Opportunities for Equitable Access to Quality Basic Education (OPEQ)

Baseline Report: Results from the Early Grade Reading Assessment, the Early Grade Math Assessment, and children’s demographic data in Katanga Province, Democratic Republic of Congo

October 2011

Prepared by:
Catalina Torrente & John Lawrence Aber
New York University

Anjuli Shivshanker
International Rescue Committee Research, Evaluation & Learning Unit

Evaluation Team:
Jeannie Annan & Tom Bundervoet
International Rescue Committee Research, Evaluation & Learning Unit

With special thanks to:
Anita Anastacio, Aissatou Balde, Willy Mpwate, Jennifer Sklar, Nina Weisenhorn, the Democratic Republic of Congo Ministry of Primary, Secondary, and Professional Education (MEPSP), and our talented data collection supervisors and enumerators

USAID
FROM THE AMERICAN PEOPLE

INTERNATIONAL RESCUE COMMITTEE

NEW YORK UNIVERSITY
This report is made possible by the generous support of the American people through the United States Agency for International Development (USAID). The contents are the responsibility of the IRC and do not necessarily reflect the views of USAID or the United States Government.

Acknowledgments

Thanks to all our dedicated research assistants and volunteers: Brian Johnston, Todd Solomon, Dirk Witteneen, Mahjabeen Raza, Marissa McCoy, Alana Gross, Emily Jacobson, Eugene Chang, Marian Tes and Claudette Carter
# TABLE OF CONTENTS

**EXECUTIVE SUMMARY** ............................................................................................. 8

1. **INTRODUCTION** ..................................................................................................... 8
2. **EVALUATING IMPACT: EVALUATION DESIGN AND GOALS** .................................. 8
3. **DATA COLLECTION AND ANALYSIS** .................................................................... 9
4. **KEY FINDINGS FROM SCHOOL, TEACHER, AND STUDENT SOCIO-DEMOGRAPHIC DATA** ................................................................. 10
5. **KEY FINDINGS: EARLY GRADE READING ASSESSMENT** ................................... 11
6. **KEY FINDINGS: EARLY GRADE MATH ASSESSMENT** ........................................ 12
7. **RELATIVE PERFORMANCE OF CHILDREN IN OPEQ BASELINE** .......................... 13
8. **LIMITATIONS** ......................................................................................................... 14
9. **CONCLUSIONS AND RECOMMENDATIONS** ......................................................... 15

**INTRODUCTION** ......................................................................................................... 18

**RESEARCH DESIGN** .................................................................................................. 19

1. **OVERVIEW** ............................................................................................................ 19
2. **INSTRUMENTS** ....................................................................................................... 19
3. **SAMPLE** ................................................................................................................. 20
4. **INFORMED CONSENT** .......................................................................................... 21
5. **ANALYTICAL METHOD** ......................................................................................... 21

**SAMPLE DESCRIPTION** .......................................................................................... 22

6. **SCHOOL CHARACTERISTICS** ............................................................................... 22
7. **TEACHER CHARACTERISTICS** ............................................................................ 25
8. **CHILD DEMOGRAPHIC CHARACTERISTICS** ....................................................... 30

**BASELINE FINDINGS – EARLY GRADE READING ASSESSMENT (EGRA)** .................. 34

1. **VARIABILITY IN CHILDREN’S EGRA SCORES** ..................................................... 34
2. **ZERO SCORES BY GRADE** ..................................................................................... 36

**SUMMARY OF EGRA DESCRIPTIVE ANALYSIS** ...................................................... 36

2. **AVERAGE CORRECT SCORES BY GENDER** ......................................................... 38
3. **AVERAGE CORRECT SCORES BY SUBDIVISION** ................................................ 39

**DESCRIPTIVE ANALYSIS BY SUBTEST** .................................................................... 40

1. **VOCABULARY** ....................................................................................................... 40
2. **INITIAL SOUND IDENTIFICATION (PHONOLOGICAL AWARENESS)** ..................... 44
3. **KNOWLEDGE OF GRAPHEMES (LETTER KNOWLEDGE)** .................................... 47
4. **FAMILIAR WORD READING (FLUENCY)** ............................................................. 51
5. **INVENTED WORD DECODING (FLUENCY)** ......................................................... 54
6. **ORAL PASSAGE READING (CONNECTED TEXT FLUENCY)** ................................ 57
7. **READING COMPREHENSION** ................................................................................. 60
8. **LISTENING COMPREHENSION** ............................................................................. 62
9. **WRITING A COMPLETE SENTENCE** ..................................................................... 65
EARLY GRADE MATH ASSESSMENT (EGMA) DESCRIPTIVE ANALYSIS .............................................................67
  1. VARIABILITY IN CHILDREN’S EGMA SCORES ..................................................................................67
  1. ZERO SCORES BY GRADE .............................................................................................................69
SUMMARY OF DESCRIPTIVE ANALYSIS – EGMA .................................................................................69
  2. AVERAGE CORRECT SCORES BY GENDER ..................................................................................70
  3. AVERAGE CORRECT SCORES BY SUBDIVISION ........................................................................71
DESCRIPTIVE ANALYSIS BY SUBTESTS .............................................................................................73
  1. NUMBER IDENTIFICATION ...........................................................................................................73
  2. NUMBER (QUANTITY) DISCRIMINATION (COMPARISON OF QUANTITIES) .............................76
  3. MISSING NUMBER (NUMBER PATTERNS) ..................................................................................79
  4. ADDITION ......................................................................................................................................82
  5. SUBTRACTION ..............................................................................................................................85
  6. MULTIPLICATION .........................................................................................................................88
  7. WORD PROBLEMS .......................................................................................................................91
  8. SHAPE RECOGNITION (SHAPE IDENTIFICATION) ......................................................................94
  9. SHAPE NAMING ..........................................................................................................................97
CONCLUSIONS AND RECOMMENDATIONS .....................................................................................100
LIMITATIONS ........................................................................................................................................102
REFERENCES .........................................................................................................................................104
ANNEX A. SCHOOL SELECTION CRITERIA ..........................................................................................106
ANNEX B. FURTHER DETAILS ON DATA ANALYSIS METHODS ......................................................107
ANNEX C. GUIDE FOR THE INTERPRETATION OF BOXPLOTS .......................................................109
Tables

Table 1. Differences in EGRA scores by raw score/percentage between OPEQ and PAQUED, Grades 2&4 ................................................................. 13
Table 2. Differences in EGMA scores by percentage between OPEQ and PAQUED, Grades 2&4 ................................................................. 13
Table 3. Number of children per school per subdivision ......................................................................................................................... 22
Table 4. Number of teachers by grade and subdivision .......................................................................................................................... 25
Table 5. Number of children by subdivision, grade and gender .................................................................................................................. 31
Table 6. Percentage of children per subdivision who answered “yes” to wealth questions ................................................................. 33
Table 7. Overview of EGRA subtests ......................................................................................................................................................... 34
Table 8. Distribution of variance in children's EGRA scores .................................................................................................................... 35
Table 9. Percent improvement in EGRA zero scores .............................................................................................................................. 37
Table 10. Differences between mean correct scores by gender, by grade ............................................................................................. 38
Table 11. Mean correct scores by subdivision; Subdivision rank by subtest .......................................................................................... 39
Table 12. Overview of EGMA subtests .................................................................................................................................................. 67
Table 13. Distribution of variance in children's EGMA scores .................................................................................................................. 68
Table 14. Percent improvement in EGMA zero scores .......................................................................................................................... 70
Table 15. Differences between mean correct scores by gender, by grade ............................................................................................. 71
Table 16. Mean correct scores by subdivision; Subdivision rank by subtest .......................................................................................... 72
Table 17. Questions included in the 2-factor solution for child-reported household wealth ................................................................. 108

Figures and Graphs

Graph 1. School religious affiliation and auspice by subdivision ............................................................................................................ 23
Graph 2. Percentage of male and female teachers by subdivision ....................................................................................................... 25
Graph 3. Percentage of male and female teachers by grade level ....................................................................................................... 26
Graph 4. Percentage of teachers’ mother tongues by subdivision ...................................................................................................... 26
Graph 5. Average teacher age by subdivision ....................................................................................................................................... 27
Graph 6. Teacher education level; secondary school distribution ....................................................................................................... 27
Graph 7. Teacher education level by subdivision ................................................................................................................................... 28
Graph 8. Average number of years of teaching experience by subdivision .......................................................................................... 29
Graph 9. Percentage of boys and girls by subdivision ............................................................................................................................ 31
Graph 10. Percentage of children’s self-reported mother tongues by subdivision .................................................................................. 32
Graph 11. EGRA Zero Scores by Grade .................................................................................................................................................. 32
Graph 12. Percentage of children per number of correctly identified words ........................................................................................... 41
Graph 13. Mean vocabulary scores by grade and gender .......................................................................................................................... 42
Graph 14. Mean Vocabulary scores by subdivision .................................................................................................................................. 42
Graph 15. Distribution of Vocabulary scores by grade and subdivision .................................................................................................. 43
Graph 16. Percentage of children per number of correctly identified sounds .......................................................................................... 44
Graph 17. Mean scores by grade and gender for Initial Sound Identification .......................................................................................... 45
Graph 18. Mean scores by subdivision for Initial Sound Identification .................................................................................................. 45
Graph 19. Distribution of Initial Sound Identification scores by grade and by subdivision ........................................................................ 46
Graph 20. Percentage of children per number of correctly identified graphemes .................................................................................... 48
Graph 21. Mean scores by grade and gender for Knowledge of Graphemes ............................................................................................. 48
Graph 22. Mean scores by subdivision for Knowledge of Graphemes ........................................... 49
Graph 23. Distribution of Knowledge of Graphemes scores by grade and subdivision ......................... 50
Graph 24. Percentage of children per number of familiar words read in one minute ............................. 51
Graph 25. Mean scores by grade and gender, Familiar Word Reading ........................................... 52
Graph 26. Mean scores by subdivision, Familiar Word Reading ..................................................... 52
Graph 27. Distribution of Familiar Word Reading Scores by subdivision and grade ............................ 53
Graph 28. Percentage of children per number of invented words read in one minute .......................... 54
Graph 29. Mean scores by grade and gender, Invented Word Decoding .......................................... 55
Graph 30. Mean scores by subdivision, Invented Word Decoding .................................................. 55
Graph 31. Distribution of Invented Word Decoding scores by subdivision and grade ............................ 56
Graph 32. Percentage of children per number of words read in an oral passage in one minute ................. 57
Graph 33. Mean scores by grade and gender, Oral Passage Reading ............................................... 58
Graph 34. Mean scores by subdivision, Oral Passage Reading ....................................................... 58
Graph 35. Distribution of Oral Passage Reading scores by subdivision and grade ................................. 59
Graph 36. Percentage of children per number of correct reading comprehension answers .................. 60
Graph 37. Mean scores by grade and gender, Reading Comprehension ........................................... 61
Graph 38. Mean scores by subdivision, Reading Comprehension .................................................... 61
Graph 39. Percentage of children per number of correct listening comprehension answers .................. 62
Graph 40. Mean scores by grade and gender, Listening Comprehension ........................................... 63
Graph 41. Mean scores by subdivision, Listening Comprehension .................................................... 63
Graph 42. Distribution of Listening Comprehension scores by subdivision and grade ............................ 64
Graph 43. Percentage of children per number of correctly written words ......................................... 65
Graph 44. Mean scores by gender, Writing a Complete Sentence .................................................... 66
Graph 45. Mean scores by subdivision, Writing a Complete Sentence .............................................. 66
Graph 46. EGMA Zero Scores by grade ......................................................................................... 70
Graph 47. Percentage of children per number of correct numbers identified ........................................ 73
Graph 48. Mean scores by grade and gender, Number Identification ............................................... 74
Graph 49. Mean scores by subdivision, Number Identification ........................................................ 74
Graph 50. Distribution of scores by subdivision and by grade, Number Identification ............................ 75
Graph 51. Percentage of children per number of correct Quantity Discrimination questions ................ 76
Graph 52. Mean scores by grade and gender, Quantity Discrimination ............................................. 77
Graph 53. Mean scores by subdivision, Quantity Discrimination ....................................................... 77
Graph 54. Distribution of scores by subdivision and grade, Quantity Discrimination ............................ 78
Graph 55. Percentage of children per number of correct answers. Missing Number ............................ 79
Graph 56. Mean scores by grade and gender, Missing Number .......................................................... 80
Graph 57. Mean scores by subdivision, Missing Number ..................................................................... 80
Graph 58. Distribution of scores by subdivision and grade, Missing Number ......................................... 81
Graph 59. Percentage of children per number of correct answers, Addition ....................................... 82
Graph 60. Mean scores by grade and gender, Addition ......................................................................... 83
Graph 61. Mean scores by subdivision, Addition .................................................................................. 83
Graph 62. Distribution of scores by subdivision and grade, Addition ................................................... 84
Graph 63. Percentage of children per number of correct answers, Subtraction .................................... 85
Graph 64. Mean scores by grade and gender, Subtraction .................................................................... 86
Graph 65. Mean scores by subdivision, Subtraction ............................................................................. 86
Graph 66. Distribution of Subtraction scores by subdivision and grade ................................................. 87
Graph 67. Percentage of children per number of correct answers, Multiplication ........................................ 88
Graph 68. Mean scores by grade and gender, Multiplication ........................................................................... 89
Graph 69. Mean scores by subdivision, Multiplication .................................................................................... 89
Graph 70. Distribution of scores by subdivision and grade, Multiplication .................................................. 90
Graph 71. Percentage of children per correct answers, Word Problems .................................................... 91
Graph 72. Mean scores by grade and gender, Word Problems ........................................................................ 92
Graph 73. Mean scores by subdivision, Word Problems ................................................................................ 92
Graph 74. Distribution of scores by subdivision and grade, Word Problems ................................................ 93
Graph 75. Percentage of children per number of shapes identified ............................................................... 94
Graph 76. Mean scores by grade and gender, Shape Identification ............................................................... 95
Graph 77. Mean scores by subdivision, Shape Identification ........................................................................ 95
Graph 78. Distribution of scores by subdivision and grade, Shape Identification ........................................... 96
Graph 79. Percentage of children per number of correct answers, Shape Naming ....................................... 97
Graph 80. Mean scores by grade and gender, Shape Naming ........................................................................ 98
Graph 81. Mean scores by subdivision, Shape Naming ................................................................................ 98
Graph 82. Distribution of scores by subdivision and grade, Shape Naming ................................................ 99

Figure 1. Guide for the interpretation of boxplots ........................................................................................... 109
Executive Summary

1. Introduction

Extreme poverty, periods of violent conflict and systemic mismanagement of resources have characterized life in the Democratic Republic of Congo (DRC) for the past three decades, with a disproportionate negative impact on an entire generation of Congo’s most vulnerable citizens – children and youth. Economic downturns and divestment in Congo have led to a two-thirds reduction in the Ministry of Education budget (World Bank, 2004) and diminished employment opportunities for young people (Ministère des Affaires Sociales, Humanitaires et Solidarité Nationale, 2009.).

Today, approximately 4.4 million school-age children are not enrolled in school and an estimated 31% of school-age children have never set foot in a classroom. Demand for education and training outstrips the available supply of educational opportunities, both formal and non-formal. An estimated 53% of the population is under the age of 18 (UNICEF and the Government of DRC, 2008), and the future economic development, security and wellbeing of DRC and its young people are in peril if the education and training needs for young people are not met. However, the development of recent stabilization initiatives in eastern Congo and increased government commitment to the education sector are strong indicators that now is the right time for increased investment in the DRC education sector.

To respond to the needs of young people and to improve their well-being, the International Rescue Committee (IRC), in partnership with Research Triangle Institute (RTI), the Flemish Association for Development Cooperation and Technical Assistance (VVOB) and the Institute of Human Development and Social Change (IHDSC) at New York University (NYU), has undertaken an initiative entitled Opportunities for Equitable Access to Quality Basic Education (OPEQ).

OPEQ aims to improve equitable access to quality basic education in North Kivu, South Kivu and Katanga provinces for more than 499,000 girls, boys, young women and young men. This will be achieved by developing and implementing teacher training systems that lead to positive outcomes on teaching and learning; strengthening community-level participation in education to ensure equitable access to education; providing out-of-school youth sustainable, relevant alternative, non-formal, vocational and livelihoods education and training; and building the institutional capacity of the Ministry of Primary, Secondary and Professional Education (MEPSP).

2. Evaluating Impact: Evaluation design and goals

In order to assess the impact of the school and community-based components of OPEQ on the quality of teaching and learning in primary school classrooms, a rigorous impact evaluation is underway. The impact evaluation has three key objectives:

1) To estimate the impact of OPEQ on teachers’ motivation and performance and children’s outcomes;
2) To examine how the impact of OPEQ varies by individual characteristics (e.g., gender, grade), schools (e.g., school size) and communities (e.g., access to resources); and
To use impact evaluation data to continually improve OPEQ over the life of the project and inform similar projects and policy efforts in DRC and internationally.

The impact evaluation is a cluster-randomized evaluation with a wait-list control design. This means that in the impact evaluation, groups of schools (clusters) are randomly assigned to begin participating in the OPEQ project in different years. While the clusters are waiting to begin OPEQ, they serve as control schools and when the clusters begin OPEQ, they become treatment schools. The wait-list control design is also beneficial from the programming perspective because the scale and novelty of OPEQ required a staggered scale-up. Using this design, all schools will have the opportunity to participate in OPEQ, with each new school year seeing an improved intervention, and the IRC and NYU will have the opportunity to collect impact evaluation data.

This report presents descriptive findings and recommendations from the baseline data collection conducted from March to May of 2011, in a random sample of 84 schools and 4,956 children in grades 2, 3 and 4 across six educational subdivisions (i.e., Kambove, Kasenga, Kalemie, Kongolo, Mutshatsha, and Lubudi) in Katanga province.

### 3. Data Collection and Analysis

Six different kinds of questionnaires were used to obtain data from three sources: school principals, teachers and children. School principals answered a short questionnaire about basic school characteristics (e.g., number of enrolled children, religious affiliation). Teachers completed a questionnaire about their backgrounds as well as their living and teaching conditions. Children provided information about their backgrounds and households through a demographic questionnaire, were assessed on their reading and math skills using the EGRA (Early Grade Reading Assessment) and EGMA (Early Grade Math Assessment), and assessed on their social-emotional learning using a novel instrument developed by NYU.

Basic descriptive analyses on children’ demographic characteristics are reported based on subdivision, grade and gender, and analyzed through different types of educational and household risks. Results of the Early Grade Reading Assessment (EGRA) and Early Grade Math Assessment (EGMA) are presented through descriptive analyses for the overall sample and comparisons are drawn between groups (grade, gender, subdivision) for each of the subtests. Finally, we draw conclusions based on these findings and offer recommendations. Definitive conclusions about the causal impact of the OPEQ initiative on children's reading, math and socio-emotional learning cannot be drawn until collection and analysis of follow-up data with children and teachers in 2012 and 2013.

Future reports of baseline data will describe (a) the backgrounds, living conditions, professional development, pedagogical beliefs and motivations of teachers, and (b) the social and emotional development of the children.
4. **Key findings from school, teacher, and student socio-demographic data**

School characteristics: 84 schools total
- 75% to 93% of schools are religiously affiliated.
- 67% to 91% of schools are supported by the MEPSP.
- Schools varied substantially in size within and between subdivisions. Between subdivisions, school size ranged from an average of 260 students/school to 485 students/school.
- Kambove, Kalemie, and Lubudi had the largest schools on average; Kasenga and Kongolo had the smallest schools on average; and Mutshatsha schools fell in the middle.

Teacher characteristics: 470 teachers total
- There are more than twice as many male teachers are female teachers in all subdivisions except for Kalemie, which was almost gender balanced.
- As grade level increases, the gender imbalance also increases. By 6\textsuperscript{th} grade, 98% of teachers sampled are male.
- 15% of teachers identify French as their mother-tongue (most come from Kalemie); 47% of teachers identify Swahili as their mother tongue.
- In Lubudi, Mutshatsha, Kasenga, and Kambove, between 56% and 90% of teachers report a language other than French or Swahili as their mother tongue.
- 96% of teachers have completed at least one year of secondary school, with 69% actually having completed secondary school.
- On average, teachers across subdivisions have 8-13 years of teaching experience. In Kambove, teachers had an average of 22 years of teaching experience.
- Kongolo and Lubudi had more teacher sampled than other subdivisions; Kambove had the fewest teachers sampled.

Children’s characteristics: 4,956 children total from grades 2-4
- The sample is fairly gender-balanced, although there are more boys than girls overall. In addition, the proportion of girls declines from grade 2 to grade 4.
- 94% to 100% of children reported speaking Kiswahili at home in all subdivisions except Kasenga. In Kasenga, most children reported speaking Kibemba.
- Only 2.7% of children overall reported speaking French at home.
- 20% of children moved households in the past year.
- 89% of children live with both parents.
- 9% of children repeated the previous grade.
- 79% of children report that their parents can read and write.
- 55% of children are sometimes late to school; approximately 50% are sometimes absent.
5. Key Findings: Early Grade Reading Assessment

The EGRA tests children’s proficiency with fundamental reading skills in French through nine subtests. All directions are completed in the child’s mother-tongue, though test questions and responses are in French. The vast majority of children have extremely low levels of proficiency in fundamental reading skills assessed through EGRA. The majority of children cannot read, write, or say the alphabet in French.

Zero scores: Zero scores measure the percent of children who could not respond correctly to even one question per subtest. High zero scores indicate that most children have no proficiency with the content.

- Seven of the nine subtests had zero scores over 50%.
- 68% of children scored zero for Oral Passage Reading; 91% of children scored zero for Reading Comprehension.
  - Reading comprehension is particularly striking as the zero score is made up only of children who successfully read at least 10 words of the reading comprehension passage.
- Only two subtests (Vocabulary and Knowledge of Graphemes) had zero scores under 5%; however, average scores for these subtests were still low.
- Zero scores improve with grade level. In some cases, zero score improve by over 80% in 4th grade; however, the average scores for these subtests remain very low.

Average scores by subtest: All nine subtests have extremely low average scores.

- Vocabulary had the highest average score: 40%.
- Reading comprehension and Invented word decoding have the lowest average scores: 0%- 10%.
- Initial sound identification, Familiar word reading, Oral passage reading, Listening Comprehension, and Knowledge of Graphemes have average scores from 11% to 20%.

Average scores by subdivision: Performance is very low across subdivisions, but there is significant variation within and between subdivisions for the majority of subtests.

- Children answered on average 1% (Reading comprehension, Kasenga) to 48% (Writing a complete sentence, Kalemie) of questions correctly.
- Each subdivision performs similarly across subtests: Kalemie is always ranked first, Kongolo is almost always ranked second, Mutshatsha and Lubudi are ranked either third or fourth, and Kambove and Kasenga are ranked fifth or sixth.
- Performance gaps between subdivisions are extremely varied: for example, for Invented word decoding, Kalemie, the highest-performing subdivision, had an average score nine times higher than the average score of the lowest performing subdivision.

Average scores by gender and by grade

- Boys outperformed girls across all grades in the majority of subtests.
- The gap between boys and girls average scores increase as children progress in grade level.
- Some gender differences in average scores seem large, but once we consider the amount of variation between children and within grades, they are actually relatively modest.
- Average scores improved as grade level improved, but the amount of variability within grades is large and important. For example, on the Vocabulary subtest, some 2nd graders perform as well as 4th graders, and some 4th graders are still scoring 0, unable to identify a single word.
6. **Key Findings: Early Grade Math Assessment**

The EGMA tests children’s proficiency with fundamental math skills in French through nine subtests. All directions are completed in the child’s mother-tongue, test questions and answers are in French. The vast majority of children have very low levels of proficiency in fundamental math skills assessed through EGMA.

Zero scores: Zero scores measure the percent of children who could not respond correctly to even one question per subtest. High zero scores are concerning because they indicate that most children have no proficiency with the content of the subtest.

- Zero scores range from less than 1% (Number Identification) to 41% (Subtraction). At least 20% of children could not respond to a single question on six of nine subtests.
- Zero scores decrease as grade levels increase, as much as 27% by 4th grade. This is true for all subtests except Word Problems, for which the zero score increases in 3rd grade and decreases in 4th grade. This may be due to the increase in subtest difficulty between 2nd and 3rd grade; however, three other subtests also increase in difficulty, yet their zero scores consistently decrease.
- About half of the subtests have zero scores below 10% by 4th grade. (Number identification, Quantity discrimination, Missing number, Addition, and Shape identification.)

Average scores by subtest: All nine subtests have low average scores.

- Shape identification and Quantity discrimination have the highest average scores: 62% and 61% respectively. Multiplication has the lowest average score at 10%.
- The remaining subtest average scores fall between 22% and 46%.

Average scores by subdivision:

- Children answered on average 5.6% (Multiplication, Kambove) to 69% (Identification of shapes, Kalemie) of questions correctly.
- Some subdivisions perform similarly across subtests. Kasenga is ranked sixth for almost all subtests. Kambove and Lubudi tend to rank towards the bottom (fourth or fifth). Mutshatsha tends to rank towards the top (first, second, or third), but the remaining two subdivisions (Kalemie and Kongolo) rank from first to fifth place, depending on the subtest.
- For four of nine subtests, the highest scoring subdivisions had average scores over twice as large as the lowest scoring subdivisions.

Average scores by gender and by grade: Boys outperformed girls across all grades in all subtests.

- Some gender differences in average scores seem large, but once we consider the amount of variation between children and within grades, they are actually relatively modest.
- The gap between boys and girls average scores increases as children progress in grade level, except for some tests that increase in difficulty between 2nd and 3rd grade (Number Identification, Quantity Discrimination, Addition, and Word Problems). Average scores increase again by 4th grade.
7. Relative Performance of Children in OPEQ Baseline

In 2010, another USAID-funded education project called PAQUED piloted EGRA and EGMA in a random sample of 256 children from 12 schools in the provinces of Bandudu, Equateur, and Province Orientale (RTI International 2011b & 2011c) in the DRC. While EGRA and EGMA have been administered in other Francophone countries, RTI stresses the importance of not comparing scores between countries. PAQUED administered EGRA and EGMA to grades 2, 4, and 6, and presented means or percentage correct for each test. Table 1 shows the difference in EGRA scores by raw score and percentage between OPEQ and PAQUED for grades 2 and 4; Table 2 shows the difference in EGMA scores by percentage between OPEQ and PAQUED for grades 2 and 4. PAQUED reported only percentage scores in the EGMA report. Table 1 shows PAQUED scores were much higher for Vocabulary and Initial Sound Identification in grades 2 and 4 and Listening Comprehension in Grade 2; otherwise, OPEQ scores tended to be modestly higher. Table 2 shows children in the PAQUED baseline perform better in almost all subtests than children in the OPEQ baseline, with the exception of Word Problems in Grade 4. Otherwise, differences tended to be larger between OPEQ and PAQUED scores in Grade 2, reducing in Grade 4.

Table 1. Differences in EGRA scores by raw score/percentage between OPEQ and PAQUED, Grades 2&4

<table>
<thead>
<tr>
<th>EGRA SUBTEST</th>
<th>Difference in Means (PAQUED-OPEQ)</th>
<th>Grade 2</th>
<th>Grade 4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Raw score</td>
<td>Percent</td>
<td>Raw Score</td>
</tr>
<tr>
<td>Vocabulary (20)</td>
<td>3.07</td>
<td>15.4%</td>
<td>2.22</td>
</tr>
<tr>
<td>Initial Sound Identification (10)</td>
<td>1.96</td>
<td>19.6%</td>
<td>3.67</td>
</tr>
<tr>
<td>Knowledge of Graphemes (100)</td>
<td>-4.85</td>
<td>-4.85%</td>
<td>-2.00</td>
</tr>
<tr>
<td>Familiar Word Reading (50)*</td>
<td>-</td>
<td>-</td>
<td>-2.98</td>
</tr>
<tr>
<td>Invented Word Decoding (50)*</td>
<td>-</td>
<td>-</td>
<td>-0.16</td>
</tr>
<tr>
<td>Oral passage reading (50)*</td>
<td>-</td>
<td>-</td>
<td>0.43</td>
</tr>
<tr>
<td>Reading comprehension (5)*</td>
<td>-</td>
<td>-</td>
<td>-0.18</td>
</tr>
<tr>
<td>Listening comprehension (5)</td>
<td>0.56</td>
<td>11.2%</td>
<td>-0.35</td>
</tr>
<tr>
<td>Writing a complete sentence (3)*</td>
<td>-</td>
<td>-</td>
<td>-0.18</td>
</tr>
</tbody>
</table>

*Not administered to Grade 2 students for both PAQUED and OPEQ

Table 2. Differences in EGMA scores by percentage between OPEQ and PAQUED, Grades 2&4

<table>
<thead>
<tr>
<th>EGMA SUBTEST</th>
<th>Difference in Means (PAQUED-OPEQ)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Grade 2</td>
</tr>
<tr>
<td>Number Identification (30)</td>
<td>20.4%</td>
</tr>
<tr>
<td>Quantity Discrimination (10)</td>
<td>14.5%</td>
</tr>
<tr>
<td>Missing Number (10)</td>
<td>7.3%</td>
</tr>
<tr>
<td>Addition (24)</td>
<td>40.42%</td>
</tr>
<tr>
<td>Subtraction (24)</td>
<td>29.1%</td>
</tr>
<tr>
<td>Multiplication (10)*</td>
<td>-</td>
</tr>
<tr>
<td>Word Problems (6)</td>
<td>14.83%</td>
</tr>
<tr>
<td>Shape Identification (4)**</td>
<td>11.2%</td>
</tr>
<tr>
<td>Shape Naming (4)**</td>
<td>-</td>
</tr>
</tbody>
</table>

* Test not given to Grade 2; ** PAQUED did not administer test in Grade 4; ***PAQUED did not administer test
8. **Limitations**

There are several important limitations to this report. First, the dearth of similar studies in DRC posed challenges to our ability to select culturally and psychometrically valid measures of children's outcomes and to contextualize the interpretation of our results. Regarding the assessment of math and reading in particular, there are no available norms for EGRA and EGMA that would allow for meaningful comparisons between our sample and other samples or populations. In addition, a formal method to summarize the performance of children across math or reading subtests is not available. The OPEQ team will continue to refine the measurement procedures and instruments based on analysis conducted with the baseline data.

Second, the OPEQ program and design and methods for the impact evaluation were being developed simultaneously during year 1 of the project. This means that not all measures and procedures utilized as part of the evaluation were fine-tuned to the needs of the program. The coordination between program and evaluation design will continue improving as the research and program mature.

Third, results presented in this report are entirely descriptive, which means that definite conclusions cannot be reached about the reasons that explain some of the patterns found. The pace at which data needs to be analyzed and fed back to program and partners prohibited conducting appropriate analysis to explore associations between children's demographic characteristics and academic performance. Future reports will include progress made in this area.

Fourth, it is important to note that while efforts were made to select a representative sample of schools in Katanga province, it was necessary to exclude some schools. Therefore, our results are only generalizable to schools with similar profiles to the ones participating in OPEQ.

Finally, the language diversity in the sample posed additional challenges for data collectors who in some cases had to do real-time translations of instructions into local languages. Ideally, assessments are to be translated and back translated prior to data collection to ensure that meaning is preserved and administration instructions are standard.

In spite of these limitations, this report offers a rich picture of 2nd to 4th graders’ reading and math performance in six educational subdivisions in Katanga province, based on a rigorous and sound research design. Results from this report are expected to inform the continuous development of OPEQ, which aims to be increasingly responsive to the specific needs of boys and girls in these regions.

Moreover, this report provides a solid baseline to evaluate the impact of OPEQ in children’s learning after 1 and 2 years of implementation. The impact evaluation of OPEQ will contribute to the knowledge base about what works for improving children’s educational outcomes, not only in the eastern DRC but, to the extent that children in other regions face similar conditions, in the country as a whole.
9. Conclusions and Recommendations

Conclusion 1: Taken together, baseline findings from the six subdivisions targeted by OPEQ in Katanga province revealed very low levels of performance in reading and math, regardless of subdivision, gender and grade. These results seem similar to other EGRA and EGMA results in other DRC provinces.

The zero scores on the Oral Reading Passage and Reading Comprehension subtests illustrate clearly that primary school children cannot read in French. 68% of children could not read even one word of a simple passage; 91% of children who were able to read the first line of the passage could not answer a simple comprehension question about the content. As Tables 1 and 2 indicate, OPEQ baseline results are similar to the PAQUED pilot studies with EGRA and EGMA. This most likely indicates that low scores are characteristic of primary school children in the DRC, and OPEQ’s baseline score are not out of the ordinary.

While in several reading subtests the majority of children obtained 0 scores, this was rare for math subtests. As suggested before, it is difficult to grasp new concepts and master new skills in a foreign language. Compared to reading in French, math may be less dependent on language proficiency and this may explain in part why children obtained better results in math than reading. In addition, children most likely encounter opportunities to use math concepts in French in their daily lives, for instance when they give or receive a telephone number or count money. They most likely have almost no exposure to written French, and as the language data shows, 98% of children have no exposure to spoken French in the home either.

One of the factors that account for the overall low levels of performance is that while children are obtaining increasingly better scores as they progress from one grade to the next, average differences between grades are rather modest in the majority of the subtests. Most likely, most children are not improving by large amounts from grade to grade because they are getting promoted to higher grade levels without demonstrating mastery of concepts at lower grade levels, as per a national policy in the DRC. Developing strategies to identify and support children who are falling behind their same-grade peers would make an important difference in children’s learning. Given the large number of children per classroom, effective and feasible strategies could rely upon group work or other forms of peer learning in which children who are performing at higher levels may provide scaffolding for peers performing at lower levels.

Other factors that may account for the small-to-modest gains children are making might be related to low exposure to the subject matter (for instance due to inefficient use of instruction time or low attendance rates); difficulties with grasping concepts in a new language (as French is for most children); and/or the use teaching strategies centered around rote learning that fail to promote knowledge that readily translates to new contexts, such as assessment situations. All these are areas that, if improved, could result in greater gains in children’s skills over the course of each grade and throughout primary school.

Conclusion 2: The most remarkable finding is the considerable variation in performance between children in the same grades and subdivisions. This calls for a deeper understanding of the factors that might account for the wide gaps between low and high performing children who live in the same regions and attend the same grades. While the typical child showed low levels of performance across
subtests for EGRA and EGMA, there were large numbers of children who performed well above or well below average for a child in their grade. When setting benchmarks, it is important to focus on closing the performance gap between high-performing and low-performing children in the same grade levels.

Understanding the factors that may explain why some children are doing much better or worse than others is critical for developing strategies to benefit all children. Low repetition rates (i.e., only 9% of children reported repeating the previous grade), which likely reflect the national policy of automatic promotion, may explain some of the significant heterogeneity found within grades, in the sense that a large number of children is promoted to the next grade without having mastered the basic skills from the previous grade. This promotion policy makes it vital to identify children who are falling behind their same-grade peers in order to provide the supports they need to boost their performance before they are promoted.

The OPEQ team will continue working on unpacking the factors that might explain such variation in the coming months. While variability within subdivisions and grades was one of the most striking findings, it is also worth noting the comparatively high average performance of children from Kalemie and the low performance of children from Kasenga. A better understanding of the reasons that explain the relative success of Kalemie, and the low average performance in Kasenga, could be pivotal in informing efforts to improve academic attainment across the board.

**Conclusion 3: There are two factors that could underpin the low performance across reading and math assessments: a) the challenge of transitioning from mother-tongue instruction to French instruction and b) the poor quality of instruction around basic math and reading skills.**

These two types of issues, one a language issue, the other an instructional issue, may hinder their learning and academic performance across subjects. At this point, we do not have enough information to determine to which degree each of these factors contributes to the overall problem. Extremely low scores across all EGRA subtests, particularly for Initial Sound Identification, Knowledge of Graphemes, and Vocabulary, seem to indicate that the majority of children are not familiar with the French language. Extremely low scores on these tests could also indicate that the curriculum does not adequately focus on the foundational skills of reading or that teachers do not have the necessary skills to efficiently teach foundational reading skills.

Sequencing lessons so that foundational skills are taught *and mastered* before more complex skills is critical for children’s learning process. For instance, children need to be aware and knowledgeable of the relationship between sounds and written symbols in order to read and write, and they need to correctly identify numbers and discriminate quantities in order to perform basic arithmetic operations. Learning is often cumulative; therefore, children who do not master the foundational reading and math skills when they are expected to do so, are likely to benefit the least from instruction on more complex skills and to continue falling behind their peers.

OPEQ is focusing on improving instructional quality through providing better in-service training and support to teachers; however, if any effort to improve children’s academic outcomes is to be effective, it should intentionally address the language barriers that most children are facing in school.
supports provided for children's acquisition of French as a second language also seems key to improving children's academic success.

**Conclusion 4: There were important differences between boys and girls, with boys outperforming girls in every reading and math subtest.**

It is important to note that there are small but non-trivial gender differences in school enrollment (or possibly attendance) and in reading and math performance. While these differences are generally small they should not be overlooked. The gap between boys' and girls' performance appears to increase as children get older and, for academics in particular, consistently favors boys over girls, which could result in substantial differences by the time children reach high school. The consistent gender differences found in this baseline can be taken as a sign of systematic, yet poorly understood, differences in the learning opportunities offered to boys and girls in these regions of DRC. Intervention efforts should be tailored to the needs and strengths of girls and boys, if they are to reduce, and not inadvertently propagate or enlarge, the gap between genders.

**Conclusion 5: Even though children showed improvements as they progressed through primary school, such improvements were rather modest.** For example, children only gained an average of 1.5 words from one grade to the next in the Vocabulary subtest, and were able to solve 1 or 2 more Addition problems than children in the grade below. If one considers the low levels of performance in 2nd grade and the modest gains between grades, it is not surprising that even 4th graders were still struggling with the most basic reading and math skills.
Introduction

The International Rescue Committee (IRC), in partnership with Research Triangle Institute (RTI), the Flemish Association for Development Cooperation and Technical Assistance (VVOB) and the Institute of Human Development and Social Change (IHDSC) at New York University (NYU), has undertaken an initiative entitled Opportunities for Equitable Access to Quality Basic Education (OPEQ) to respond to the needs of young people and to improve their well-being. OPEQ aims to improve equitable access to quality basic education in North Kivu, South Kivu and Katanga provinces of the Democratic Republic of Congo (DRC), for more than 499,000 girls, boys, young women and young men.

As part of the IRC’s commitment to gathering evidence about the impact of our interventions, IRC and NYU are conducting a cluster-randomized impact evaluation of OPEQ. The impact evaluation will gather evidence about the impact of OPEQ on teachers' motivation and performance and children’ outcomes; examine whether the impact of OPEQ varies by individuals’ characteristics (e.g., gender, grade), schools (e.g., school size) and communities (e.g., access to resources); make improvements over the life of the project; and inform similar projects and policy efforts in DRC and internationally.

Baseline data were collected from March to May, 2011, in Katanga province to assess children's current living conditions, math and reading performance and socio-emotional wellbeing. Data were also collected on teachers' living and work conditions, and other factors that may impact their motivation and performance (e.g., teacher accountability, perceived obstacles to teaching).

The current report presents an overview of the research design and sample, including school, teacher and child characteristics and descriptive findings on children’s baseline living conditions, reading and math performance. Basic descriptive analyses on children’ demographic characteristics are reported based on subdivision, grade and gender, and analyzed through household risks and wealth. Results of the Early Grade Reading Assessment (EGRA) and Early Grade Math Assessment (EGMA) are presented through descriptive analyses for the overall sample and comparisons are drawn between groups (grade, gender, subdivision) for each of the subtests. Lastly, we draw conclusions based on these findings and offer recommendations. Definitive conclusions about the causal impact of the OPEQ initiative on children's reading, math and socio-emotional learning cannot be drawn until collection and analysis of follow-up data with children and teachers in 2012 and 2013.

Future reports of baseline data will describe (a) the backgrounds, living conditions, professional development, pedagogical beliefs and motivations of teachers, and (b) the social and emotional development of the children.
Research Design

1. Overview

A total of 203 schools grouped in 54 geographically defined clusters were selected to participate in the OPEQ intervention in Katanga province. School selection criteria are described in Annex A. The OPEQ team, with support from the DRC Ministry of Education (MEPSP), attempted to select a highly representative sample of schools in order to improve the generalizability of the evaluation findings.

The cluster-randomized impact evaluation took advantage of the staggered scale-up of the OPEQ project to create a wait-list control evaluation design, in which schools waiting to begin OPEQ serve as the control groups. In Katanga province, the 54 clusters of selected schools were randomly split into three groups of 20, 17 and 17 clusters that will start OPEQ in consecutive years. In May 2011, after the baseline was conducted, public lotteries were held in Katanga province so that representatives from each school cluster could be involved in assigning the school clusters. This academic year (2011-12), OPEQ will be piloted in the first group of 20 clusters. The remaining two groups of 17 clusters will receive the intervention in academic years 2012-13 or 2013-14.

Thus, the impact evaluation in 2012-13, year 2 of the OPEQ intervention, will use an experimental wait-list control group design with 20 treatment and 17 control clusters. In 2013-14, year 3 of the OPEQ intervention, OPEQ will be implemented in the previous year’s wait-list control group clusters and the previous year’s experimental condition for a second year (37 clusters total).

2. Instruments

School characteristics were collected through surveys administered to school directors during the school selection phase. The survey included questions such as number of children enrolled by grade and gender, number and gender breakdown of teachers, auspice, and religious affiliation.

Teacher characteristics were assessed through oral interviews administered to teachers. The interview included questions about education and teaching experience, personal health, safety and security, motivation and school working conditions, among others. The maximum time slotted for each teacher interview was 90 minutes.

Children’s demographic characteristics were assessed through structured oral interviews in which children were asked questions about their home language, family structure, health and illness, current and last year’s grade level, means and time of transportation to school, absenteeism and tardiness, child labor, household composition, household mobility and parents’ literacy and education. To assess household wealth, children were asked a series of questions about household assets and the characteristics of the house’s construction materials.

Children’s literacy skills were assessed using the Early Grade Reading Assessment (EGRA; RTI, 2009). The EGRA consists of eight subtests: Vocabulary (Cronbach’s $\alpha = 0.92^{2a}$), Initial Sound Identification ($\alpha =$

---

1 All interviews and assessments were conducted in French (the official language) and Kiswahili (the most common local language in most subdivisions) by local staff trained in data collection procedures by the OPEQ team.

2 $1a$ (Piper, 2009) $1b$ (Crouch & Korda, 2008).
0.92\(^{1a}\), Knowledge of Graphemes (\(\alpha = 0.94\)^{1a}), Familiar word reading (\(\alpha = 0.92\)^{1a}), Invented word reading (\(\alpha = 0.80\)^{1b}), Oral passage reading fluency with comprehension (\(\alpha = 0.94\)^{1a}), Listening comprehension (\(\alpha = 0.82\)^{1b}), and Writing of a complete sentence (\(\alpha\) not applicable). The EGRA takes about 15 minutes per child to administer, all of the subtests are orally administered, and some subtests are timed. The assessment Overall reliability for the EGRA has been found to be very high with practitioners reporting Cronbach’s \(\alpha\) values in the range of 0.80\(^{1b}\)-0.95\(^{1a}\). The EGRA has been used in 31 different countries including DRC.

**Children’s math skills** were tested using the Early Grade Mathematics Assessment (Reubens & Kline, 2009) The EGMA consists of nine sub-tests: Number Identification (\(\alpha = 0.94\)), Quantity Discrimination (\(\alpha = 0.68\)), Missing Number/pattern completion (\(\alpha = 0.86\)), Word Problems (\(\alpha = 0.55\)), addition/subtraction/multiplication/division\(^3\) problems (\(\alpha = 0.87\)-0.91), Shape Identification/geometry (\(\alpha = 0.43\)), and Shape Naming/geometry (\(\alpha = 0.63\)). The EGMA takes about 15 minutes per child to administer, all of the subtests are orally administered, and some subtests are timed. Overall reliability for the EGMA has been generally found to be high (\(\alpha = 0.9\)). The EGMA is still under development but has been successfully used in Kenya, the DRC, Malawi, and Rwanda.

### 3. Sample

The baseline sample is a random selection of 84 of the 203 schools targeted by OPEQ in Katanga province. The schools were distributed across six educational sub-divisions, the smallest administrative district in the MEPSP structure: Kambove, Kasenga, Kalemie, Mutshatsha, Lubudi, and Kongolo. Using simple cluster sampling, at least one school from every cluster was selected. At the time of sampling, cluster size ranged from three to six schools. Due to uneven numbers of schools per cluster, if a cluster had three or less schools, one school was selected for the baseline evaluation; if a cluster had four or more schools, then two schools were selected for the baseline evaluation.

In the 84 schools randomly selected to participate in the baseline, 6,311 children in 2\(^{nd}\), 3\(^{rd}\) and 4\(^{th}\) grades, and 470 teachers in grades 1 through 6 were also randomly selected to participate.\(^4\) Children who participated in the baseline were also randomly assigned to complete different pairs of assessments in addition to a demographic questionnaire (e.g., math and reading, math and socio-emotional well being) in order to reduce the burden on children. This *purposive missing data design* means that 4,965 of the 6,311 children who participated in the baseline were administered either EGRA or EGMA.

Even though all children were expected to complete the demographic survey and 2 other instruments, there are 264 (3.8\%) children who are missing all demographic information and 291 (4.2\%) children who only answered the demographic survey. For the purposes of this report, children with missing demographic data, children with only demographic data, and children who did not take EGRA or EGMA were excluded from the current analysis.

\(^3\) Note that “division” was not included in the version of EGMA used in the current study.

\(^4\) Sampling of children was stratified within schools and grade levels; sampling of teachers was always stratified within schools, and when numbers allowed, within grades. While teachers in grades 2-4 were prioritized, teachers from all grades were often invited to participate in order to reach the target sample size.
4. **Informed Consent**

Assent and consent were requested from all children and teachers at the time of data collection. Refusal to participate was very rare. The Ministry of Education of DRC and OPEQ’s field team widely advertised the evaluation in each school and community to ensure that parents were fully informed and had the opportunity to ask any questions, raise any concerns and opt out.

5. **Analytical Method**

Analysis of children’s demographic, reading, and math data focused on describing differences in results between groups (gender, grade, subdivision), measuring the size of these differences, and describing how results varied within each group.

Descriptive group comparisons were made for gender, grade and subdivision by examining mean scores, standard deviations, frequencies, percentages and boxplots for EGRA, EGMA, and demographic data.

For a more detailed description of the analytical methods used to assess between-group differences or develop aggregate measures of household economic status, please see Annex B.
Sample Description

6. School Characteristics

To describe the sample, we will start with describing the 84 schools selected to participate in the baseline by size and by religious auspice and affiliation. Table 3 shows the number of children per school, per subdivision. On average, schools in the six educational subdivisions varied substantially in size. Schools in Kambove, Kalemie and Lubudi were on average the largest; schools in Kasenga and Kongolo were on average the smallest; and schools in Mutshatsha fell in the middle. There is a similar trend when comparing the largest schools in each subdivision. For instance, the largest school in Kalemie had 1,221 children whereas the largest school in Kongolo had only 591 children.

<table>
<thead>
<tr>
<th>Subdivision</th>
<th># Schools</th>
<th>Number of Children (1-6 grade)</th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
<th>Std. Dev</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kambove</td>
<td>22</td>
<td></td>
<td>114</td>
<td>1213</td>
<td>475.09</td>
<td>245.71</td>
</tr>
<tr>
<td>Kasenga</td>
<td>30</td>
<td></td>
<td>78</td>
<td>651</td>
<td>298.67</td>
<td>145.36</td>
</tr>
<tr>
<td>Kalemie</td>
<td>32</td>
<td></td>
<td>199</td>
<td>1221</td>
<td>484.38</td>
<td>252.12</td>
</tr>
<tr>
<td>Kongolo</td>
<td>43</td>
<td></td>
<td>120</td>
<td>591</td>
<td>259.74</td>
<td>105.22</td>
</tr>
<tr>
<td>Mutshatsha</td>
<td>34</td>
<td></td>
<td>82</td>
<td>1240</td>
<td>399.85</td>
<td>268.07</td>
</tr>
<tr>
<td>Lubudi</td>
<td>39</td>
<td></td>
<td>122</td>
<td>1130</td>
<td>459.74</td>
<td>252.00</td>
</tr>
</tbody>
</table>

Graph 1 shows that similar to many schools in the DRC, a large majority of schools in the six subdivisions are religiously affiliated (75% to 93%). Most schools are also state-supported. With the exception of Kasenga, within each subdivision the percentages of state-supported and religiously-affiliated schools mirror each other closely. Kasenga had 23% more religiously-affiliated than state-supported schools.
Graph 1. School religious affiliation and auspice by subdivision

<table>
<thead>
<tr>
<th>Subdivision</th>
<th>Secular</th>
<th>Religious</th>
<th>No state support</th>
<th>State supported</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kambove</td>
<td>10%</td>
<td>90%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>Kasenga</td>
<td>30%</td>
<td>80%</td>
<td>20%</td>
<td>0%</td>
</tr>
<tr>
<td>Kalemie</td>
<td>20%</td>
<td>80%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>Kongolo</td>
<td>0%</td>
<td>100%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>Mutshatsha</td>
<td>50%</td>
<td>30%</td>
<td>20%</td>
<td>10%</td>
</tr>
<tr>
<td>Lubudi</td>
<td>0%</td>
<td>100%</td>
<td>0%</td>
<td>0%</td>
</tr>
</tbody>
</table>
## Summary of Teacher Characteristics

### Number of teachers
- Largest amount of teachers sampled: Kongolo (95) and Lubudi (96).
- Smallest amount of teachers sampled: Kambove (48).

### Gender balance
- Almost all subdivisions have at least twice as many male teachers as female teachers.
  (Exception: Kalemie)
- The gender imbalance increases with grade level, with over 98% teachers being male by grade 6.

### French as mother-tongue
- Kongolo (28%) and Kalemie (48%) have the largest percentage of teachers who reported French as their mother tongue.
- Less than 2% of teachers in Lubudi, Mutshatsha, Kambove, and Kasenga reported French as their mother tongue.

### Age
- The age range of teachers is very large, from 17-74 years.
- Average ages by subdivision ranges from 33 to 45 years.

### Education
- Over 95% over teachers have completed at least one year of secondary education, with at least 50% in each subdivision having completed secondary school.

### Years of teaching experience
- The range of years of teaching experience is very large, from 0-54 years.
- Teachers in Kambove report the largest average number of years of teaching experience; teachers in Kongolo report the smallest average number of years.
- Inside subdivisions, there was often substantial variability (one to two decades) in years of teaching experience.
2. Teacher Characteristics

470 teachers were sampled across the 84 schools selected to participate in the baseline from grades 1 to 6. The number of teachers sampled was fairly balanced across grade levels, even when broken into subdivisions; however, more teachers were sampled in Kongolo and Lubudi than in any of the other subdivisions. Kambove had the least amount of teachers sampled (see Table 4).

Table 4. Number of teachers by grade and subdivision

<table>
<thead>
<tr>
<th>Grade</th>
<th>Kambove</th>
<th>Kasenga</th>
<th>Kalemie</th>
<th>Kongolo</th>
<th>Mutshatsha</th>
<th>Lubudi</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st</td>
<td>8</td>
<td>8</td>
<td>13</td>
<td>15</td>
<td>14</td>
<td>16</td>
<td>74</td>
</tr>
<tr>
<td>2nd</td>
<td>9</td>
<td>13</td>
<td>16</td>
<td>21</td>
<td>14</td>
<td>16</td>
<td>88</td>
</tr>
<tr>
<td>3rd</td>
<td>9</td>
<td>11</td>
<td>16</td>
<td>18</td>
<td>11</td>
<td>20</td>
<td>85</td>
</tr>
<tr>
<td>4th</td>
<td>8</td>
<td>14</td>
<td>17</td>
<td>15</td>
<td>14</td>
<td>19</td>
<td>87</td>
</tr>
<tr>
<td>5th</td>
<td>7</td>
<td>9</td>
<td>10</td>
<td>13</td>
<td>12</td>
<td>13</td>
<td>64</td>
</tr>
<tr>
<td>6th</td>
<td>7</td>
<td>5</td>
<td>10</td>
<td>13</td>
<td>12</td>
<td>12</td>
<td>59</td>
</tr>
<tr>
<td>Total</td>
<td>48</td>
<td>60</td>
<td>81</td>
<td>95</td>
<td>77</td>
<td>96</td>
<td>457</td>
</tr>
</tbody>
</table>

The distribution of teachers sampled across gender was extremely unbalanced, with almost three times as many males (71.1%) as females (27.2%) overall. Some subdivisions, such as Kalemie, were almost gender-balanced with 52% male teachers, while others were less balanced such as Kasenga with 83% male teachers (Graph 2). Notably, the higher the grade level, the higher the percentage of male teachers (Graph 3).

Graph 2. Percentage of male and female teachers by subdivision

![Graph showing percentage of male and female teachers by subdivision](image-url)
Data were also collected on the teachers' mother tongue (see Graph 4). The three languages most frequently reported were Kiswahili (47.2%), French (14.5%), and Tshiluba (2.1%). French is the official language of instruction in DRC beginning in 3rd grade, yet less than one-fifth of sampled teachers reported French as their mother tongue. Future analysis will examine the proportion of teachers who reported speaking French as a second language. A large proportion of teachers reported speaking a language other than one of the five national languages of DRC (Lingala, Kikongo, Kiswahili, Tshiluba, or French). The questionnaire asked them to specify which language they spoke at home if they selected “Other,” but this disaggregated data is not presented here.
The teachers that were sampled spanned a large age range (17-74 years old). Kasenga had the youngest teachers on average at 33.23 years of age while Kambove had the highest with 44.58 years. Kambove also had the largest variability in teacher ages (see Graph 5).

With regards to their own education, the majority of teachers (95.5%) reported having completed at least one year of secondary school but having gone no further. Most of the teachers reported completing 6th grade of secondary school, and the majority of the remaining teachers reported completing 4th grade of secondary school (see Graph 6).

When broken down, the subdivisions followed the same secondary education trend, with most teachers completing 4th or 6th grade of secondary school (see Graph 7). Kambove had slightly lower teacher education levels than the other subdivisions while Kasenga showed a relatively higher rate of teachers completing secondary school.
Lastly, data were collected on previous years of teaching experience. Teachers across subdivisions reported a wide range of years of experience, ranging from 0-54 years. Both Kambove and Mutshatsha reported a maximum of 54 years. Graph 8 shows the average years of teaching experience per subdivision. Kambove had the highest average amount of prior experience reported by the teachers sampled (22.1 years), while Kongolo has the lowest (8.8 years). In this case, the mean should be examined alongside the standard deviation.

Kambove had the greatest amount of variation in teacher years of experience (Standard Deviation\(^5\): 17.25), while Kongolo reports of teacher experience seemed to be more tightly clustered (SD: 8.67). Large standard deviations indicate that the range of answers was also large, so while the mean is descriptive, the standard deviation is complementary information that gives more insight into the characteristics of the sample. For example, the mean implies that teachers in Kambove are more experienced on average than the teachers in Kongolo; however, the majority of teachers in Kambove actually reported having from 4.9 years – 39.4 years of teaching experience.\(^6\) For program planning, this means that within one cluster of schools, there will be as many teachers with 5 years of teaching experience as with 40 years of teaching experience, making it a challenge to develop in-service teacher training that is of interest to all levels.

\(^5\) "SD" stands for Standard Deviation, a measure of the dispersion of the data around the mean.

\(^6\) This range is calculated by adding and subtracting the standard deviation of 17.25 from the mean of 22.12.
Graph 8. Average number of years of teaching experience by subdivision

- Kambove: 22.12
- Kasenga: 11.11
- Kalemie: 13.03
- Kongolo: 8.81
- Mutshatsha: 13.13
- Lubudi: 12.95
3. Child Demographic Characteristics

The number of children sampled was fairly balanced across grades, but fairly unbalanced across subdivisions (see Table 5). Kambove and Kasenga have a smaller numbers of sampled children relative to Kalemie, Kongolo, Mutshatsha, and Lubudi.

Gender distributions by grade indicate a greater number of boys than girls in every grade (see Table 5), but while there is a greater percentage of boys than girls in the sample as a whole, all subdivisions are fairly gender-balanced (see Graph 9). Kalemie is the most gender-balanced subdivision (50% male, 50% female) whereas Kasenga is the least gender-balanced subdivision (58% male, 42% female).

The proportion of girls declines from 2nd to 4th grade in all subdivisions; this decline seems particularly drastic in Kasenga. Kalemie was the only subdivision where the percentage of girls surpassed that of boys during 3rd grade. The declining tendency in the percentage of girls who attend school has been previously reported by studies in other regions of DRC (RTI, 2011a).
Table 5. Number of children by subdivision, grade and gender

<table>
<thead>
<tr>
<th>Subdivision</th>
<th>Boys</th>
<th>Girls</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2nd Grade</td>
<td>3rd Grade</td>
</tr>
<tr>
<td>Kambove</td>
<td>98</td>
<td>84</td>
</tr>
<tr>
<td>Kasenga</td>
<td>137</td>
<td>120</td>
</tr>
<tr>
<td>Kalemie</td>
<td>141</td>
<td>134</td>
</tr>
<tr>
<td>Kongolo</td>
<td>166</td>
<td>194</td>
</tr>
<tr>
<td>Mutshatsha</td>
<td>159</td>
<td>157</td>
</tr>
<tr>
<td>Lubudi</td>
<td>183</td>
<td>195</td>
</tr>
<tr>
<td>SUBTOTAL</td>
<td>884</td>
<td>884</td>
</tr>
</tbody>
</table>

TOTAL CHILDREN: 4941

In all of the subdivisions except for Kasenga, over 93% of children reported speaking Kiswahili at home (see Graph 10). In Kasenga, 87% of children reported speaking a language other than French or another national language of DRC. The majority of these children reported speaking Kibemba at home, a language that is also spoken in neighboring Zambia.

An average of 2.7% children reported speaking French at home, which is noteworthy considering that French is the official language of instruction in the DRC beginning in 3rd grade, with reports that parents pressure teachers into beginning French-language instruction in earlier grades.
Household and educational risks and protective factors

Children report an important number of household and educational risks across all subdivisions, which may hinder their ability to perform in school; however, the results are difficult to interpret. Almost half of children report going to bed hungry at least once in the past month, but 54% of children reported never going to bed hungry in the past month. About 20% of children moved households in the past year. The majority of children reported living with both parents; only 3.3% of children reported not living with either parent, and 8.8% of children reported living with one parent. Approximately 79% of children reported that both parents can read and write, although they did not specify the language. Only about 9% of children repeated the previous grade, although in the DRC, the MEPSP has a policy that encourages social promotion, regardless of whether grade-appropriate skills have been mastered. Approximately 55% of children reported they are sometimes late to school, although 40% of children report they are never late to school. The same pattern is present for absenteeism as well.

Household wealth

To measure household wealth, children were asked a series of “yes” or “no” questions about whether someone in their household owns certain types of assets, whether there was electricity at their house, and from what materials their house is constructed. Based on these results, we hypothesize that there are two groups of assets that describe different kinds of wealth: Factor 1, made up of livestock and a bicycle, indicates rural wealth, and Factor 2, made up of electric appliances and motor transportation assets, indicates peri-urban wealth. See Annex B for more details regarding the analysis.

When children were asked about languages spoken at home, they were allowed to select more than one language. Therefore the overall percentage for the languages spoken for each subdivision is greater than 100%. 

---

7 When children were asked about languages spoken at home, they were allowed to select more than one language. Therefore the overall percentage for the languages spoken for each subdivision is greater than 100%.
Table 6 shows the percentage of children who reported that their families have certain types of wealth. The subdivisions with the wealthiest households were Mutshatsha and Lubudi, followed by Kalemie and Kambove.

Relative to households in other subdivisions, households in Kongolo and Kasenga appear to be more rural, having less access to assets such as electronic appliances, motorcycles and cars and to sturdiest construction materials such as concrete and brick.

These patterns may also suggest differential access to education across subdivisions. It may be that in Mutshatsha and Lubudi, the two subdivisions that appear wealthiest, children in rural areas face more obstacles to access education, which could result in an overrepresentation of relatively wealthier (or more peri-urban) children in those subdivisions. At the same time, there could be conditions in Kasenga and Kongolo that enable a greater proportion of children living in rural areas to attend school.

<table>
<thead>
<tr>
<th>Factor 1</th>
<th>Kambove</th>
<th>Kasenga</th>
<th>Kalemie</th>
<th>Kongolo</th>
<th>Mutshatsha</th>
<th>Lubudi</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Cattle</td>
<td>.40</td>
<td>4.60</td>
<td>.60</td>
<td>1.30</td>
<td>1.00</td>
<td>.70</td>
</tr>
<tr>
<td>2 Goats/sheep/pigs</td>
<td>21.10</td>
<td>65.30</td>
<td>22.40</td>
<td>60.10</td>
<td>40.50</td>
<td>31.20</td>
</tr>
<tr>
<td>3 Chickens</td>
<td>77.80</td>
<td>72.70</td>
<td>54.60</td>
<td>85.80</td>
<td>77.20</td>
<td>76.20</td>
</tr>
<tr>
<td>4 Bicycle</td>
<td>83.20</td>
<td>81.30</td>
<td>52.20</td>
<td>74.20</td>
<td>84.30</td>
<td>81.00</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Factor 2</th>
<th>Kambove</th>
<th>Kasenga</th>
<th>Kalemie</th>
<th>Kongolo</th>
<th>Mutshatsha</th>
<th>Lubudi</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Motorcycle</td>
<td>4.70</td>
<td>6.70</td>
<td>8.30</td>
<td>1.30</td>
<td>19.70</td>
<td>12.40</td>
</tr>
<tr>
<td>2 Radio</td>
<td>73.10</td>
<td>59.70</td>
<td>87.50</td>
<td>68.10</td>
<td>85.00</td>
<td>71.60</td>
</tr>
<tr>
<td>3 Telephone / Cell Phone</td>
<td>65.10</td>
<td>29.10</td>
<td>69.80</td>
<td>26.50</td>
<td>71.30</td>
<td>68.60</td>
</tr>
<tr>
<td>4 Electricity</td>
<td>11.70</td>
<td>5.00</td>
<td>37.30</td>
<td>.30</td>
<td>20.10</td>
<td>32.40</td>
</tr>
<tr>
<td>5 Television</td>
<td>16.80</td>
<td>6.00</td>
<td>27.60</td>
<td>.70</td>
<td>23.20</td>
<td>31.30</td>
</tr>
<tr>
<td>6 Refrigerator</td>
<td>3.10</td>
<td>1.90</td>
<td>6.00</td>
<td>.10</td>
<td>6.00</td>
<td>14.20</td>
</tr>
<tr>
<td>7 Indoor toilet</td>
<td>.80</td>
<td>2.20</td>
<td>.80</td>
<td>.10</td>
<td>1.30</td>
<td>2.90</td>
</tr>
<tr>
<td>8 Car, Truck, 4x4</td>
<td>1.40</td>
<td>.50</td>
<td>.60</td>
<td>.00</td>
<td>1.50</td>
<td>2.20</td>
</tr>
<tr>
<td>9 Concrete floors</td>
<td>14.20</td>
<td>14.40</td>
<td>18.00</td>
<td>1.90</td>
<td>26.40</td>
<td>33.60</td>
</tr>
<tr>
<td>10 Cured brick walls</td>
<td>26.30</td>
<td>22.40</td>
<td>5.50</td>
<td>4.90</td>
<td>46.50</td>
<td>46.90</td>
</tr>
<tr>
<td>11 Aluminum/tin roof</td>
<td>63.50</td>
<td>15.20</td>
<td>41.80</td>
<td>4.70</td>
<td>73.50</td>
<td>74.30</td>
</tr>
</tbody>
</table>
Baseline Findings – Early Grade Reading Assessment (EGRA)

This section starts with an overview of the EGRA subtests and the grade levels to which they were administered. Then, we describe how much of the variability in children's scores can be explained by differences between subdivisions, clusters of schools and between children's themselves. Finally, we present descriptive analyses for the overall sample and by comparisons between groups (grade, gender, subdivision) for each EGRA subtest. Table 7 summarizes the competencies measured by each subtest, and which grade levels took each subtests.

Table 7. Overview of EGRA subtests

<table>
<thead>
<tr>
<th>Subtest</th>
<th>Competencies Measured</th>
<th>Administered to</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Vocabulary</td>
<td>Basic vocabulary and understanding of commands.</td>
<td>All grades (2-4)</td>
</tr>
<tr>
<td>2. Initial Sound Identification (Phonological awareness)</td>
<td>Understanding of the relationship between sounds and letters.</td>
<td>All grades (2-4)</td>
</tr>
<tr>
<td>3. Knowledge of Graphemes (letter knowledge)</td>
<td>Ability to read the letters of the alphabet, or say their sounds. Automaticity of letter recognition in addition to letter knowledge. (Timed).</td>
<td>All grades (2-4)</td>
</tr>
<tr>
<td>4. Familiar word reading (fluency)</td>
<td>Ability to read high-frequency words quickly. (Timed).</td>
<td>3rd &amp; 4th grades</td>
</tr>
<tr>
<td>5. Invented word decoding (fluency)</td>
<td>Ability to decode made up words fluently. (Timed).</td>
<td>3rd &amp; 4th grades</td>
</tr>
<tr>
<td>6. Oral passage reading (connected text fluency)</td>
<td>Ability to read a short passage (50 words long), that tells a story. (Timed).</td>
<td>3rd &amp; 4th grades</td>
</tr>
<tr>
<td>7. Reading comprehension</td>
<td>Ability to answer up to five questions based on the portion of the oral passage that the child could read.</td>
<td>3rd &amp; 4th grades</td>
</tr>
<tr>
<td>8. Listening comprehension</td>
<td>Ability to follow and understand a simple oral story.</td>
<td>All grades (2-4)</td>
</tr>
<tr>
<td>9. Writing a complete sentence</td>
<td>Ability to correctly write a sentence (3 words). Scores take into account content and spelling.</td>
<td>4th grade</td>
</tr>
</tbody>
</table>

1. Variability in children's EGRA scores

Originally, we had hypothesized that variation in EGRA scores can be caused by differences in the children themselves (e.g. demographic differences, household differences, cognitive differences), differences in school clusters within subdivisions, and differences between subdivisions. Table 7 shows that on average, 79% of the variation in children's EGRA scores can be explained by differences between children themselves; 12% of the variation can be explained by differences in subdivisions; 8.5% of the variation can be explained by differences between school clusters within subdivisions.

Individual factors that can contribute to the explanation of between-child variation in reading scores range from demographic characteristics such as gender and household socio-economic status, to differences in children’s core cognitive abilities such as working memory and attention control. Differences between schools clusters could include differential access to human resources (e.g., more or...
less educated teachers), variation in the quality of school infrastructure and support from the community, the state and other agents.

Table 8 also shows that many subtests have large amounts of variability. As discussed earlier, this is important because while the average scores for each subtest are descriptive, they do not fully illustrate the EGRA results. Large amounts of variance means that even though EGRA average scores might be very low on most subtests, there are students that are scoring much higher and much lower. This will be a challenge, and perhaps not possible, to develop a curriculum that helps the highest and lowest performing students remain engaged and improve over the course of the year. The amount of variance is also important to consider when setting targets around the average scores.

Table 8. Distribution of variance in children's EGRA scores

<table>
<thead>
<tr>
<th>Subtest (number of items)</th>
<th>Percentage of variance***</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total variance</td>
</tr>
<tr>
<td>Vocabulary (20)</td>
<td>13.77</td>
</tr>
<tr>
<td>Initial Sound Identification (10)</td>
<td>4.66</td>
</tr>
<tr>
<td>Knowledge of graphemes (100)</td>
<td>196.57</td>
</tr>
<tr>
<td>Familiar word reading (50)*</td>
<td>69.04</td>
</tr>
<tr>
<td>Invented word decoding (50)*</td>
<td>58.66</td>
</tr>
<tr>
<td>Oral passage reading (50)*</td>
<td>135.52</td>
</tr>
<tr>
<td>Reading comprehension (5)*</td>
<td>0.50</td>
</tr>
<tr>
<td>Listening Comprehension (5)</td>
<td>0.67</td>
</tr>
<tr>
<td>Writing a complete sentence (3)**</td>
<td>0.92</td>
</tr>
<tr>
<td>Mean</td>
<td>12.26</td>
</tr>
<tr>
<td>Standard Deviation (SD)</td>
<td>4.75</td>
</tr>
</tbody>
</table>

*Not administered to 2nd graders, ** Administered to 4th graders only
***All variance estimates were significant at $p < .05$, which means there was significant variation in children's scores at each level examined (i.e., subdivision, cluster, individual)
1. **Zero scores by grade**

The zero score per subtest is the percentage of children who could not respond correctly to even one question. The zero score is a measure of competencies that children do not have at all and can indicate a potential need for increased curricular focus. Graph 11 shows that for only two subtests, (Vocabulary and Knowledge of Graphemes) the majority of children are able to respond correctly to at least one question. The remaining seven subtests range from zero scores of 54% (Writing a Complete Sentence) to 91% (Reading Comprehension), indicating that more than half of to almost all children have no competency at all in most of the skills that must be mastered before becoming a functional reader.

Table 9 shows that zero scores decrease as grade levels increases, implying that children are learning as they progress in school. Reading Comprehension and Initial Sound Identification show the least average progress between grades, improving by only 13-14% each, while Vocabulary and Knowledge of Graphemes show the most progress between grades, improving by 85% and 88% respectively. Yet, zero scores in 4th grade remain high, ranging from 35% (Listening Comprehension) to 78% (Reading Comprehension). In the DRC, the national curriculum does not break down reading into smaller sets of skills; it only requires that children can read. This means it is likely that the subcomponents of reading skills, like sound identification or decoding words, are not taught in primary schools; moreover, as we saw previously, in 3rd and 4th grades, all instruction is French-language. The zero scores may remain high across grades because children do not become much more familiar with French in 4th grade than they were in 2nd grade.
Graph 11. EGRA Zero Scores by Grade

*Familiar word reading, Invented word decoding, Oral passage reading, Reading comprehension were only administered to 3rd and 4th Grades; Writing a complete sentence was only administered to 4th Grade.

Table 9. Percent improvement in EGRA zero scores

<table>
<thead>
<tr>
<th>Subtest</th>
<th>Improvement by 3rd Grade</th>
<th>Improvement by 4th Grade</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vocabulary</td>
<td>29.17%</td>
<td>87.50%</td>
</tr>
<tr>
<td>Initial sound identification</td>
<td>7.92%</td>
<td>13.95%</td>
</tr>
<tr>
<td>Knowledge of Graphemes</td>
<td>88.24%</td>
<td>85.29%</td>
</tr>
<tr>
<td>Familiar word reading *</td>
<td>-</td>
<td>27.95%</td>
</tr>
<tr>
<td>Invented word decoding*</td>
<td>-</td>
<td>23.15%</td>
</tr>
<tr>
<td>Oral passage reading *</td>
<td>-</td>
<td>23.83%</td>
</tr>
<tr>
<td>Reading comprehension*</td>
<td>-</td>
<td>13.37%</td>
</tr>
<tr>
<td>Listening comprehension</td>
<td>25.00%</td>
<td>48.05%</td>
</tr>
<tr>
<td>Writing a complete sentence*</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

* Only administered to 3rd and 4th grades, so percent improvement is from 3rd grade to 4th grade; ** Only administered to 4th grade, so there is no percent improvement.
2. Average correct scores by gender

Boys outperformed girls across all grades in the majority of subtests, with the gap between boys’ and mean correct scores increasing as children progress in grade level. Table 10 shows the differences between the mean correct score for boys and girls by grade. For example, looking at the Vocabulary subtest, in 2nd grade, boys knew on average 0.3 more words than girls; in 3rd grade, boys knew 0.65 more words than girls; and in 4th grade, boys knew 0.51 more words than girls. In all subtests, except Vocabulary and Listening Comprehension, the differences between boys and girls mean scores increased as grades increased. In these two tests, the differences in mean scores increased from 2nd to 3rd grade, but decreased from 3rd to 4th grade. In all tests except Listening Comprehension, Vocabulary, and Reading Comprehension, the magnitude of the gap between boys and girls mean scores increases. While some of these differences appear large, the gender differences are actually relatively modest. The more interesting aspect is how systematically they occur across subtests.

<table>
<thead>
<tr>
<th>EGRA Subtest</th>
<th>△ Means by Gender (Boys Mean Correct Score – Girls Mean Correct Score)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2nd grade</td>
</tr>
<tr>
<td>Vocabulary (20)</td>
<td>0.3</td>
</tr>
<tr>
<td>Initial sound identification (10)</td>
<td>-0.02</td>
</tr>
<tr>
<td>Knowledge of Graphemes (100)</td>
<td>1.34</td>
</tr>
<tr>
<td>Familiar word reading (50)*</td>
<td></td>
</tr>
<tr>
<td>Invented word decoding (50)*</td>
<td></td>
</tr>
<tr>
<td>Oral passage reading (50)*</td>
<td></td>
</tr>
<tr>
<td>Reading comprehension (5)*</td>
<td></td>
</tr>
<tr>
<td>Listening comprehension (5)</td>
<td>0.01</td>
</tr>
<tr>
<td>Writing a complete sentence (3)**</td>
<td></td>
</tr>
</tbody>
</table>

* Only administered to 3rd and 4th grades; ** Only administered to 4th grade
3. Average correct scores by subdivision

As noted, differences between subdivisions account for a portion of the variance seen in EGRA scores, although not to the degree that differences between children do. Table 11 shows the mean correct score by subtest and by subdivision, as well as the rank of each subdivision by subtest. For example, for Initial Sound Identification, Kambove had a mean correct score of 0.15 sounds out of 10 possible sounds and ranked 5th out of 6 subdivisions. The mean correct scores across subdivisions show that across all subtests, performance is incredibly low, with children answering on average 1% (Reading Comprehension, Kasenga) to 48% (Writing a Complete Sentence, Kalemie) of questions correctly.

The subdivisions perform similarly across all subtests. Kalemie and Kongolo are almost always ranked first and second, Mutshatsha and Lubudi are ranked third or fourth, and Kambove and Kasenga are ranked fifth or sixth. Even though subdivisions seem to hang together in rank order, the performance gaps between subdivisions is extremely varied. For many subtests, there is often a large gap between higher and lower performing subdivisions. For example, in Familiar Word Reading, we see Kalemie (10.95) vastly outperforming Kongolo (6.53) and Mutshatsha (5.22), but we see Kongolo and Mutshatsha vastly outperforming Lubudi (2.65), Kambove (2.47) and Kasenga (1.72).

Table 11. Mean correct scores by subdivision; Subdivision rank by subtest

<table>
<thead>
<tr>
<th>EGRA Subtest</th>
<th>Kambove</th>
<th>Kasenga</th>
<th>Kalemie</th>
<th>Kongolo</th>
<th>Mutshatsha</th>
<th>Lubudi</th>
<th>Subdivision Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vocabulary (20)</td>
<td>7.82</td>
<td>5.69</td>
<td>8.91</td>
<td>8.57</td>
<td>8.31</td>
<td>8.1</td>
<td>4</td>
</tr>
<tr>
<td><strong>Subdivision Rank</strong></td>
<td>5</td>
<td>6</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Initial Sound Identification (10)</td>
<td>0.15</td>
<td>0.06</td>
<td>2.9</td>
<td>1.66</td>
<td>0.33</td>
<td>0.49</td>
<td></td>
</tr>
<tr>
<td><strong>Subdivision Rank</strong></td>
<td>5</td>
<td>6</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Knowledge of Graphemes (100)</td>
<td>10.88</td>
<td>8.2</td>
<td>24.95</td>
<td>20.42</td>
<td>20.06</td>
<td>13.62</td>
<td></td>
</tr>
<tr>
<td><strong>Subdivision Rank</strong></td>
<td>5</td>
<td>6</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Familiar word reading (50)</td>
<td>2.47</td>
<td>1.72</td>
<td>10.95</td>
<td>6.53</td>
<td>5.22</td>
<td>2.65</td>
<td></td>
</tr>
<tr>
<td><strong>Subdivision Rank</strong></td>
<td>5</td>
<td>6</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Invented word decoding (50)</td>
<td>1.66</td>
<td>1.04</td>
<td>9.4</td>
<td>5.02</td>
<td>4.59</td>
<td>2.29</td>
<td></td>
</tr>
<tr>
<td><strong>Subdivision Rank</strong></td>
<td>5</td>
<td>6</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Oral passage reading (50)</td>
<td>3.02</td>
<td>1.92</td>
<td>12.99</td>
<td>8.95</td>
<td>6.22</td>
<td>2.76</td>
<td></td>
</tr>
<tr>
<td><strong>Subdivision Rank</strong></td>
<td>4</td>
<td>6</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Reading comprehension (5)</td>
<td>0.07</td>
<td>0.05</td>
<td>0.65</td>
<td>0.48</td>
<td>0.19</td>
<td>0.08</td>
<td></td>
</tr>
<tr>
<td><strong>Subdivision Rank</strong></td>
<td>5</td>
<td>6</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Listening comprehension (5)</td>
<td>0.28</td>
<td>0.31</td>
<td>1.03</td>
<td>1.02</td>
<td>0.53</td>
<td>0.54</td>
<td></td>
</tr>
<tr>
<td><strong>Subdivision Rank</strong></td>
<td>6</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Writing a complete sentence (3)</td>
<td>0.64</td>
<td>0.59</td>
<td>1.44</td>
<td>0.67</td>
<td>0.82</td>
<td>0.6</td>
<td></td>
</tr>
<tr>
<td><strong>Subdivision Rank</strong></td>
<td>4</td>
<td>6</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>5</td>
<td></td>
</tr>
</tbody>
</table>
Descriptive Analysis by Subtest

Summary of Vocabulary Results (all grades)

- Children have low levels of competency with French vocabulary; on average, children could correctly identify fewer than half of the test words.
- The zero score for Vocabulary is low (1.5%).
- Almost half of children correctly identified between 5 and 9 words out of 20 possible words; almost half of children could not correctly identify even half of the words.
- 6% of children identified 15 out of 20 words correctly; a score of 75%.
- The average number of correctly identified words slowly increases with grade (1.5 words per grade level).
- Children’s scores vary considerably within grades and within subdivisions.
  - Some 2nd graders are performing as well as 4th graders; some 4th graders still cannot identify even one vocabulary word correctly.
- Boys performed better than girls overall, though the difference is modest.
- Kasenga has the lowest average score (5.69); the other subdivisions had comparable average scores, ranging from 7.8 to 8.9 correctly identified words.

1. Vocabulary

This subtest targets children’s knowledge of and ability to identify vocabulary words from three categories: body parts, words for everyday surroundings and spatial terms. Test administrators read aloud 20 words to children. After reading a word from each category, children are asked to point to the corresponding body part, classroom object or to perform a simple action with a pencil, for example, to put it “on the paper” or “behind you.” If a child does not respond within 3 seconds, the child is asked to move to the next item. The score is based on the number of words the child is able to correctly identify.

As shown in Graph 12, only 0.1% of children were able to correctly identify all 20 words presented; 1.5% of children were unable to identify a single word correctly. The majority of children (70%) were able to identify between 6 and 15 words.

---

8 All subtest descriptions taken from RTI (2011a-b).
As expected, the average number of correctly indentified words increased with grade. However, children only gained an average of 1.5 of words with each grade increment (average $d$: 0.43), indicating slow rate of growth in vocabulary during the primary school years. This finding suggests that grade is not the main source of variability in children's vocabulary scores. Instead, as shown in Graph 13, there is considerable variability in children's performance within grades. For example, there are a number of children in 2$^{nd}$ grade who obtained scores as high as those of their peers in 3$^{rd}$ or 4$^{th}$ grades. Similarly, there are children in 4$^{th}$ grade who are still unable to correctly identify a single word. There is a diversity of factors in children's school and home experiences that may explain within-grade variability in children performance, which merits further exploration.

This task was explained to children in the local language when the test administrator judged it necessary, but the target vocabulary word was provided in French. Therefore, low scores in this subtest are indicative of inadequate mastery of French vocabulary across all grades, which is understandable if one considers that French is a secondary language for the vast majority of these children. It is possible that children who are performing at higher levels than the majority of their peers have more exposure to French in the home. This is a hypothesis that will be explored in future analysis.

Regarding gender, boys outperformed girls in every grade. On average, boys were able to correctly identify 0.6 slightly more words than girls ($d$: 0.16), which is an important difference considering that children's scores improved at a rate of 1.5 words per grade. Interestingly, boys gained more between 2$^{nd}$ to 3$^{rd}$ grade ($d$: 0.22) than from 3$^{rd}$ to 4$^{th}$ grade ($d$: 0.16), and girls gained more from 3$^{rd}$ to 4$^{th}$ grade ($d$: 0.21) than from 2$^{nd}$ to 3$^{rd}$ grade ($d$: 0.15). Overall, grade increments were small to modest. Gender differences within grade were trivial for this subtest ($d$ ranges from 0.045 to 0.053).
Finally, children's average performance in the vocabulary subtest did not appear to vary substantively as a function of subdivision, with the exception of Kasenga, which had a much lower average relative to other subdivisions. This large difference may be attributable to language differences between Kasenga and the other subdivisions included in this study.

While mean scores do not suggest important differences between subdivisions, except for the case of Kasenga, there is considerable variability in children's vocabulary within grades and subdivisions (see Graph 4).

---

Please note that for illustration purposes the y-axis in this and other graphs may be truncated to emphasize group differences in observed data. In other words, the y-axis may range from the minimum possible score to the maximum score actually observed in the sample, as opposed to the maximum achievable score.
Graph 15. For example, looking at Kalemie, while median scores increased with grade, the minimum scores in 2nd and 3rd grade are the same. In addition, the range of scores in each grade is very large. In 4th grade, the range encompasses almost all possible scores. While half of 4th graders in Kalemie scored in a 5-point range or so, the other 50% scored much higher and much lower. A better understanding of the factors underlying these sizable differences is critical to inform strategies aimed at reducing the gap between low and high performing children. In Kasenga, 25% of 4th graders were unable to identify more than 5 words, which is more typical of children in 2nd and 3rd grades in the other subdivisions.

Graph 15. Distribution of Vocabulary scores by grade and subdivision

How to interpret boxplots:
- The box (blue) shows the range of scores for 50% of children.
- The horizontal line inside the box is the median score (50% of children scored above this number and 50% of children scored below this number).
- The lines (or whiskers) coming out of the box on either end show the full range of scores (the maximum and the minimum).
- The line coming from the top of the box represents the top 25%; the line coming from the bottom of the box represents the bottom 25%.
- The “longer” the box and whiskers appear, the more variability is present.
- The dots and stars outside of the lines (or whiskers) show outlier scores.

10 For a graphical description of boxplot components refer to Appendix C.
2. Initial Sound Identification (Phonological awareness)

This subtest examines phonemic awareness by measuring a child’s ability to identify the first sound of a word. Test administrators orally present children with 10 monosyllabic words. Children are required to identify the first sound of each word presented. For example, the first sound of the word *soup* would be ‘sss.’ The child is given 3 seconds to respond before moving to the next item. Grasping the connection between written letter symbols and their sounds is a fundamental skill for early reading acquisition (RTI, 2011a).

The results indicated that regardless of subdivision, gender and grade, children had substantial difficulties identifying the first sounds of words. As shown in Graph 16, the majority of children (78%) was unable to identify a single sound.

<table>
<thead>
<tr>
<th>Percentage of children</th>
<th>Number of sounds identified</th>
</tr>
</thead>
<tbody>
<tr>
<td>78.6</td>
<td>0.00</td>
</tr>
<tr>
<td>1.6</td>
<td>1.00</td>
</tr>
<tr>
<td>1.8</td>
<td>2.00</td>
</tr>
<tr>
<td>3.7</td>
<td>3.00</td>
</tr>
<tr>
<td>3.7</td>
<td>4.00</td>
</tr>
<tr>
<td>3.3</td>
<td>5.00</td>
</tr>
<tr>
<td>3.1</td>
<td>6.00</td>
</tr>
<tr>
<td>2.0</td>
<td>7.00</td>
</tr>
<tr>
<td>0.7</td>
<td>8.00</td>
</tr>
<tr>
<td>0.5</td>
<td>9.00</td>
</tr>
<tr>
<td>1.0</td>
<td>10.00</td>
</tr>
</tbody>
</table>

While the number of correctly identified sounds increased with grade, differences between grades were small ($d: 0.15$), with children identifying about 0.3 sounds more as they move from one grade level to
the next (see Graph 17). Boys performed slightly better than girls ($d: 0.09$), but considering how low scores are across the board, children of both genders have considerable room for improvement.

Graph 17. Mean scores by grade and gender for Initial Sound Identification

As Graph 18 and Graph 19 show, there are large differences between subdivisions, with children in Kalemie and Kongolo outperforming children in all other subdivisions ($d: 0.94$). The factors explaining these differences will be explored in future analysis.

Graph 18. Mean scores by subdivision for Initial Sound Identification

Graph 19 in particular shows that in Kambove, Kasenga, Mutshatsha, and Lubudi, 50% of children scored 0 (the horizontal black line represents the median), with the other 50% distributed along the range of possible answers, in some cases up to the maximum number of possible answers. This means that in 4th grade in Lubudi, for example, while 50% of children scored 0 on Initial Sound Identification, the other
50% scored from 1 to 10 (a perfect score). It will be very challenging to teach a classroom where half of the children have no familiarity with the sounds of French letters and the other half of the children includes some who know the sounds very well and others who are almost as unfamiliar with them as the children who scored 0.

Graph 19. Distribution of Initial Sound Identification scores by grade and by subdivision

<table>
<thead>
<tr>
<th>Kambove</th>
<th>Kasenga</th>
<th>Kalemie</th>
<th>Kongolo</th>
<th>Mutshatsha</th>
<th>Lubudi</th>
</tr>
</thead>
</table>

How to interpret boxplots:
- The box (blue) shows the range of scores for 50% of children.
- The horizontal line inside the box is the median score (50% of children scored above this number and 50% of children scored below this number).
- The lines (or whiskers) coming out of the box on either end show the full range of scores (the maximum and the minimum).
- The line coming from the top of the box represents the top 25%; the line coming from the bottom of the box represents the bottom 25%.
- The “longer” the box and whiskers appear, the more variability is present.
- The dots and stars outside of the lines (or whiskers) show outlier scores.
3. **Knowledge of graphemes (letter knowledge)**

In this subtest, children are shown a page with 100 letters. The test administrator points to a letter and the child is asked to give the name or sound. For example, the administrator points to ‘ch’ in the example line. The child is required to read it as /ch/ as in the word “chat.” This subtest is timed and children are given one minute to complete it. Final scores represent the number of letters correctly named per minute, or “correct letters per minute” (clpm).

Children appear to have limited knowledge of graphemes in French: 1.1% of children were unable to identify a single grapheme and about half of the sample was able to correctly identify 1 to 14 graphemes per minute (see Graph 20). The average child was able to correctly identify 18 graphemes per minute but there was considerable variation around the mean (SD: 14.01). Only 3.4% of children were able to identify more than 50 graphemes per minute correctly.
The average number of graphemes children were able to identify increased substantially with grade. On average, children gained about 6 graphemes per grade level ($d: 0.45$).

Results on this subtest showed the same pattern of gender differences found in other subtests, with boys performing slightly better than girls in every grade ($d: 0.17$). On average, boys were able to identify 2.4 more graphemes than girls (see Graph 21). Moreover, the difference between boys and girls was greater in the upper grades. Even though the effect size is small, this is an important difference in light of the number of letters children gained per grade.

Results suggest substantial variation in children’s performance across subdivisions, with children in Kasenga and Kambove performing worse than their peers from other subdivisions (see Graph 22). While Kasenga’s low performance could be attributed to language differences, it is unclear what might explain the low performance of children from Kambove. The reasons for these vast differences require further

Graph 20. Percentage of children per number of correctly identified graphemes

Graph 21. Mean scores by grade and gender for Knowledge of Graphemes
exploration. It would also be interesting to better understand the reasons for Kalemie's comparatively high performance.

Graph 22. Mean scores by subdivision for Knowledge of Graphemes

Finally, as with other subtests, there is substantial variability in children's scores within grades and subdivisions. This variability is represented by the boxplots in Graph 23. One can clearly see the relatively condensed scores of Kasenga and Kambove for all grades as illustrated by the very short boxes. Not only do the subdivision means on this subtest indicate poor performance in these two specific subdivisions, the range of their scores was relatively small when compared to the other subdivisions. Potential factors in the school and home that may account for such variation will be explored in future analysis.
Graph 23. Distribution of Knowledge of Graphemes scores by grade and subdivision

How to interpret boxplots:
- The box (blue) shows the range of scores for 50% of children.
- The horizontal line inside the box is the median score (50% of children scored above this number and 50% of children scored below this number).
- The lines (or whiskers) coming out of the box on either end show the full range of scores (the maximum and the minimum).
- The line coming from the top of the box represents the top 25%; the line coming from the bottom of the box represents the bottom 25%.
- The “longer” the box and whiskers appear, the more variability is present.
- The dots and stars outside of the lines (or whiskers) show outlier scores.
Summary of Familiar Word Reading Results (3rd and 4th grade only)

- Most children are unable to read familiar words in French; considering French is only the official language of instruction beginning in 3rd grade, these results are not surprising.
- The zero score was moderate in comparison to other subtests: 56%, but the mean was low.
- 3.5% of children could correctly read 26 words or more per minute. In other words, less than 5% of children could correctly read at least 50% of the words in the subtest in one minute.
- The average number of correctly identified graphemes increased quickly with grade; children gained about 6 graphemes per grade level.
- Children’s scores vary considerably across grade levels and subdivisions.
  - 4th graders read on average 3.6 more words per minute than 3rd graders.
  - The majority of 4th graders in the worst-performing subdivisions (Lubudi, Kambove, Kasenga) performed more poorly than half of 3rd graders in the best-performing subdivision (Kalemie)
- Boys performed better than girls overall, though the difference is modest.

4. Familiar Word Reading (Fluency)

In this subtest, children are given a list of 50 commonly used words. They are asked to read these words and to stop after one minute. The final score measures the number of words that were read correctly in one minute. This subtest was administered to 3rd and 4th graders only.

Graph 24 indicates that 56% of the children were unable to read a single familiar word correctly. Only a minority of children (3.5%) was able to read more than 26 words in one minute. As French is only introduced as the official language of instruction in 3rd grade, these results are not surprising.
There are substantial differences in the performance of 3rd versus 4th graders ($d: 0.44$), with 4th graders reading an average of 3.6 more words per minute than 3rd graders (see Graph 25). There are also small gender differences, which become greater as children move from 3rd ($d: 0.13$) to 4th grade ($d: 0.21$).

### Graph 25. Mean scores by grade and gender, Familiar Word Reading

![Bar graph showing average number of familiar words read by grade and gender](image)

Similar to other subtests, on average children in Kalemie perform substantially better than children in all other subdivisions, while children in Kambove, Kasenga, and notably, Lubudi, perform much worse than all other children (see Graph 26).

### Graph 26. Mean scores by subdivision, Familiar Word Reading

![Bar chart showing average number of familiar words read by subdivision](image)

As shown in Graph 27, the vast majority of 3rd grade children in Kambove (87.8%), Kasenga (75.8%) and Lubudi (81.4%), were unable to read a single familiar word, whereas more than 50% of 3rd grade children in the remaining subdivisions were able to read at least 1 word correctly. This is illustrated by the lack of a box or whiskers in the boxplot, signifying that the majority of children scored 0.
Even though there are important gains from 3rd to 4th grade in these subdivisions, a majority (75%) of 4th graders in Kambove, Kasenga and Lubudi performed worse than 50% of 3rd graders in Kalemie (compare the horizontal black lines, which represent the means). The reasons for these marked differences in performance across subdivisions warrant exploration beyond Kasenga's linguistic profile. It is possible that compared to other subdivisions, instruction in reading starts later in Kambove, Kasenga and Lubudi.

Graph 27. Distribution of Familiar Word Reading Scores by subdivision and grade

How to interpret boxplots:
- The box (blue) shows the range of scores for 50% of children.
- The horizontal line inside the box is the median score (50% of children scored above this number and 50% of children scored below this number).
- The lines (or whiskers) coming out of the box on either end show the full range of scores (the maximum and the minimum).
- The line coming from the top of the box represents the top 25%; the line coming from the bottom of the box represents the bottom 25%.
- The “longer” the box and whiskers appear, the more variability is present.
- The dots and stars outside of the lines (or whiskers) show outlier scores.
5. Invented Word Decoding (Fluency)

Invented words measure assembling-decoding skills, a building block of reading acquisition (RTI, 2011a-b). This subtest requires that children use their emerging decoding skills to identify invented words using the rules of decoding. Children are presented with a sheet of 50 invented words and asked to stop after one minute. The final score is a measure of the number of invented words read correctly in a minute. This subset was administered to 3rd and 4th grades only.

Considering the strong correlation between this subtest and familiar word reading \((r = 0.77**)\), it is not surprising that two thirds of children were unable to decode a single invented word (see Graph 28).

Graph 28. Percentage of children per number of invented words read in one minute
Again, sizeable differences were found between grades ($d$: 0.40) and small yet notable differences were found between boys and girls ($d$: 0.19) (see Graph 29).

Comparisons by subdivision parallel the results from the familiar word reading subtest, with the lowest average scores for children in Kambove, Kasenga and Lubudi and the highest mean scores for children in Kalemie (see Graph 30). However, as expected, more children struggled with decoding invented as opposed to familiar words, as shown by the lower means in this subtest.

The range and variability per grade and subdivision are indicated by the box plots below (Graph 31). The range of these scores is larger for the high performing subdivisions (Kalemie, Kongolo, Mutshatsha), which in part explains the large mean differences between subdivisions in the average scores for this subtest.
Graph 31. Distribution of Invented Word Decoding scores by subdivision and grade

How to interpret boxplots:

- The box (blue) shows the range of scores for 50% of children.
- The horizontal line inside the box is the median score (50% of children scored above this number and 50% of children scored below this number).
- The lines (or whiskers) coming out of the box on either end show the full range of scores (the maximum and the minimum).
- The line coming from the top of the box represents the top 25%; the line coming from the bottom of the box represents the bottom 25%.
- The “longer” the box and whiskers appear, the more variability is present.
- The dots and stars outside of the lines (or whiskers) show outlier scores.
6. Oral Passage Reading (Connected Text Fluency)

In this subtest, 3rd and 4th grade children were presented with a short reading passage consisting of 50 words. The oral reading score is obtained by calculating the number of words the child is able to read correctly per minute.

Along the lines of the two previous subtests (i.e., familiar and invented word reading), about two thirds of children were unable to read a single word (see Graph 32). Somewhat unexpectedly, 9.7% of children were able to read 25 words or more. Remember in Familiar Word Reading and Invented Word Reading, less than 5% of children could read 25 words or more. This apparent discrepancy is most likely due to the nature of the words included in each subtest. Longer and more complex words are included in the familiar and invented word subtests, whereas more articles, connectors and pronouns are found in the oral passage reading subtest.

Graph 32. Percentage of children per number of words read in an oral passage in one minute

Summary of Oral Passage Reading Results (3rd and 4th grades only)

- The zero score for Oral Passage Reading is also high: 67.9%.
- 9.7% of children could score at least 50% - this is larger than expected considering that less than 5% of children could score at least 50% on familiar word reading and invented word reading.
- There are substantial differences between 3rd and 4th grades, and smaller but still significant differences between boys and girls. 4th graders performed better than 3rd graders, and boys performed better than girls.
- There were considerable differences between subdivisions, with Kalemie outperforming all other subdivisions (average score=12.99). Kasenga (average=1.92), Lubudi (average=2.76), and Kambove (average=3.02) again scored the lowest.
Again, differences between 3<sup>rd</sup> and 4<sup>th</sup> graders were substantial ($d$: 0.41) and differences between boys and girls were small but not trivial ($d$: 0.24). Moreover, differences between genders increased in magnitude from 3<sup>rd</sup> to 4<sup>th</sup> grade (see Graph 33).

**Graph 33. Mean scores by grade and gender, Oral Passage Reading**

Comparisons across subdivision show the same pattern found in familiar and invented word reading (see Graph 34). The performance of Kalemie, Kongolo, and Mutshatsha is considerably higher than Kambove, Kasenga, and Lubudi.

**Graph 34. Mean scores by subdivision, Oral Passage Reading**

This pattern is also similar when looking at the variability and range (see boxplots in Graph 35). The range of both 3<sup>rd</sup> graders and 4<sup>th</sup> graders was extremely small in the high performing subdivisions of Kalemie and Kongolo. The only exception was the narrow range of Mutshatsha’s 3<sup>rd</sup> graders, which shows that the relatively high mean for this subdivision is mostly accounted for by 4<sup>th</sup> graders.
Graph 35. Distribution of Oral Passage Reading scores by subdivision and grade

How to interpret boxplots:

- The box (blue) shows the range of scores for 50% of children.
- The horizontal line inside the box is the median score (50% of children scored above this number and 50% of children scored below this number).
- The lines (or whiskers) coming out of the box on either end show the full range of scores (the maximum and the minimum).
- The line coming from the top of the box represents the top 25%; the line coming from the bottom of the box represents the bottom 25%.
- The “longer” the box and whiskers appear, the more variability is present.
- The dots and stars outside of the lines (or whiskers) show outlier scores.
### Summary of Reading Comprehension Results (3rd and 4th grades only)

- Of children who were able to read at least the first line of the Oral Passage correctly, 85% were unable to answer one reading comprehension question.
- No child could answer all 5 reading comprehension questions correctly.
- 4th graders answered more questions than 3rd graders, and boys performed slightly better than girls.
- The pattern of scores between subdivisions mirrored Oral Passage Reading. Kalemie outperformed other subdivisions, with an average score of 0.65. Kasenga, Kambove, and Lubudi were the worst performing subdivisions, with average scores of 0.05, 0.07, and 0.08 respectively.
- When the results of Oral Passage Reading and Reading Comprehension are viewed together, it is clear that the majority of children across grades and subdivisions are not able to read in French.

### 7. Reading Comprehension

After children take the Oral Passage Reading subtest, they answer up to 5 reading comprehension questions about the material in the Oral Passage subtest. Reading comprehension is measured by the child’s ability to answer this set of five questions. Children are only asked questions related to text they actually read. Child proficiency on the reading comprehension subtest is of particular interest, as being able to read and understand connected text is the ultimate goal of reading acquisition (RTI, 2011a-b).

Since 65.8% of children were unable to read a single word of the passage in the Oral Passage subtest, the majority of children were not asked any reading comprehension questions. Of the children who were able to read at least the first line correctly (i.e. 10 words), almost 85% were unable to answer a single question. Furthermore, no one in the entire sample was able to answer all 5 questions correctly (see Graph 36).

Graph 36. Percentage of children per number of correct reading comprehension answers

<table>
<thead>
<tr>
<th>Number of correct reading comprehension answers</th>
<th>Percentage of Children</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>84.8</td>
</tr>
<tr>
<td>1</td>
<td>6.8</td>
</tr>
<tr>
<td>2</td>
<td>5.3</td>
</tr>
<tr>
<td>3</td>
<td>2.2</td>
</tr>
<tr>
<td>4</td>
<td>.8</td>
</tr>
</tbody>
</table>
Even though scores are remarkably low across the board, there are meaningful differences between grades ($d: 0.31$), and rather small differences between boys and girls ($d: 0.14$). Specifically, $4^{th}$ grade children answered on average one fifth of a question more than $3^{rd}$ grade children, and boys answered about one tenth of a question more than girls (see Graph 37).

Mean scores by subdivision follow an almost identical pattern to Oral Passage Reading. Kalemie again outperforms all of the other subdivisions, with Lubudi, Kambove, and Kasenga ranking near the bottom (see Graph 38). More important than the difference between the subdivision mean scores, however, is the magnitude of the mean scores. No subdivision achieved a mean score of even 1 reading comprehension question answered correctly. In other words, Kalemie might out-perform the other subdivisions, but even in the best-performing subdivision, the average child who can read at least one line of the passage, cannot answer even one question about the passage correctly.
8. Listening Comprehension

In this subtest, a short story is read twice, in French, to the child by the test administrator. The children are then asked to respond to five questions that test listening comprehension. The total score is the number of questions correctly answered out of five. Administrators were trained to judge whether the responses were correct in content, regardless of sentence construction or language used. This subtest was administered to children in 2nd to 4th grades.

As shown in the graph below (Graph 39), while the percentage of children unable to answer a single question was still high (51%), it was considerably lower than the percentage for reading comprehension. This suggests that in part, children's difficulties with comprehension are related to difficulties with decoding and not necessarily with other skills such as the ability to understand the content or maintain the story in memory. Difficulties with this task may be mostly attributable to children's unfamiliarity with French.

Graph 39. Percentage of children per number of correct listening comprehension answers

Summary of Listening Comprehension Results (all grades)

- The zero score is moderate relative to the other subtests: 51%; however, 37% of children could only respond to one question correctly, for a total of 88% of children who have little or no listening comprehension skills.
  - This appears to illustrate that children are also unfamiliar with spoken French in addition to written French (as tested by Oral Passage Reading and Reading Comprehension).
- Children improved slightly in listening comprehension as they move up in grade levels; however, the improvement is overshadowed by the overall poor performance.
- Boys outperformed girls at every grade level, although the differences were small.
- Average scores from subdivisions grouped together in pairs on this subtest. Kalemie and Kongolo, Mutshatsha and Lubudi, and Kambove and Kasenga each had identical scores.
- The uniformity of performance in pairs of subdivisions, along with the uniform improvement in the upper grades, suggests that these results may be a reflection of different approaches and emphases in the reading curriculum across subdivisions.
As depicted in Graph 40, there were small but non-trivial differences in the performance of children in different grade levels ($d$: 0.28), with scores increasing in the expected direction. On average, children were able to answer one-fifth of a question more as they moved up a grade level. In addition, there were small differences between boys and girls, with boys slightly outperforming girls in every grade ($d$: 0.10). In spite of average growth across grades, it is important to emphasize that the majority of children across all grades (87.8%) was either unable to answer a single question or could answer only 1 out of 5 questions. Averages for this subtest suggest that even older children were still struggling with this type of task.

Graph 40. Mean scores by grade and gender, Listening Comprehension

While comparisons across subdivisions reveal similar patterns than those found for other subtests, note the similarity of scores by pair of subdivisions in this subtest (Graph 41). The reasons why Kalemie and Kongolo, Mutshatsha and Lubudi, and Kambove and Kasenga scored so closely to each other are not clear as they did not have such similar average scores on the remaining subtests or demographic profiles.

Graph 41. Mean scores by subdivision, Listening Comprehension
Graph 42 suggests that the large differences between high-performing (i.e. Kalemie and Kongolo) and average-performing subdivisions (i.e. Mutshatsha and Lubudi) are due in part to the jump in positive performance rates between 3rd and 4th grade. While the average performance of children in Mutshatsha and Lubudi was relatively constant from 2nd to 4th grade, children in Kalemie and Kongolo were making considerable improvements, especially from 3rd to 4th grade.

The majority of 2nd and 3rd graders in Kambove and the majority of 2nd graders in Kasenga were unable to answer any listening comprehension questions accurately, suggesting that the emphasis of instruction in these 2 subdivisions in early primary school may differ in important ways from instruction in the other subdivisions. Importantly, by 4th grade children in Kambove had reached the level of 2nd and 3rd graders in Kalemie and Kongolo, and by 3rd grade children in Kasenga had done so. The uniformity of performance in these two subdivisions, along with the uniform improvement in the upper grades, suggests that these results may be a reflection of different approaches and emphases in the reading curriculum across subdivisions.

**How to interpret boxplots:**
- The box (blue) shows the range of scores for 50% of children.
- The horizontal line inside the box is the median score (50% of children scored above this number and 50% of children scored below this number).
- The lines (or whiskers) coming out of the box on either end show the full range of scores (the maximum and the minimum).
- The “longer” the box and whiskers appear, the more variability is present.
- The dots and stars outside of the lines (or whiskers) show outlier scores.
Summary of Writing a Complete Sentence Results (4th grade only)

- Even by 4th grade, over half of children were unable to write one of three simple words correctly.
- Overall mean scores were very low, so even though there are differences between genders and subdivisions, the emphasis should remain on the fact that average scores across all categories were extremely low.
  - Boys performed slightly better than girls, but the average score of both genders was less than 1 correct word.
  - Kalemie was again the highest performing subdivision, with an average score of 1.44 words. All other subdivisions had average scores of less than 1 correct word.

9. Writing a Complete Sentence

For this subtest, children were asked to write down a simple sentence, which was dictated three times by the test administrator. The final score represents the number of correctly written words out of maximum possible score of 3, including correct spelling. The second part of the assessment looks at whether or not the child capitalized the initial word in the sentence, wrote from left to right, and put the correct amount of spaces between words. This subtest was administered to 4th graders only. As shown in Graph 43, over half of 4th grade children was unable to write a single word correctly, and only 3.3% was able to write 3 words correctly.

Graph 43. Percentage of children per number of correctly written words
As with every other EGRA subtest, boys did better than girls ($d: 0.29$) (see Graph 44) and children in Kalemie outperformed children in all other subdivisions (see Graph 45).

**Graph 44. Mean scores by gender, Writing a Complete Sentence**

**Graph 45. Mean scores by subdivision, Writing a Complete Sentence**
Early Grade Math Assessment (EGMA) Descriptive Analysis

We start with an overview of EGMA subtests and the grade levels to which they were administered. This is followed by a description of how variance in children's scores was distributed across subdivisions, clusters of schools and children. Lastly, we present descriptive analyses for the overall sample and compare children's scores in each EGMA subtest by grade, gender and subdivision.

Table 12. Overview of EGMA subtests

<table>
<thead>
<tr>
<th>Subtest</th>
<th>Competencies Measured</th>
<th>Administered to</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Number Identification</td>
<td>Ability to identify written numeric symbols.</td>
<td>All grades (2-4) Differences in difficulty level of items administered to 2nd vs. 3rd and 4th grades.</td>
</tr>
<tr>
<td>2. Number (Quantity) Discrimination (Comparison of Quantities)</td>
<td>Ability to make judgments about differences in quantities represented by numbers.</td>
<td>All grades (2-4) Differences in difficulty level of items administered to 2nd vs. 3rd and 4th grades.</td>
</tr>
<tr>
<td>3. Missing Number (Number Patterns)</td>
<td>Ability to discern and complete number patterns.</td>
<td>All grades (2-4) Differences in difficulty level of items administered to 2nd vs. 3rd and 4th grades.</td>
</tr>
<tr>
<td>4. Addition</td>
<td>Procedural competency in basic operations of addition.</td>
<td>All grades (2-4)</td>
</tr>
<tr>
<td>5. Subtraction</td>
<td>Procedural competency in basic operations of subtraction.</td>
<td>All grade (2-4)</td>
</tr>
<tr>
<td>6. Multiplication</td>
<td>Procedural competency in basic operations of multiplication.</td>
<td>3rd &amp; 4th grades</td>
</tr>
<tr>
<td>7. Word Problems</td>
<td>Ability to conceptually understand basic operations when reading a word problem.</td>
<td>All grades (2-4) Differences in difficulty level of items administered to 2nd vs. 3rd and 4th grades.</td>
</tr>
<tr>
<td>8. Shape Attributes (Shape Identification)</td>
<td>Ability to recognize geometric shapes.</td>
<td>All grades (2-4)</td>
</tr>
<tr>
<td>9. Shape Naming</td>
<td>Ability to name geometric shapes.</td>
<td>All grades (2-4)</td>
</tr>
</tbody>
</table>

1. Variability in children's EGMA scores

Originally, we had hypothesized that variation in EGMA scores can be caused by differences in the children themselves (e.g. demographic differences, household differences, cognitive differences), differences in school clusters within subdivisions, and differences between subdivisions. Table 13 shows that on average, 82% of the variation in children's EGMA scores can be explained by differences between children themselves; 8.4% of the variation can be explained by differences in subdivisions; 8.3% of the variation can be explained by differences between school clusters within subdivisions.

For instance, 3.8% of variability in children's performance on the Missing Numbers Subtest was due to differences between clusters while 11% of variability in children's performance on the Shape Identification Subtest was due to differences between clusters of schools.
Individual factors that can contribute to the explanation of between-child variation in reading scores range from demographic characteristics such as gender and household socio-economic status, to differences in children’s core cognitive abilities such as working memory and attention control. Differences between schools clusters could include differential access to human resources (e.g., more or less educated teachers), variation in the quality of school infrastructure and support from the community, the state and other agents.

Table 13 also shows that many subtests have large amounts of variability. As discussed earlier, this is important because while the average scores for each subtest are descriptive, they do not fully illustrate the EGRA results. Large amounts of variance means that even though EGMA average scores might be low on most subtests, there are students that are scoring much higher and much lower. This will be a challenge, and perhaps not possible, to develop a curriculum that helps the highest and lowest performing students remain engaged and improve over the course of the year. The amount of variance is also important to consider when setting targets around the average scores.

<table>
<thead>
<tr>
<th>Subtest (number of items)</th>
<th>Percentage of variance**</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total variance</td>
</tr>
<tr>
<td>Number Identification (30)</td>
<td>55.807</td>
</tr>
<tr>
<td>Quantity Discrimination (10)</td>
<td>6.925</td>
</tr>
<tr>
<td>Missing Number (10)</td>
<td>5.284</td>
</tr>
<tr>
<td>Addition (24)</td>
<td>28.836</td>
</tr>
<tr>
<td>Subtraction (24)</td>
<td>24.242</td>
</tr>
<tr>
<td>Multiplication (10)*</td>
<td>2.931</td>
</tr>
<tr>
<td>Word Problems (6)</td>
<td>2.057</td>
</tr>
<tr>
<td>Shape Identification (4)</td>
<td>1.710</td>
</tr>
<tr>
<td>Shape Naming (4)</td>
<td>1.548</td>
</tr>
<tr>
<td><strong>Mean</strong></td>
<td>8.41</td>
</tr>
<tr>
<td><strong>Standard deviation (SD)</strong></td>
<td>3.52</td>
</tr>
</tbody>
</table>

Note: *Not administered to 2nd graders, ***All variance estimates were significant at p < .05, which means there is significant variation in children’s scores at each level examined (i.e., subdivision, cluster, individual)
1. **Zero scores by grade**

The zero score per subtest is the percentage of children who could not respond correctly to even one question. The zero score is a measure of competencies that children do not have at all and can indicate a potential need for increased curricular focus. Graph 46 shows that for all subtests except Multiplication, the majority of children are able to respond correctly to at least one question. This is the opposite of EGRA findings. Zero scores on these tests range from less than 1% (Number Identification, 4th grade) to 41% (Subtraction, 2nd grade), indicating that while much improvement is still needed, children appear to have more basic math skills than basic reading skills.

Zero scores decrease as grade levels increase, implying that children are learning as they progress in school. The only zero score that increases is Word Problems between 2nd and 3rd grade (see Table 14). The difficulty of the test increases from 2nd to 3rd grade, so it is not surprising that the zero score also increases; however, the test difficulty increases on a number of other subtests, yet all other zero scores decrease. Number Identification and Quantity Discrimination show the least progress between grades, improving by 2.9% and 4.1% each by 4th grade, while Subtraction shows the most progress between grades, improving 27.1% by 4th grade. More than half of subtests (Number Identification, Quantity Discrimination, Missing Number, Addition, and Shape Identification) have zero scores below 10% by the 4th grade.

---

### Summary of Descriptive Analysis – EGMA

- Performance is very low, and while average scores for EGMA are higher than EGRA, children are still not performing well enough to be considered competent in many areas of math.
- Children answer on average between 6% (Multiplication, Kambove) and 69% (Shape Identification, Kalemie) of questions correctly.
- The majority of EGMA subtests had moderate zero scores, ranging from 16% (Shape Identification) to 41% (Subtraction). Multiplication had an extremely high zero score (72%) and Number Identification and Quantity Discrimination had zero scores less than 10%. Tests with low zero scores tended to have higher average scores.
- Zero scores improved with grade level, up to 27%; however, this improvement is not reflected in average scores for subtests. This could be because the EGMA for 3rd and 4th grade is more difficult than the EGMA for 2nd grade.
- Boys outperformed girls across all grades in the majority of subtests, with the gap between boys’ and mean correct scores increasing as children progress in grade level except for tests that were more difficult for 3rd/4th graders than 2nd graders.
- Some subdivisions perform similarly across all subtests. For example, Kasenga is ranked sixth for almost all subtests. Kambove and Lubudi tend to rank towards the bottom (fourth or fifth). Mutshatsha tends to rank towards the top (first, second, or third), but the remaining two subdivisions (Kalemie and Kongolo) rank from first to fifth place, depending on the subtest.
2. Average correct scores by gender

Boys outperformed girls across all grades in all subtests. The gap between boys’ and mean correct scores increases as children progress in grade level except for some tests that became more difficult for 3rd/4th graders than they were for 2nd graders (Number Identification, Quantity Discrimination, Addition, Word Problems). Table 15 shows the differences between the mean correct score for boys and girls by grade. For example, looking at the Number Identification subtest, in 2nd grade, on average boys could identify 0.9 more numbers than girls; in 3rd grade, boys could identify 0.48 more numbers than girls; and in 4th grade, boys could identify 1.18 more numbers than girls.
Table 15. Differences between mean correct scores by gender, by grade

<table>
<thead>
<tr>
<th>EGMA Subtest</th>
<th>2nd grade</th>
<th>3rd grade</th>
<th>4th grade</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number Identification</td>
<td>0.9</td>
<td>0.48</td>
<td>1.18</td>
</tr>
<tr>
<td>Quantity Discrimination</td>
<td>0.41</td>
<td>0.01</td>
<td>0.23</td>
</tr>
<tr>
<td>Missing Number</td>
<td>0.27</td>
<td>0.35</td>
<td>0.3</td>
</tr>
<tr>
<td>Addition</td>
<td>0.72</td>
<td>0.64</td>
<td>0.92</td>
</tr>
<tr>
<td>Subtraction</td>
<td>0.4</td>
<td>0.73</td>
<td>1.44</td>
</tr>
<tr>
<td>Multiplication*</td>
<td>-</td>
<td>0.3</td>
<td>0.61</td>
</tr>
<tr>
<td>Word Problems</td>
<td>0.2</td>
<td>0.13</td>
<td>0.29</td>
</tr>
<tr>
<td>Shape Identification</td>
<td>0.07</td>
<td>0.10</td>
<td>0.15</td>
</tr>
<tr>
<td>Shape Naming</td>
<td>0.18</td>
<td>0.45</td>
<td>0.28</td>
</tr>
</tbody>
</table>

* Only administered to 3rd and 4th grades

3. Average correct scores by subdivision

As noted, differences between subdivisions account for a portion of the variance seen in EGMA scores, although not to the degree that differences between children do. Table 16 shows the mean correct score by subtest and by subdivision, as well as the rank of each subdivision by subtest. For example, for Missing Number, Kambove had a mean score of 2.66 missing numbers correctly identified out of 10 possible numbers and ranked 5th out of 6 subdivisions. The mean correct scores across subdivisions show that across all subtests, performance is low, with children answering on average 6% (Multiplication, Kambove) to 69% (Shape Identification, Kalemie) of questions correctly.

Some subdivisions perform similarly across all subtests. For example, Kasenga is ranked sixth for almost all subtests. Kambove and Lubudi tend to rank towards the bottom (fourth or fifth). Mutshatsha tends to rank towards the top (first, second, or third), but the remaining two subdivisions (Kalemie and Kongolo) rank from first to fifth place, depending on the subtest.
Table 16. Mean correct scores by subdivision; Subdivision rank by subtest

<table>
<thead>
<tr>
<th>EGMA Subtest</th>
<th>Kambove</th>
<th>Kasenga</th>
<th>Kalemie</th>
<th>Kongolo</th>
<th>Mutshatsha</th>
<th>Lubudi</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number Identification (30)</td>
<td>12.16</td>
<td>10.23</td>
<td>18.01</td>
<td>13.26</td>
<td>16.05</td>
<td>12.97</td>
</tr>
<tr>
<td>Subdivision Rank</td>
<td>5</td>
<td>6</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>Quantity Discrimination (10)</td>
<td>6.73</td>
<td>5.01</td>
<td>6.43</td>
<td>5.34</td>
<td>6.54</td>
<td>6.53</td>
</tr>
<tr>
<td>Subdivision Rank</td>
<td>1</td>
<td>6</td>
<td>4</td>
<td>5</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Missing Number (10)</td>
<td>2.66</td>
<td>2.04</td>
<td>4.54</td>
<td>3.04</td>
<td>3.69</td>
<td>3.14</td>
</tr>
<tr>
<td>Subdivision Rank</td>
<td>5</td>
<td>6</td>
<td>1</td>
<td>4</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Addition (24)</td>
<td>6.75</td>
<td>4.95</td>
<td>8.55</td>
<td>8.71</td>
<td>8.94</td>
<td>6.6</td>
</tr>
<tr>
<td>Subdivision Rank</td>
<td>4</td>
<td>6</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>Subtraction (24)</td>
<td>3.54</td>
<td>2.96</td>
<td>5.87</td>
<td>6.83</td>
<td>6.86</td>
<td>4.09</td>
</tr>
<tr>
<td>Subdivision Rank</td>
<td>5</td>
<td>6</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>Multiplication (10)</td>
<td>0.56</td>
<td>0.84</td>
<td>1.13</td>
<td>1.47</td>
<td>1.1</td>
<td>0.82</td>
</tr>
<tr>
<td>Subdivision Rank</td>
<td>6</td>
<td>4</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>Word Problems (6)</td>
<td>2.23</td>
<td>1.29</td>
<td>3.07</td>
<td>2.04</td>
<td>3.31</td>
<td>2.48</td>
</tr>
<tr>
<td>Subdivision Rank</td>
<td>4</td>
<td>6</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Shape Identification (4)</td>
<td>2.65</td>
<td>1.56</td>
<td>2.74</td>
<td>2.53</td>
<td>2.65</td>
<td>2.57</td>
</tr>
<tr>
<td>Subdivision Rank</td>
<td>3</td>
<td>6</td>
<td>1</td>
<td>5</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>Shape Naming (4)</td>
<td>1.34</td>
<td>1.25</td>
<td>2.37</td>
<td>1.39</td>
<td>1.5</td>
<td>1.27</td>
</tr>
<tr>
<td>Subdivision Rank</td>
<td>4</td>
<td>6</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>5</td>
</tr>
</tbody>
</table>
Descriptive Analysis by subtests

Summary of Number Identification Results (all grades; increase in difficulty for 3rd/4th grades)

- This subtest has a low zero score (1.95%), but only 2.4% of children could identify all 30 numbers. Instead, the children are approximately evenly distributed across the possible test scores, with at least half of children able to identify at least half of the numbers.
- Scores do improve with grade, although there are only improvements from 2nd to 4th grade, and from 3rd to 4th grade. Scores decrease from 2nd to 3rd grade, most likely because the 3rd/4th grade subtest is harder than the 2nd grade subtest.
- Kalemie performed the best (average score = 18.01) and Kasenga performed the worst (10.23); however, there was a wide range of scores within subdivisions.

1. Number Identification

This subtest targets the child’s knowledge and identification of written symbols. It assesses the child’s recognition and understanding that each of the numbers is a constant with one number-word associated with it, and that the child knows the number word(s) associated with the number symbol.

In this task children are asked to point to each number and identify it. Children are given one minute to correctly identify as many numbers as possible out of a total of 30 numbers. The score is based on the last correct number the child says previous to making an error or at the end of a minute. In Grade 2 children were asked to identify both single and double-digit numbers. In Grade 3 and 4 children encountered single, double and triple digit numbers as well as numbers with decimals out to the third place.

Results show that on average children were able to identify approximately 14 of the 30 numbers in this task. As can be seen from Graph 47, only 2.43% of children were able to identify all 30 numbers in this task; 1.95% was not able to identify any number and just under 16% was able to identify between 1 and 6 numbers. Half of the sample (52.09%) could identify at least 14 numbers.

Graph 47. Percentage of children per number of correct numbers identified
Unlike most of the results from the EGRA there was not an upward trend associated with grade when it came to proficiency in number identification (see Graph 48). The average number that was identified decreased from 14.19 to 12.15 from Grade 2 to Grade 3 and then increased to 15.72 for Grade 4. This could possibly be attributed to the fact that for 2nd graders the number identification task involved 1 and 2 digit numbers while for 3rd and 4th graders it involved 1 to 4 digit numbers. Therefore, the change in task difficulty from grade 2nd to 3rd could lead to what looks like a decrease in performance for 3rd graders.

In line with results on EGRA, the average number of answers correct was only slightly higher for boys than girls across all grades ($d$: 0.12).

Graph 48. Mean scores by grade and gender, Number Identification

Graph 49 shows children in Kasenga performed worse than children in all other subdivisions, while children from Kalemie performed the best.

Graph 49. Mean scores by subdivision, Number Identification
As shown by the boxplot (Graph 50), there was a wide range of performance within grade and subdivision. For example, 2\textsuperscript{nd} graders in all subdivisions except Kasenga scored the full range of possible scores, from 0 to 30, yet the median of each subdivision is markedly different. It is worth highlighting the improvement from 2\textsuperscript{nd} and 3\textsuperscript{rd} grades to 4\textsuperscript{th} grade in Kasenga. While younger children in this subdivision were doing much worse than children in the same grades in other subdivisions, the gap was reduced by the time children reached 4\textsuperscript{th} grade. While 50\% of 4\textsuperscript{th} grade children in Kasenga were still getting scores comparable to those of children in the bottom 25\textsuperscript{th} percentile in Kalemie and Mutshatsha, the distribution of scores resembles those of the remaining distributions.

Graph 50. Distribution of scores by subdivision and by grade, Number Identification

How to interpret boxplots:
- The box (blue) shows the range of scores for 50\% of children.
- The horizontal line inside the box is the median score (50\% of children scored above this number and 50\% of children scored below this number).
- The lines (or whiskers) coming out of the box on either end show the full range of scores (the maximum and the minimum).
- The line coming from the top of the box represents the top 25\%; the line coming from the bottom of the box represents the bottom 25\%.
- The “longer” the box and whiskers appear, the more variability is present.
- The dots and stars outside of the lines (or whiskers) show outlier scores.
2. Number (Quantity) Discrimination (Comparison of Quantities)

Number discrimination in EGMA measures children’s ability to make judgments about differences by comparing quantities. Quantity discrimination in the early grades demonstrates a critical link to an effective and efficient counting strategy for problem solving. Without this ability, children are more apt to make errors or use less efficient strategies.

In this subtest, each item consists of two numbers. Children are asked to identify the larger number (e.g., “Which one is bigger?”). In grade 2 the number pairs involve numbers below 100. In grades 3 & 4 the number range is expanded to include numbers pairs in the hundreds, including both fractions and decimals.

The results revealed that children were able to correctly discriminate about 6.09 out of the 10 number pairs on average. As can be seen from Graph 51, only 6.73 % of children was able to correctly discriminate all 10 number pairs on the subtest, 4.41% of children was unable to discriminate between any number pair and 17.25 % of children was able to discriminate between 0 to 3 number pairs only.

Graph 51. Percentage of children per number of correct Quantity Discrimination questions

Summary of Quantity Discrimination Results (all grades; increase in difficulty for 3rd/4th grades)

- This subtest has a low zero score (4.41%), but only 6.73% of children could correctly answer all 10 questions. Instead, the children are distributed normally across the possible test scores, with 65% of children able to answers at least half of the questions.
- Boys in general performed better than girls, except in 3rd grade, where scores were equal.
- Scores do improve with grade, although there are only improvements from 2nd to 4th grade, and from 3rd to 4th grade. Scores decrease from 2nd to 3rd grade, most likely because the 3rd/4th grade subtest is harder than the 2nd grade subtest.
- Kambove performed the best (average score = 6.73), although Kalemie, Mutshatsha, and Lubudi had similar scores. Kasenga and Kongolo performed the worst.
When examining the results by grade, number discrimination followed the same pattern as task number identification (see Graph 52). The average decreased from 6.29\textsuperscript{12} to 5.56 from grades 2 to 3 and then increased to 6.43 from grades 3 to 4. Effect sizes for these comparisons were fairly small but not trivial, with \(d\) of 0.27 for the changes from grade 2 to 3 and \(d\) of 0.37 for the changes from grade 3 to 4. The decrease in performance from 2\textsuperscript{nd} to 3\textsuperscript{rd} grade can be attributed to the higher difficulty in items administered to 3\textsuperscript{rd} compared to 2\textsuperscript{nd} graders.

Surprisingly, when looking at results by gender and grade, the mean number of correct quantity discrimination items in 3\textsuperscript{rd} grade was almost equal for both genders, while the general pattern of boys outperforming girls held true in the other grades.

For the quantity discrimination subtest, with the exception of Kasenga and Kongolo, the remaining subdivisions had very similar means (see Graph 53).

\textsuperscript{12} Note that these means are the average of the means for boys and girls in each grade.
It is important to note, however, that in spite of similarities in means, there were important differences across subdivisions.

As shown by the length of the whiskers in Graph 54, Kongolo had children in all grades spread out across the full range of the scale (with the exception of 4th grade in which no child got a score higher than 9), which indicates that an important percentage of children in the upper grades were still getting scores at the bottom of the scale. Regarding Kasenga, it is noteworthy that 50% of the children in this subdivision obtained lower scores than 75% of the children in Kalemie, Mutshatsha and Lubudi.

Graph 54. Distribution of scores by subdivision and grade, Quantity Discrimination

How to interpret boxplots:
- The box (blue) shows the range of scores for 50% of children.
- The horizontal line inside the box is the median score (50% of children scored above this number and 50% of children scored below this number).
- The lines (or whiskers) coming out of the box on either end show the full range of scores (the maximum and the minimum).
- The line coming from the top of the box represents the top 25%; the line coming from the bottom of the box represents the bottom 25%.
- The “longer” the box and whiskers appear, the more variability is present.
- The dots and stars outside of the lines (or whiskers) show outlier scores.
3. **Missing Number (Number Patterns)**

In this subtest, children are shown three to four numbers in a number sequence and a place-holder for a next or missing number. The child is asked to name the missing number. Being able to recognize number patterns and counting in patterns (by ones, by twos, by tens, counting backwards, etc.) lays the foundation for multiplication and, leads into algebra.

For all grades, pattern completion involved counting forwards and backwards in ones, twos, fives, tens and in one hundreds. 3rd and 4th graders had to count in ones in a series of three hundred numbers (i.e. 347, 348, 349, 350) while 2nd graders only counted in ones in series including numbers up to ninety two.

For the missing number subtest, results indicated, that on average, children were able to name 3.26 out of the 10 missing numbers. As Graph 55 indicates, only 0.68% of children were able to correctly name all missing numbers; 14.24% was unable to name any of the 10 missing numbers, and 56.49% was only able name 0 to 3 of the missing numbers.

<table>
<thead>
<tr>
<th>Number of Correct Answers</th>
<th>Percentage of Children</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.96%</td>
</tr>
<tr>
<td>1</td>
<td>2.32%</td>
</tr>
<tr>
<td>2</td>
<td>5.02%</td>
</tr>
<tr>
<td>3</td>
<td>9.05%</td>
</tr>
<tr>
<td>4</td>
<td>11.89%</td>
</tr>
<tr>
<td>5</td>
<td>13.59%</td>
</tr>
<tr>
<td>6</td>
<td>11.10%</td>
</tr>
<tr>
<td>7</td>
<td>10.13%</td>
</tr>
<tr>
<td>8</td>
<td>7.34%</td>
</tr>
<tr>
<td>9</td>
<td>23.80%</td>
</tr>
<tr>
<td>10</td>
<td>14.24%</td>
</tr>
</tbody>
</table>

Graph 55. Percentage of children per number of correct answers. Missing Number
In contrast with the number identification and quantity discrimination subtests, proficiency with this subtest increased by grade (see Graph 56). Children's scores changed substantially from 2\textsuperscript{nd} to 3\textsuperscript{rd} grade ($d$: 0.45) and there were smaller yet important changes from grades 3 to 4 ($d$: 0.31). Boys slightly outperformed girls in every grade ($d$: 0.16).

These results show that Kalemie was once again at the top for this subtest. As shown in Graph 57, Kalemie was the only subdivision in which 4th graders were no longer represented at the bottom of the scale.
When examining the distributions of this subtest a few things stand out (Graph 58). One is the low medians of 2nd and 3rd graders in Kasenga, which indicate that 50% of children had less than 1 or 2 correct questions, respectively. Another is that about 75% of 2nd and 4th graders in Kasenga scored lower than 75% of 2nd and 4th graders in Kalemie, illustrating how wide the gap is between these two subdivisions.

Graph 58. Distribution of scores by subdivision and grade, Missing Number

How to interpret boxplots:

- The box (blue) shows the range of scores for 50% of children.
- The horizontal line inside the box is the median score (50% of children scored above this number and 50% of children scored below this number).
- The lines (or whiskers) coming out of the box on either end show the full range of scores (the maximum and the minimum).
- The line coming from the top of the box represents the top 25%; the line coming from the bottom of the box represents the bottom 25%.
- The “longer” the box and whiskers appear, the more variability is present.
- The dots and stars outside of the lines (or whiskers) show outlier scores.
Basic Operations

In these subtests children are presented with basic math problems. The items are presented in written format and children are tested in their ability to perform the calculations mentally. Addition and subtraction were assessed in all grades using both single digit and double digit numbers with all numbers and answers below 25. Multiplication was assessed in only grades 3 and 4.

Summary of Addition Results (all grades)

- This subtest has a zero score 13.3%, with less than 1% of children able to answer all 24 questions correctly. Children's scores were distributed approximately normally over the remaining possible scores.
- Scores improve with grade, and boys outperform girls at all grade levels.
- Mutshatsha performed the best (average score = 8.94) and Kasenga performed the worst (4.95); however, while 2nd graders in Kasenga have very low scores, 4th graders in Kasenga perform more similarly to their peers in other subdivisions.

4. Addition

When tested on addition, children across all grades were able to answer 7.58 out of 24 questions on average (see Graph 59).

The number of addition problems answered correctly improved substantially with each successive grade (average $d$ across grades: 0.62) (see Graph 60). From 2nd to 3rd grade the mean increased from 4.80 to 7.53 and from 3rd to 4th grade it reached 10.72.

The mean number of addition questions that were correct followed the trend of previous subtests with boys doing only slightly better than girls across all grades ($d$: 0.17).
Children in Kalemie, Kongolo, and Mutshatsha on average performed better than children in Kambove, Kasenga, and Lubudi (see Graph 61).

The most striking result of the distributions presented in the boxplots (Graph 62) is the low median of Kasenga’s 2nd graders (close to 0). The range of these 2nd graders is also extremely narrow when compared to the ranges in other subdivisions. Yet it is interesting to see that Kasenga’s 4th graders performed at the level of most other subdivisions – even displaying a higher median than Kambove and Kongolo.
Graph 62. Distribution of scores by subdivision and grade, Addition

How to interpret boxplots:
- The box (blue) shows the range of scores for 50% of children.
- The horizontal line inside the box is the median score (50% of children scored above this number and 50% of children scored below this number).
- The lines (or whiskers) coming out of the box on either end show the full range of scores (the maximum and the minimum).
- The line coming from the top of the box represents the top 25%; the line coming from the bottom of the box represents the bottom 25%.
- The “longer” the box and whiskers appear, the more variability is present.
- The dots and stars outside of the lines (or whiskers) show outlier scores.
Summary of Subtraction Results (all grades)

- This subtest has a zero score 27.4%, almost double that of addition, with no children able to answer more than 20 out of 24 questions correctly. Children’s scores were skewed towards zero.
- Scores improve with grade, and boys outperform girls at all grade levels.
- Like the Addition subtest, Mutshatsha performed the best (average score = 6.86) and Kasenga performed the worst (2.96). Kasenga also showed less variation in scores for 2nd graders in comparison to other subdivisions.

5. Subtraction

The mean number of correct responses on subtraction questions was 5.24 out of 24 questions. As Graph 63 indicates, a majority of children (55.98%) was only able to answer between 0 to 5 subtraction questions.

Graph 63. Percentage of children per number of correct answers, Subtraction

Graph 64 shows that on average, children made important improvements from one grade to the next on the number of subtraction questions answered correctly (average d across grades: 0.48). In grade 2, the mean was 3.30 while in grade 3 the mean was 5.03 and in grade 4, the mean was 7.66. Girls over all grades received a slightly lower mean number of correct subtraction answers than boys (d: 0.19).
Children in Kongolo and Mutshatsha performed the best, with an average of 7 correct subtraction questions, while Kasenga was at the bottom with an average of 3 subtraction questions answered correctly (see Graph 65).
The vast majority of 2\textsuperscript{nd} grade children in Kasenga obtained scores equal or very close to 0, while 2\textsuperscript{nd} graders in other subdivisions showed greater variability in performance as shown by the size of the boxes and the length of the upper bound whiskers (see Graph 66).

**Graph 66. Distribution of Subtraction scores by subdivision and grade**

How to interpret boxplots:
- The box (blue) shows the range of scores for 50\% of children.
- The horizontal line inside the box is the median score (50\% of children scored above this number and 50\% of children scored below this number).
- The lines (or whiskers) coming out of the box on either end show the full range of scores (the maximum and the minimum).
- The line coming from the top of the box represents the top 25\%; the line coming from the bottom of the box represents the bottom 25\%.
- The “longer” the box and whiskers appear, the more variability is present.
- The dots and stars outside of the lines (or whiskers) show outlier scores.
Summary of Multiplication Results (3rd/4th grades only)

- This subtest has a very high zero score 72.1%, dramatically higher than any other EGMA subtest.
- Scores improve with grade, although even in 4th grade, the average correct score remains low.
- Boys slightly outperform girls at all grade levels.
- As with subtraction and addition, children in Kalemie, Kongolo, and Mutshatsha on average performed better than children in Kambove, Kasenga, and Lubudi on the multiplication subtest. Kongolo performed the best (average score = 1.47) and Kambove performed the worst (0.56).
- Uniformity in children’s scores in Kambove, Kasenga and Lubudi may be indicative of different curricular emphases across subdivisions.

6. Multiplication

Children on average could only answer 1.03 multiplication questions out of 10. Furthermore, 72.11 percent of the sample was unable to answer any of the 10 questions on multiplication, as shown on Graph 67.

Graph 67. Percentage of children per number of correct answers, Multiplication

Graph 68 shows the mean number of correct responses almost doubled from grade 3 to 4 increasing from .69 to 1.38 ($d$: 0.41). Similar to other subsets, boys outperformed girls in the mean number of multiplication problems answered correctly, and this was the case in 3rd and 4th grade (average $d$ across grades: 0.28). As shown in the graph, the advantage of boys over girls is larger in 4th grade (difference in $d$: 0.10).
As with subtraction and addition, children in Kalemie, Kongolo, and Mutshatsha on average performed better than children in Kambove, Kasenga, and Lubudi on the multiplication subtest (Graph 69).
Graph 70 shows that the vast majority of 3rd graders in Kambove, Kasenga and Lubudi were unable to answer a single multiplication question (demonstrated by the lack of a box and whiskers), whereas at least 25% of their counterparts in Kalemie, Kongolo and Mutshatsha were able to answer 1 or more questions. Uniformity in children's scores in Kambove, Kasenga and Lubudi may be indicative of different curricular emphasis across subdivisions.

Graph 70. Distribution of scores by subdivision and grade, Multiplication

How to interpret boxplots:
- The box (blue) shows the range of scores for 50% of children.
- The horizontal line inside the box is the median score (50% of children scored above this number and 50% of children scored below this number).
- The lines (or whiskers) coming out of the box on either end show the full range of scores (the maximum and the minimum).
- The line coming from the top of the box represents the top 25%; the line coming from the bottom of the box represents the bottom 25%.
- The “longer” the box and whiskers appear, the more variability is present.
- The dots and stars outside of the lines (or whiskers) show outlier scores.
Summary of Word Problems Results (all grades; increase in difficulty for 3rd/4th grades)

- This subtest has a zero score of 16.7%. Children’s scores were otherwise normally distributed, with the most children (20.6%) able to respond to half of the questions.
- Scores improve with grade, although even in 4th grade, the average correct score only reaches approximately 50%.
- Boys slightly outperform girls at all grade levels.
- Mutshatshe performed the best (average score = 3.31) and Kasenga performed the worst (1.29); however, the means conceal a significant amount of variation within grades and between subdivision.

7. Word Problems

Word problems help us to assess children’s informal concepts of addition and subtraction and their ability to apply mathematics concepts to analyze and solve problems. In this section, children are presented with short problems that cover different problem types for each mathematical procedure (addition and subtraction for grade 2 as well as multiplication and division for grades 3 and 4). In Grade 2 the word problems cover the three key addition and subtraction problem types: joining/separating, combining, and comparing. In Grades 3 and 4 the word problems cover the three key addition and subtraction problem types: joining/separating, combining, and comparing; as well as sharing, grouping (division) and multiplication situations.

Children in the sample had a mean of 2.46 correct answers out of 6 word problems (see Graph 71). Approximately 17% of children answered 0 word problems correctly while about 8% was able to answer all 6 word problems presented to them.

Graph 71. Percentage of children per correct answers, Word Problems
The mean proficiency on the word problems subtest increased with each grade (Graph 72). The differences between grades were fairly small from 2nd to 3rd grade ($d: 0.19$), and moderate from 3rd to 4th grade ($d: 0.40$). On average, children in 2nd grade could only answer 2.02 word problems while children in 3rd grade could answer 2.35 word problems, and children in grade 4 averaged 3.06 correct answers. Resembling the results found in other subtests, boys had a slightly higher mean than girls on correctly answered word problems ($d: 0.14$).

Graph 72. Mean scores by grade and gender, Word Problems

On average, children in Kalemie and Mutshatsha obtained the highest scores on this subtest compared to the rest of the subdivisions (Graph 73).

Graph 73. Mean scores by subdivision, Word Problems
As shown in the boxplot (Graph 74), the median scores of all children in Kasenga are well below those of other subdivisions. We see a similar pattern as in other math tests. Almost 75% of the 2nd, 3rd, and 4th graders in Kasenga performed worse than 75% of their grade counterparts in Kalemie. Furthermore, even though Lubudi has a mean above average (overall mean = 2.48) on this subtest, 50% of 2nd and 3rd graders obtained scores below 2. Also interesting to note is that the median scores across several grades and subdivisions is very similar. For example, grades 2-4 in Kambove appear to have the same mean as grade 4 in Kasenga, grades 2-4 in Kongolo, and grades 2 and 3 in Lubudi. Grades 2 and 3 in Kasenga have the same median, as do grades 2 and 3 in Kalemie and grade 2 in Mutshatsha. Grade 4 in Kalemie and Grade 4 in Lubudi have the same median as Mutshatsha.

Graph 74. Distribution of scores by subdivision and grade, Word Problems

How to interpret boxplots:

- The box (blue) shows the range of scores for 50% of children.
- The horizontal line inside the box is the median score (50% of children scored above this number and 50% of children scored below this number).
- The lines (or whiskers) coming out of the box on either end show the full range of scores (the maximum and the minimum).
- The line coming from the top of the box represents the top 25%; the line coming from the bottom of the box represents the bottom 25%.
- The “longer” the box and whiskers appear, the more variability is present.
- The dots and stars outside of the lines (or whiskers) show outlier scores.
8. Shape Recognition (Shape Identification)

In this subtest children are given different stimulus sheets. The child is given a description of a shape and asked to identify the shape on the sheet that does not meet the criteria. For example, when given the triangle sheet the child is told: “Look at these shapes. These shapes should all have three sides and three angles or points. Point to the shape that does not have three sides and three angles.” The shapes tested are circles, squares, triangles and rectangles.

As shown in Graph 75, about 10% of children were unable to correctly identify a single geometric shape out of the 4 shapes presented. Almost 26% of children correctly identified all shapes that were presented to them. In addition, about 29% was able to identify three out of the total four shapes.

On average, children's ability to identify shape attributes slightly improved with each grade (average $d$ across grades: 0.20) (Graph 76). Second grade children were able to answer 2.21 questions, 3rd graders were able to answer 2.52 ($d$ for 2nd vs. 3rd grade: 0.24), and 4th graders answered an average of 2.70 questions ($d$ for 3rd vs. 4th grade: 0.15)
Notably, across all grades on this subtest, the average number of shapes correctly identified by both males and females were very similar ($d: 0.09$).

Graph 76. Mean scores by grade and gender, Shape Identification

Based on Graph 77, most of the subdivisions were about the same in terms of the average number of shapes identified with the exception of Kasenga, which performed at a lower level than everyone else. This sharp deviation of Kasenga is visible in the graph below. When looking at the average number of shapes correctly identified in each subdivision, it can be stated that the difference between Kasenga and all other subdivisions is about one shape.

Graph 77. Mean scores by subdivision, Shape Identification
As shown in the boxplots (Graph 78), the median scores of all children in Kasenga are well below those of other subdivisions. The difference between the 2nd graders of Kasenga and their counterparts in the other five subdivisions is most striking. About 75% of the 2nd graders in Kasenga performed worse than 75% of the 2nd graders in Kalemie and Kongolo, and also worse than at least 50% of their counterparts in Kambove, Mutshatsha, and Lubudi.

Graph 78. Distribution of scores by subdivision and grade, Shape Identification

How to interpret boxplots:
- The box (blue) shows the range of scores for 50% of children.
- The horizontal line inside the box is the median score (50% of children scored above this number and 50% of children scored below this number).
- The lines (or whiskers) coming out of the box on either end show the full range of scores (the maximum and the minimum).
- The line coming from the top of the box represents the top 25%; the line coming from the bottom of the box represents the bottom 25%.
- The “longer” the box and whiskers appear, the more variability is present.
- The dots and stars outside of the lines (or whiskers) show outlier scores.
9. **Shape Naming**

This subtest is a continuation of subtest 8 (i.e., shape identification). After children are presented with a stimulus sheet and asked to identify the shape that does not meet the criteria (see Shape Identification), they are asked to give the name of the shapes that meet the criteria – for example, those that have three sides and three angles: “What’s the name of these shapes?”

Similar to the other subtests, there is a substantial percentage of children that were unable to even answer a single question correctly. About 75% of children were able to name between 0 to 2 shapes correctly (see Graph 79). Only 7.7% of the sample could name all 4 shapes correctly.

[Graph 79. Percentage of children per number of correct answers, Shape Naming]

The average number of shapes that was named correctly followed the trend of the other subtests (Graph 80). Boys slightly outperformed girls in all grades ($d$: 0.11) and performance improved importantly with grade (average $d$ across grades: 0.22).
As can be seen from Graph 81, the average number of correct shapes named is about the same for all the subdivision except for Kalemie, which has an average that is substantially larger than all the other subdivisions.

Graph 81. Mean scores by subdivision, Shape Naming

As shown in the boxplot (Graph 82), Kalemie had higher median scores than all other subdivisions in all grades. The performance of 4th graders was remarkably high; about 75% of Kalemie’s 4th graders outperformed 75% of their counterparts in Kasenga and Lubudi and 50% in Kambove and Kongolo. Despite the exceptional scores of Kasenga, one can see that Mutshatsha had greater variability than the other low performing subdivisions. These differences are clearer among 3rd and 4th graders. For example,
about 50% of the 3rd graders in Mutshatsha had higher scores than 75% of the 3rd graders in Lubudi, Kongolo, Kasenga, and Kambove. Again, there are interesting similarities in the medians of particular grades and subdivisions.

Graph 82. Distribution of scores by subdivision and grade, Shape Naming

How to interpret boxplots:
- The box (blue) shows the range of scores for 50% of children.
- The horizontal line inside the box is the median score (50% of children scored above this number and 50% of children scored below this number).
- The lines (or whiskers) coming out of the box on either end show the full range of scores (the maximum and the minimum).
- The line coming from the top of the box represents the top 25%; the line coming from the bottom of the box represents the bottom 25%.
- The “longer” the box and whiskers appear, the more variability is present.
- The dots and stars outside of the lines (or whiskers) show outlier scores.
Conclusions and Recommendations

Conclusion 1: Taken together, baseline findings from the six subdivisions targeted by OPEQ in Katanga province revealed very low levels of performance in reading and math, regardless of subdivision, gender and grade. These results seem similar to other EGRA and EGMA results in other DRC provinces.

The zero scores on the Oral Reading Passage and Reading Comprehension subtests illustrate clearly that primary school children cannot read in French. 68% of children could not read even one word of a simple passage; 91% of children who were able to read the first line of the passage could not answer a simple comprehension question about the content. As Tables 1 and 2 indicate, OPEQ baseline results are similar to the PAQUED pilot studies with EGRA and EGMA. This most likely indicates that low scores are characteristic of primary school children in the DRC, and OPEQ's baseline score are not out of the ordinary.

While in several reading subtests the majority of children obtained 0 scores, this was rare for math subtests. As suggested before, it is difficult to grasp new concepts and master new skills in a foreign language. Compared to reading in French, math may be less dependent on language proficiency and this may explain in part why children obtained better results in math than reading. In addition, children most likely encounter opportunities to use math concepts in French in their daily lives, for instance when they give or receive a telephone number or count money. They most likely have almost no exposure to written French, and as the language data shows, 98% of children have no exposure to spoken French in the home either.

One of the factors that accounts for the overall low levels of performance is that while children are obtaining increasingly better scores as they progress from one grade to the next, average differences between grades are rather modest in the majority of the subtests. Most likely, most children are not improving by large amounts from grade to grade because they are getting promoted to higher grade levels without demonstrating mastery of concepts at lower grade levels, as per a national policy in the DRC. Developing strategies to identify and support children who are falling behind their same-grade peers would make an important difference in children's learning. Given the large number of children per classroom, effective and feasible strategies could rely upon group work or other forms of peer learning in which children who are performing at higher levels may provide scaffolding for peers performing at lower levels.

Other factors that may account for the small-to-modest gains children are making might be related to low exposure to the subject matter (for instance due to inefficient use of instruction time or low attendance rates); difficulties with grasping concepts in a new language (as French is for most children); and/or the use teaching strategies centered around rote learning that fail to promote knowledge that readily translates to new contexts, such as assessment situations. All these are areas that, if improved, could result in greater gains in children's skills over the course of each grade and throughout primary school.
Conclusion 2: The most remarkable finding is the considerable variation in performance between children in the same grades and subdivisions. This calls for a deeper understanding of the factors that might account for the wide gaps between low and high performing children who live in the same regions and attend the same grades.

While the typical child showed low levels of performance across subtests for EGRA and EGMA, there were large numbers of children who performed well above or well below average for a child in their grade. When setting benchmarks, it is important to focus on closing the performance gap between high-performing and low-performing children in the same grade levels.

Understanding the factors that may explain why some children are doing much better or worse than others is critical for developing strategies to benefit all children. Low repetition rates (i.e., only 9% of children reported repeating the previous grade), which likely reflect the national policy of automatic promotion, may explain some of the significant heterogeneity found within grades, in the sense that a large number of children is promoted to the next grade without having mastered the basic skills from the previous grade. This promotion policy makes it vital to identify children who are falling behind their same-grade peers in order to provide the supports they need to boost their performance before they are promoted.

The OPEQ team will continue working on unpacking the factors that might explain such variation in the coming months. While variability within subdivisions and grades was one of the most striking findings, it is also worth noting the comparatively high average performance of children from Kalemie and the low performance of children from Kasenga. A better understanding of the reasons that explain the relative success of Kalemie, and the low average performance in Kasenga, could be pivotal in informing efforts to improve academic attainment across the board.

Conclusion 3: There are two factors that could underpin the low performance across reading and math assessments: a) the challenge of transitioning from mother-tongue instruction to French instruction and b) the poor quality of instruction around basic math and reading skills.

These two types of issues, one a language issue, the other an instructional issue, may hinder their learning and academic performance across subjects. At this point, we do not have enough information to determine to which degree each of these factors contributes to the overall problem. Extremely low scores across all EGRA subtests, particularly for Initial Sound Identification, Knowledge of Graphemes, and Vocabulary, seem to indicate that the majority of children are not familiar with the French language. Extremely low scores on these tests could also indicate that the curriculum does not adequately focus on the foundational skills of reading or that teachers do not have the necessary skills to efficiently teach foundational reading skills.

Sequencing lessons so that foundational skills are taught and mastered before more complex skills is critical for children's learning process. For instance, children need to be aware and knowledgeable of the relationship between sounds and written symbols in order to read and write, and they need to correctly identify numbers and discriminate quantities in order to perform basic arithmetic operations. Learning is often cumulative; therefore, children who do not master the foundational reading and math skills when they are expected to do so, are likely to benefit the least from instruction on more complex skills and to continue falling behind their peers.
OPEQ is focusing on improving instructional quality through providing better in-service training and support to teachers; however, if any effort to improve children’s academic outcomes is to be effective, it should intentionally address the language barriers that most children are facing in school. Improving the supports provided for children's acquisition of French as a second language also seems key to improving children's academic success.

**Conclusion 4:** There were important differences between boys and girls, with boys outperforming girls in every reading and math subtest.

It is important to note that there are small but non-trivial gender differences in school enrollment (or possibly attendance) and in reading and math performance. While these differences are generally small they should not be overlooked. The gap between boys' and girls' performance appears to increase as children get older and, for academics in particular, consistently favors boys over girls, which could result in substantial differences by the time children reach high school. The consistent gender differences found in this baseline can be taken as a sign of systematic, yet poorly understood, differences in the learning opportunities offered to boys and girls in these regions of DRC. Intervention efforts should be tailored to the needs and strengths of girls and boys, if they are to reduce, and not inadvertently propagate or enlarge, the gap between genders.

**Conclusion 5:** Even though children showed improvements as they progressed through primary school, such improvements were rather modest. For example, children only gained an average of 1.5 words from one grade to the next in the Vocabulary subtest, and were able to solve 1 or 2 more Addition problems than children in the grade below. If one considers the low levels of performance in 2nd grade and the modest gains between grades, it is not surprising that even 4th graders were still struggling with the most basic reading and math skills.

**Limitations**

There are several important limitations to this report. First, the dearth of similar studies in DRC posed challenges to our ability to select culturally and psychometrically valid measures of children's outcomes and to contextualize the interpretation of our results. Regarding the assessment of math and reading in particular, there are no available norms for EGRA and EGMA that would allow for meaningful comparisons between our sample and other samples or populations. In addition, a formal method to summarize the performance of children across math or reading subtests is not available. The OPEQ team will continue to refine the measurement procedures and instruments based on analysis conducted with the baseline data.

Second, the OPEQ program and design and methods for the impact evaluation were being developed simultaneously during year 1 of the project. This means that not all measures and procedures utilized as part of the evaluation were fine-tuned to the needs of the program. The coordination between program and evaluation design will continue improving as the research and program mature.

Third, results presented in this report are entirely descriptive, which means that definite conclusions cannot be reached about the reasons that explain some of the patterns found. The pace at which data needs to be analyzed and fed back to program and partners prohibited conducting appropriate analysis.
to explore associations between children’s demographic characteristics and academic performance. Future reports will include progress made in this area.

Fourth, it is important to note that while efforts were made to select a representative sample of schools in Katanga province, it was necessary to exclude some schools. Therefore, our results are only generalizable to schools with similar profiles to the ones participating in OPEQ.

Finally, the language diversity in the sample posed additional challenges for data collectors who in some cases had to do real-time translations of instructions into local languages. Ideally, assessments are to be translated and back translated prior to data collection to ensure that meaning is preserved and administration instructions are standard.

In spite of these limitations, this report offers a rich picture of 2nd to 4th graders’ reading and math performance in six educational subdivisions in Katanga province, based on a rigorous and sound research design. Results from this report are expected to inform the continuous development of OPEQ, which aims to be increasingly responsive to the specific needs of boys and girls in these regions.

Moreover, this report provides a solid baseline to evaluate the impact of OPEQ in children’s learning after 1 and 2 years of implementation. The impact evaluation of OPEQ will contribute to the knowledge base about what works for improving children’s educational outcomes, not only in the eastern DRC but, to the extent that children in other regions face similar conditions, in the country as a whole.
References


Annex A. School selection criteria

<table>
<thead>
<tr>
<th>Community Primary School Selection Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>for OPEQ Implementation</td>
</tr>
<tr>
<td>1. Official MEPSP school with a registered number (public, subsidized, non-subsidized)</td>
</tr>
<tr>
<td>2. No other agency is supporting the school for similar OPEQ activities (private, local or international)</td>
</tr>
<tr>
<td>3. At least 4 classrooms / 4 teachers / 120 children</td>
</tr>
<tr>
<td>4. Minimum of 12 teachers in the cluster; maximum of 40 teachers in the cluster</td>
</tr>
<tr>
<td>5. 2 to 6 schools in close proximity to form a cluster (+/- 10 km OR 1 hours of walking to the central school of the cluster)</td>
</tr>
<tr>
<td>6. Accessible: cluster can be reached by motorbike with not more than one hour walking</td>
</tr>
<tr>
<td>7. Willingness of school and community to participate in OPEQ</td>
</tr>
<tr>
<td>8. Prioritize communities where there is IRC programming (Tuungane / Health)</td>
</tr>
<tr>
<td>9. In a secure zone (no movement of armed groups, main roads in the area are secure, and the zone has been approved by IRC’s security office)</td>
</tr>
</tbody>
</table>
Annex B. Further details on data analysis methods

To assess the magnitude of particular between-group differences, we used effect size comparisons, defining a “practically” significant effect size as 0.41 (Ferguson, 2009). Effect size comparisons were preferred over other types of significance testing (e.g. t-test, ANOVA, Chi Square) for two reasons: 1) effect size comparisons adjust for the nested structure of the data (i.e., children nested in clusters, clusters nested in subdivisions), and 2) effect sizes, unlike significance tests, are independent of sample size and are thus more accurate (Ferguson, 2009). Future analyses to examine association between variables will be conducted in a multi-level model framework to adjust for nesting of the data.

Variability in children’s reading and math outcomes was assessed using unconditional multilevel models (HLM, 6.06, Raundebush & Bryk, 2002) to determine the proportion of variability in that can be explained by differences between children, school clusters and subdivisions.

To create an aggregate measure of household economic status, an Exploratory Factor Analysis (EFA) with Oblique rotation was conducted using Mplus version 5.2 (Muthen & Muthen, 1998-2007). The WLSMV (weighted least squares with robust standard errors, mean and variance adjusted) estimator was employed to adjust for the nesting of children in subdivisions and to robustly and precisely estimate parameters in models that contain binary (i.e., yes/no) indicators (Muthen & Muthen, 1998-2007; Heck & Thomas, 2009).

In order to prevent multicollinearity problems arising from high correlations between household construction materials, only items indicating higher wealth (i.e., concrete floors, brick walls, and aluminum roofs) were included in the analysis. A total of 16 items were included (see table 1) but only 15 were retained based on EFA findings. The best fitting solution resulted in 2 factors, with 4 and 11 items each. An item about canoe ownership was excluded as it had very low loadings in both factors. A 2 factor solution means that, based on the correlations between answers given by participants to all questions, household wealth can be thought as consisting of two different underlying dimensions: Factor 1 and Factor 2. This means that the assets included in Factor 1 are more strongly associated with each other than they are with assets included in Factor 2.

Total scores were computed for each factor by adding the number of "yes" answers across all items belonging to the same factor. In order to reduce the loss of data due to missing information, children's scores were prorated by multiplying the total number of items in the factor (i.e., 4 or 11) by the proportion of "yes" answers out of the total number of answers provided by the child.

For illustration purposes, factor scores were split by the median and children were categorized as having a "high" or "low" score in each factor. "High" scores indicate the child total score is equal to, or above the median of the sample, and "Low" scores indicate the child score is below the median of the sample. Subdivisions were compared in terms of the percentages of children with high scores in both factors, low scores in both factors, and combinations of high and low scores in both factors, with the aim of exploring differences in the distribution of household wealth across the sample. These descriptive analyses are presented in the sample description section.

13 Cohen’s $d$ effect sizes with pooled standard deviations were employed (Cohen, 1998).
14 Details on EFA results are available upon request.
<table>
<thead>
<tr>
<th>Question</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factor 1 – Livestock/manual transport</td>
</tr>
<tr>
<td>1. Is there someone in your household who owns cattle?</td>
</tr>
<tr>
<td>2. Is there someone in your household who owns goats/sheep/pigs?</td>
</tr>
<tr>
<td>3. Is there someone in your house who owns chickens?</td>
</tr>
<tr>
<td>4. Does someone at your house own a bicycle?</td>
</tr>
<tr>
<td>Factor 2 – Electric Appliances/motor transport</td>
</tr>
<tr>
<td>5. Does someone at your house own a motorcycle?</td>
</tr>
<tr>
<td>6. At your house is there a radio?</td>
</tr>
<tr>
<td>7. At your house is there a telephone / cell phone?</td>
</tr>
<tr>
<td>8. At your house is there electricity?</td>
</tr>
<tr>
<td>9. At your house is there a television?</td>
</tr>
<tr>
<td>10. At your house is there a fridge / refrigerator?</td>
</tr>
<tr>
<td>11. At your house is there an indoor toilet?</td>
</tr>
<tr>
<td>12. At your house is there a: car, truck, 4x4?</td>
</tr>
<tr>
<td>13. Is the floor of your house made of concrete?</td>
</tr>
<tr>
<td>14. Are the walls of your house made of cured brick?</td>
</tr>
<tr>
<td>15. Is the roof of your house made of aluminum/tin?</td>
</tr>
</tbody>
</table>

Table 17. Questions included in the 2-factor solution for child-reported household wealth
Annex C. Guide for the interpretation of boxplots

Figure 1. Guide for the interpretation of boxplots

- **OUTLIER** More than 3/2 times of upper quartile
- **MAXIMUM** Greatest value, excluding outliers
- **UPPER QUARTILE** 25% of data greater than this value
- **MEDIAN** 50% of data is greater than this value; middle of dataset
- **LOWER QUARTILE** 25% of data less than this value
- **MINIMUM** Least value, excluding outliers
- **OUTLIER** Less than 3/2 times of lower quartile

15 Diagram taken from http://flowingdata.com