results show that Grade 3 students have significantly ($p<.001$) higher Filipino dictation scores than Grade 2 learners. In Grade 2, as seen in the figure below, nearly two-thirds of students scored over 60%; roughly a quarter (28%) scored between 80-100%. For Grade 3, three-quarters of learners scored over 60% on the dictation subtest, with nearly half (45%) scoring between 80-100%.

Figure 54. Filipino Dictation Results (SY 2015/16), by Grade



Breaking down the dictation composite, scores for spelling ranged from zero to 12, with a mean of 7.5 for Grade 2 students and 8.8 for Grade 3 students. With regard to the conventions of text, Grade 3 students on average scored 2.5 out of 4, while Grade 2 students scored on average 2.1 out of possible 4. Most students used spacing and the direction of the text correctly. However, only one third of Grade 2 students and half of Grade 3 students capitalized correctly. Only 15.1% of Grade 2 students and roughly a quarter (24.8%) of Grade 3 students used a period at the end of the sentence.

Figure 55. Filipino Dictation Breakdown (Out of Possible 16 Points) (SY 15/16), by Grade



Disaggregating by sex, girls significantly outperform boys in dictation. This is not surprising given girls have outperformed boys in all other subtests.

GRADES 2 AND 3 ENGLISH SUBTEST RESULTS

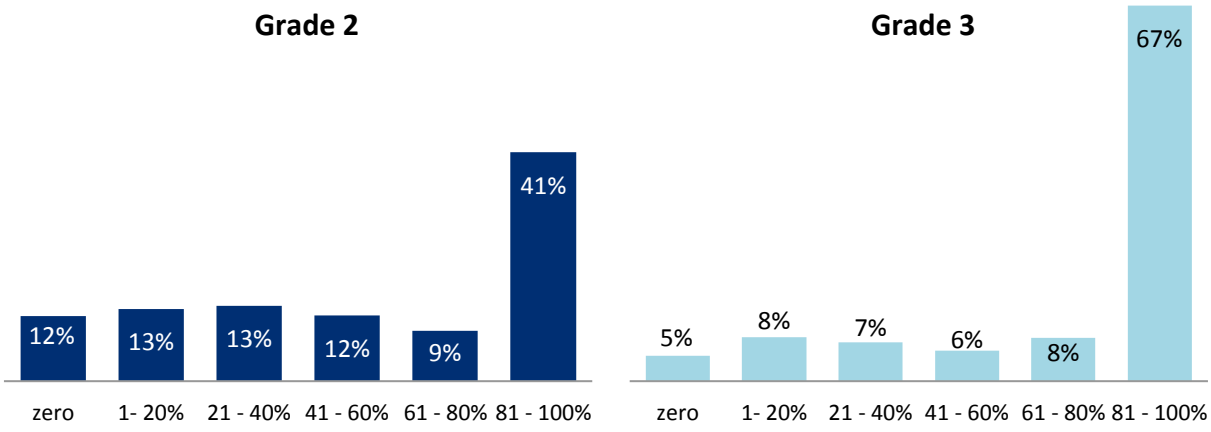
The following section provides detailed results for English EGRA subtests for Grade 3 students compared to Grade 2 students in Cohort 3 (SY 2015/16).

FAMILIAR WORDS

On the **familiar word identification** subtest, students are given a list of 50 English words that children are expected to be able to read at their grade level. Learners are given 60 seconds and are asked to read as many familiar words out loud as they can. Responses ranged from zero to 50 familiar words read correctly. On average, results showed that Grade 3 learners performed significantly ($p<.001$) better than Grade 2 learners in English familiar word reading, with a mean of 37.9 (75.8%) for Grade 3 learners and a mean of 28.4 (56.7%) for Grade 2 students. The graphs below show that the distribution of Grade 2 and Grade 3 English familiar word scores are both

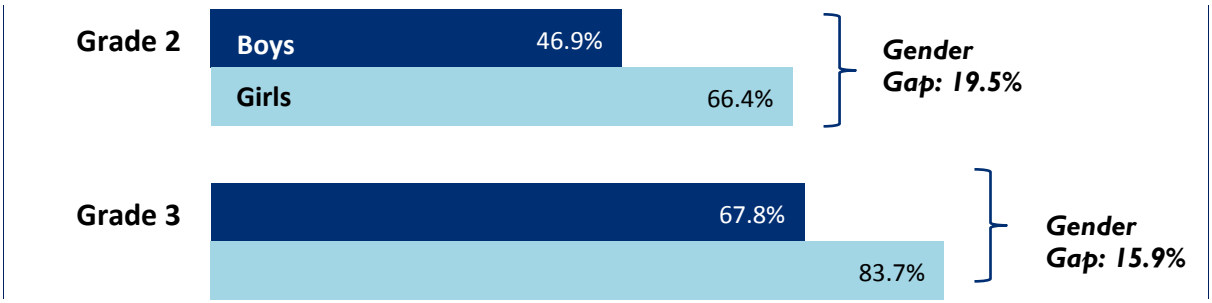
skewed to the left. As seen in the figure, by Grade 3, the majority learners (75.0%) are able to read between 61 and 100% of the words correctly compared to half of learners at Grade 2.

Figure 56. English Familiar Word Identification Subtest Results, by Grade



Analysis by sex showed that girls in both grades significantly outperform boys ($p<.001$). In fact, boys appear to be falling substantially behind. Results show that, on average, Grade 3 boys are performing similarly to Grade 2 girls in English familiar word identification. Despite this fact, the gender gap appears to be shrinking from Grade 2 to Grade 3, in which the gap decreased from 19.5% to 15.9%.

Figure 57. Average Percent of English Familiar Words Identified (SY 15/16), by Grade and Sex



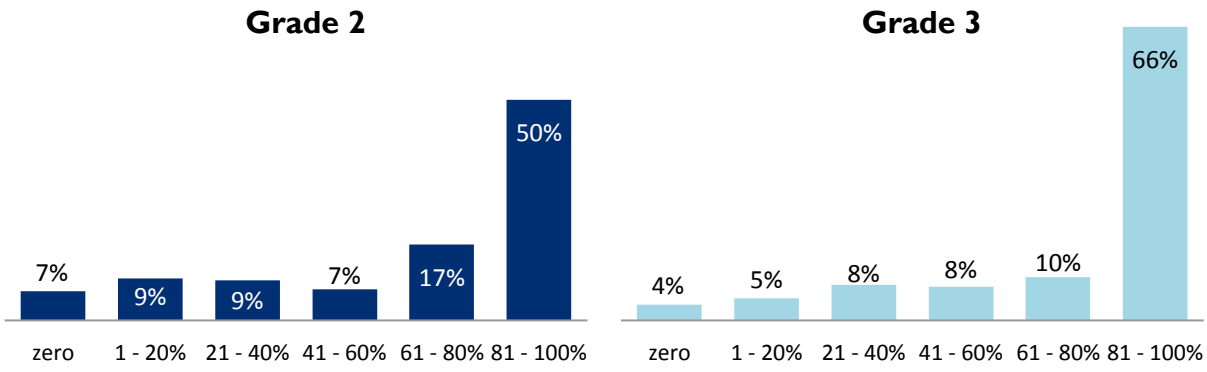
ORAL READING FLUENCY AND READING COMPREHENSION

To measure oral reading fluency, the ability to read quickly and accurately with proper expression, and reading comprehension, Grade 2 and Grade 3 learners were asked to read aloud a grade-level text in English and to answer reading comprehension questions. On the English passage reading and comprehension subtest, students were scored on the number of words they read correctly in the passage (total possible 59 words for Grade 2 and 61 words for Grade 3), reading comprehension (total possible 5), and prosody (total possible 4).



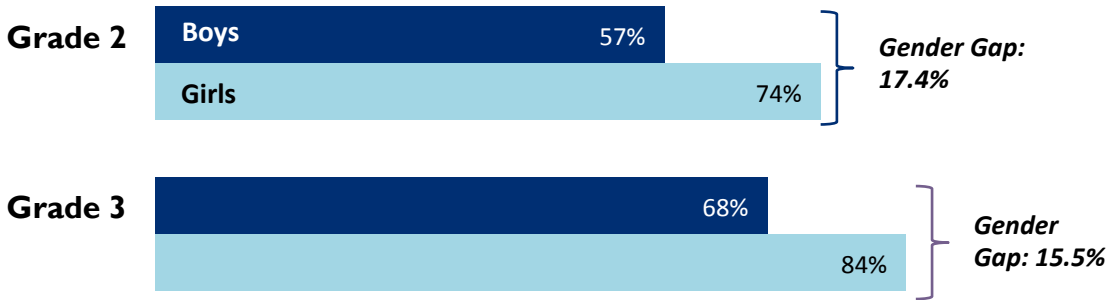
Oral Passage Reading Percent Correct. Oral Passage Reading scores in English ranged from zero to 100% with a mean of 65.5% for Grade 2 students and a mean 76.0% for Grade 3 students. Comparisons across grade show that there is a slight improvement in English oral passage reading scores from Grade 2 to Grade 3. The results presented in the graphs below show that by Grade 3 roughly two-thirds (66%) of learners are able to read the majority of the grade-level oral passage reading text correctly compared to half of Grade 2 learners.

Figure 58. English Oral Passage Reading Results, Percent of Words Read Correctly, by Grade



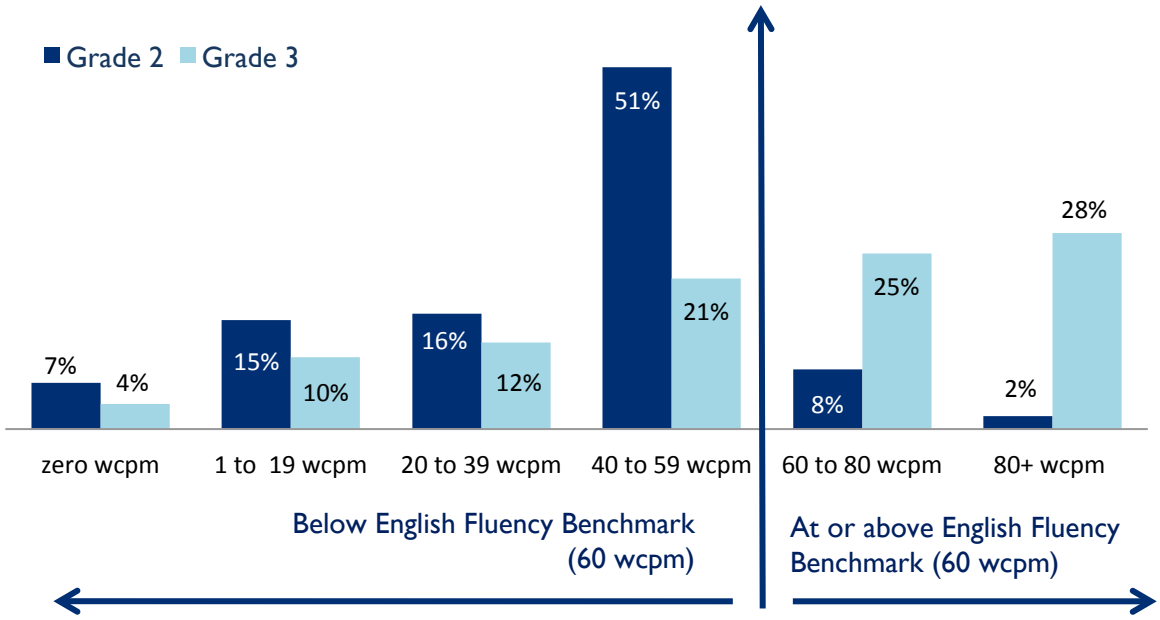
Similar to the familiar words subtests, girls in both Grades 2 and 3 significantly ($p<.001$) outperformed boys in the oral reading subtest. However, as seen below, the difference in boys and girls oral passage reading scores is shrinking slightly from Grade 2 to Grade 3. Girls in Grade 2, read, on average, 17.4% more of the text compared to boys; in Grade 3, girls, on average, read 15.5% more of the text than boys. This suggests that the gender gap in English oral passage reading is getting slightly smaller as students advance from Grade 2 to Grade 3.

Figure 59. English Oral Passage Reading Average Percent Correct (SY 15/16), by Grade and Sex



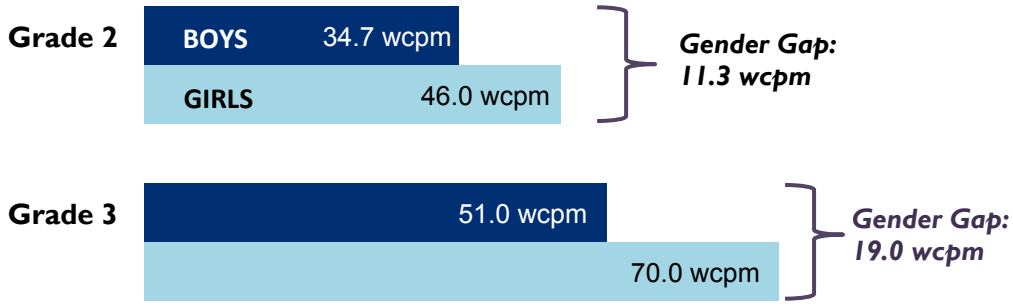
Fluency (Words correctly read per minute). Students were timed on reading the text, with the limit of 60 seconds. The number of words read correctly was divided by the seconds it took to read and then multiplied by 60 seconds to find the number of correct words per minute, which is the standard fluency measure used to measure USAID reading interventions. Grade 3 students on average read 60.5 words correct per minute, compared to 40.4 words correct per minute for Grade 2 students. The graph below shows the distribution of English fluency scores by grade. As seen below, 53% of students in Grade 3 were reading at the proposed proficiency level of 60 wcpm, compared to only 10% for students in Grade 2.

Figure 60. English Oral Passage Reading Subtest Results for Cohort 3 (SY 15/16) by Grade



There were significant differences (p<.001) in fluency rates between boys and girls for both Grade 2 and Grade 3 students, in which, on average, girls read at a faster rate in English than boys. Analysis suggests that girls are not only demonstrating higher fluency rates than boys but are also improving in fluency faster than boys from Grade 2 to Grade 3. As seen in the figure below, the gender gap between boys and girls in English oral reading fluency nearly doubled from Grade 2 to Grade 3. These findings suggest that the gender gap in English oral reading fluency is widening from Grade 2 to Grade 3.

Figure 61. English Fluency Measure (wcpm) for Cohort 3 (SY 15/16), by Sex and Grade



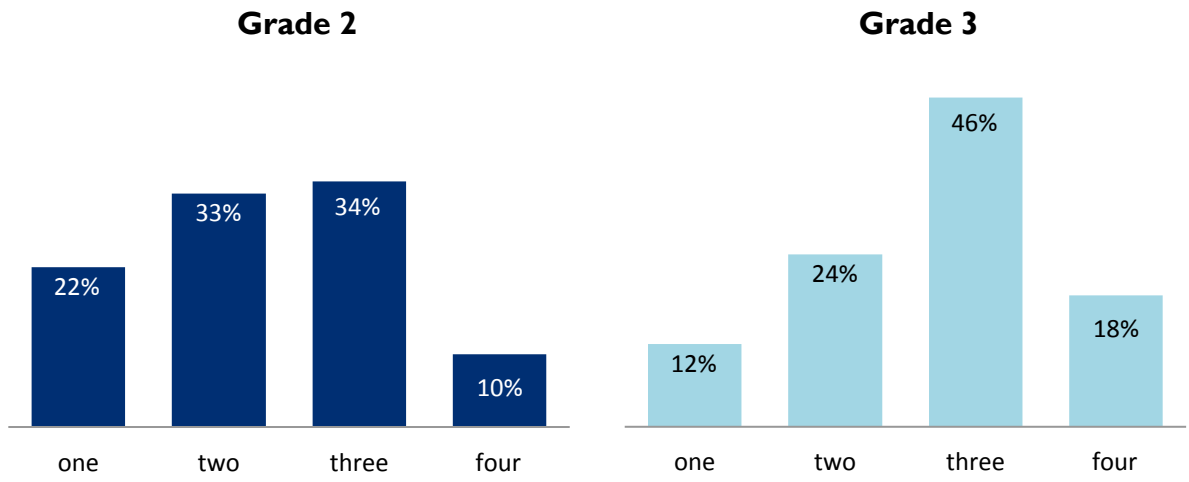
The **prosody scores** measured the degree to which students were able to read with intonation and expression. Prosody measures were divided into four categories and scored accordingly:

1 = word-by-word, slow, laborious;
2 = small chunks, awkward;
3 = fluent, but does not mark punctuation; incorrect phrase groups, no expression;
4 = fluent, with expression to mark punctuation and/or direct speech.

Students reading the English passages ranged from one to four in prosody, with a mean of 2.7 for Grade 3 students and 2.3 for Grade 2 students. The graphs below show the distribution in students'

prosody for both cohorts. As seen below, improvements can be seen in prosody from Grade 2 to Grade 3. The distribution shows for Grade 3 that nearly two-thirds (64%) of students read fluently (scored 3 or 4), compared to Grade 2, in which only 44% read fluently. By Grade 3, only 12% of students read laboriously (score of 1), which is a reduction from Grade 2, where roughly a fifth of students read laboriously in English.

Figure 62. English Oral Passage Reading - Prosody Score (SY 15/16), by Grade

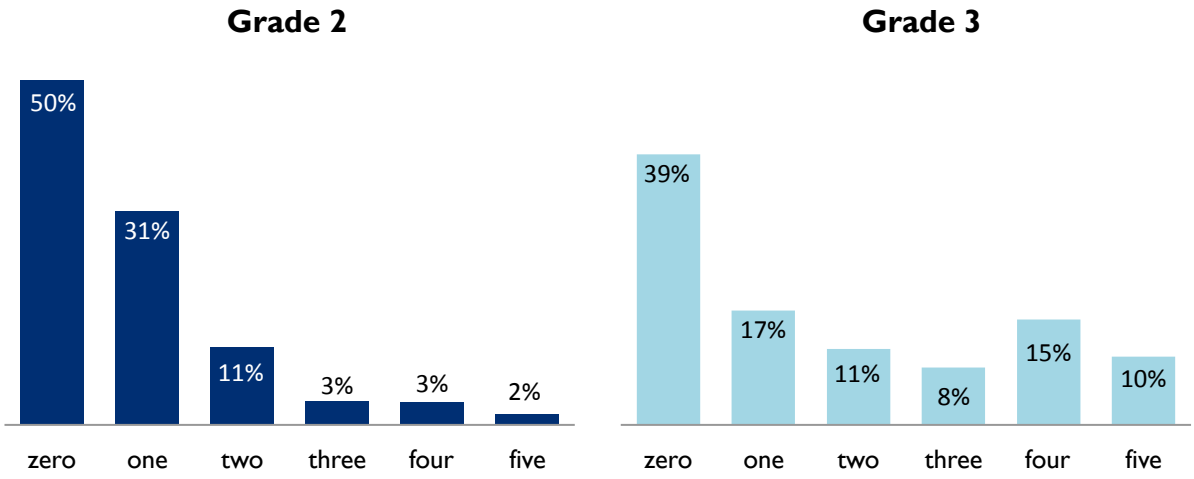


Disaggregation by sex showed that girls from both groups read the Filipino passage with greater prosody than boys, which is consistent with the overall pattern of girls demonstrating higher reading proficiency.

Reading Comprehension. To measure reading comprehension in English, students were asked five questions about the text after reading a grade-level passage. Learners were not allowed to look back at the text to help them answer questions. Analysis shows that despite proficient oral reading fluency and prosody scores, learners in both Grades 2 and 3, on average, demonstrated relatively low performance in English reading comprehension. Learners in Grade 3 are able to answer more reading comprehension questions on average than Grade 2 students. For Grade 2, the total number of reading comprehension correct ranged from zero to five, with a mean of 0.8 questions correct (16.8%). Grade 3 learners on average answered 1.7 (34.7%) reading comprehension questions correctly.

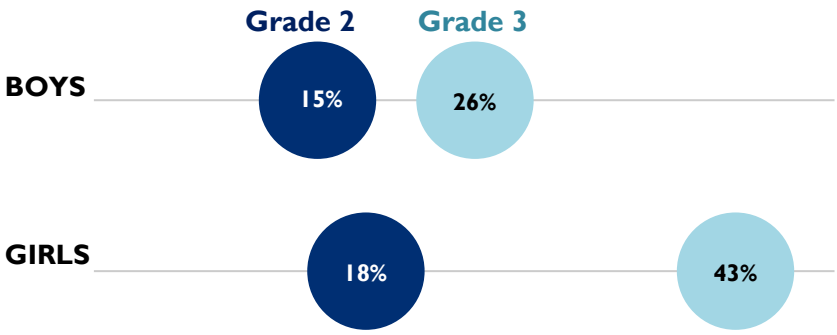
The figures below show the distribution of English reading comprehension results by grade. Students in Grade 3 were able to answer more reading comprehension questions correctly than students in Grade 2. Half of Grade 2 students and 39% of Grade 3 students were unable to answer a single English reading comprehension question. Less than a fifth (19%) of Grade 2 students were able to answer more than one reading comprehension question compared to 44% of Grade 3 students. Very few (5%) Grade 2 students were able to answer four or more questions; while a quarter of Grade 3 students demonstrated reading comprehension and could answer four or more questions.

Figure 63. Reading Comprehension Results, Cohort 3 (SY 15/16) English EGRA, by Grade



Comparisons by sex show that in Grade 2, girls significantly ($p<.01$) outperformed boys in English reading comprehension. However, as seen in the figure below, in Grade 3, the difference between girls and boys average reading comprehension results were substantial, in which the gender gap more than doubled from Grade 2 to Grade 3. Grade 3 girls on average scored 25% higher than boys in English reading comprehension, compared to 11% in Grade 2. Detailed results by sex are found in Annex 3.

Figure 64. Average Percent Correct on English Reading Comprehension (SY 15/16), by Sex and Grade

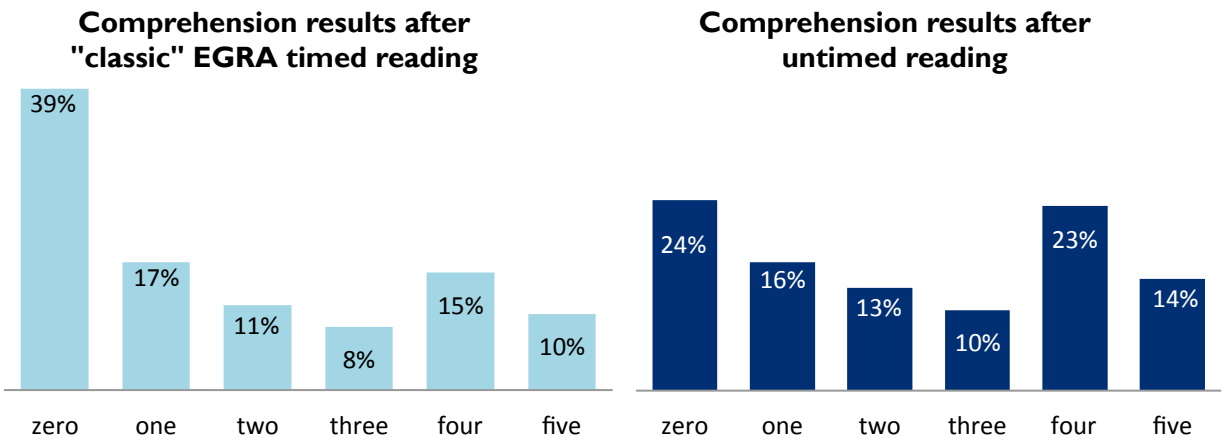


Similar to the Filipino EGRA, reading comprehension questions were administered in two rounds. The first round the administration followed the standard EGRA administration procedures. The second round immediately followed the first round. During the second round, assessors gave the text back to the students and allowed them to finish reading the passage[17] (if they haven't done so already), and then asked the student comprehension questions without taking the text away.

The results of the comparison between the two models of testing Grade 3 students' comprehension skills are presented in Figure 65. As seen in the figure, the percent of students who were able to meet comprehension proficiency benchmark (4 or more comprehension questions correctly) increased by twelve percentage points (from 37% to 25%) from Grade 2 to Grade 3.

[17] Students who answered four or five comprehension questions during the first round were not asked to read the passage again nor answer reading comprehension questions again given that they demonstrated reading comprehension during the first around of administration.

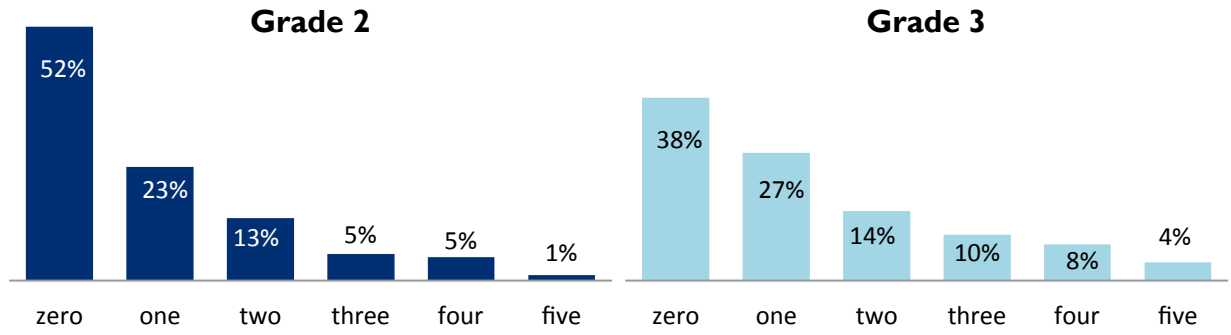
Figure 65. Comparison of Grade 3 Comprehension Results Using Timed and Untimed Reading (SY 15/16)[18]



LISTENING COMPREHENSION AND DICTATION

On the **listening comprehension subtest**, Grade 2 and Grade 3 students were read a passage and asked five comprehension questions. On average, results showed significant ($p<.001$) increases in listening comprehension in English from Grade 2 to Grade 3. Grade 3, students answered 26.7% of listening comprehension correctly compared to 18.1% for Grade 2. The figure below shows distributions of the number of listening comprehension questions Grade 2 and Grade 3 students were able to answer correctly. Analysis of distribution results showed that in both Grade 2 and Grade 3, scores were skewed to the right, with the majority of learners answering less than two listening comprehension questions correctly. In Grade 2, roughly half of learners were unable to answer a single listening comprehension question correctly; this percentage decreased for Grade 3 learners, where roughly a third of learners had a zero score.

Figure 66. English Listening Comprehension Questions – Number Correct, by Grade (SY 15/16)



Listening comprehension results by sex showed that in Grade 2, girls and boys performed similarly on the listening comprehension subtests. However, for Grade 3, as seen in other subtests, girls

[18] In order to compare results from the first (timed) round and the second (untimed) round, untimed comprehension results include data for students who, during the second round, were allowed to read the passage (untimed) and answer the comprehension questions, as well as, results for students who answered four or more comprehension questions correctly during the first round, and as a result did not participate in the second, untimed, round.

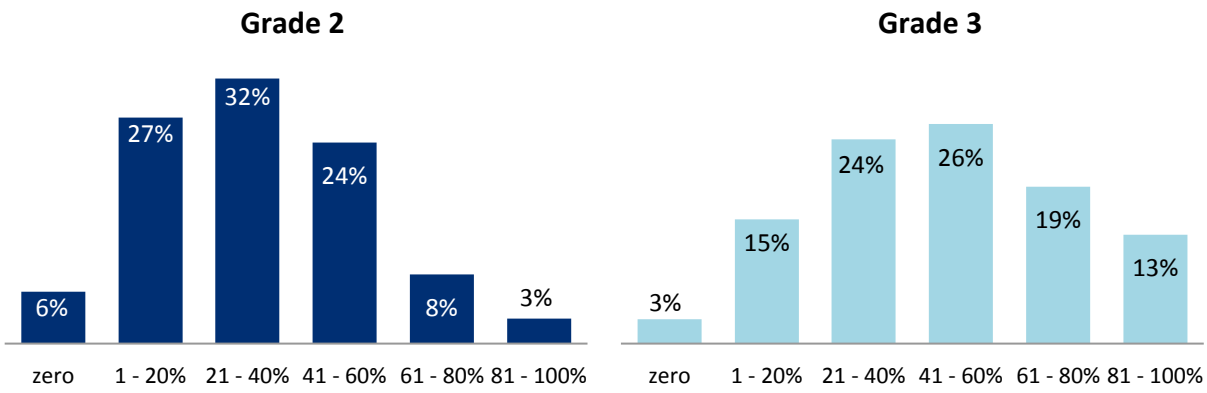
performed better than boys; the difference between boys and girls was statistically significant (p<.001).

On the **dictation subtest** (total possible 17 correct answers for the English test), students were read a passage once, given a pencil and paper, and read the passage a second time with pauses, then read the entire passage a third time. Students wrote the words on the paper. The English dictation subtest was the same for both Grades 2 and 3. Dictation scores were broken up into two subtests:

- Number of words spelled correctly (total possible 13)
- Other items relating to conventions of text in writing including, spacing, text direction, capital letter, and using a period at the end of a sentence (total possible 4)

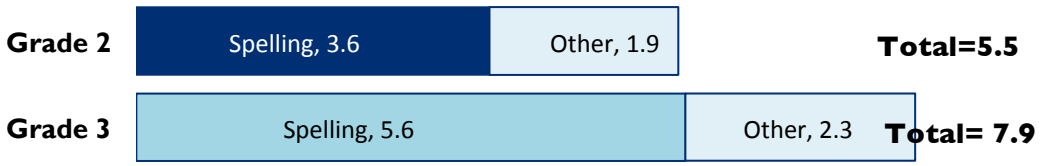
Overall English dictation results showed Grade 3 students have significantly (p<.001) higher dictation scores than Grade 2 students. The number of correct answers for the English dictation subtest ranged from zero to 17, with a mean of 5.5 (32.5%) for Grade 2 students and a mean of 7.9 (46.7%) for Grade 3 students. Figure 67 below shows the distribution of dictation results for Grade 2 and Grade 3. In Grade 2, as seen in the figure below, dictation results were skewed to the right, with roughly two-thirds of students who scored 40% or less; only 11% of learners scored between 61-100%. However, for Grade 3, improvements in English dictation results can be seen with the distribution of results shifting slightly, in which, nearly a third of learners scored between 61-100%.

Figure 67. English Dictation Results (SY 2015/16), by Grade



Breaking down the dictation composite, scores for spelling ranged from zero to 13, with a mean of 3.6 for Grade 2 students and 5.6 for Grade 3 students. With regard to the conventions of text, Grade 3 students on average scored 2.3 out of 4, while Grade 2 students scored on average 1.9 out of possible 4. Most students used spacing and the direction of the text correctly. One third of Grade 2 students and roughly half of Grade 3 students capitalized correctly. Very few learners used a period at the end of the sentence. Only 11.9% of Grade 2 students and roughly a fifth (21.3%) of Grade 3 students used a period at the end of the sentence.

Figure 68. English Dictation Breakdown (SY 15/16) (Out of possible 17), by Grade



Disaggregating by sex, girls significantly outperform boys in dictation. This is not surprising given girls have outperformed boys in all other subtests.

5. IMPACT OF CONTEXTUAL FACTORS ON ACHIEVEMENT

LEARNER CHARACTERISTICS AND EGRA RESULTS

Various factors from the student context interview examined for association with key outcomes: Filipino and English fluency and comprehension. Bivariate statistical analysis found small, but statistically significant, correlations between learner results in Filipino and English oral reading fluency and comprehension, and several learner context interview questions. Correlation analysis results are shown in the table below.[19] These findings are consistent with the results of the SY 2014/15 (Cohort 2) evaluation which found similar correlations between home environment, school and teacher, and socio-economic status composite variables.

Table 14. Correlations between Student Context Interview Results and Cohort 3 (SY15/16) EGRA Results

Student Context Interview Questions	Grade 2 Learners				Grade 3 Learners			
	Filipino		English		Filipino		English	
	fluency	comp.	fluency	comp.	fluency	comp.	fluency	comp.
Total household possessions	.081**	.120**	.121**	.170**	.095**	.126**	.139**	.216**
Does someone at home check your assignments/homework? (1=Yes)	.063*	.055*	.067**		.056*	.005	.071**	
Are you allowed to take books home from school? (1=Yes)								
Do you ever take books from school to read a home? (1=Yes)								
Do you have books at home? (1=Yes)						.093**	.023	.092**
Does your mother know how to read and write? (1=Yes)					.099**	.057*	.103**	.075**
Does your father know how to read and write? (1=Yes)		.061*	.052*					
Do you get to choose the story books you read at school? (1=Yes)	.049*		.059*	.026	.104**	.079**	.087**	
Did you have something to eat today before you came to school? (1=Yes)								
Did you go to kinder? (1=Yes)	.055*	.069**	.063*					
How often do you miss school? (1=Never, 2+Rarely, 3=Sometimes, 4=Often)								
Do you watch TV in English at home? (1=Yes)	.050*	.077**	.076**	.103**	.073**	.104**	.105**	.152**
Does someone read stories with you at home? (1=Yes)	-.055*	-.066*	-.065**					-.056*

*Correlations are significant at the >0.05 level (2-tailed)
**Correlations are significant at the >0.01 level (2-tailed)
Blanks denote no statistically significant association between variables.

[19] In social science research correlations below .2 are not considered to be of high importance. Correlations between .2 and .4 are considered small, correlations between .4 and .6 are moderate, and above .6 they are large.

Analysis showed that the strongest relationship between learner characteristics and performance on the Filipino and English EGRA, was found with the number of household possessions of the learner, which is a proxy for socioeconomic status. In that, there was a significant positive relationship between learners with more household possessions and higher achievement on Filipino and English EGRA for both Grade 2 and Grade 3. However, as seen in the table above the relationship was weak.



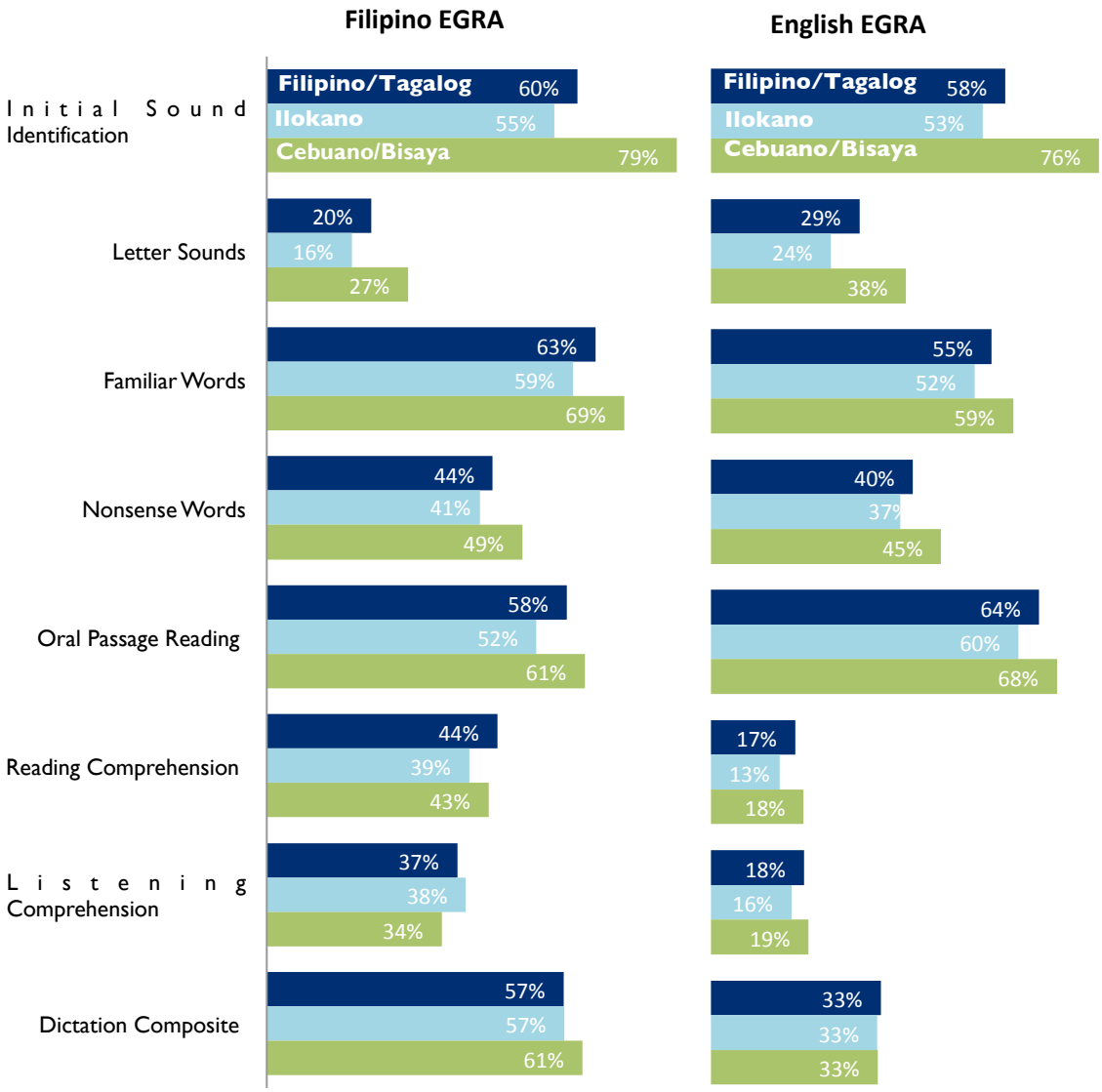
Parental literacy was also found to have a significant positive association with student performance. As Table 14 shows, maternal literacy, in particular, is associated with better results in both Filipino and English fluency and comprehension for Grade 3 learners; for Grade 2 learners, only paternal literacy had a significant correlation with Filipino comprehension and English fluency results.

Data analysis also revealed several weak positive relationships between higher assessment results in Filipino and English fluency and comprehension measures and other factors such as kindergarten attendance, having books at home, getting to choose the story books to read at school, watching TV in English. Of interest, kindergarten attendance was found to have a positive relationship with Filipino ORF and comprehension and English ORF for only Grade 2; attending kindergarten did not have a significant relationship with higher Grade 3 ORF and comprehension results. The availability of books at home had a weak correlation with Grade 3 Filipino and English comprehension results; however, did appear to have a significant relationship with Grade 2 results. Overall, watching TV in English, had one of the strongest relationships with higher Filipino and English oral reading fluency and comprehension measures for both grades; however, as seen in the table above, the relationships were weak. Interestingly, having someone read to learners at home had a significant negative relationship with Grade 2 Filipino ORF and comprehension and English ORF results; however, the relationship was very weak.

IMPACT OF HOME LANGUAGE

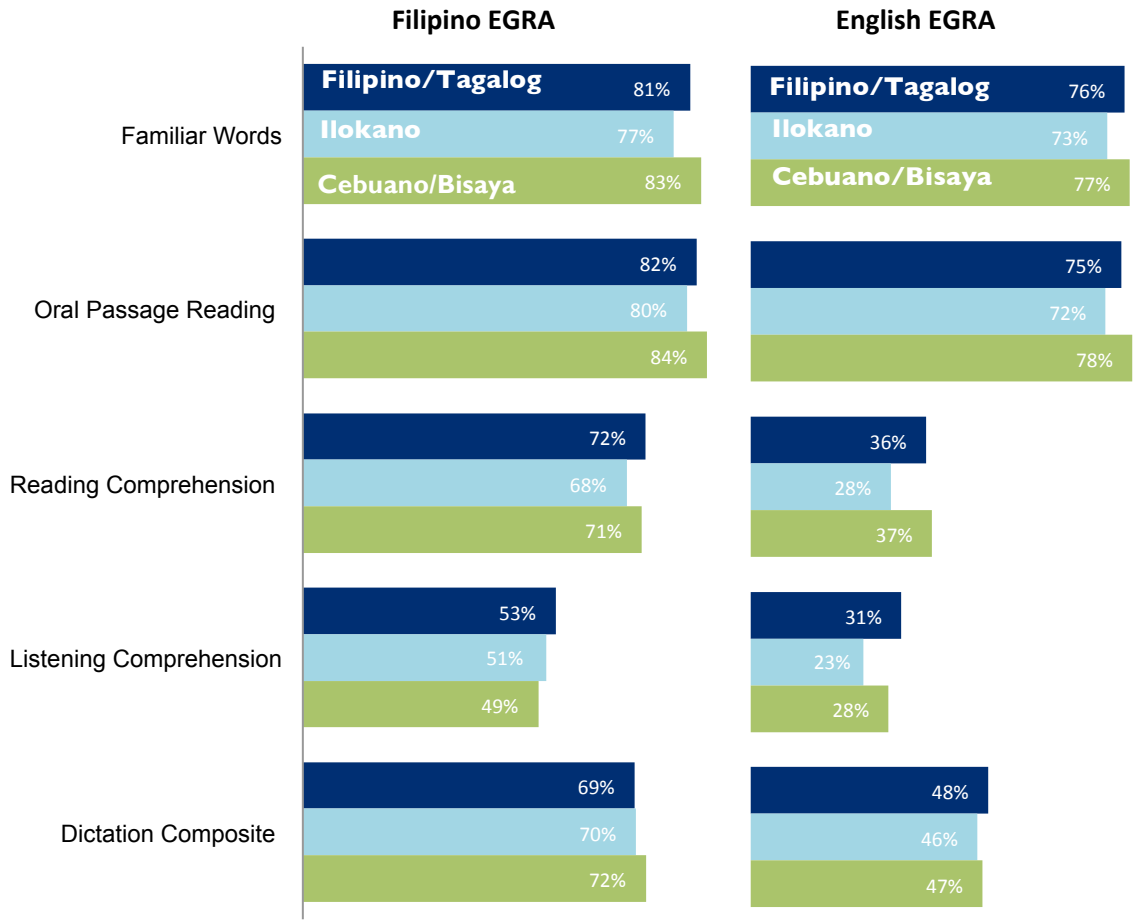
When Grade 2 EGRA results were disaggregated by language spoken at home, students who spoke Cebuano/Bisaya or Filipino/Tagalog at home performed better than students who reported speaking Ilokano at home on nearly all subtests on both the Filipino and English EGRA tests. Cebuano/Bisaya speakers performed the best in nearly every subtest on both tests. Differences between Cebuano/Bisaya speakers and Ilokano speakers were significant at the $p<.05$ level for both tests, with the exception of the English dictation subtest. Interestingly, Cebuano/Bisaya speakers also performed significantly better than Filipino/Tagalog speakers in 5 out of 8 Filipino subtests and 3 out of 8 English subtests. For both the Filipino and English EGRA, the largest differences between Filipino/Tagalog speakers and Cebuano/Bisaya speakers were in Initial Sound Identification and Letter Sounds.

Figure 69. Grade 2 EGRA Results for Cohort 3 (SY14/15) Disaggregated by Language Spoken at Home



For Grade 3, results were similar to Grade 2 in that Cebuano/Bisaya speakers largely performed better than Ilokano and Filipino/Tagalog speakers on both the Grade 3 Filipino and English EGRAs. Cebuano/Bisaya speakers performed significantly better ($p<.05$) than Ilokano speakers in all subtests with the exception of Filipino Listening Comprehension and Dictation and English Dictation. Cebuano/Bisaya speakers on average performed slightly better than Filipino speakers, however, differences were only statistically significant for the Filipino Familiar Words subtest. Filipino/Tagalog speakers only performed significantly better than Ilokano speakers in one subtest – English Listening Comprehension.

Figure 70. Grade 3 EGRA Results for Cohort 3 (SY14/15) Disaggregated by Language Spoken at Home



Ilokano's orthography has less similarities with Filipino than Cebuano. Therefore, Cebuano speakers may find it easier to transfer some aspects of phonics skills to Filipino and English than Ilokano speakers. This could explain the relatively higher scores for Cebuano speakers on initial sound identification, letter sounds and nonsense words. It may also explain how students are attaining high fluency scores despite the fact that they only started formal instruction in reading and writing Filipino at the start of grade two and English in the middle of second grade. However, vocabulary knowledge, which is crucial to both listening and reading comprehension, are not transferable across languages especially to English. Students need time to develop vocabulary in order to understand what they hear and read in Filipino and English, as a result they are not performing as well on comprehension tests in Grade 2. As seen in the figures, for both Filipino and English, comprehension scores have improved by Grade 3, which is likely due to additional instructional time in Filipino and English, which is needed to continue to develop vocabulary knowledge.

IMPACT OF SCHOOL LEADERSHIP

Bivariate statistical analysis found few statistically significant correlations between learner Filipino and English results and school leadership characteristics and practices. As seen in the table below, whether a principal has attended LAC training, self-reports sufficient knowledge to provide support to teachers and having time to provide instructional support to teachers had statistically significant positive relationships with nearly all Filipino and English fluency and comprehension measures, particularly for Grade 2.

Table 15. Correlations between School Leadership Practices and EGRA Results, Cohort 3 (SY15/16)

	Grade 2 Learners				Grade 3 Learners			
	Filipino		**English**		**Filipino**		**English**	
Classroom Observation[20]	fluency	comp.	fluency	comp.	fluency	comp.	fluency	comp.
Have you attended the School Head Orientation Training? (1=Yes)	.090**		.070*					
Have you attended the LAC Training? (1=Yes)	.069*	.104**	.080**		.078**	.071*	.056*	
I have sufficient knowledge to provide instructional support to my teachers on teaching reading. (5=Strongly agree, 4=agree, 3=neither agree or disagree, 2=disagree, 1=strongly disagree)	.059*	.057*	.062*	080**		.064*		.079**
I oftentimes have time to provide instructional support to my teachers on teaching reading. (5=Strongly agree, 4=agree, 3=neither agree or disagree, 2=disagree, 1=strongly disagree)	.103**	.103**	.097**	.066*		.061*	.059*	.078**

*Correlations are significant at the >0.05 level (2-tailed)
**Correlations are significant at the >0.01 level (2-tailed)
Blanks denote no statistically significant association between variables.

IMPACT OF CLASSROOM ENVIRONMENT

To better understand variation in learner scores, we also looked at the differences in the classroom environment. Various classroom variables were analyzed for relationships with Grade 2 and Grade 3 EGRA results in Filipino and English. A variety of factors in the classroom environment were found to have statistically significant positive relationships with fluency and comprehension scores, albeit the relationships were very weak. As seen in the table, statistically significant relationships were found with several observed classroom infrastructure and environment variables. Having observed comfort rooms for boys and girls had the strongest relationship with EGRA fluency results, however, the relationship was weak.

[20] Classroom observations were scored by assessors using the following scale:(3=Yes, completely, 2= Yes, somewhat, 1=No).

Table 16. Correlations of Classroom Infrastructure and Environment with Filipino and English EGRA Results for Cohort 3 (SY 15/16)

	Grade 2 Learners				Grade 3 Learners			
	Filipino		**English**		**Filipino**		**English**	
Classroom Observation[21]	fluency	comp.	fluency	comp.	fluency	comp.	fluency	comp.
Comfort rooms for boys and girls are clean and functional.	.085**		.089**		.091**		.105**	
The roof is in good condition and the classroom does not show signs of water leakage.				.075**	-.050*			-.055*
The blackboard in the classroom is functional (i.e. Condition allows for legible writing).	.067**							
Classroom is clean and well ventilated.		-.058*					.072**	
Classroom has good lighting, so pupils can easily see their work.	.053*		.057*				.072**	.080**
There are sufficient and appropriate desks for pupils.		.074**						
There are sufficient writing materials for students (i.e. Notebooks, paper).								
There is a reading corner or classroom library where pupils can access books on their own.					.055*		.075**	

*Correlations are significant at the >0.05 level (2-tailed) **Correlations are significant at the >0.01 level (2-tailed)
Blanks denote no statistically significant association between variables.

Analysis of the relationship between self-reported teaching practices and EGRA results showed only a few statistically significant relationships, which as seen in the table below, were weakly correlated and on a whole, rather inconsistent. Use of revised Teacher's Guides provided by Basa was not consistently associated with higher assessment results; neither was the use of old DepEd Teacher's guides. Interesting, whether learners brought home Learner Manuals and Leveled Readers was found to have a negative relationship with EGRA results, particularly for Grade 2 learners, however, as noted previously, the relationships were weak.

Table 17. Correlations of Teacher Practices with Filipino and English EGRA Results, Cohort 3 (SY 15/16)

	Grade 2 Learners				Grade 3 Learners			
	Filipino		**English**		**Filipino**		**English**	
Teacher Practices	fluency	comp.	fluency	comp.	fluency	comp.	fluency	comp.
I only use the revised Teacher's Guide provided by Basa (1=Yes)	.050*							-.059*
I only use the old DepEd Teacher's Guide. (1=Yes)			.055*	.070**	-.055*	-.069**	-.057*	
Have learners brought home learner manuals? (1=Yes)	-.101**	-.073**	-.092**				-.056*	
Have learners brought home leveled readers?(1=Yes)	-.109**	-.085**	-.096**	-.106**		-.068*		-.074**

*Correlations are significant at the >0.05 level (2-tailed) **Correlations are significant at the >0.01 level (2-tailed)
Blanks denote no statistically significant association between variables.

[21] Classroom observations were scored by assessors using the following scale:(3=Yes, completely, 2= Yes, somewhat, 1=No).

6. CHANGES IN TEACHER PRACTICE

Standards-based Classroom Observation Protocol for Educators (SCOPE) was developed by EDC to measure the quality of classroom instruction across grades and subject matters. The observation protocol has subsequently been adapted to capture the quality of instructional practices with specific focus on literacy. The *SCOPE Literacy* tool has two major sections: *Classroom Structure* and *Language Literacy Instruction* and a total of 13 items describing specific literacy-related instructional practices.

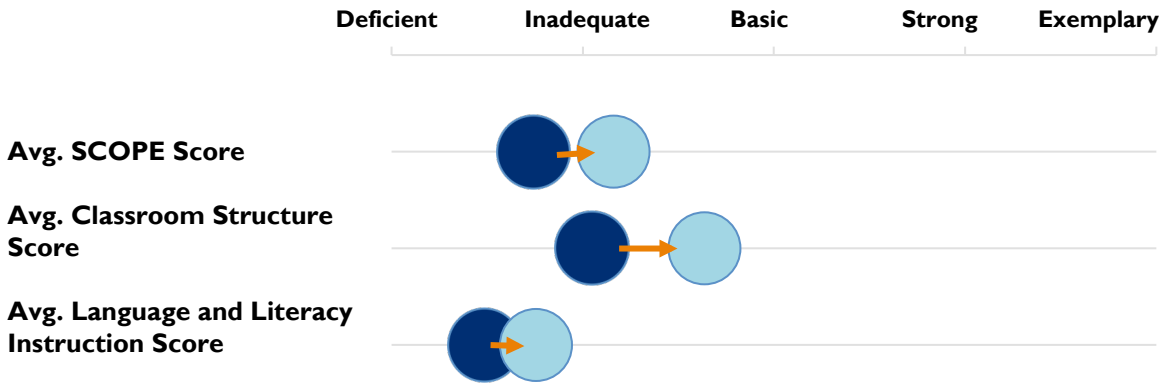
Based on the observation of a class on literacy, each item is scored on a scale from 1 to 5:

(Rating 1) **Deficient.** There is minimal or no evidence of the practice.
(Rating 2) **Inadequate.** There is limited evidence of the practice.
(Rating 3) **Basic.** There is some evidence of the practice.
(Rating 4) **Strong.** There is ample evidence of the practice.
(Rating 5) **Exemplary.** There is compelling evidence of the practice.

Grade two teachers from a sample of Basa schools in Cebu and La Union were observed during a Filipino reading lesson and scored longitudinally using the *SCOPE Literacy* tool in November/December 2013, and again in December 2015 to measure changes in teaching practices as a result of the Basa intervention. The original sample was forty teachers, however, only twenty-eight teachers were able to be tracked longitudinally. This section presents the results for only the longitudinally tracked teachers.

As Figure 71 demonstrates, overall, improvements were seen in observed teachers' practices in classroom structure and language literacy instruction from SY 2013/14, before Basa began, to SY 2015/16, after 2 years of Basa intervention. The largest gains were seen in *Classroom Structure* teaching practices. Improvements from 2013 to 2015 suggest that teachers are starting to apply new teaching practices. It should be noted that ratings of 4 and 5, or "strong" and "exemplary," are not easy to attain. However, a "basic" rating of 3 is quite positive in the context of introducing new literacy instruction techniques.

Figure 71. Changes in Average SCOPE Scores from SY 2013/14 to SY 2015/16 (n=28)

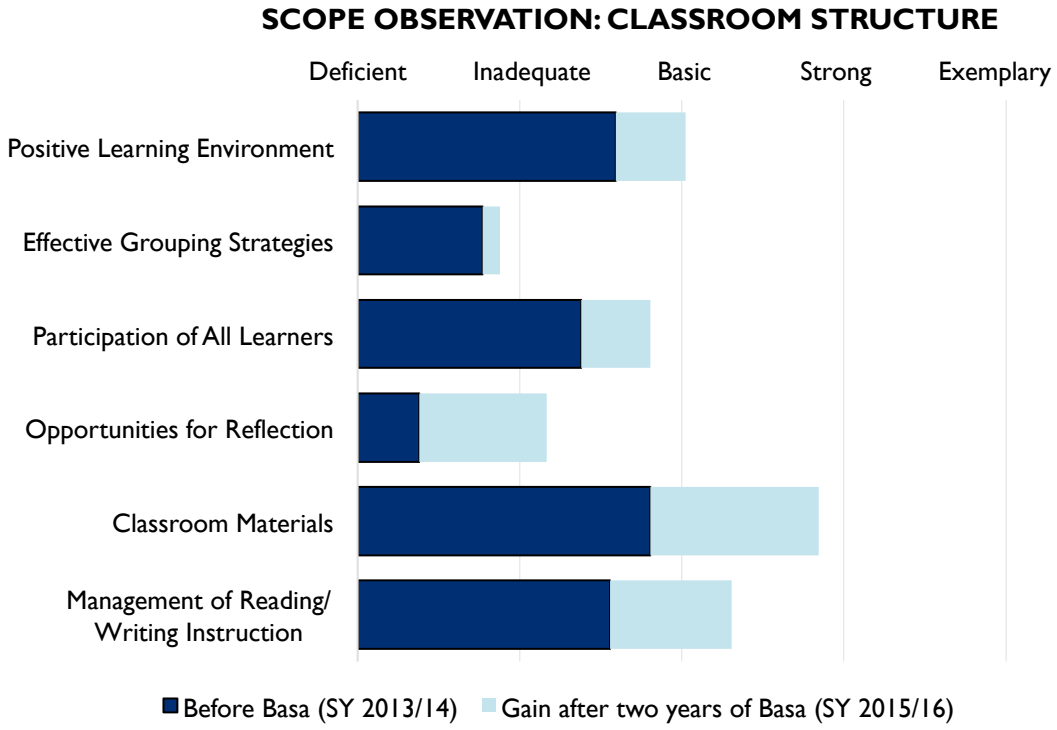


CLASSROOM STRUCTURE RESULTS

The *Classroom Structure* section of the *SCOPE-Literacy* tool contains six items: supportive learning environment, effective grouping strategies, participation of all learners, opportunities for reflection, classroom materials, and manages reading and writing instruction. By December 2015,

improvements were seen in nearly all teaching practices observed; however, the largest gains were seen in the *Classroom Structure* domain. In the *Classroom Structure* domain, teachers saw the largest improvements in "ensuring accessible classroom materials," "providing opportunities for learners to reflect on their reading and writing" and "effective management of reading and writing instruction" items of the *SCOPE Literacy* tool; these improvements were statistically significant ($p<.01$). In fact, in 2015, more than half (62%) of observed teachers were scored "strong" in "classroom materials" and nearly a third (30%) in "management of reading and writing instruction." This is important given that Basa has introduced a large number of new reading materials for teachers to manage in the classroom. Ensuring access to the materials is key for student learning, as is management of reading and writing instruction in the classroom. Teachers didn't score as highly in the areas of "effective grouping strategies" and "opportunities for reflection," both items that require more advanced classroom management skills. Teachers who can effectively group students are better at providing differentiated learning opportunities as well as opportunities for reflection that can deepen students' understanding of text.

Figure 72. Change in Average Classroom Structure Scores for SCOPE Observation from SY13/14 to SY15/16 (n=28)

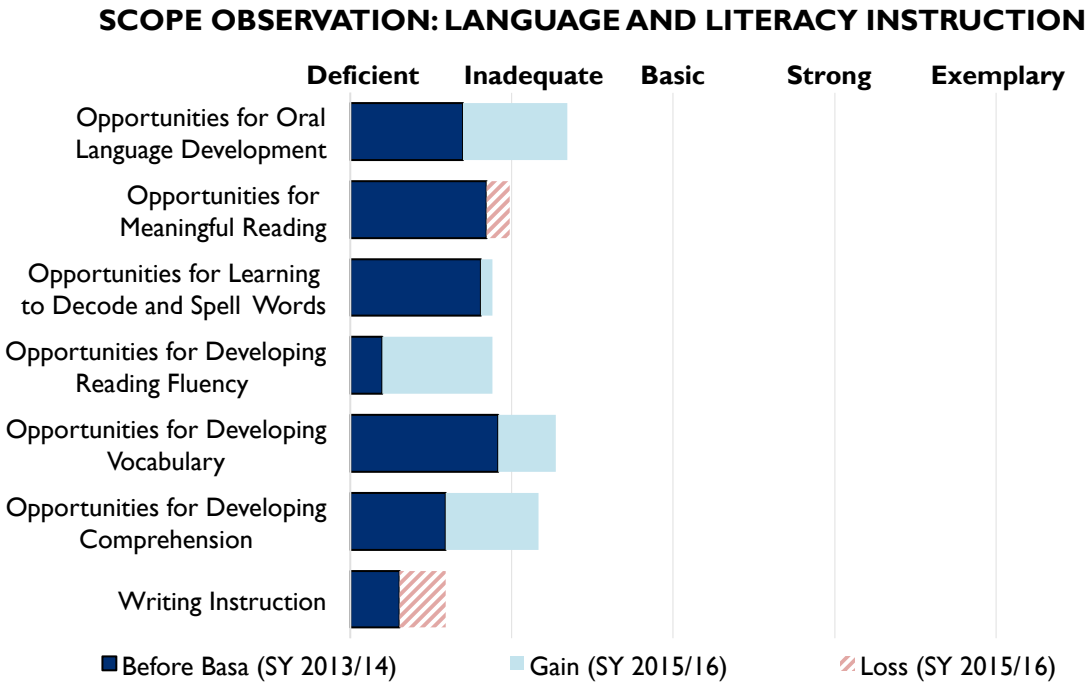


LANGUAGE AND LITERACY INSTRUCTION RESULTS

The *Language and Literacy Instruction* section of the *SCOPE-Literacy* tool includes 7 items: opportunities for oral language development, opportunities for meaningful reading, opportunities for learning to decode and spell words, develops reading fluency, opportunities for developing vocabulary, opportunities for developing reading comprehension, and writing instruction. While baseline scores were lower overall in the domain of *Language and Literacy Instruction*, teachers showed improvement in five out of seven observed practices. This is not unexpected, as teachers may not have had much exposure to teaching reading prior to the Basa intervention. For the *Language and Literacy Instruction* domain, teachers significantly ($p<.05$) improved literacy instruction in the areas of oral language development, developing reading fluency, and developing comprehension. The improvement in "opportunities for developing reading fluency" was particularly striking, as we saw almost no

evidence of this practice in the first observation. Oral language development is also a key skill for teachers, particularly in a multi-lingual context where bridging opportunities from one language to another needs to be intentionally planned by teachers.

Figure 73. Change in Average Language and Literacy Instruction Scores for SCOPE Observation from SY13/14 to SY15/16 (n=28)



One general caveat, Basa teachers follow an instructional sequence in which all 14 domains of the K-12 curriculum are not taught everyday but over a period of 5 days for Filipino and 10 days for English. This is by design to allow adequate time for pupils to complete skills related tasks in their second and third language. Depending on the lesson plan for the day, it would not be expected that teachers teach all domains. In addition, domains such as phonics take on additional meaning in a language such as English which has an opaque orthography versus Filipino, a syllabic language that has a transparent orthography. Since the teachers were observed during their Filipino class, one possible reason for the little to no increase in the score for "opportunities for learning to decode and spell words" is because by Grade 2 quarter 3, majority of the students have already learned to decode and spell in Filipino.



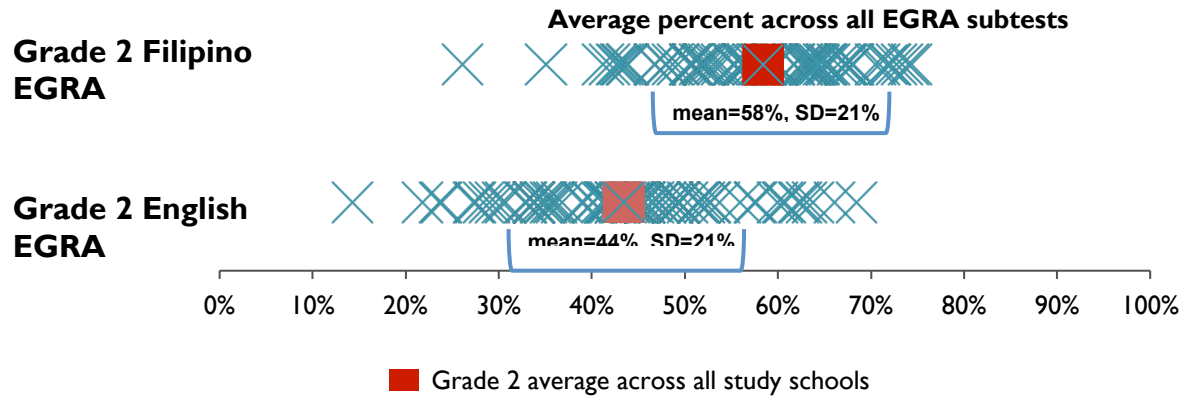
Bivariate statistical analysis showed that *SCOPE* scores were significantly ($p<.001$) positively correlated with average Grade 2 Filipino EGRA scores, however, the relationship was relatively weak ($r=.222$). Providing opportunities for reflection, oral language development, developing vocabulary and comprehension had the strongest correlations with average Filipino EGRA results (Pearson's *r* ranged from 0.186 to 0.20 for these domains).

7. BEST PRACTICES STUDY (BPS)

DESCRIPTION OF THE STUDY

During SY 2014/15 (Cohort 2) EGRA assessment in 2015, the Basa team noted a large variance in Grade 2 average EGRA results from school to school. Results showed that in sampled schools average Filipino EGRA results by school ranged from 26.1% to 74.4% and from 14.3% to 68.4% for average English EGRA results. Figure 74 shows scatterplots of average percent correct across all EGRA subtests for Grade 2 students in Cohort 2 (SY 14/15) by school. Each cross on the graphs represents average EGRA results for each study school. As seen in the figure, results showed a wide variance in EGRA scores across sample schools. For both the Filipino and English EGRA, groups of outlier schools are seen that have extreme EGRA results: either performing substantially worse than other schools or performing substantially better.

Figure 74. Average Percent Correct on Filipino and English EGRA Subtests for Cohort 2 (SY 14/15) Learners in Sampled Schools (n=79)[22]



In order to explore differentiating factors contributing to the significant variance in Grade 2 student achievement on Filipino and English reading assessments in the sampled schools, a Best Practices Study (BPS) was conducted in five top performing and five bottom performing schools identified using literacy assessment results in 2015 by Basa.

STUDY METHODOLOGY

The study methodology involved school visits by the Basa study team. During the school visits, the team conducted the following data collection activities:

- Teacher Questionnaire;
- Principal Survey;[23]and
- Structured observation of a Filipino reading lesson of participating teachers using the Timed Classroom Observation Protocol and *SCOPE*.

[22] The figure shows the average percent correct across subtests for both Filipino and English by school. Average EGRA results for each student are averaged at the school level. Results in the figure are unweighted.
[23] Note that the Principal's Survey used was modeled on the Snapshot for School Effectiveness tools used in a number of different countries. Basa adapted a previous Principal's Survey that was used by the project in previous data collections.

Schools identified for the study were selected using average Filipino and English EGRA results by school from the 2015 reading assessments in which 5 top performing schools and 5 lowest performing schools were selected.[24] Data was collected in February and March 2016. Analysis of Cohort 3 (SY 2015/16) EGRA results by school showed that all but one school remained in either the top or bottom quintile. However, one school moved from the top quintile to the bottom quintile from SY 2014/15 to SY 2015/16; this was likely due to the fact that this school had a new Grade 2 teacher in SY 2015/16, which could explain why EGRA scores changed so drastically from one school year to the next. As a result, this school was categorized as a bottom performing school for analysis. The table below shows the study sample used in the analysis in the following section. In all, 4 Top Performing schools and 6 Bottom Performing schools were included in the study.

Table 18. Best Practices Study Sample

Category	# of Schools	# of Principals Surveyed[25]	# of Teachers Surveyed	# of Classrooms Observed
Top Performing	4	3	4	4
Bottom Performing	6	6	6	6
Total	10	9	10	10

The study design is intended only to provide a broader understanding of potential differentiating factors between top and bottom performing schools; it is not intended to make valid inferences about the larger Basa population in terms of factors associated with improved Filipino or English EGRA scores.

STUDY FINDINGS

The school's context, environment and management is critical to understanding the teaching and learning that is taking place in the school as well as variances in the learning outcomes of students. In all, 9 principals and 10 teachers were surveyed/observed. Data was collected to provide and overall picture of the school and learning environment; and to determine differentiating factors between top and bottom performing Basa schools, particularly in terms of: 1) the community context; 2) principal and teacher demographics; 3) school leadership and management; 4) the learning environment; and 5) teacher practices and beliefs.

BARANGAY SOCIOECONOMIC PROFILE

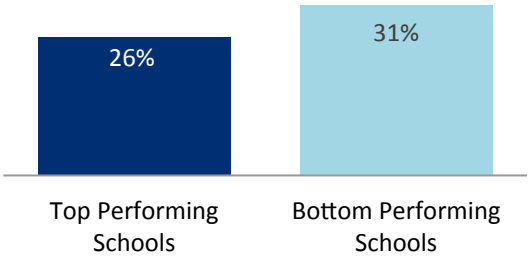
As a proxy for the socioeconomic profile of communities, barangays, that schools are located in, data was collected on the percent of K-3 learners who participate in the 4P's program- "Pantawid Pamilyang Pilipino Program"- in 9 sample schools included in the BPS. The 4Ps program is a social assistance program that provides cash grants to families with specific conditions, which include regular attendance at school for their children. The program targets poor families. It is reasoned that schools who have larger percentages of students participating in the 4Ps program would be a good indicator of whether a school is located in a more impoverished community.

[24] Schools were ranked in order by Filipino and English EGRA results for Oral Passage Reading percent correct, reading comprehension (timed and untimed), and ORF. The five schools that were consistently in the top performing or bottom performing group of schools for all these subtests were selected to be included in the study.

[25] The principal survey was administered to four top performing schools and five bottom performing schools. The school head in one of the top performing schools had retired, and as a result, the Principal survey was not administered in this school.

As seen in Figure 75, results showed that bottom performing schools had on average a slightly larger percentage of K-3 students participating in the 4P program than top performing schools, which suggests that bottom performing schools tend to be located in slightly poorer communities. The effect size, or magnitude of difference, between the percentage of K-3 students in the 4P program in top and bottom performing schools was small to moderate (d=0.40). Correlation analysis found a small correlation (r=0.25) between the percent of students participating in the 4P program and whether a school was a top or bottom performing school. Additionally, a statistically significant (p<.01) negative relationship was found between the percentage of K-3 students in the 4P program and Filipino fluency and comprehension scores (r=-.31).

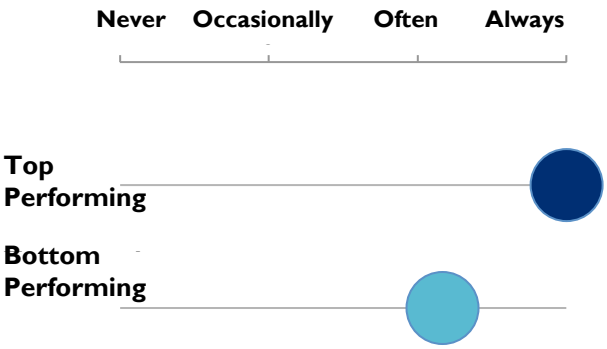
Figure 75. Percent of K-3 Students in 4P Program in School, by School Performance



PRINCIPAL AND TEACHER DEMOGRAPHICS

Surveyed principals/head teachers from top and bottom performing schools were largely similar in terms of key demographics (education, years of experience, etc.). All surveyed principals reported participating in school management training; however, principals in top performing schools reported being able to apply school management training in their work more frequently than principals in bottom performing schools. Common ways principals reported using management training included, conducting classroom observations, providing coaching and instruction to teachers in terms of classroom management, as well as, planning various school improvement projects. Despite similarities in how principals applied school management training, one key difference in top and bottom performing schools is that all principals in top performing schools mentioned the development and use of improvement plans within their schools, including Results-based Performance Management System (RPMS) and School Improvement Plans (SIPs); the use of improvement plans was only mentioned by one bottom performing school.

Figure 76. Frequency Principals Apply School Management Training to Work (n=9)



All three principals from top performing schools reported having attended some form of gender awareness training, while four out of five principals from Bottom Performing schools did; the majority of the trainings had been facilitated by DepEd.

SCHOOL ENVIRONMENT

Overall, results from principal interviews in study schools showed that top and bottom performing schools were largely similar in terms of many school environment factors. For instance, study schools reported comparable numbers of teachers employed at their school, with between 10 to 11 teachers employed on average, whom were largely female. Classroom size was also largely consistent among top and bottom performing schools. Results showed that, on average, Grade 2

classrooms in top performing schools can be expected to have between 32 and 40 learners per one teacher, with a mean of 35.3 compared to bottom performing schools which ranged from 26 to 47 learners per one teacher, with a mean of 35.0. Additionally, in terms of school closures, nearly all top and bottom performing schools were closed at least once during the 2015/16 school year. For those schools that were closed, the average number of days schools were closed was largely consistent across the two groups.



In regards to whether the physical infrastructure was both boy and girl friendly, principals were asked whether their schools had classroom or communal toilets. The majority of principals responded that their schools had classroom toilets instead of communal toilets. Of the schools who had communal toilets, all schools had separate communal toilets for boys and girls.

In terms of attendance, principal and teacher attendance was consistent across top and bottom performing schools. However, in terms of student absenteeism, teachers in bottom performing schools reported higher levels of student absenteeism. On average, teachers in bottom performing schools stated that roughly 20% of learners in their classroom do not come to school every day compared to 4.5% of learners in top performing schools. Very few learners were repeaters; only one teacher in a bottom performing school had a learner that was a repeater.

Teachers in bottom performing schools reported **higher levels** of student absenteeism and roughly a **half hour less of class time**, on average, for students in Grades 1, 2 and 3.

On average, bottom performing schools reported roughly a half hour less of class time (amount of time students are in school, excluding time in assemblies, lunch and breaks) for Grades 1, 2 and 3 students in their schools.

SCHOOL MANAGEMENT

Principals were also asked about their school management and leadership practices: procedures for tracking attendance and enrollment, monitoring and supporting teachers in their schools, as well as tracking student progress. In terms of procedures for tracking data in their schools, surveyed principals were largely consistent. Principals reported using official forms for collecting data in their schools, which suggests that monitoring procedures are clearly defined and followed for both top and bottom performing schools, especially for teacher and student attendance.

Table 19. Use of an Official Form (n=9)

Type of Attendance Data	% of Top Performing schools	% of Bottom Performing schools
Student attendance	100%	100%
Teacher attendance	100%	100%
Student enrollment	100%	83.3%
School census	100%	100%

In nearly every sampled school, the teacher attendance records were easily available for the interviewer to review. Results were largely consistent across bottom and top performing schools in that the majority of schools appeared to complete teacher attendance records daily for the previous month. In terms of student attendance data, top performing schools were observed to record student attendance more frequently than bottom performing schools, in which all top performing schools collected student attendance data 1-2 times a month compared to roughly two-thirds of bottom performing schools. One bottom performing school collected student attendance only once a year, while another never collected student attendance at the school level.

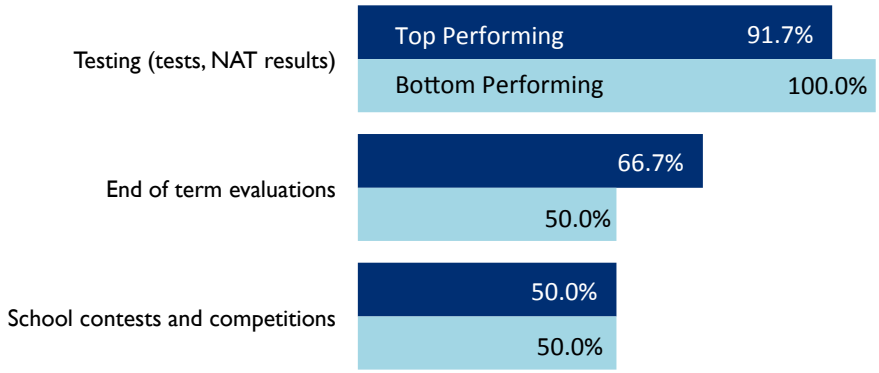
Overall, in terms of monitoring and providing continued support to teachers in their schools, principals in top performing schools appeared be more engaged than principals in bottom performing schools. For instance, results showed that principals in top performing schools reported checking teacher's lesson plans and observing classrooms more frequently than low performing schools. Analysis showed that two out of three principals from top performing schools reported checking lesson plans 2-3 times a week or more, compared to only a third of low performing schools. In terms of classroom observation, 66% of principals in top performing schools observed classrooms 2-3 times a week, while no principal in a bottom performing school observed classes that frequently. The majority (83.3%) of principals from bottom performing schools reported observing classes once a week or 1-2 times per month. Despite these differences, the majority of teachers in both top and bottom performing schools reported receiving enough support from school administration to effectively teach.

In terms of monitoring and providing continued support to teachers in their schools, principals in top performing schools reported **checking teacher's lesson plans and observing classrooms more frequently** than low performing schools.

All schools reported that teachers in their school participate in Learning Action Cells (LACs). There was little difference in how often teachers convene for the LACs, in which the majority indicated that they convened 1-2 times a month.

Results were largely similar in terms of how principals monitor student progress. The most common method used by principals in both top and bottom performing schools was through testing, which includes end of lesson tests, as well as high stakes testing such as the National Achievement Test (NAT).

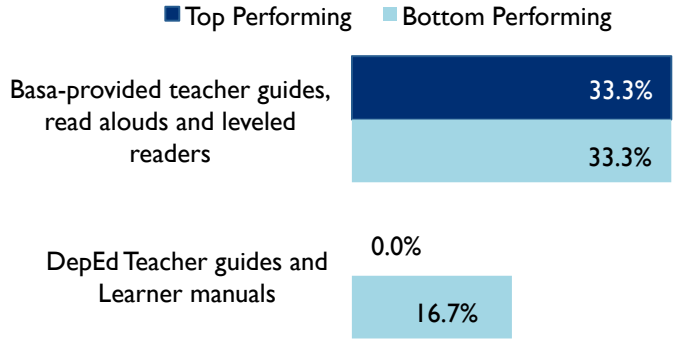
Figure 77. How do you know whether your students are progressing academically? (n=9, multiple response)



READING ENVIRONMENT

Adequate levels and timeliness of receipt of learning materials is a challenge for both top and bottom performing schools. However, principals from top performing schools reported having more of a challenge receiving adequate number of DepEd learning materials than Bottom Performing schools in that more top performing schools either did not receive or only partially received DepEd learning materials at the beginning of the year. Although top performing schools were more likely to report not receiving sufficient learning materials than bottom performing schools, top performing schools did report receiving materials more quickly than bottom performing schools, in which two-thirds of schools reported getting learning materials within 1-5 months. Conversely, in bottom performing schools, 2 out of 6 schools never received the missing textbooks and/or learner manuals; another third stated that it took a year to receive them.

Figure 78. Percent of Schools That Have Received Appropriate Number of Learning Materials at Beginning of School Year (n=9)

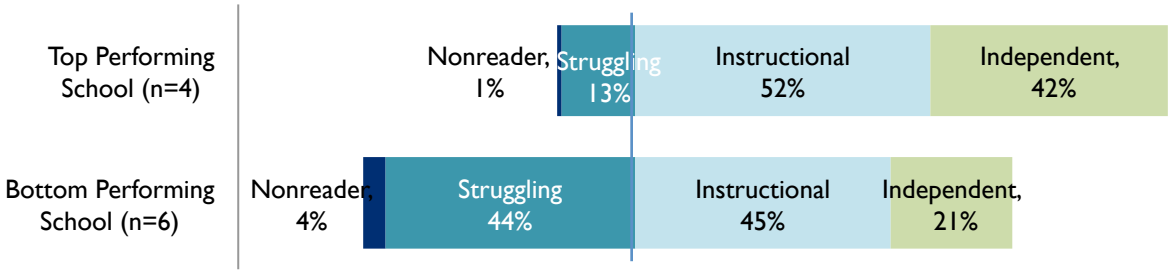


Access to reading resources at school was relatively the same between top and bottom performing schools; two-thirds of schools reported having a school library, in which classes visited the school library daily or weekly. Similarly, both top and bottom performing schools indicated that learners were able to read library books in the library, the classroom and at home. Bivariate statistical analysis found that for schools that did have libraries, there was a moderate relationship between the frequency that students visited the library and reading Filipino comprehension scores (r=.538, p<.01), in that schools where learners visited the school library more frequently tended to have higher Filipino comprehension scores.

Teachers were asked the number of learners in their classroom that were nonreaders, struggling readers, instructional readers and independent readers. As seen in the figure below, as expected, top performing schools had larger percentages of instructional and independent readers than bottom

performing schools. Bottom performing schools reported larger percentages of struggling and nonreaders in their classrooms than top performing schools.

Figure 79. Teacher Reported Percent of Readers in the Classroom (n=10)



TEACHER BELIEFS AND MOTIVATION

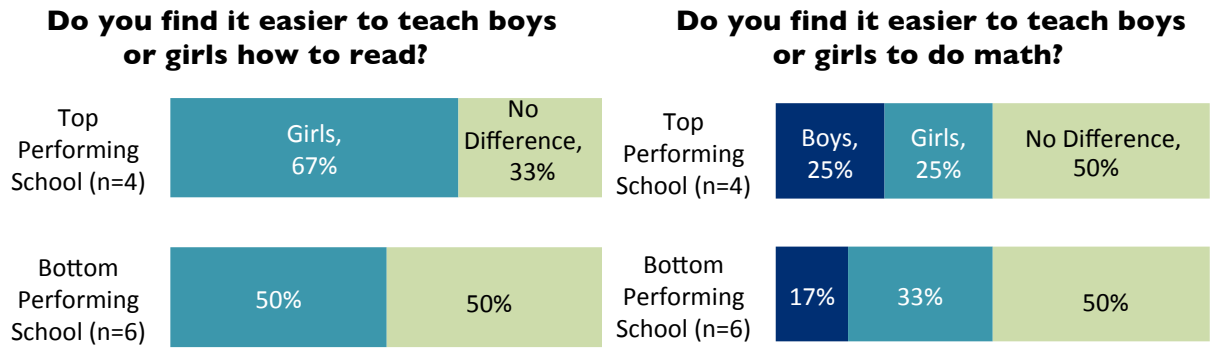
Surveyed teachers were also asked about their beliefs around literacy instruction and their motivation to teach. Overall, teachers in top performing schools reported that teaching reading was easier than teachers in bottom performing schools, this may be due to larger numbers of instructional and independent readers in their classrooms than in bottom performing schools. Half of the teachers in top performing schools said that teaching reading was "mostly easy" compared to only a third of teacher in bottom performing schools. In fact, one teacher in the sampled bottom performing schools stated that teaching reading was "not easy at all." The figure below shows the distribution of responses.

Figure 80. Do you find it easy to teach reading? (n=10)



To assess gender bias in literacy beliefs and practices, teachers were asked their opinion related to gendered beliefs around literacy instruction. Several sampled teachers appear to have a gender bias in that teachers feel that it is easier to teach girls to read than boys. As seen in the figure below, teachers in top performing schools were slightly more gender biased than teachers in bottom performing schools, in which more teachers in top performing schools responded that they felt it was easier to teach girls how to read. Teachers in bottom performing schools were more likely to say that there was "no difference" in teaching boys or girls how to read than their counterparts in top performing schools. Teachers were also asked about their beliefs about teaching math to boys and girls. Half of surveyed teachers in both top and bottom performing schools reported that they felt that there was "no difference." For those teachers that reported a gender bias, slightly more teachers in bottom performing schools reported that they felt girls were easier to teach math than boys, while for top performing schools, responses were split.

Figure 81. Teacher Beliefs on Teaching Reading and Math (n=10)

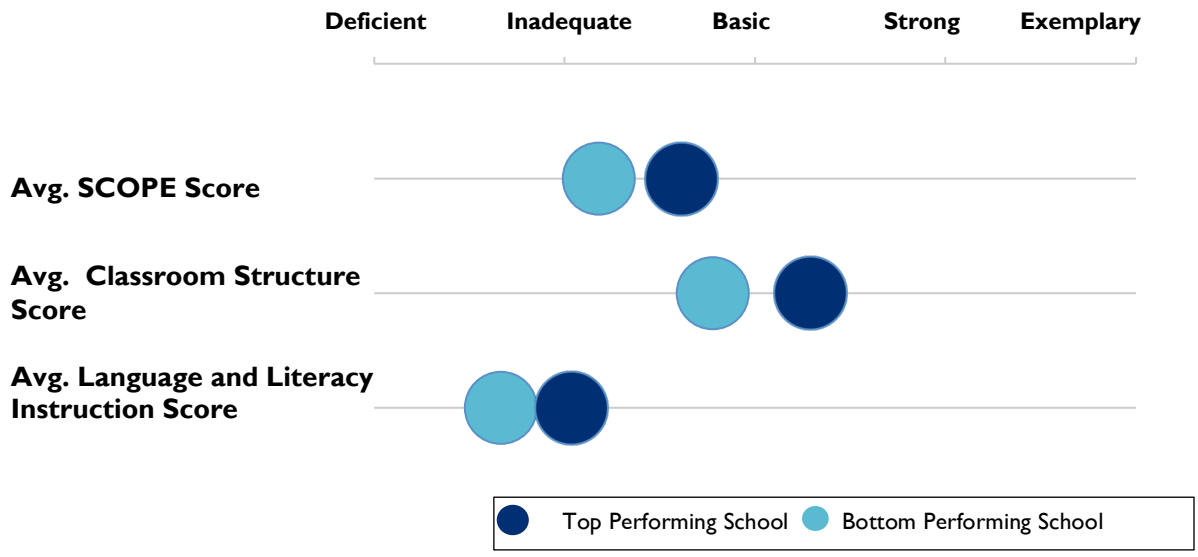


Teachers in top performing schools were found to be slightly more motivated to teach than teachers in bottom performing schools. Three-quarters of sampled teachers in top performing schools stated that they felt "very motivated" to teach compared to only 33% of teachers in bottom performing schools. The remainder of teachers reported that they were "motivated" to teach.

TEACHING PRACTICES

Teaching practices delve into how teachers support the reading and writing environment in their classroom. Teachers in study schools were observed during a Filipino reading lesson and scored using the *SCOPE Literacy* tool which assesses teacher practices in the domains of *Classroom Structure* and *Language and Literacy Instruction*. Assessors observed Filipino reading lessons and scored teachers on thirteen key practices on a scale of 1 to 5, with 1 being "Deficient" and 5 being "Exemplary." Overall, teachers in top performing schools had higher average *SCOPE* scores as seen in the figure below. Teachers in top performing schools were scored higher particularly in the *Classroom Structure* domain. Differences in average scores between top and bottom performing schools for the *Language and Literacy Instruction* domain were less pronounced.

Figure 82. Average SCOPE Score for Top Performing and Bottom Performing Schools (n=10)

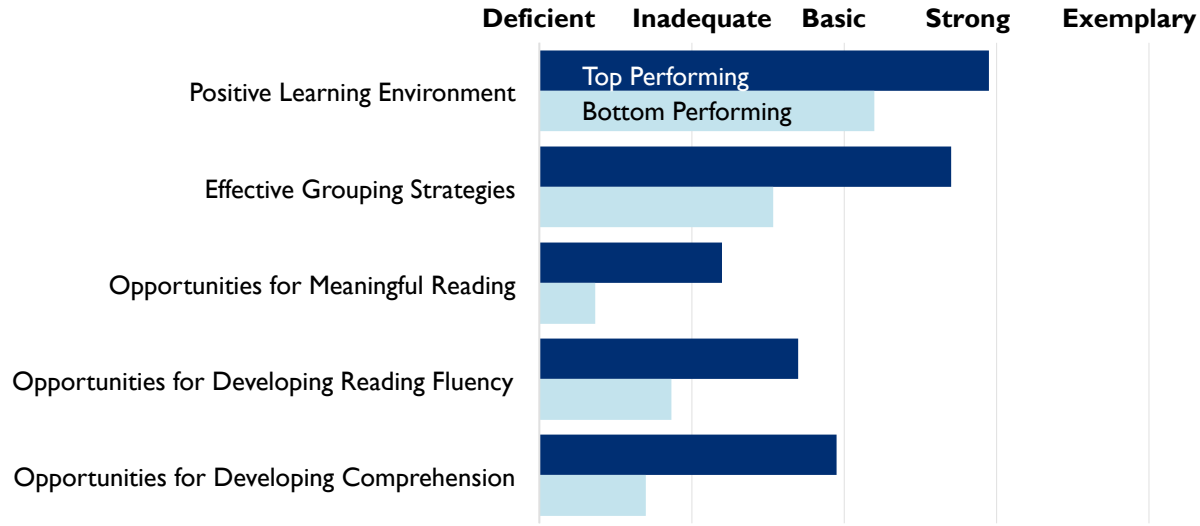


Across both groups, teachers demonstrated consistent and effective practices in "ensuring accessible classroom materials," "ensuring a positive learning environment," "ensuring participation of all learners" and "effective management of reading and writing instruction." In fact, on average, teachers were scored between "basic" and "strong" in these domains. Conversely, teachers from both top and bottom performing schools received lower scores in "writing instruction" and "providing opportunities for learning to decode and spell words," in which teachers were largely scored "deficient." In addition to these domains, teachers in bottom performing schools also scored poorly in "providing opportunities for meaningful reading and comprehension."



Overall, observed teachers in top performing schools were scored higher on average than teachers in bottom performing schools in ten out of the thirteen key reading and literacy instruction domains on the *SCOPE* tool. Observed teachers in bottom performing schools only received higher scores in "managing reading and writing instruction," "providing opportunities for learning to decode and spell words" and "providing opportunities for developing vocabulary." the figure below highlights the largest differences in average *SCOPE* scores between teachers in top and bottom performing schools.

Figure 83. Average Scores for Select *SCOPE* Domains for Top and Bottom Performing Schools (n=10)



As seen above, on average, observed teachers in top performing schools showed substantially stronger teaching practices than teachers in bottom performing schools in "ensuring a positive learning environment," "effective grouping strategies," "providing opportunities for meaningful reading," "providing opportunities for developing reading fluency" and "providing opportunities for developing comprehension." Analysis showed three-quarters of observed teachers in top performing classrooms ensured a positive learning environment for learners in which teachers were rated by assessors as "exemplary" or "strong" in this skill. Conversely only one third of teachers in bottom performing classrooms were rated "strong" in this domain, the remaining teachers were rated "basic" or "inadequate."

In terms of effective grouping strategies, half of teachers in top performing classrooms were rated "strong" or "exemplary." Among teachers in bottom performing schools, the majority (66%) of observed teachers were rated "inadequate" in this practice. Assessors noted that in all top performing schools, teachers used a mix of whole class instruction and group and individual work. Conversely, in bottom performing schools, five out of six observed teachers largely focused on whole-class instruction, with little to no time for group or individual work.

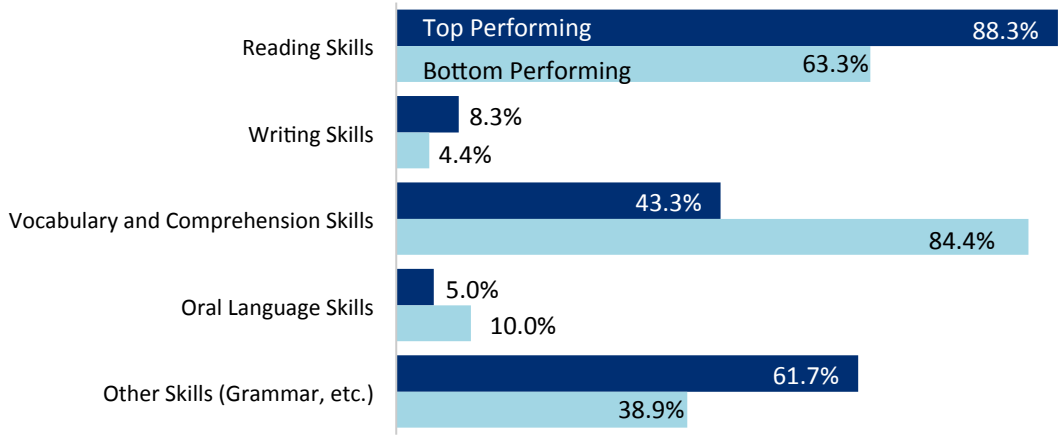
A notable difference was also observed for teachers in "providing opportunities for developing comprehension." Half of teachers in top performing classrooms were rated "strong" in this practice, compared to only one teacher in a bottom performing school. In fact, 3 out of 6 teachers from bottom performing classrooms were rated "deficient" in providing opportunities for developing comprehension compared to only one out of four teachers in top performing schools.

TIME ON TASK IN READING

Teachers in the study sample were also observed to determine the amount of time teachers spend teaching core reading skills during Filipino reading lessons. A Timed Classroom Observation form was used to collect the data.[26] Figure 84 summarizes the amount of time teachers from top and bottom performing schools spent teaching four main reading components—reading, writing, vocabulary and comprehension, and oral language. A fifth category of "other" activities observed in the classroom is also included ("other" activities include, activities such as grammar, classroom discussion, individual seatwork, group work, and other reading or non-reading activities).

Of the total Filipino class time, for both top and bottom performing schools, most time was spent on teaching specific reading skills (letter sounds, syllables, words, passage/story reading) and vocabulary and comprehension (vocabulary, reading comprehension and image interpretation) skills. Top performing schools spent more time on reading skills, while teachers in bottom performing schools spent the most time on vocabulary and comprehension skills. Very little time was spent teaching writing or oral language skills in either group.

Figure 84. Average Percent of Reading Lesson Teacher Spends on Core Reading Skills During Filipino Reading Lesson (n=10)[27]



[26] The Timed Observation tool is adapted from the Nigeria Reading and Access Research Activity (RARA) Baseline Data Collection by RTI, as published on www.eddataglobal.org.

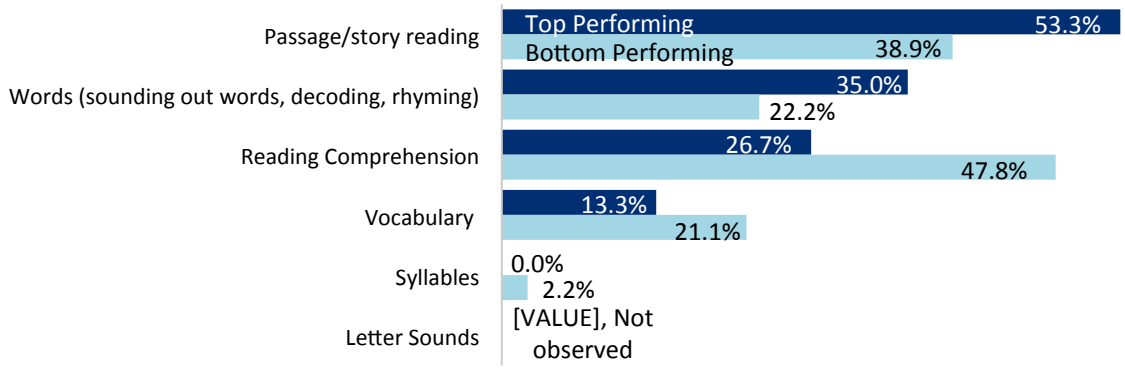
[27] Percentage of lesson can exceed 100%. Data collectors were allowed to select all reading sub-skills that applied during classroom observation.

Teachers were also observed how they delivered the lesson. As seen during the *SCOPE* observation, teachers' method for delivering lesson content in bottom performing schools was overwhelmingly teacher-directed and not focused on active student engagement. Teachers spent the large majority of the reading lesson presenting and explaining the lesson to the whole class and only a small fraction of the lesson (roughly 20%) was spent for independent or small group work. Overall, 50% of observed teachers from bottom performing schools did not devote any time to independent or small group work during the Filipino lesson. Conversely, teachers in top performing schools, on average, had a slightly more balanced lesson delivery, in which roughly two-thirds of the class time was spent instructing the whole class, and the other third was devoted to independent or small group work. Neither group had significant portions of the reading lesson where there was no instruction. In fact, only one instance was observed where a teacher in a bottom performing school was not instructing for a few minutes during the lesson. These findings were substantiated by bivariate statistical analysis. A moderate negative relationship was found between the percent of time teachers spend instructing the whole class during a Filipino reading lesson and lower Filipino oral reading fluency scores (r=-.419). Conversely, a small statistically significant (p<.01) relationship was found between the percent of time teachers devote to independent or group work and higher Filipino oral reading fluency scores (r=.259).

Teachers' method for delivering lesson content in bottom performing schools was overwhelmingly teacher-directed and not focused on active student engagement. Conversely, teachers in top performing schools, had a more balanced lesson delivery, in which roughly two-thirds of the class time was spent instructing the whole class, and the other third was devoted to independent or small group work.

In order to understand the distinct skills that fall under the overall reading and vocabulary and comprehension skills category, Figure 85 breaks down the amount of time spent teaching specific sub-skills. Teachers in top performing schools focused the most time on passage/story reading and decoding words. During the observation, only about a quarter of the reading lesson was devoted to reading comprehension. Bottom performing schools focused more time on reading comprehension than top performing schools, however, as noted previously, reading comprehension activities were primarily focused on whole class instruction and not independent or small group work. Top performing schools were observed to have more balanced instruction of reading comprehension activities in which students spent time both in small groups answering reading comprehension questions as well as time discussing, as a whole class, group work and reading comprehension answers. In both groups, very little time was devoted to syllables; no time was devoted to the instruction of letter sounds.

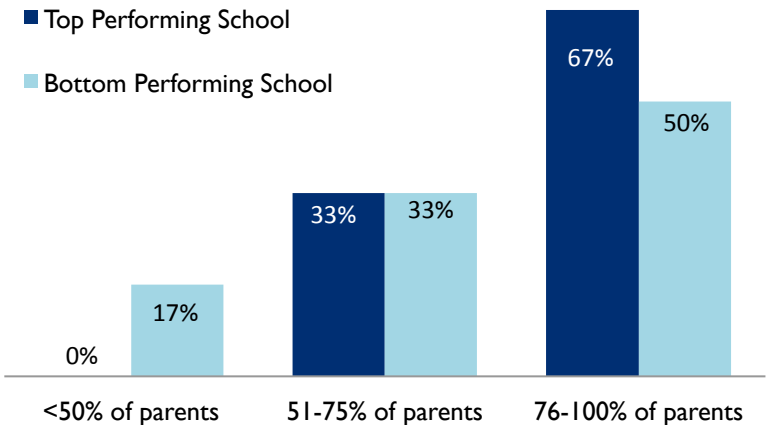
Figure 85. Percentage of Instructional Time Spent Teaching Specific Reading and Comprehension Skills (n=10)



PARENTAL AND COMMUNITY INVOLVEMENT

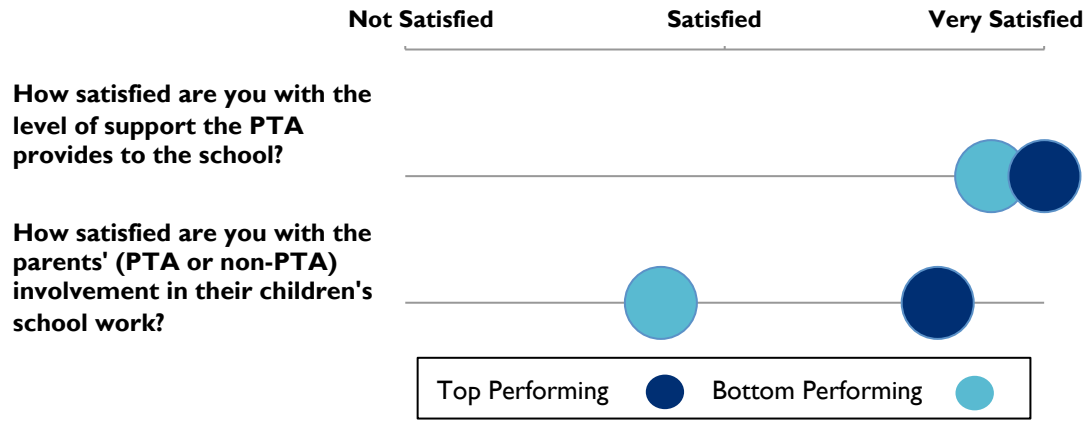
All study schools reported having parent teacher associations (PTAs) that met at various intervals during the year. On average, principals in bottom performing schools reported that the PTA met five to six times a year, while PTAs at top performing schools met on average four to five times a year. Top performing schools reported meeting with a larger percentage of parents/guardians during the school year than bottom performing schools, in which 67% of top performing schools met with three quarters or more of parents during the school year, compared to 50% of bottom performing schools.

Figure 86. On average, what percent of the pupils' parents and guardians did you meet with during the school year? (n=9)



In terms of satisfaction of PTA involvement, principals in top performing schools reported being slightly more satisfied than bottom performing schools. This was similarly the case for parental involvement, where principals in top performing schools also reported being more satisfied with parent involvement in their children's school work than bottom performing schools. Results show that two-thirds of principals from top performing schools said they were "very satisfied" with parental involvement (PTA or non-PTA). In contrast, only one principal from a bottom performing school mentioned being "very satisfied" with parental involvement; in fact, 40% of principals from bottom performing schools reported being "not satisfied" with parental involvement.

Figure 87. Principals' Satisfaction with PTA and Parental Involvement (n=9)



Additional analysis also showed a moderate positive relationship between principal satisfaction of parental involvement and higher Filipino oral reading fluency scores (r=.405, p<.001). The relationship between principal satisfaction of PTA involvement was not quite as strong a relationship, however, it was statistically significant (p<.001; r=.320).

Interestingly, teachers felt differently in terms of parental involvement in children's education. Teachers in bottom performing schools reported being more satisfied with parental involvement in their children's education than teachers in top performing schools. In all, 50% of teachers in top performing schools reported being either "very satisfied" or "satisfied" with parental involvement in their children's schoolwork, compared to 66% of teachers in bottom performing schools.



ANNEXES

ANNEX 1. METHODOLOGY

DESIGN

As part of its scheduled monitoring and evaluation (M&E) activities, Basa conducted three phases of outcome evaluations over the life of the project to measure changes at different levels of the intervention: at the student level (reading), teacher and classroom level (beliefs, pedagogical and instructional practices) and school level (environment and management). The purpose of this final outcome evaluation study is as follows: 1) to measure changes in Grade 2 student achievement associated with the Basa intervention; 2) to measure growth in reading performance from Grade 2 to Grade 3 in Filipino and English; 3) to measure improvement in teachers' instructional practices in reading over the course of the Basa intervention; and 4) to examine various school, classroom, and pupil factors that are associated with higher early grade reading performance in Filipino and English in study schools (e.g., gender, time devoted to reading). Specifically, the evaluation study was designed to answer the following questions:

1. After two years of the Basa intervention, do students, both male and female, demonstrate improved reading and comprehension skills in Filipino at the end of Grade 2?
2. Do students, both male and female, demonstrate improvement in reading and comprehension skills in Filipino and English from Grade 2 to Grade 3?
3. Have teachers' instructional practices in reading (in the Filipino language) improved in target schools?
4. What are the best practices in early grade reading instruction that are associated with better early grade reading performance of Grade 2 students in Filipino and English?

To answer the *first and second questions* on student progress, the evaluation study followed a quasi-experimental, cross-sectional design, which is recommended by the USAID Education Strategy. Cross-sectional design belongs to the family of quasi-experimental designs and is widely accepted as a rigorous methodology by education researchers worldwide. In the cross-sectional design, samples of learners from different cohorts are tested in the same time in the school year and results are compared. The Basa outcome evaluation was designed and carried out in three stages:

- *Cohort 1*: Before the Basa intervention (SY 13/14);
- *Cohort 2*: after one year of Basa intervention (SY 14/15); and
- *Cohort 3*: after two years of Basa intervention (SY 15/16).

In 2013/14 before the full Basa intervention began, a comparison cohort (**Cohort 1**) of Grade 2 students in a sample of schools in Cebu and La Union was assessed in reading (Filipino) at the end of the school year to measure reading performance prior to the Basa intervention. In total, 40 schools were selected to participate in the evaluation. The subsequent year (**Cohort 2** - SY 2014/15), after one year of the Basa intervention, the evaluation was expanded to additional schools in Cebu and La Union, as well as to schools in Ilocos Norte, Ilocos Sur and Bohol to provide a more complete picture of Basa outcomes. In total, 80 schools were randomly selected to participate in the 2014/15 evaluation study, in which Grade 2 learners were assessed in Filipino and English. Finally, during SY 2015/16 (**Cohort 3**), after two years of Basa intervention, a random sample of Grade 2 and Grade 3 students (Cohort 3) were assessed in Filipino and English in the same sample of schools assessed during SY 2013/14 and SY 2014/15 (120 schools). Student assessment results are analyzed cross-sectionally comparing data from Cohort 1 (SY 2013/14) which was collected prior to the Basa intervention to data collected after two years of Basa intervention (Cohort 3 -SY 2015/16).

To answer evaluation *question #3* regarding changes in teachers' instructional practices in reading in the Filipino language in target schools, SCOPE data are used. A longitudinal design was utilized, in which, to the extent possible, the same teachers were observed annually. The *SCOPE* data, collected from 40 schools in Cebu and La Union, were used to analyze changes in instructional practices among teachers trained by Basa. This report will provide findings from November/December 2013 and December 2015 for twenty-eight longitudinally matched teachers.

Lastly, to answer the *final evaluation question*, what are the best practices in early grade reading instruction that are associated with better early grade reading performance of Grade 2 students in Filipino and English, in February and March 2016 a Best Practices Study (BPS) was implemented to examine factors potentially impacting learner achievement in Grade 2 Filipino and English reading performance. The Best Practices Study (BPS) was conducted in four top performing and six bottom performing schools identified using literacy assessment results in 2015 by Basa. The study explores factors that may or may not contribute to schools' status as a top or bottom performer in terms of Filipino and English EGRA results.

The table below summarizes the evaluation study design and tools used in the evaluation.

Evaluation Question	Tools Used
1. After two years of the Basa intervention, do students, both male and female, demonstrate improved reading and comprehension skills in Filipino at the end of Grade 2?	Electronic Early Grade Reading Assessment (eEGRA) tests in Filipino with a random sample of Grade 2 learners in study schools in Cohort 1 (SY 13/14), Cohort 2 (SY 14/15) and Cohort 3 (SY 15/16)
2. Do students, both male and female, demonstrate continued improvement in reading and comprehension skills in Filipino and English at the end of Grade 3?	In Cohort 3 (SY15/16), Electronic Early Grade Reading Assessment (eEGRA) tests in Filipino and English with a random sample of Grade 2 and Grade 3 learners in sample schools
3. Have teachers' instructional practices in reading (in the Filipino language) improved in target schools?	Teacher observation protocol (Standard Classroom Observation Protocol-SCOPE Literacy) to gauge changes in pedagogical practices. Administered longitudinally with a sample of Cohort 1 (SY13/14) teachers (Cebu and La Union).
4. What are the best practices in early grade reading instruction that are associated with better early grade reading performance of Grade 2 students in Filipino and English?	A package of tools (Principal's SSME survey, Teacher Questionnaire, SCOPE, and Timed Classroom Observation Tool) will be administered with 5 Top Performing and 5 Bottom Performing schools to determine potential factors that are associated with better early grade reading performance in Grade 2.

Additionally, a Classroom Monitoring Checklist which comprised of a school observation, and brief principal and teacher survey was administered in SY 2015/16 (Cohort 3) to provide contextual information about the school and classroom environment.

SAMPLE

Sampling was conducted at three levels: 1) school, 2) classrooms, and 3) student. The **school sample** used for the Basa evaluation activities was drawn randomly from the project's five provinces: Cebu, La Union, Bohol and Ilocos Norte/Sur, excluding schools with multi-grade or combination classrooms. For Cohort 1 (2013/14), 20 schools in both Cebu and La Union were randomly selected to be included in the evaluation. For Cohort 2 (2014/15), only schools receiving direct interventions from the Basa project with activities in teachers training, curriculum

development, provision of learning materials, and other targeted support were included in the sample. In all, 20 intervention schools from each province, La Union, Cebu, Ilocos Norte/Sur and Bohol were randomly selected to be included in the evaluation. For the Cohort 3 (2015/16) sample, the same sample of 20 schools selected for Cohort 1 and the 80 schools selected for Cohort 2 were included in the evaluation. In Cohort 3, a principal survey was administered with each principal/head teacher in each sampled school.

At the second level of sampling, for Cohorts 1, 2 and 3, one Grade 2 **classroom** was selected randomly from the total number of Grade 2 classrooms at the school (i.e. one grade 2 from 5 total classrooms). Additionally, during Cohort 3, one Grade 3 classroom was selected randomly from each school. The teacher from selected classrooms were interviewed[28]. For Cohort 1, the 40 teachers randomly selected were also observed in the classroom using the SCOPE-Literacy tool. SCOPE data, collected from 40 schools in Cohort 1, in Cebu and La Union were tracked longitudinally, in which the same teachers to the extent possible were tracked each year (SY 13/14, SY 14/15, and SY 15/16). For those teachers who could not be tracked, due to attrition, replacements were selected. For the purpose of analysis, this report provides findings from November/December 2013 and December 2015 for only longitudinally matched teachers (twenty-eight teachers).

Finally, a random sample of **students** will be selected from within the sample classroom. Using Optimal Designs software, the student sample was determined by the following sampling parameters: Independent t-test, two tail, effect size .25 (small), alpha = 0.025, Power = 80% and Attrition=15%. Analyses revealed a needed sample size of 20 schools/classrooms per province. For Cohort 1 (SY 2013/14), a minimum of 10 students were selected from each of the 40 sample classrooms, while for Cohort 2 (SY 2014/15), 13-15 students were selected from each of the 80 sample classrooms to account for expected attrition. For Cohort 3 (SY 2015/16), 12-14 students were randomly selected from each of the 120 Grade 2 classrooms. Additionally, 12-14 students were randomly selected in the selected Grade 3 classroom to be included in the evaluation. Equal numbers of male and female students were selected as much as possible. The table below maps out the complete student sample by cohort and grade.

Table A-1. Student Sample by Cohort and Grade

Cohort	Evaluation Sample				
	Divisions	Test	# of schools	# of Grade 2 students	# of Grade 3 students
Cohort 1 (SY 13/14)	Cebu & La Union	Filipino EGRA	40	469	--
Cohort 2 (SY 14/15)	Bohol, Cebu, Ilocos Norte, Ilocos Sur & La Union	Filipino & English EGRA	80	1,216	--
Cohort 3 (SY 15/16)	Bohol, Cebu, Ilocos Norte, Ilocos Sur & La Union	Filipino & English EGRA	118	1,658	1,597

[28] For Cohorts 1 and 2, teachers were interviewed using the BIPI Survey. For Cohort 3, teachers were interviewed using a Classroom Monitoring Checklist protocol that included a brief teacher interview.

Additionally, during Cohort 3 (SY 2015/16), surveys were conducted with principals/head teachers and Grade 2 and Grade 3 teachers. Grade 2 classroom observations were also conducted in a sample of 40 schools in Cebu and La Union. The table below shows the final sample evaluation sample for Cohort 3 (SY 2015/16) data collection activities reported in this report.

Table A-2. Cohort 3 (SY 2015/16) Evaluation Sample, by Province

Province	Schools	Principals surveyed	GRADE 2			GRADE 3	
			Classrooms Observed	Grade 2 teachers	Grade 2 students	Grade 3 teachers	Grade 3 students
Bohol	20	15	--	20	283	20	267
Cebu	38	32	20	37	538	38	534
Ilocos Norte/ Sur	20	18	--	19	276	20	275
La Union	40	31	20	38	561	39	521
TOTAL	**118**	**96**	**40**	**114**	**1,658**	**117**	**1,597**

DATA COLLECTION TOOLS

In order to gather data needed to answer the specified research questions, several tools were used to collect data used in this study. The timeframe and number of administrators varied by the tool. Detailed descriptions of the data collection tools can be found below.

OUTCOME EVALUATION TOOLS

Classroom Monitoring Checklist. The Classroom Monitoring Checklist captures data on the classroom and school environment, as well as teacher and school management practices and beliefs. The checklist was administered during school visits during the administration of the EGRA in February/March 2016. The tool is comprised of three components: 1) Observation of the Classroom Infrastructure and Environment; 2) Teacher Survey; and 3) Principal/Head Teacher Survey. For the Classroom infrastructure and environment observation, assessors were to observe and rate Grade 2 and Grade 3 classrooms that they were conducting EGRA assessments in for quality of infrastructure and resources (books, desks, chalkboards, Basa-provided materials etc.). The second component of the Classroom Monitoring Checklist is the teacher survey which was completed during the class break. Teachers were asked to provide details about any training they have attended, their current teaching practices, including lesson planning and preparation as well as comments on Basa provided teaching and learning materials. Lastly, the Principal Survey was administered one-on-one with principals in which they were asked whether they had attended Basa training, their opinions on the literacy instruction in their schools as well as their abilities to provide instructional support to teachers in their schools in teaching reading. Classroom Monitoring Checklist data was collected and processed using SurveyToGo software.

Student Context Interview. In order to collect basic demographic data as well as information about a pupil's educational background and opportunities for reading, a student context interview was administered prior to administration of the EGRA. The interview protocol included questions in the following subject areas: language(s) spoken at home and at school; household items and parental occupation; availability of books in the pupil's home and their subject areas and languages;

availability of reading support at home from a parent or other adult or family member; opportunities for reading and oral listening in school; and educational background, specifically whether the pupil attended kindergarten. The student context interview provides potentially useful information. However, the information obtained must be considered with care as it is based on self-reports of young children. Data were collected electronically using tablets and was processed using SurveyToGo software.

Standard Classroom Observation Protocol in Education-Literacy (SCOPE-Literacy). The quality of literacy instruction was observed using *SCOPE-Literacy*, a tool developed by EDC and tested in several countries. The tool includes two sections with a total of 13 items; each item is rated on a scale from one to five. SCOPE is organized into two major sections: 1) Classroom Structures and 2) Language and Literacy Instruction. The thirteen dimensions of literacy practice and indicators reflecting the dimensions are displayed in Figure A1.

Figure A-1. SCOPE-Literacy Dimensions and Indicators

Section I. **Classroom Structure**	Section II. **Language and Literacy Instruction**
1. Supportive Learning Environment • Understanding of rules routines. • Environment supports pupil language and literacy learning • Teacher management of conflicts and non-compliance	**7. Opportunities for Oral Language Development** • Learner talk • Teacher language • Direct instruction • Discussion
2. Effective Grouping Strategies • Grouping strategies • Learner participation • Learner cooperation and collaboration	**8. Opportunities for Meaningful Reading** • Text choice • Opportunity to read individually • Print resources
3. Participation of All Learners • Learners prior knowledge and interests • Strategies that support learner inclusion • Practice that provides learner access to learning	**9. Opportunities for Learning to Decode and Spell Words** • Direct instruction • Adaptations for individuals • Strategies for decoding
4. Opportunities for Reflection • Opportunities to self-assess reading and writing • Tools to support learner reflection and self-assessment • Ongoing assessment	**10. Develops Reading Fluency** • Modeling fluency • Varied instructional strategies • Activities to build automaticity
5. Classroom Materials • Print-rich environment • Classroom materials to support literacy learning • Use of books in instruction	**11. Opportunities for Developing Vocabulary** • Teacher modeling • Vocabulary selection • Varied approaches to vocabulary instruction • Strategies for learning word meanings independently
6. Manages Reading and Writing Instruction • Lesson planning • Patterns of instruction • Directions to support learner	**12. Opportunities for Developing Reading Comprehension** • Learner thinking • Instructional strategies • Questioning • Direct Instruction
	13. Writing Instruction • Opportunities for self-expression • Writing process • Direct Instruction

For the outcome evaluation, during observation, a pair of data collectors was sent to observe teachers. Each pair observed sampled teachers' instruction separately and then agreed on a score for

each of the items, following a discussion. The agreement score was used in the analysis. Individual observer scores were used for the inter-rater reliability (IRR) analysis, to measure how well individual assessors understood the observation criteria and to what extent they scored the same teacher practices similarly. Since the SCOPE rating data are categorical and not nominal, the interclass correlation coefficient (ICC) was used in the IRR analysis. The ICC assesses rating reliability by comparing the variability of different ratings of the same subject to the total variation across all ratings of all subjects, and results in a coefficient between 0 and 1 where 0 = complete lack of agreement across raters on how to score a particular item, and 1 = perfect agreement. It is considered to be desirable to achieve ICC of .8 or higher. The IRR analysis of observers showed high rates of agreement among observers, with an average ICC of .939 in SY 2013/14, 0.989 in SY 2014/15 and .9642 in SY 2015/16.

A reliability analysis of the *SCOPE-Literacy* found that the test reliability (internal consistency reliability) was high (Cronbach's alpha = .891 in SY 2013/14; .828 in SY 2014/15; and .765 in SY 2015/16). The item showing lower internal consistency to the overall *SCOPE-Literacy* is the "writing instruction" domain.

Table A-3. SCOPE Literacy Reliability, SY 2015/16

SCOPE Literacy Items	Item-Total Correlation	Cronbach's Alpha if Item Deleted
Positive Learning Environment	.628	.720
Effective Grouping Strategies	.358	.753
Participation of All Learners	.613	.725
Opportunities for Reflection	.293	.759
Classroom Materials	.518	.737
Management of Reading & Writing Instruction	.568	.731
Opportunities for Oral Language Development	.313	.758
Opportunities for Meaningful Reading	.292	.759
Opportunities for Learning to Decode & Spell Words	.333	.756
Opportunities for Developing Reading Fluency	.457	.744
Opportunities for Developing Vocabulary	.292	.764
Opportunities for Developing Comprehension	.268	.764
Writing Instruction	.087	.771

The SCOPE tool was organized into two sections: Classroom Structure and Language and Literacy Instruction. A reliability analysis of the two SCOPE sections in SY 2015/16 showed strong reliability of the Classroom Structure section (Cronbach's Alpha = .802) and a moderate reliability for the Language and Literacy Instruction section (Cronbach's Alpha = .576).

Table A-4. SCOPE Classroom Structure Items Reliability, SY 2015/16

SCOPE Classroom Structure Items	Item-Total Correlation	Cronbach's Alpha if Item Deleted
Positive Learning Environment	.661	.746
Effective Grouping Strategies	.467	.790
Participation of All Learners	.601	.761
Opportunities for Reflection	.475	.789

Classroom Materials	.560	.771
Management of Reading & Writing Instruction	.589	.764

Table A-5. SCOPE Language and Literacy Instruction Reliability, SY 2015/16

SCOPE Language and Literacy Instruction Items	Item-Total Correlation	Cronbach's Alpha if Item Deleted
Opportunities for Oral Language Development	.266	.549
Opportunities for Meaningful Reading	.211	.567
Opportunities for Learning to Decode & Spell Words	.244	.557
Opportunities for Developing Reading Fluency	.605	.428
Opportunities for Developing Vocabulary	.295	.545
Opportunities for Developing Comprehension	.388	.499
Writing Instruction	.084	.591

Early Grade Reading Assessment (EGRA). In order to assess student reading proficiency, Basa utilized the Early Grade Reading Assessment (EGRA) tool. EGRA is a standardized reading test that assesses early reading skills, from basic skills such as phonemic awareness and letter sounds, to more sophisticated reading measures such as fluency and comprehension. Basa utilized Grade 2 EGRA tests developed in English by RTI in this evaluation. The Grade 2 EGRA[29] adapted for the Filipino language by RTI was used at the end of the school year in SY 2013/14; SY 2014/15 and SY 2015/16. Two versions of the Grade 2 Filipino EGRA were developed; version A was administered in SY 2013/14 and SY 15/16 and version B was administered in SY 2014/15. The Grade 2 English EGRA was administered in SY 2014/15 (Cohort 2) and SY 2015/16 (Cohort 3).

Additionally, in SY 2015/16 (Cohort 3) the Basa team developed and used a shortened version of the EGRA[30] tool in Filipino and English that was modified from the Grade 2 tests which included 5 subtests per language, for Grade 3 learners rather than the full version, which includes 7 subtests per language. For the Filipino Grade 3 test, all subtests were different than the Grade 2 EGRA, with the exception of the dictation subtest. The Grade 3 reading passage was adapted from a passage developed by RTI in English and was adapted for the Filipino language to align with a Grade 3 Leveled Reader at the I/J level as per curriculum standards. The English EGRA for Grade 2 and Grade 3 were the same for three out of five subtests (Familiar Word Reading, Listening Comprehension and Dictation). Similar to the Grade 3 EGRA, the English reading passage was adapted to align with a Grade 3 Leveled Reader. Grade 3 EGRA tests were piloted in January 2014.

EGRA is a diagnostic instrument designed to assess the foundation skills for literacy acquisition for the early grades according to the following subtests:

1. *Initial sound identification* assessed student's phonemic awareness (the ability to explicitly identify and manipulate the sounds of language). Phonemic awareness has been found to be one of the most robust predictors of reading acquisition and is often used to identify students at risk for reading difficulties in the primary grades in developed countries. In this

[29] See Annex 2 for the summary of EGRA subtests.
[30] Grade 3 EGRA tests in Filipino and English only included the following subtests: Familiar Word Reading, Oral Passage Reading and Comprehension, Listening Comprehension and Dictation.

subtask, students were asked to listen to a word and identify the first sound in that word. After two practice items, students were given ten test items. This subtest was not timed.

2. *Letter sounds* assessed students' knowledge of the sounds that the letters of Filipino alphabet make. Students were presented with a random mix of 100 upper case and lower case letters of the alphabet, and asked to identify what sounds those letters make. Only letter sounds, not letter names, constituted correct answers. The test was timed at 60 seconds; the score was the number of correct letters per minute.

3. *Familiar word reading* assessed student's skill at reading high-frequency words. Recognizing familiar words is critical for developing reading fluency. In this timed subtask, students were presented a chart of 50 familiar words. Students were asked to read as many words as they could. The subtest was timed at 60 seconds and yielded a score of percent correct and correct words per minute.

4. *Simple non-word decoding* assessed student's skills in decoding words they could not have memorized. Tested students were asked to decode a list of 50 pronounceable nonsensical words that followed legal spelling patterns of Filipino for the Filipino EGRA test or English for the English EGRA test. Students were asked to decode as many invented words as they could within 60 seconds. The scores were percent correct and correct words per minute.

5. *Oral passage reading* assessed student's fluency in reading a simple connected text aloud and their ability to understand what they had read. The passage length varied, with 55 words in the Filipino version and 60 words in the English version. The subtest was timed at 60 seconds and yielded a score of correct words per minute. In addition to determining the fluency of reading, data collectors marked a prosody score for each student on a four-point scale, from "word by word, slow, laborious" (1) to "fluent, with expression to mark punctuation and/or direct speech" (4).

6. *Reading comprehension* indicates how well the students understood what they read. After the students finished reading the oral reading passage, or the minute ended, the passage was removed and students were asked five questions with varying difficulty about the passage they just read, but primarily locator.

7. *Listening comprehension* is considered to be an important skill for reading comprehension. In this subtask, the test administrator read a passage to students. Students were then asked questions about that passage. The number of questions asked varied by test; three questions were asked in the Filipino version and five questions were asked in the English version. This subtest was not timed.

8. *Dictation* was designed to assess student's skill at spelling and basic writing rules, such as capitalization, punctuation, text direction, and spacing between words. The data collector read a short sentence to the students and students attempted to write the sentence. The data collector scored the dictation results after the child was finished with the test. This subtest was not timed.

EGRA was programmed into tablets using SurveyToGo software, and sampled students were tested on a one-on-one basis by a trained assessor using a tablet.

EGRA Reliability Analysis. A statistical analysis of test reliability is used to describe an internal consistency of the test, and is based on the correlations between different items (subtests). Internal consistency of the test is measured with Cronbach's alpha which is the result of pairwise correlations between items. Cronbach's alpha ranges from zero to 1, where zero denotes an absence of any correlation across items on the test, and 1 denotes a perfect correlation across

items. A typical and acceptable range for Cronbach's alpha is above .8. A good internal consistency of a literacy assessment means that a child who scores higher on some items would also score higher on other items in the test. A test of internal consistency of *EGRA* in SY 2015/16, found that the overall test reliability was relatively high (Cronbach's alpha = .879 for both the Grade 2 Filipino and English tests). Similarly, for Grade 3, the test reliability was relatively high with a Cronbach's alpha of .832 for the Grade 3 Filipino test and .852 for the Grade 3 English test.

For Grade 2, the item level analysis showed that for both Filipino and English tests listening comprehension did not correlate well with other items. If removed from the test, the Cronbach's alpha would go up to .884 and .880 for the Filipino and English tests respectively. For the Grade 2 Filipino test, other subtests that were least correlated were the initial sound identification and letter sounds subtests. On the Grade 2 English EGRA, phonemic awareness, reading comprehension and letter sounds also did not correlate well the other subtests (Cronbach's alpha = .559, .474 and .599 respectively). Remaining items on the Filipino and English EGRA tests correlated very well with the rest of the test. For Grade 3, item level analysis showed that for both the Filipino and English test, listening comprehension was the least correlated with the other items. If removed from the test, the Cronbach's alpha would go up to .864 and .860 respectively. Similar to Grade 2, the remainder of items correlated well with the rest of the test.

Overall, bivariate correlations between subtests for each test (English and Filipino) were found to be statistically significant for both Grades 2 and 3. For both grades, oral passage reading and Familiar Word reading was found to be highly correlated on both Filipino and English tests. Additionally, for Grade 2, on both tests, Familiar Word and Nonsense Word reading was found to be highly correlated; Nonsense words was also highly correlated with oral passage reading. Tables A-8 through A-11 below show the results.

Table A-6. Grade 2 EGRA Reliability, Cohort 3 (SY 2015/16)

	Filipino EGRA		English EGRA	
EGRA Subtests	**Item-Total Correlation**	**Cronbach's Alpha if Item Deleted**	**Item-Total Correlation**	**Cronbach's Alpha if Item Deleted**
1. Initial sound identification	.540	.879	.559	.876
2. Letter sounds	.559	.876	.599	.871
3. Familiar word reading	.828	.844	.833	.842
4. Nonsense word reading	.804	.849	.806	.846
5. Oral passage reading	.812	.845	.791	.848
6. Reading comprehension	.599	.873	.474	.880
7. Listening comprehension	.430	.884	.477	.880
8. Dictation	.730	.857	.673	.864

Table A-7. Grade 3 EGRA Reliability, Cohort 3 (SY 2015/16) by Grade

	Filipino EGRA		English EGRA	
EGRA Subtests	**Item-Total Correlation**	**Cronbach's Alpha if Item Deleted**	**Item-Total Correlation**	**Cronbach's Alpha if Item Deleted**
Familiar word reading	.753	.765	.724	.805
Oral passage reading	.753	.767	.748	.802
Reading comprehension	.616	.806	.664	.828
Listening comprehension	.419	.864	.511	.860
Dictation	.691	.788	.725	.809

Table A-8. Correlations of Grade 2 Filipino EGRA Subtest Cohort 3 (SY 2015/16)

Filipino EGRA Subtests	ISI	Letter Sounds	Familiar Words	Nonsense Words	Oral Reading	Reading Comp.	Listening Comp.	Dictation
ISI	1							
Letter sounds	.636**	1						
Familiar words	.583**	.557**	1					
Nonsense word reading	.558**	.552**	.914**	1				
Oral passage reading	.551**	.518**	.935**	.913**	1			
Reading comprehension	.334**	.343**	.516**	.484**	.530**	1		
Listening comprehension	.319**	.322**	.346**	.333**	.348**	.522**	1	
Dictation	.548**	.471**	.768**	.734**	.754**	.466**	.377**	1

** Correlation is significant at the p<.001 level

Table A-9. Correlations of Grade 2 English EGRA Subtests for Cohort 3 (SY 2015/16)

English EGRA Subtests	ISI	Letter Sounds	Familiar Words	Nonsense Words	Oral Reading	Reading Comp.	Listening Comp.	Dictation
ISI	1							
Letter sounds	.664**	1						
Familiar words	.590**	.588**	1					
Nonsense word reading	.560**	.622**	.890**	1				
Oral passage reading	.601**	.573**	.914**	.827**	1			
Reading comprehension	.259**	.307**	.380**	.396**	.356**	1		
Listening comprehension	.350**	.356**	.414**	.389**	.366**	.478**	1	
Dictation	.410**	.399**	.641**	.631**	.592**	.408**	.425**	1

** Correlation is significant at the p<.001 level

Table A-10. Correlations of Grade 3 Filipino EGRA Subtest Cohort 3 (SY 2015/16)

Filipino EGRA Subtests	Familiar Words	Oral Reading	Reading Comp.	Listening Comp.	Dictation
Familiar words	1				
Oral passage reading	.945**	1			
Reading comprehension	.497**	.508**	1		
Listening comprehension	.322**	.314**	.506**	1	
Dictation	.773**	.763**	.451**	.385**	1

** Correlation is significant at the p<.001 level

Table A-11. Correlations of Grade 3 English EGRA Subtests for Cohort 3 (SY 2015/16)

English EGRA Subtests	Familiar Words	Oral Reading	Reading Comp.	Listening Comp.	Dictation
Familiar words	1				
Oral passage reading	.951**	1			
Reading comprehension	.487**	.501**	1		
Listening comprehension	.348**	.350**	.594**	1	
Dictation	.682**	.685**	.580**	.460**	1

** Correlation is significant at the p<.001 level

Student Assessment Intra-Class Correlation (ICC) Analysis. While there are many uses for the intra-class correlation coefficient (ICC), we looked specifically at its use to determine the amount of clustering in a nested design. The present assessment drew students from classrooms, so we looked at students nested within classrooms model.

As students within classrooms tend to be similar given the same instruction and teacher, their test scores cannot always be considered as statistically independent from each other. This violates the assumption built into OLS regression models that cases are indeed independent. This leads to false precision in the estimated standard errors of the regression coefficients.

To examine the extent to which clustering impacts the data (so we can determine which model to use) we calculated the intra-class correlation coefficient to calculate the clustering effect. The results of the EGRA assessment data analysis for Cohort 2 (SY 2013/14) revealed the ICC for Filipino EGRA to be 0.1458, and for English EGRA to be 0.1762. For Cohort 3, results show ICC values of 0.1494 and 0.1503 for Grade 2 Filipino and English EGRA respectively. For Grade 3, ICC results were 0.1518 for the Filipino EGRA and 0.1769 for the English EGRA. ICC of this size suggest that a good portion of variance can be explained at the classroom level.

BEST PRACTICE STUDY (BPS) EVALUATION TOOLS

Principal (School Snapshot) Survey. The school's environment and management is critical to understanding the teaching and learning that is taking place in the school. The Principal Survey[31] was designed to capture information on: 1) the physical infrastructure, 2) the teaching and learning demographic, 3) the school manager's background and characteristics, 4) school policies, practices, and monitoring, 5) the reading environment, 6) parent and community involvement, and 7) disabilities and gender policies and practices. For the Best Practices Study, the survey was administered one-on-one with 9 principals (or other school leaders) included in the study that were present at the time of data collection in Feb/March 2016 in study schools. The survey data was collected on tablets and processed in the SurveyToGo system. Detailed Principal Survey findings can be found Chapter 7: Best Practices Study Findings.

Teacher Questionnaire. The teacher questionnaire is a brief survey that captures data on student attendance and repeaters, teacher's beliefs about teaching reading and math as well as

[31] Note that the Principal's Survey used was modeled on the Snapshot for School Effectiveness tools used in a number of different countries. RTI was not using a Principals' Survey in Philippines at the time of data collection, so Basa developed its own tool.

satisfaction of teachers with school leadership and parental support. The survey was conducted in Feb/March 2016 with 10 teachers in study schools using tablets by trained partner and Basa staff and processed using the SurveyToGo system. Detailed teacher questionnaire findings can be found in the Best Practices Study Findings section of this report.

Standard Classroom Observation Protocol in Education-Literacy (SCOPE-Literacy). The quality of literacy instruction was observed using *SCOPE-Literacy*, a tool developed by EDC and tested in several countries. The tool includes two sections with a total of 13 items; each item is rated on a scale from one to five. SCOPE is organized into two major sections: 1) Classroom Structures and 2) Language and Literacy Instruction. Data was collected on paper and processed using SPSS software. For the Best Practices Study, in Feb/March 2016, 10 Grade 2 teachers in BPS study schools were observed during a Filipino lesson. Results from the SCOPE observation tool in BPS study schools can be found in the Best Practices Study Findings section of this report.

Timed Classroom Observation Tool. The timed classroom observation tool is a companion instrument to the non-timed *SCOPE-Literacy* instrument. The timed classroom observation instrument was designed to capture information regarding how much time teachers spent teaching reading, writing, vocabulary, comprehension, and oral language skills. Every three minutes, assessors indicated which specific skills (and sub-skills) were being taught, and whether activities were being conducted by the whole class, pairs, or independently. The timed classroom observation tool was administered in the same 10 classrooms as the SCOPE in Feb/March 2016. Assessors observed time on task during the Filipino reading lesson. Data was collected on tablets and processed using the SurveyToGo system. Detailed results from the timed classroom observation can be found in the Best Practices Study Findings section of the report.

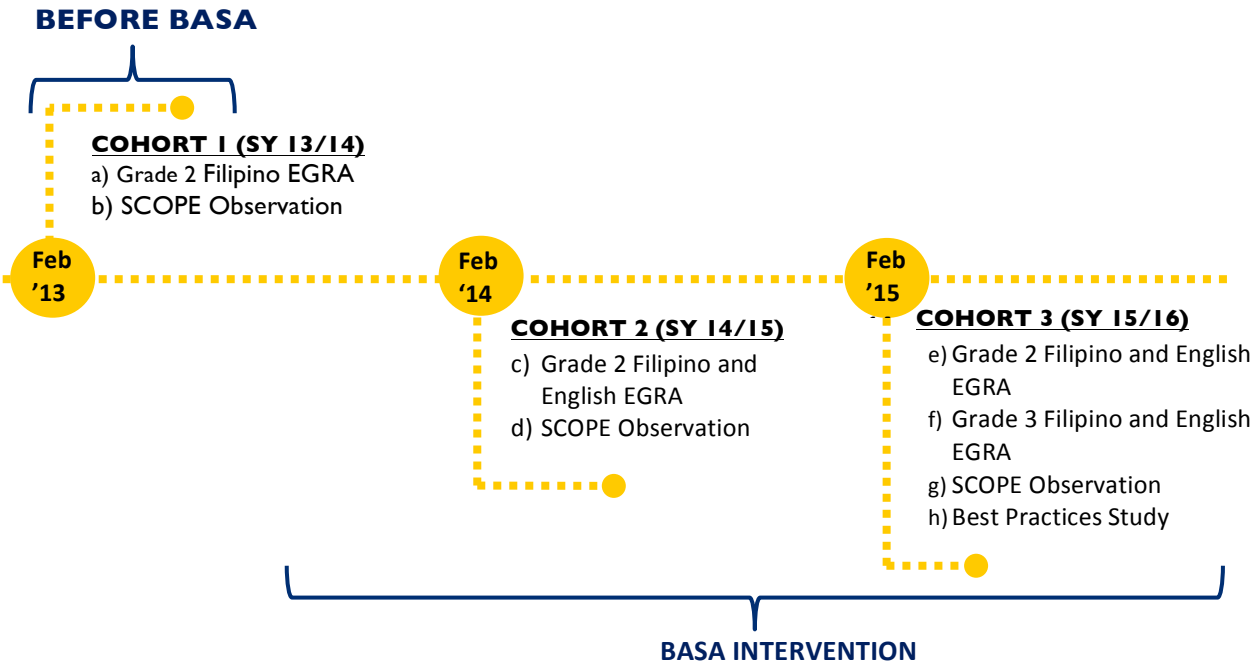
DATA COLLECTION

To answer the project's research questions, data were conducted at three points of time:

- *Cohort 1*: Before the Basa intervention (SY 13/14);
- *Cohort 2*: After one year of Basa intervention (SY 14/15); and
- *Cohort 3*: After two years of Basa intervention (SY 15/16).

The figure below shows the timeline of evaluation activities for data presented in this report.

Figure A-2. Timeline of Basa Evaluation Activities Across Cohort



SCOPE DATA COLLECTION

In SY 13/14, baseline data were collected in November 2013 to February 2014[32] in Cebu and La Union by trained observers working in pairs. Subsequently in SY13/14 and SY 15/16 the same teachers were observed in December 2014 and December 2015, respectively. Given the longitudinal design, Basa attempted to observe the same teachers each year of data collection, however, given changes in teaching staff, only twenty-eight teachers were able to be followed longitudinally (original sample was 40 teachers) from SY 13/14 to SY 15/16. In SY13/14, while teachers had received some preliminary training on the DepEd curriculum through Basa, the full intervention[33] had not been delivered as of the time of observation. By SY 14/15, the full intervention had been rolled out.

[32] Note that the scheduled observation was November 2013, but there were some delays experienced, notably scheduled school holidays and difficulty in getting to some of the school due to emergency relief situations involving the 2013 typhoon.

[33] The Basa intervention consists of a set of materials made up of Read-Aloud stories, leveled readers and teacher guides as well as trainings on how to use the materials.

EGRA DATA COLLECTION

The *EGRA* was administered to Cohort 1 (SY 2013/14) during February 2014. Subsequently in February 2015 the EGRA was administered to Cohort 2 (SY 2014/15) and in February/March 2016 to Cohort 3 (SY 2015/16). See Assessment schedule below.

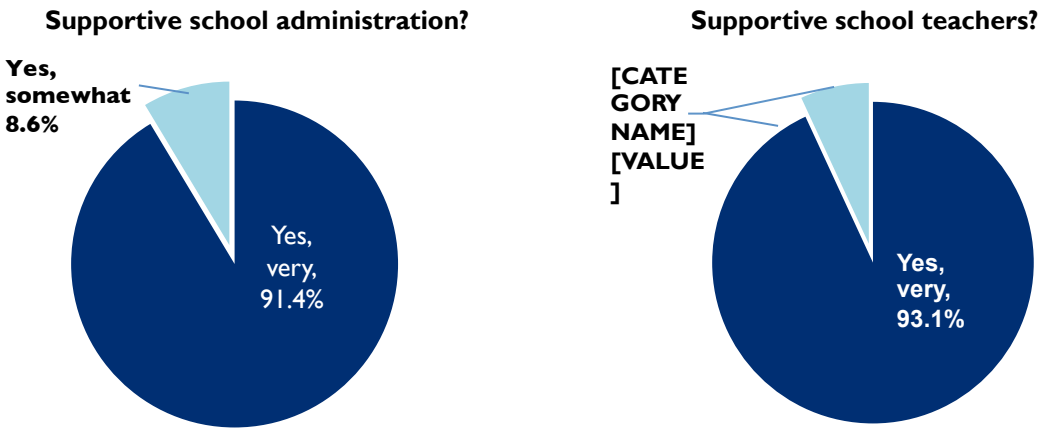
Table A-12. Assessment Schedule

	Grade	Assessment	Feb 2014	Feb 2015	Feb 2016
Cohort 1 (SY 13/14)	Grade 2	Filipino EGRA	X		
Cohort 2 (SY 14/15)	Grade 2	Filipino & English EGRA		X	
Cohort 3 (SY 15/16)	Grade 2 & Grade 3	Filipino & English EGRA			X

Assessors, supervised by BASA staff, participated in data collection. All assessors attended training in data collection procedures, including random selection of boys and girls from the classrooms for the student assessment.

Assessors were asked whether schools and teachers were supportive of data collection. The majority of assessors reported that the school administration and teachers were supportive of data collection. When asked about school administration specifically, 91% of assessors reported that administrators were very supportive, with 9% indicating administrators were somewhat supportive. School teachers were rated as very supportive of data collection by 93% of assessors, with 7% saying that teachers were somewhat supportive.

Figure A-3. School/Teacher are supportive of data collection (n=116)



The majority (62.1%) of assessors reported that they were able to conduct interviews with a few interruptions by teachers or other learners walking into the room where the assessment was being conducted. Only a third of assessors (34.5%) experienced no interruptions. Four assessors (3.4% of respondents) reported experiencing several interruptions.

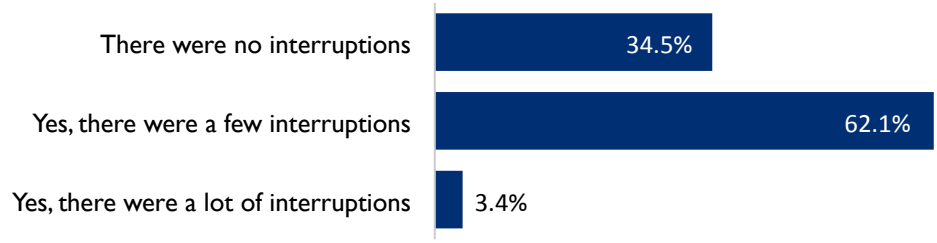
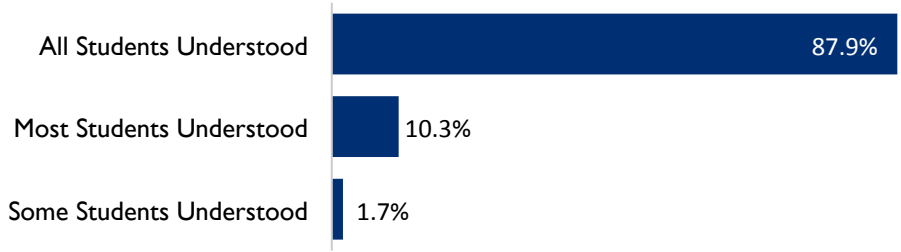


Figure A-4. Assessments were conducted with Few interruptions (n=116)

Assessors were asked whether the students they assessed were able to understand the language they were speaking. Nearly all assessors (98.3%) of assessors said that all or most of the students were able to understand the language they were speaking; only 2 assessors (1.7%) indicated that only some students were able to understand the language they were speaking. The most common languages reported being used by assessors were Cebuano/Bisaya, Ilokano and Filipino; only 5 assessors reported using English.

Figure A-5: Percent of Students Who Understood the Language the Assessor Used (n=116)



Assessors were also asked whether they experienced problems during data collection. The most common problems assessors faced were disruptions during assessments, difficulty accessing schools), limited available space to conduct EGRAs and the absence of school heads/administrators and students. Others mentioned a variety of isolated incidents including having problems with a tablet.

To measure how well individual administrators graded the sub-tests similarly, inter-rater reliability (IRR) exercises were conducted during the training. All administrators took part in IRR exercises. During the group role play, administrators scored the mock child respondent and the trainer noted the variances in the scores for each of the subtests. Administrators with consistent discrepancies were given additional training, monitoring and support. Items with larger discrepancies were furthered reviewed with the larger group during practice sessions. In addition, during the practice testing with actual children, two administrators were paired together to score the same child respondent. Each administrator scored the respondent separately. At the end of the testing, the administrators compared scoring data and discussed discrepancies with the oversight of the trainers. Those administrators that were not consistent in their scoring by the end of training were not allowed to participate in the testing.

In addition, IRR during the actual data collection was conducted for all EGRA data collection activities. Overall, results from all three Cohorts showed strong reliability among data assessors. For Cohort 1 (SY 2013/14), IRR data was collected for 30 students who were tested by two

assessors. The mean ICC score when EGRA data was collected at the end of the school year was .830, and the median was 1, which indicates very strong reliability.

For Cohort 2 (SY 2014/15), IRR was conducted on 166 students. IRR was consistent for both Grade 2 EGRA tests, which suggests that assessors consistently scored the same behaviors similarly regardless of the language the test was administered in, English or Filipino. The median was .995 for Filipino and .992 for English which again indicates very strong reliability.

For Cohort 3 (SY 2015/16), IRR was conducted with 235 Grade 2 learners and 137 Grade 3 learners in both Filipino and English during data collection. For Grade 2, the mean ICC score was 0.972 (median=0.999) for Filipino and 0.976 (median = 0.998) for English, which demonstrates strong reliability among assessors. Similarly, strong reliability among assessors was seen for Grade 3 EGRA data collection; results showed a mean ICC score of 0.979 (median=0.999) for Filipino and 0.979 (median =0.999) for English.

BEST PRACTICES STUDY

To explore differentiating factors contributing to the significant variance in Grade 2 student achievement on Filipino and English reading assessments in the sampled schools, data was collected for Best Practices Study (BPS) in February and March 2016 in 10 study schools. Data collection activities involved school visits by the Basa study team. Schools were visited in pairs by the Basa study team. During the school visits, the team conducted the following data collection activities:

- Teacher Questionnaire;
- Principal Survey;[34]and
- Structured observation of a Filipino reading lesson of participating teachers using the Timed Classroom Observation Protocol and SCOPE.

Basa study team members were trained on data collection instruments one-on-one. Training was delivered by the Basa M&E and Research Specialist. All data collection was recorded using paper forms and encoded into SurveyToGo software for processing and analysis.

DATA ANALYSIS

All collected data were cleaned by EDC M&E staff and analyzed using standard statistical techniques such as univariate and bivariate statistics as needed for different analytical purposes. The results were disaggregated by sex, and province, as appropriate. Univariate and multivariate statistical analyses for were conducted. Central tendency analysis (e.g. mean, median) were conducted for continuous demographic variables. Comparison of means statistical tests (independent samples *t*-test) were conducted to estimate differences between groups such as province and sex, where appropriate. Bivariate statistical analyses (e.g., correlations) were conducted to examine the relationship between different variables.

LIMITATIONS

The assessment had some limitations. Since the evaluation design does not include random assignment of teachers and students into participant and non-participant groups to assess the true impact of the program, the attribution of the observed outcomes to the program will be limited

[34] Note that the Principal's Survey used was modeled on the Snapshot for School Effectiveness tools used in a number of different countries. Basa adapted a previous Principal's Survey that was used by the project in previous data collections.

since other factors may have contributed to the changes in the studied outcomes. In cross-sectional designs, major threats to validity[35] involve selection-history (when other events occur between cohorts that may impact one group but not the other), and selection-instrumentation (when the test used with cohorts is slightly different). Basa attempted to control for the selection-instrumentation bias by extensive pilot testing. The other two threats relate to the passage of time and external events outside of control or knowledge of the study team. It is therefore unknown to what extent external factors may impact different cohorts.

Other limitations originate from the assessment's sampling strategy, particularly the fact that the Cohort 1 (SY 2013/14) sample was selected only from Cebu and La Union, while the samples for Cohort 2 (SY 2014/15) and Cohort 3 (SY 2015/16) were selected from all five provinces (Bohol, Cebu, Ilocos Norte, Ilocos Sur and La Union). To justify the use of Cohort 1 (data collected only Cebu and La Union) as a comparison for Cohorts 2 and 3 (data from all five provinces), a baseline equivalence analysis was conducted. According to WWC guidelines, the effect size difference (Cohen's d) between intervention and comparison group means at baseline should fall between 0 and 0.05 to satisfy baseline equivalence or between 0.05 and 0.25 (requires statistical adjustment to satisfy baseline equivalence. Analysis of Cohort 1 (SY 13/14) and Cohort 2 (SY 15/16) data showed that for all but one subtest (Letter Sounds) the effect size difference falls between 0.05 and 0.25, and as such meets baseline equivalence. Given these results Cohort 1 (SY 13/14) was deemed a legitimate comparison group for comparison with Cohort 2 (SY 14/15) and Cohort 3 (SY 15/16) results in all five provinces.

Table A-13. Effect Size (ES) Difference between Intervention (All Provinces) and Comparison Cohort (Cebu/La Union) Means at Baseline

Subtest	Effect size of baseline mean differences
Phonemic Awareness	0.2008
Letter Sounds	0.4759
Familiar Word Reading	0.0949
Nonsense Word Reading	0.1103
Oral Passage Reading	0.13
Reading Comprehension	-0.0512
Listening Comprehension	0.0719
Dictation	0.1471

Lastly, while the sample was stratified by province to ensure adequate representation of students across different provinces in the country, the province-level sub-samples are not large enough to be treated as separate samples. A much larger sample size would be required to enable such analyses.

[35] W. Trochim, Research Methods Knowledge Base. Cornell University, 2006.

ANNEX 2. EGRA SUBTESTS

Table A-14. Grade 2 Filipino EGRA Subtests

#	Description (Instrument)	Tasks/Max. Pts.	Reported %/wcpm	Timed
1	Initial sound identification	10 letters/sounds	percent: value / 10 * 100%	No
2	Letter Sound Knowledge	100 letters	percent: value * 1% lcpm: value / timer * 60	Yes (60 sec.)
3	Familiar Word Identification	50 words	percent: value / 50 * 100% wcpm: value / timer * 60	Yes (60 sec.)
4	Simple Non-word decoding	50 words	percent: value / 50 * 100% wcpm: value / timer * 60	Yes (60 sec.)
5 A	Oral Passage Reading	64 words (Cohort 1 and 3- SY 2013/14 and SY 2015/16)	percent: value / # of words in the text * 100% wcpm: value / timer * 60	Yes (60 sec.)
5 B	Oral Reading Comprehension (after	5 questions	percent: value / 5 * 100%	No
5 C	Oral Reading Comprehension (after	5 questions	percent: value / 5 * 100%	No
6	Listening Comprehension	3 questions (Cohort 1 and 2 – SY13/14 and SY14/15) 5 questions (Cohort 3 – SY	percent: value / 3 * 100% percent: value / 5 * 100%	No
7 A	Dictation (spelling)	12 words (Filipino)	percent: value / 16 * 100%	No
7 B	Dictation (conventions of text)	4		No

Table A-15. Grade 2 English EGRA Subtests[36]

#	Description (Instrument)	Tasks/Max. Pts.	Reported %/wcpm	Timed
1	Initial sound identification	10 letters/sounds	percent: value / 10 * 100%	No
2	Letter Sound Knowledge	100 letters	percent: value * 1% lcpm: value / timer * 60	Yes (60 sec.)
3	Familiar Word Identification	50 words	percent: value / 50 * 100% wcpm: value / timer * 60	Yes (60 sec.)
4	Simple Non-word decoding	50 words	percent: value / 50 * 100% wcpm: value / timer * 60	Yes (60 sec.)
5 A	Passage Reading	60 words (Cohorts 2 and 3 – SY13/14 and SY15/16)	percent: value / 60 * 100% wcpm: value / timer * 60	Yes (60 sec.)
5 B	Oral Reading Comprehension (after timed	5 questions	percent: value / 5 * 100%	No
5 C	Oral Reading Comprehension (after	5 questions	percent: value / 5 * 100%	No
6	Listening Comprehension	5 questions	percent: value / 5 * 100%	No
7 A	Dictation (spelling)	13 words	percent: value / 17 * 100%	No
7 B	Dictation (conventions of text)	4		No

[36] The Grade 2 English EGRA was only administered for Cohort 2 (SY 2014/15) and Cohort 3 (SY 2015/16) at the end on the school year. The English EGRA was not administered for Cohort 1 (SY 2013/14).

Table A-16. Grade 3 Filipino EGRA Subtests[37]

#	Description (Instrument)	Tasks/Max. Pts.	Reported %/wcpm	Timed
1	Familiar Word Identification	50 words	percent: value / 50 * 100%	Yes (60 sec.)
2A	Oral Passage Reading	55 words	percent: value / # of words in the text * 100% wcpm: value / timer * 60	Yes (60 sec.)
2B	Oral Reading Comprehension (after timed reading	5 questions	percent: value / 5 * 100%	No
2C	Oral Reading Comprehension (after untimed reading)	5 questions	percent: value / 5 * 100%	No
3	Listening Comprehension	5 questions	percent: value / 5 * 100%	No
4A	Dictation (spelling)	12 words (Filipino)	percent: value / 16 * 100%	No
4B	Dictation (conventions of text)	4		No

Table A-17. Grade 3 English EGRA Subtests

#	Description (Instrument)	Tasks/Max. Pts.	Reported %/wcpm	Timed
1	Familiar Word Identification	50 words	percent: value / 50 * 100%	Yes (60 sec.)
2A	Passage Reading	61 words	percent: value / 60 * 100% wcpm: value / timer * 60	Yes (60 sec.)
2B	Oral Reading Comprehension (after	5 questions	percent: value / 5 * 100%	No
2C	Oral Reading Comprehension (after untimed reading)	5 questions	percent: value / 5 * 100%	No
3	Listening Comprehension	5 questions	percent: value / 5 * 100%	No
4A	Dictation (spelling)	13 words	percent: value / 17 * 100%	No
4B	Dictation (conventions of text)	4		No

[37] Grade 3 EGRA was only administered in SY 2015/16 (Cohort 3).

ANNEX 3. SUMMARY EGRA RESULTS

In the summary tables below, mean values are shown for each subtest for Grade 2 and Grade 3 Filipino and English EGRAs for Cohorts 1, 2 and 3. A *95% confidence interval* is shown for average gains across cohort. *Mean* refers to the percentage of items answered correctly. *Mean for non-zero* refers to the percentage of items answered correctly among students who scored above zero. Additionally, the percent of students that had zero scores on that sub-test is also shown. Note that design weights were applied to calculate EGRA results presented in this Annex in order to compensate for oversampling/under sampling at the provincial level and to ensure that results are representative of all provinces in the sample.

GRADE 2 EGRA RESULTS

DETAILED GRADE 2 FILIPINO EGRA RESULTS

Table A-18. Overall Grade 2 Filipino EGRA Results, by Subtest and Cohort

Grade 2 Filipino EGRA Subtests	Cohort 1 (SY 2013/14)	Cohort 2 (SY 2014/15)	Cohort 3 (SY 2015/16)	Gain (Cohort 1 to Cohort 3)	Effect Size (Cohort 1, Cohort 3)
Initial Sound Identification (pct correct)	57.6%	78.4%	72.1%	14.5% (±3.6%)	0.42
Letter Sounds (pct correct)	19.1%	29.5%	24.1%	5.0% (±1.5%)	0.32
Letter Correct (per min)	19.2	29.5	24.3	5.1 (±1.5)	0.31
Familiar Words(pct correct)	68.4%	73.4%	65.9%	-2.5% (±3.3)	-0.08
Familiar Words Correct (per min)	37.5	41.9	35.3	-2.2 (±2.0)	-0.11
Nonsense Words (pct correct)	46.2%	51.4%	46.8%	0.6% (±2.6%)	0.02
Nonsense Words Correct (per min)	23.2	26.0	23.6	0.40 (±1.4)	0.02
Oral Passage Reading (pct correct)	56.1%	75.1%	58.7%	2.6% (±3.0%)	0.09
Words Correct in a Text (per min)	37.0	48.1	39.4	2.4 (±2.1)	0.11
Prosody score	2.1	3.0	2.4	0.3 (±0.1)	0.33
Reading Comprehension: timed (pct correct)	28.1%	42.4%	41.8%	13.7% (±2.9%)	0.43
Reading Comprehension: untimed (pct correct)	-	54.5%	50.9%	-	-
Listening Comprehension (pct correct)	49.1%	63.7%	34.7%	-14.4% (±3.7%)	-0.49
Dictation Composite (pct correct)	45.6%	61.9%	59.7%	14.1% (±2.8%)	0.52

Table A-19. Grade 2 Filipino EGRA Subtests – Percent of Learners with Zero Scores

Percent of Learners with Zero Scores on Grade 2 Filipino EGRA Subtests				
Grade 2 Filipino Subtest	**Cohort 1 (SY 13/14)**	**Cohort 2 (SY 14/15)**	**Cohort 3 (SY 15/16)**	**Change in Zero Scores (Cohort 1, Cohort 3**
Initial Sound Identification (pct correct)	12.9%	4.7%	8.7%	-4.2% (±3..0%)
Letter Sounds (pct correct)	6.2%	1.7%	6.6%	0.3% (±2.5%)
Letter Correct (per min)	6.2%	1.7%	6.6%	0.3% (±2.5%)
Familiar Words(pct correct)	3.4%	3.6%	4.3%	0.9% (±2.0%)
Familiar Words Correct (per min)	3.4%	3.6%	4.3%	0.9% (±2.0%)
Nonsense Words (pct correct)	6.8%	7.5%	8.2%	1.4% (±2.8%)
Nonsense Words Correct (per min)	6.8%	7.5%	8.2%	1.4% (±2.8%)
Oral Passage Reading (pct correct)	4.4%	3.3%	4.7%	0.3% (±2.2%)
Words Correct in a Text (per min)	4.4%	3.3%	4.7%	0.3% (±2.2%)
Prosody score	--	--	--	--
Reading Comprehension: timed (pct correct)	29.9%	14.3%	20.8%	-9.1% (±4.4%)
Reading Comprehension: untimed (pct correct)	--	11.6%	9.7%	--
Listening Comprehension (pct correct)	28.3%	18.2%	22.3%	-6.0% (±4.4%)
Dictation Composite (pct correct)	5.3%	1.8%	5.4%	0.1% (±2.3%)

Table A-20. Overall Grade 2 Filipino EGRA Results, by Sex and Cohort

	Cohort 1 (SY 13/14)		**Cohort 2 (SY14/15)**		**Cohort 3 (SY 15/16)**		**Gain Cohort 1 to Cohort 3**		**Effect Size (d)**	
Filipino Subtest	**Boys**	**Girls**	**Boys**	**Girls**	**Boys**	**Girls**	**Boys**	**Girls**	**Boys**	**Girls**
Initial Sound Identification (percent correct)	55.7%	59.6%	74.3 %	82.7%	67.6%	76.6%	11.9%	17.0 %	0.33	0.53
Letter Sounds(percent correct)	17.2%	21.0%	26.4%	32.8%	22.0%	26.3%	4.8%	5.3%	0.30	0.32
Letter Correct (per min)	17.5	21.0	26.4	32.6	22.1	26.6	4.6	5.6	0.29	0.31
Familiar Words(percent correct)	61.4%	75.4%	66%	81.2%	57.4%	74.4%	-4.0%	-1.0%	-0.12	-0.04
Familiar Words Correct(per min)	33.1	42.0	36.2	47.8	30.5	40.0	-2.6	-2.0	-0.13	-0.11
Nonsense Words (percent correct)	40.5%	52.0%	44.4%	58.8%	39.9%	53.6%	-0.6%	1.6%	-0.02	0.06
Nonsense Words Correct (per min)	20.5	26.0	22.3	29.8	20.1	27.0	-0.4	1.0	-0.03	0.07
Oral Passage Reading (percent correct)	47.9%	64.3%	67.7%	82.9%	50.4%	66.9%	2.5%	2.6%	0.08	0.09
Words Correct in a Text (per min)	31.3	42.8	40.6	55.8	33.7	45.2	2.4	2.4	0.10	0.11
Prosody score	1.8	2.3	2.8	3.2	2.1	2.6	0.3	0.3	0.33	0.32
Reading Comprehension	24.9%	31.2%	39.6%	45.4%	41.0%	42.5%	16.1%	11.3	0.47	0.34

(percent correct)								%		
Reading Comprehension: untimed (pct correct)	--	--	51.9%	57.2%	50.4%	51.4%	--	--	1.48	1.68
Listening Comprehension (percent correct)	50.0%	48.1%	62.4%	65.1%	32.5%	36.9%	-17.5%	-11.2 %	-0.68	-0.43
Dictation Composite (percent correct)	40.0%	51.3%	56.0%	68.0%	53.4%	65.9%	13.4%	14.6 %	0.47	0.61

Table A-21. Grade 2 EGRA Filipino Subtests Summary, by Province and Cohort

Descriptive Statistics for Grade 2 Filipino Subtests, by Province				
Province	**Subtest**	**Cohort 1 (SY13/14)**	**Cohort 2 (SY14/15)**	**Cohort 3 (SY15/16)**
BOHOL				
	Initial Sound Identification (percent correct)	-	82.1%	72.8%
	Letter Sounds (percent correct)	-	35.2%	27.7%
	Familiar Words (percent correct)	-	75.3%	68%
	Nonsense Words (percent correct)	-	53.5%	48.4%
	Oral Passage Reading (percent correct)	-	76.5%	60.9%
	Words Correct in a Text (per min)	-	49.9	41.7
	Prosody	-	2.9	2.4
	Reading Comprehension (percent correct)	-	51.1%	47.9%
	Listening Comprehension (percent correct)	-	68.7%	35.1%
	Dictation Composite (percent correct)	-	63.7%	55.3%
CEBU				
	Initial Sound Identification (percent correct)	61.4%	81.0%	84.1%
	Letter Sounds (percent correct)	19.4%	27.4%	26.1%
	Familiar Words (percent correct)	70.6%	76.1%	69.2%
	Nonsense Words (percent correct)	47.9%	53.3%	49.6%
	Oral Passage Reading (percent correct)	58.1%	78.0%	61.7%
	Words Correct in a Text (per min)	38.5	50.5	40.9
	Prosody	2.1	3.0	2.4
	Reading Comprehension (percent correct)	27.7%	43.9%	37.7%
	Listening Comprehension (percent correct)	44.3%	56.1%	32.0%
	Dictation Composite (percent correct)	48.2%	64.4%	66.2%
ILOCOS NORTE	Initial Sound Identification (percent correct)	-	68.7%	46.4%
	Letter Sounds (percent correct)	-	21.3%	12.5%
	Familiar Words (percent correct)	-	68.7%	54.7%
	Nonsense Words (percent correct)	-	48.4%	38%
	Oral Passage Reading (percent correct)	-	70.1%	46.3%
	Words Correct in a Text (per min)	-	44.6	30.1
	Prosody	-	3.0	2.2
	Reading Comprehension (percent correct)	-	49.3%	26.8%
	Listening Comprehension (percent correct)	-	71.1%	32.3%
	Dictation Composite (percent correct)	-	59.3%	46.8%
Province	**Subtest**	**Cohort 1**	**Cohort 2**	**Cohort 3**

		(SY13/14)	(SY14/15)	(SY15/16)
ILOCOS SUR	Initial Sound Identification (percent correct)	-	68.3%	55.5%
	Letter Sounds (percent correct)	-	23.1%	15.3%
	Familiar Words (percent correct)	-	68.7%	54%
	Nonsense Words (percent correct)	-	48.0%	38.5%
	Oral Passage Reading (percent correct)	-	70.4%	48.6%
	Words Correct in a Text (per min)	-	43.3	32.6
	Prosody	-	3.1	2.4
	Reading Comprehension (percent correct)	-	43.6%	32%
	Listening Comprehension (percent correct)	-	59.8%	33.6%
	Dictation Composite (percent correct)	-	56.9%	56.6%
LA UNION	Initial Sound Identification (percent correct)	46.6%	71.8%	56.7%
	Letter Sounds (percent correct)	18.3%	27.2%	18.4%
	Familiar Words (percent correct)	61.9%	65.7%	62.8%
	Nonsense Words (percent correct)	41.3%	44.1%	42.8%
	Oral Passage Reading (percent correct)	50.2%	68.4%	55.2%
	Words Correct in a Text (per min)	32.7	40.4	36.6
	Prosody	1.8	3.0	2.3
	Reading Comprehension (percent correct)	29.2%	46.9%	48.0%
	Listening Comprehension (percent correct)	62.8%	71.3%	43.3%
	Dictation Composite (percent correct)	38.1%	54.0%	60.9%

DETAILED GRADE 2 ENGLISH RESULTS

Table A-22. Overall Grade 2 English EGRA Results, by Cohort

Grade 2 English EGRA Subtests	Cohort 1 (SY 2013/14)[38]	Cohort 2 (SY 2014/15)	Cohort 3 (SY 2015/16)	Change Cohort 2 to Cohort 3	Effect Size (Cohort 2, Cohort 3)
Initial Sound Identification (pct correct)	--	73.1%	69.5%	-3.6% (±2.4)	-0.11
Letter Sounds(percent correct)	--	39.1%	34.2%	-4.9% (±1.5)	-0.25
Letter Correct (per min)	--	39.1	34.2	-4.8 (±1.5)	-0.24
Familiar Words(pct correct)	--	63.0%	56.7%	-6.3%(±2.7)	-0.17
Familiar Words Correct(per min)	--	39.5	34.0	-5.5 (±2.0)	-0.21
Nonsense Words (pct correct)	--	49.2%	42.7%	-6.5% (±2.2)	-0.21
Nonsense Words Correct (per min)	--	25.9	21.9	-3.9 (±1.2)	-0.23
Oral Passage Reading (pct correct)	--	72.4%	65.5%	-6.9% (±2.4)	-0.22
Words Correct in a Text (per min)	--	58.1	40.4	-17.8 (±2.0)	-0.65
Prosody score	--	2.8	2.3	-0.5 (±0.1)	-0.51
Reading Comprehension: timed (pct correct)	--	15.9%	16.8%	0.9% (±1.7)	0.04
Reading Comprehension: untimed (pct correct)	--	27.5%	27.9%	0.5% (±2.1)	0.02
Listening Comprehension (pct correct)	--	17.6%	18.1%	0.5% (±1.8)	0.02
Dictation Composite (pct correct)	--	36.2%	32.5%	-3.7% (±1.6)	-0.17

[38] English EGRA results were not collected for Cohort 1 (SY 2013/14); results were only collected during SY 2014/15 and SY 2015/16.

Table A-23. Grade 2 English EGRA Subtests – Percent of Learners with Zero Scores

Percent of Learners with Zero Scores on Grade 2 English EGRA Subtests				
Grade 2 English Subtest	**Cohort 1 (SY 13/14)**	**Cohort 2 (SY 14/15)**	**Cohort 3 (SY 15/16)**	**Change in Zero Scores (Cohort 1, Cohort 3**
Initial Sound Identification (pct correct)	--	7.4%	12.3%	4.9% (±2.2%)
Letter Sounds (pct correct)	--	1.2%	5.8%	4.6% (±1.4%)
Letter Correct (per min)	--	1.2%	5.9%	4.6% (±1.4%)
Familiar Words(pct correct)	--	9.1%	11.6%	2.5% (±2.3%)
Familiar Words Correct(per min)	--	9.1%	11.6%	2.5% (±2.3%)
Nonsense Words (pct correct)	--	8.4%	11.3%	2.9% (±2.2%)
Nonsense Words Correct (per min)	--	8.4%	11.3%	2.9% (±2.2%)
Oral Passage Reading (pct correct)	--	2.0%	6.5%	4.6% (±1.5%)
Words Correct in a Text (per min)	--	2.0%	6.5%	4.6% (±1.5%)
Prosody score	--	--	--	--
Reading Comprehension: timed (pct correct)	--	49.5%	49.8%	0.2% (±3.8%)
Reading Comprehension: untimed (pct correct)	--	33.1%	29.1%	-4.0% (±3.5%)
Listening Comprehension (pct correct)	--	56.3%	52.3%	-4.0% (±3.7%)
Dictation Composite (pct correct)	--	2.4%	6.2%	3.9% (±1.5%)

Table A-24. Grade 2 English EGRA Results, by Sex and Cohort

	Cohort 1 (SY13/14)		**Cohort 2 (SY14/15)**		**Cohort 3 (SY15/16)**		**Gain Cohort 1 to Cohort 3**		**Effect Size (d)**	
English Subtest	**Boys**	**Girls**	**Boys**	**Girls**	**Boys**	**Girls**	**Boys**	**Girls**	**Boys**	**Girls**
Initial Sound Identification (pct correct)	--	--	69.3%	77.1%	64.7%	74.4%	-4.6%	-2.7%	-0.13	-0.09
Letter Sounds (pct correct)	--	--	35.4%	43.0%	30.9%	37.5%	-4.5%	-5.5%	-0.22	-0.27
Letter Correct (per min)	--	--	35.4	42.9	30.9	37.5	-4.5	-5.4	-0.22	-0.26
Familiar Words(percent correct)	--	--	53.8%	72.7%	46.9%	66.4%	-6.9%	-6.3%	-0.18	-0.19
Familiar Words Correct(per min)	--	--	31.6	47.7	28.0	39.9	-3.6	-7.8	-0.14	-0.32
Nonsense Words (pct correct)	--	--	42.1%	56.6%	36.3%	49.1%	-5.8%	-7.5%	-0.19	-0.25
Nonsense Words Correct (per min)	--	--	21.7	30.2	18.6	25.3	-3.1	-4.9	-0.19	-0.31
Oral Passage Reading (percent correct)	--	--	65.8%	79.4%	56.7%	74.2%	-9.1%	-5.2%	-0.25	-0.18
Words Correct in a Text (per min)	--	--	48.9	67.8	34.7	46.0	-14.2	-21.8	-0.62	-1.12
Prosody score	--	--	2.6	3.0	2.1	2.5	-0.5	-0.5	-0.48	-0.61
Reading Comprehension (pct correct)	--	--	14.5%	17.4%	15.1%	18.4%	0.6%	1.0%	0.03	0.04
Reading Comprehension: untimed (pct correct)	--	--	24.2%	30.8%	24.3%	31.2%	0.1%	0.4%	0.00	0.01
Listening Comprehension (pct correct)	--	--	14.5%	20.8%	17.4%	18.7%	2.9%	-2.1%	0.12	-0.08
Dictation Composite (pct correct)	--	--	30.3%	42.4%	27.8%	37.2%	-2.5%	-5.2%	-0.12	-0.24

Table A-25. Grade 2 English EGRA Subtests Summary, by Province and Cohort

Descriptive Statistics for Grade 2 English Subtests, by Province				
Province	**Subtest**	**Cohort 1 (SY13/14)**	**Cohort 2 (SY14/15)**	**Cohort 3 (SY15/16)**
BOHOL	Initial Sound Identification (percent correct)	-	76.2%	68.2%
	Letter Sounds (percent correct)	-	45.0%	39.1%
	Familiar Words (percent correct)	-	64.7%	58%
	Nonsense Words (percent correct)	-	51.1%	42.6%
	Oral Passage Reading (percent correct)	-	73.3%	65.3%
	Words Correct in a Text (per min)	-	59.2	41.1
	Prosody	-	2.0	2.3
	Reading Comprehension (percent correct)	-	15.8%	19.1%
	Listening Comprehension (percent correct)	-	21.0%	16.4%
	Dictation Composite (percent correct)	-	35.9%	28.4%
CEBU	Initial Sound Identification (percent correct)	-	76.4%	83.4%
	Letter Sounds (percent correct)	-	37.5%	36.9%
	Familiar Words (percent correct)	-	66.6%	59.7%
	Nonsense Words (percent correct)	-	54.2%	47.3%
	Oral Passage Reading (percent correct)	-	76.3%	69.9%
	Words Correct in a Text (per min)	-	62.8	42.5
	Prosody	-	2.8	2.4
	Reading Comprehension (percent correct)	-	16.6%	16.7%
	Listening Comprehension (percent correct)	-	22.1%	21.4%
	Dictation Composite (percent correct)	-	40.0%	36.9%
ILOCOS NORTE	Initial Sound Identification (percent correct)	-	60.2%	43.7%
	Letter Sounds (percent correct)	-	27.3%	16.6%
	Familiar Words (percent correct)	-	57.6%	46.2%
	Nonsense Words (percent correct)	-	42.2%	33.1%
	Oral Passage Reading (percent correct)	-	66.7%	54.4%
	Words Correct in a Text (per min)	-	53.5	32.5
	Prosody	-	2.8	2.3
	Reading Comprehension (percent correct)	-	16.4%	10.6%
	Listening Comprehension (percent correct)	-	6.4%	14.3%
	Dictation Composite (percent correct)	-	33.4%	24.8%
ILOCOS SUR	Initial Sound Identification (percent correct)	-	65.1%	53.4%
	Letter Sounds (percent correct)	-	32.4%	22.1%
	Familiar Words (percent correct)	-	56.1%	45.9%
	Nonsense Words (percent correct)	-	38.7%	32.8%
	Oral Passage Reading (percent correct)	-	63.8%	56.4%
	Words Correct in a Text (per min)	-	49.6	34.5
	Prosody	-	2.9	2.4
	Reading Comprehension (percent correct)	-	13.0%	12%
	Listening Comprehension (percent correct)	-	5.4%	13.9%
	Dictation Composite (percent correct)	-	28.3%	29.5%

Province	Subtest	Cohort 1 (SY13/14)	Cohort 2 (SY14/15)	Cohort 3 (SY15/16)
LA UNION	Initial Sound Identification (percent correct)	-	65.6%	54.6%
	Letter Sounds (percent correct)	-	36.5%	26.6%
	Familiar Words (percent correct)	-	55.2%	55.8%
	Nonsense Words (percent correct)	-	39.4%	40.7%
	Oral Passage Reading (percent correct)	-	67.2%	63.9%
	Words Correct in a Text (per min)	-	49.4	38.9
	Prosody	-	2.9	2.4
	Reading Comprehension (percent correct)	-	16.3%	15.8%
	Listening Comprehension (percent correct)	-	7.7%	17.6%
	Dictation Composite (percent correct)	-	32.8%	37.3%

GRADE 3 EGRA RESULTS

DETAILED GRADE 3 FILIPINO EGRA RESULTS

Table A-26. Grade 3 Filipino EGRA Results for Cohort 3 (SY15/16)

Grade 3 Filipino EGRA Subtests	Cohort 3 (SY 2015/16)
Familiar Words (pct correct)	81.5%
Familiar Words Correct (per min)	47.8
Oral Passage Reading (pct correct)	83.0%
Words Correct in a Text (per min)	57.8
Prosody score	2.9
Reading Comprehension: timed (pct correct)	69.8%
Reading Comprehension: untimed (pct correct)	80.0%
Listening Comprehension (pct correct)	49.4%
Dictation Composite (pct correct)	70.8%

Table A-27. Grade 3 Filipino EGRA Results – Percent of Learners with Zero Scores

Percent of Learners with Zero Scores on Grade 3 Filipino EGRA Subtests	
Grade 3 Filipino EGRA Subtests	**Cohort 3 (SY 2015/16)**
Familiar Words (pct correct)	1.7%
Familiar Words Correct (per min)	1.7%
Oral Passage Reading (pct correct)	1.6%
Words Correct in a Text (per min)	1.6%
Prosody score	--
Reading Comprehension: timed (pct correct)	3.6%
Reading Comprehension: untimed (pct correct)	1.4%
Listening Comprehension (pct correct)	12.5%
Dictation Composite (pct correct)	1.7%

Table A-28. Grade 3 Filipino EGRA Results by for Cohort 3 (SY 15/16) by Sex

Subtest	Boys	Girls	Gender Gap	*p* (sig)
Familiar Words(pct correct)	74.7% (±2.0%)	88.2% (±2.0%)	13.5% (±2.5%)	.000
Familiar Words Correct(per min)	42.0 (±1.4)	53.4 (±1.4)	11.4 (±1.9)	.000
Oral Passage Reading (pct correct)	76.6% (±2.0%)	89.4% (±1.4%)	12.8% (±2.5%)	.000
Words Correct in a Text (per min)	50.0 (±1.8)	65.5 (±1.7)	15.5 (±2.4)	.000
Prosody score	2.7 (±0.1)	3.2 (±0.1)	5.1 (±2.7)	.000
Reading Comprehension: timed (pct correct)	67.2% (±2.0%)	72.3% (±1.9%)	0.5% (±0.1%)	.000
Reading Comprehension: untimed (pct correct)	77.7% (±1.6%)	82.2% (±1.4%)	4.5% (±2.2%)	.000
Listening Comprehension (pct correct)	46.7% (±1.9%)	52.1% (±2.0%)	5.4% (±2.8%)	.000
Dictation Composite (pct correct)	64.6% (±1.7%)	76.9% (±1.4%)	12.3% (±2.2%)	.000

Table A-29. Grade 3 Filipino EGRA Subtests Summary for Cohort 3 (SY 15/160, by Province

Descriptive Statistics for Grade 3 Filipino Subtests, by Province		
Province	**Subtest**	**Cohort 3 (SY15/16)**
BOHOL	Familiar Words (percent correct)	84.9%
	Oral Passage Reading (percent correct)	85.7%
	Words Correct in a Text (per min)	62.5
	Prosody	2.9
	Reading Comprehension (percent correct)	76.0%
	Listening Comprehension (percent correct)	51.4%
	Dictation Composite (percent correct)	70.1%
CEBU	Familiar Words (percent correct)	81.2%
	Oral Passage Reading (percent correct)	82.8%
	Words Correct in a Text (per min)	56.3
	Prosody	2.9
	Reading Comprehension (percent correct)	65.2%
	Listening Comprehension (percent correct)	46.7%
	Dictation Composite (percent correct)	72.8%
ILOCOS NORTE	Familiar Words (percent correct)	72.5%
	Oral Passage Reading (percent correct)	76.3%
	Words Correct in a Text (per min)	50.3
	Prosody	2.9
	Reading Comprehension (percent correct)	62.4%
	Listening Comprehension (percent correct)	45.2%
	Dictation Composite (percent correct)	66.8%
ILOCOS SUR	Familiar Words (percent correct)	75.2%
	Oral Passage Reading (percent correct)	79.2%
	Words Correct in a Text (per min)	52.0
	Prosody	3.0
	Reading Comprehension (percent correct)	62.0%
	Listening Comprehension (percent correct)	48.3%
	Dictation Composite (percent correct)	69.5%
LA UNION	Familiar Words (percent correct)	80.3%
	Oral Passage Reading (percent correct)	81.3%
	Words Correct in a Text (per min)	55.4
	Prosody	3.0
	Reading Comprehension (percent correct)	73.7%
	Listening Comprehension (percent correct)	54.5%
	Dictation Composite (percent correct)	69.1%

DETAILED GRADE 3 ENGLISH EGRA RESULTS

Table A-30. Grade 3 English EGRA Results for Cohort 3 (SY 2015/16)

Grade 3 English EGRA Subtests	Cohort 3 (SY 2015/16)
Familiar Words (pct correct)	75.8% (±1.6%)
Familiar Words Correct(per min)	51.7 (±1.4)
Oral Passage Reading (pct correct)	76.0% (±1.4%)
Words Correct in a Text (per min)	60.6 (±1.6)
Prosody score	2.7 (±0.0)
Reading Comprehension: timed (pct correct)	34.7% (±1.8%)
Reading Comprehension: untimed (pct correct)	46.9% (±1.8%)
Listening Comprehension (pct correct)	26.7% (±1.4%)
Dictation Composite (pct correct)	46.7% (±1.3%)

Table A-31. Grade 3 English EGRA Results for Cohort 3 (SY 2015/16) – Percent of Learners with Zero Scores

Percent of Learners with Zero Scores on Grade 3 English EGRA Subtests	
Grade 3 English EGRA Subtests	**Cohort 3 (SY 2015/16)**
Familiar Words (pct correct)	4.5%
Familiar Words Correct(per min)	4.5%
Oral Passage Reading (pct correct)	3.5%
Words Correct in a Text (per min)	3.5%
Prosody score	--
Reading Comprehension: timed (pct correct)	39.1%
Reading Comprehension: untimed (pct correct)	23.9%
Listening Comprehension (pct correct)	38.0%
Dictation Composite (pct correct)	2.9%

Table A-32. Grade 3 English EGRA Results for Cohort 3 (SY 15/16), by Sex

Subtest	Boys	Girls	Gender Gap	*p* (sig)
Familiar Words(pct correct)	67.8% (±2.5%)	83.7% (±1.8%)	15.9% (±3.1%)	.000
Familiar Words Correct(per min)	43.8 (±2.0)	59.5 (±1.9)	15.7% (±2.7%)	.000
Oral Passage Reading (pct correct)	68.2% (±2.2%)	83.7% (±1.7%)	15.5% (±2.8%)	.000
Words Correct in a Text (per min)	51.0 (±2.1)	70.0 (±2.1)	19.0% (±3%)	.000
Prosody score	2.4 (±0.1)	3.0 (±0.1)	0.5 (±0.1)	.000
Reading Comprehension: timed (pct correct)	25.7% (±2.3%)	43.3% (±2.6%)	17.6% (±3.5%)	.000
Reading Comprehension: untimed (pct correct)	37.9% (±2.5%)	55.4% (±2.5%)	17.5% (±3.5%)	.000
Listening Comprehension (pct correct)	22.6% (±1.9%)	30.8% (±2.1%)	8.2% (±2.8%)	.000
Dictation Composite (pct correct)	38.6% (±1.6%)	54.7% (±1.7%)	16.1% (±2.4%)	.000

Table A-33. Grade 3 English EGRA Subtests Summary for Cohort 3 (SY 15/160, by Province

Descriptive Statistics for Grade 3 English Subtests, by Province		
Province	**Subtest**	**Cohort 3 (SY15/16)**
BOHOL	Familiar Words (percent correct)	77.8%
	Oral Passage Reading (percent correct)	78.3%
	Words Correct in a Text (per min)	64.8
	Prosody	2.7
	Reading Comprehension (percent correct)	41.5%
	Listening Comprehension (percent correct)	29.4%
	Dictation Composite (percent correct)	44.9%
CEBU	Familiar Words (percent correct)	76.3%
	Oral Passage Reading (percent correct)	76.9%
	Words Correct in a Text (per min)	60.9
	Prosody	2.7
	Reading Comprehension (percent correct)	32.6%
	Listening Comprehension (percent correct)	26.9%
	Dictation Composite (percent correct)	49.5%
ILOCOS NORTE	Familiar Words (percent correct)	67.8%
	Oral Passage Reading (percent correct)	67.9%
	Words Correct in a Text (per min)	52.6
	Prosody	2.8
	Reading Comprehension (percent correct)	34.2%
	Listening Comprehension (percent correct)	23.8%
	Dictation Composite (percent correct)	45.4%
ILOCOS SUR	Familiar Words (percent correct)	69.1%
	Oral Passage Reading (percent correct)	69.6%
	Words Correct in a Text (per min)	51.2
	Prosody	2.8
	Reading Comprehension (percent correct)	22.6%
	Listening Comprehension (percent correct)	20.4%
	Dictation Composite (percent correct)	41.7%
LA UNION	Familiar Words (percent correct)	76%
	Oral Passage Reading (percent correct)	74.5%
	Words Correct in a Text (per min)	57.0
	Prosody	2.7
	Reading Comprehension (percent correct)	30.4%
	Listening Comprehension (percent correct)	24.3%
	Dictation Composite (percent correct)	48.5%

ANNEX 4. SUMMARY SCOPE RESULTS

Overall, average SCOPE scores decreased from SY 2014/15 to SY 2015/16; this was consistent for both the average Classroom Structure and Language and Literacy Instruction scores. As seen in the table below, SCOPE scores of observed teachers decreased in 3 out of 6 observed Classroom Structure domains and 6 out of seven observed Language and Literacy Instruction domains.

Figure A- 6. Average SCOPE Scores from SY 2014/15 to SY 2015/16 (n=27)

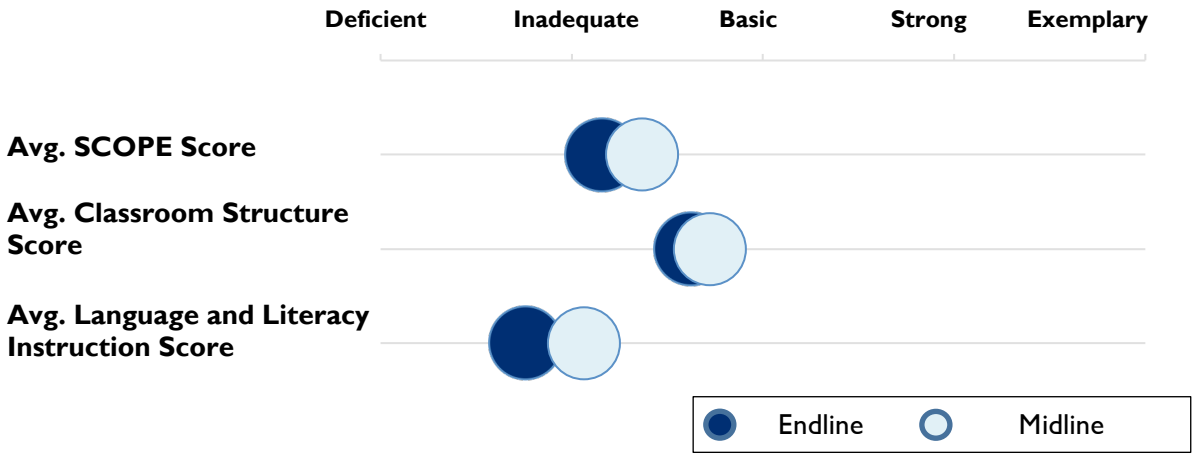


Table A-34. Average SCOPE Scores from SY 2013/14, SY 2014/15 and SY 2015/16 (n=27) for SCOPE Domains

SCOPE Domain	Baseline	Midline	Endline	Difference Midline/ Endline	Effect Size (Cohen's d)[39]
CLASSROOM STRUCTURE					
Positive Learning Environment	2.4	2.9	2.8	-0.1	-0.04
Effective Grouping Strategies	1.6	2.2	1.7	-0.4	-0.52
Participation of All Learners	2.2	3.0	2.6	-0.3	-0.48
Opportunities for Reflection	1.2	1.7	2.0	0.4	0.41
Classroom Materials	2.6	3.5	3.6	0.2	0.15
Management of Reading/ Writing Instruction	2.4	3.1	3.1	0.0	-0.04
LANGUAGE AND LITERACY INSTRUCTION					
Opportunities for Oral Language Development	1.5	2.6	2.1	-0.5	-0.46
Opportunities for Meaningful Reading	1.6	1.9	1.6	-0.3	-0.27
Opportunities for Learning to Decode and Spell Words	1.6	1.7	1.7	0.0	0.00
Opportunities for Developing Reading Fluency	1.0	2.1	1.7	-0.4	-0.56
Opportunities for Developing Vocabulary	1.7	2.4	2.1	-0.3	-0.31
Opportunities for Developing Comprehension	1.4	2.4	2.0	-0.4	-0.40
Writing Instruction	1.1	1.4	1.1	-0.3	-0.50

[39] Effect size is a statistical measure that is used to estimate the magnitude of difference between two measures. Effect sizes are interpreted as follows, according to Cohen (1998): "small, d = .2," "medium, d = .5," and "large, d = .8".