



Cohort 1 District and TTI Partner Schools Baseline Monitoring Report Volume 3: Early Grade Reading and Mathematics

Assessments (EGRA and EGMA)

November 2018

EARLY GRADE READING AND MATHEMATICS ASSESSMENT (EGRA AND EGMA)

Prepared for

The Tanoto Foundation

Jalan M.H. Thamrin No. 31 Jakarta Pusat 10230

Prepared by Myriad Research The Boulevard Office Tower UG Floor, Unit BI Jalan Fachrudin Raya No. 5 Tanah Abang Bukit Jakarta Pusat 10250

November 2018

Cover Photo: An enumerator conducting the Early Grade Reading Assessment with a grade 3 student

Table of Contents

Table of Contents	ii
List of Tables	iv
List of Figures	v
Chapter I. Introduction	I
I.I Background	I
I.2 Purposes of the Assessments	2
Chapter 2. Methodology	5
2.1 Study Location and Sample Size	5
2.2 Data Collection Implementation	6
2.3 Data Analysis	7
2.4 EGRA & EGMA Subtask Descriptions	8
Chapter 3. EGRA Results	
3.1 EGRA Student Profile	
3.2 Students' Reading Abilities	
3.3 EGRA Zero Scores	13
3.4 Overall EGRA Scores by Subtask	14
3.5 Letter Sound Identification	15
3.6 Word Reading	16
3.7 Non-word Reading	17
3.8 Oral Passage Reading & Comprehension	18
3.9 Listening Comprehension (% Correct)	20
3.10 Dictation (% Correct)	20
3.11 Students' Backgrounds vs. ORF & Comprehension	21

Chapt	er 4. EGMA Results	24
4. I	EGMA Student Profile	24
4.2	Overall EGMA Scores by Subtask	24
4.3	EGMA Zero Scores	25
4.4	Number Identification	26
4.5	Addition and Subtraction	27
4.6	Quantity Discrimination	29
4.7	Missing Numbers	30
4.8	Word Problems	31
4.9	Observed Student Calculation Methods	32
4.10	Students' Backgrounds vs. Basic Addition and Subtraction	33
Chapt	er 5. Conclusion	35

List of Tables

Table 2.1: Location of the Study	5
Table 2.2: Planned vs. Achieved Samples	6
Table 2.3: EGRA Subtask Description	8
Table 2.4: EGMA Subtask Description	9
Table 3.1: EGRA Student Profile	
Table 3.2: EGRA Zero Scores	14
Table 3.3: Overall EGRA Scores by Subtask	15
Table 3.4: Students' Backgrounds vs. ORF & Comprehension	22
Table 4.1: EGMA Respondent Profile	24
Table 4.2: Overall EGMA Performance	25
Table 4.3: Zero Score Percentage	26
Table 4.4: Student Calculation Methods	32
Table 4.5: Students' Backgrounds vs. Basic Addition and Subtraction	33

List of Figures

Figure 3.1: Overall Students' Reading Abilities	12
Figure 3.2: Overall Students' Reading Abilities	
Figure 3.3: Letter Identification Results Per Item	
Figure 3.4: Letter Identification Results Per Item	
Figure 3.5: Non-Word Results Per Item	
Figure 3.6: Oral Passage Reading	
Figure 3.7: Oral Passage Reading	
Figure 3.8: Correlation between Reading Comprehension and Reading Fluency	
Figure 3.9: Listening Comprehension Results Per Item	20
Figure 3.10: Dictation Results Per Item	21
Figure 4.1: Number Identification Results Per Item	27
Figure 4.2: Level I Addition Results Per Item	
Figure 4.3: Level I Subtraction Results Per Item	
Figure 4.4: Level 2 Addition Results Per Item	29
Figure 4.5: Quantity Discrimination Results Per Item	
Figure 4.6: Missing Number Results Per Item	31
Figure 4.7: Word Problem Results Per Item	32



Chapter 1. Introduction

Chapter I. Introduction

I.I Background

Having made impressive progress in widening access to basic education, Indonesia must now develop an education system that will better support the needs of a rapidly emerging economy. OECD (2015)¹ recommended the first priority for the Indonesian government is to raise the quality of education and ensure that all learners acquire the skills they need to succeed in life. The second priority is to widen participation, requiring a concerted effort to improve access for disadvantaged children.

According to UNICEF (2017)², there has been significant progress in Indonesia towards achieving universal basic education. Primary school enrollment rates reached 99 percent in 2015, and in 2016 some 27 million children attended primary school. Most children, according to the UNICEF data, do complete primary education. The drop-out rates during the first six years of education have fallen to under 1 percent. Yet, there are few differences between enrollment rates of girls and boys at the primary level, and overall little difference between urban and rural areas. The gap between the province at the lower end- Papua- and the best performing province- Aceh- is about 15 percentage points. According to UNICEF, the quality of learning remains a cause of concern as it leads to high repetition rates. One in 10 children has to repeat the first year of primary education to attain the required standards, and 5 percent repeat the second year. Although the situation has improved in recent years, only 81 percent of primary school teachers hold the minimum qualifications required by the government.

In line with the UNICEF analysis, the recent research that the RTI International conducted on the National Early Grade Reading Assessment (EGRA) and the Snapshot of School Management Effectiveness (SSME) across Indonesia in 2014, has revealed inconsistencies in terms of quality of basic education services across islands. For example, early grade students from Java-Bali significantly outperformed all other regions on oral reading fluency (ORF).

¹ OECD (2015), Education in Indonesia: Rising to the Challenge

² UNICEF (2017), The Children, the School Years

Sumatera was the second best, with around 12 words per minute less than students in Java-Bali. On the other hand, Maluku, Nusa Tenggara, and Papua showed the lowest ORF performance – which is below the 30% percentile nationally and approximately half the level of Java-Bali.

The above data shows the inequality of access to education and the quality of students, teachers, and schools across Indonesia. The Tanoto Foundation as a humanitarian organization fully realizes that children's access to quality basic education should be widely opened, and this is in line with the foundation founder's belief that every human being deserves to have a decent life. The foundation works in poverty alleviation in Indonesia through three pillars: education, empowerment, and enhancement of the quality of life. In the education sector, the foundation focuses on the issues of quality and access to education. In addressing the issue of quality of education, the Tanoto Foundation has been running a program called Pelita Pendidikan since 2010.

In the next few years, Pelita Pendidikan will be working in five provinces, namely: North Sumatera, Riau, Jambi, Central Java, and East Kalimantan, to improve the quality of primary and junior secondary education. The program will work with schools (SD, MI, SMP, MTs) in 14 districts in 2018 and plans to expand to 30 districts in 2019. It will also work in 10 teacher training institutions (TTIs) and their partner schools. In order to measure the impact of the program on student performance, the program intends to conduct a number of student assessments including an Early Grade Reading Assessment (EGRA) and Early Grade Mathematics Assessment (EGMA).

I.2 Purposes of the Assessments

The purposes of the assessments are as follows:

- To adapt an appropriate research design and instruments for implementing the EGRA and EGMA for the Pelita Pendidikan school partners.
- b. To conduct EGRA and EGMA according to the agreed upon and approved designs and methods to measure the level of students' literacy and numeracy skills among the Pelita Pendidikan school partners.

c. To analyze the students' literacy and numeracy skills and to benchmark the results with similar assessments at international and national levels.



Chapter 2. Methodology

Chapter 2. Methodology

2.1 Study Location and Sample Size

This study was conducted in 5 provinces of Indonesia, namely: North Sumatera, Riau, Jambi, Central Java, and East Kalimantan. The districts within each province in which the study was carried out are shown in Table 2.1. From this table, it can be seen that the study was conducted in 19 districts, which on average around 4 schools were sampled. Apart from schools managed by MOEC, schools managed by MORA were also sampled.

The EGRA and EGMA were each administered to different students, meaning that one student was only assessed with either the EGMA or the EGRA, based on a random selection. The target for each school was to assess 12 students for each EGRA and EGMA. Thus, a total of 24 students per school was targeted.

		# Sampled Schools		Total Schools	# Students	Total Number of	
Province	District	SD	MI	Sampled	EGRA: 3rd Grade	EGMA: 3rd Grade	Students
	Karo	3	1	4	48	48	96
North	Batu Bara	3	1	4	48	48	96
Sumatera	Pematang Siantar	3	1	4	48	48	96
	Medan	2	1	3	36	36	72
	Siak	3	1	4	48	48	96
Diau	Dumai	3	1	4	48	48	96
Kidu	Bengkalis	3	1	4	48	48	96
	Pekanbaru	5	2	7	84	84	168
	Batanghari	3	1	4	48	48	96
lambi	Tanjung Jabung Barat	3	1	4	48	48	96
Jannoi	Tanjung Jabung Timur	3	1	4	48	48	96
	Jambi	2	1	3	36	36	72
	Wonogiri	3	1	4	48	48	96
Control lovo	Kendal	3	1	4	48	48	96
Central Java	Surakarta	2	0	2	24	24	48
	Semarang	0	1	1	12	12	24
L	Kutai Kerta Negara	3	1	4	48	48	96
EdSL	Balikpapan	3	1	4	48	48	96
Kallingingu	Samarinda	2	1	3	36	36	72
	Total	52	19	71	852	852	1704

Table 2.1: Location of the Study

2.2 Data Collection Implementation

The data collection started on 17 September 2018 and ended on 28 September 2018. Each selected school was scheduled for a one-day data collection. Each day the enumerators would visit the schools and randomly select the third grade class in the school and then randomly selected the students within the selected classrooms. The selected students were assessed one by one in a different room provided by each school. In total, 71 schools were assessed, consisting of 52 primary schools (SD) and 19 Islamic primary schools (Madrasah Ibtidaiyah or MI).

The number of achieved samples compared to the planned samples can be seen in Table 2.2 below. The discrepancy in the number of students was due to the fact that not all schools had enough 3rd graders to meet the quota of 24 students per school.

Drowinco		EGRA	EGMA		
Province	Target Achievement		Target	Achievement	
Jambi	180	180	180	181	
Central Java	132	128	132	127	
East Kalimantan	132	132	132	132	
Riau	216	190	216	190	
North Sumatera	192	184	192	185	
Total	852	814	852	815	

Table 2.2: Planned vs. Achieved Samples

The step-by-step activities that each assessor team carried out in a school were as follows:

- 1. The following mechanism was applied in selecting the classroom(s):
 - a. If the school had more than one third grade classroom, and the total number of students in each classroom was more than 24, one third grade classroom was selected to achieve 24 students in a random manner.
 - b. If the school only had one third grade classroom, and the total number of students in the classroom was more than 24, then the students from this classroom would be randomly selected.
 - c. If the school only had one third grade classroom and the total number of students was less than 24, all the students would participate as samples.

- 2. The number of boys and girls were distributed equally, unless the student population in the school or class could not meet this gender-balanced criterion.
- 3. In classrooms with more than 24 students, a simple interval sampling based on seating position was used to sample the students.

Prior to the data collection, a 3-day enumerator training was conducted in each province, during the period of 10 to 14 of September 2018. The 3-day training session consisted of a 2-day classroom session, and a 1-day field pilot at a nearby school. The training was arranged in classical, grouped lecturing, interview paired practical, and round robin sessions. An IRR inter-rater reliability (IRR) test was conducted on the third day to assess the consistency among the enumerators.

The assessors themselves were recruited and selected from local universities, mainly from the Faculty of Education, majoring in Primary School Teacher- Indonesian Language Education (for EGRA assessors) and Mathematics Education (for EGMA assessors). The recruitment targeted students in the last two semesters of their study. Each team consisted of 3 assessors, in which one of them would be appointed as a team supervisor. Consequently, 60 EGRA and 60 EGMA assessors were recruited to conduct the study in the 5 provinces.

2.3 Data Analysis

Before the data was analyzed, data cleaning was first conducted. This process involved checking the data on a daily basis during the fieldwork. Logic checks and any missing data were addressed during this stage, so that everything went accordingly by the end of the data collection. This data check was done centrally in Jakarta, which was also part of the quality control process. Once all the data was collected, a final check was done to ensure the data sets were all consistent in structure before the analysis process.

2.4 EGRA & EGMA Subtask Descriptions

Prior to the data collection, Myriad along with Tanoto Foundation together discussed the subtasks that would be used in EGRA and EGMA. While various versions of EGRA had been conducted in Indonesia, it was the first for EGMA. As such, RTI International recommended to use the subtasks in the core of EGMA.

Each subtask description used in this study can be seen in Table 2.3 for EGRA and Table 2.4 for EGMA.

Subtasks	Skills	Description
Letter-sound identification	Alphabetic principle—letter- sound correspondence	Say the sound each letter makes, while looking at a printed page of 100 letters of the alphabet in random order, upper and lower case. Timed, one minute
Word reading	Alphabetic principle—letter- sound correspondence; and Fluency— automatic decoding	Read a number of words taken from the corpus, which are words that are often used. These words are arranged randomly.
Non word reading	Alphabetic principle—letter- sound correspondence; and Fluency— automatic decoding	Read a list of 50 non words printed on a page. Timed. Words were constructed from actual orthography, but were not real words. For example 'ipat', 'napum'. Timed, one minute.
Oral reading fluency	Fluency—automatic word reading in context	Read a grade-level-appropriate short story out loud from a printed page. Timed, one minute.
Reading comprehension	Comprehension	Orally respond to 5 questions that the assessor asks about the short story. Untimed.
Listening comprehension	Oral language comprehension and vocabulary	Listen to a story that the assessor reads out loud, then orally answer 3 questions about the story. Untimed.
Dictation	Verbal understanding; writing abilities; the alphabet process	Write a sentence that is said out loud by the evaluator. The sentence is read 3 times and the answer is evaluated based on the word appropriateness and language order.

Table 2.3: EGRA Subtask Description

Table 2.4: EGMA Subtask Description

Subtasks	Targeted Skills	Description
Number identification	The ability to identify written number symbols	State the names of numbers presented on a page with 20 numbers (i.e., items). The numbers had one, two or three digits. Timed (one minute).
Addition and Subtraction level I (basic facts)	Knowledge of and confidence with basic addition and subtraction	Solve addition/subtraction problems, with sums/differences below 20, without the aid of paper and pencil. The items ranged from problems with single digits only to problems that involved the bridging of the ten.18 Twenty items for each addition and subtraction subtask (40 total items). Timed (one minute).
Quantity discrimination (number comparison)	The ability to make judgments about differences by comparing quantities represented by numbers	Identify the larger of a pair of numbers. The number pairs used ranged from a pair of single-digit numbers to five pairs of double-digit numbers and four pairs of three-digit numbers. Ten items. Not timed.
Missing number (number patterns)	The ability to discern and complete number patterns	Determine the missing number in a pattern of four numbers. Patterns used included counting forward and backward by ones, by fives, by tens and by twos. Ten items. Not timed.
Addition and subtraction level 2	The ability to use and apply the procedural addition and subtraction knowledge assessed in the level 1 subtask (sometimes referred to as the 'basic facts') to solve more complicated addition and subtraction problems	Solve addition/subtraction problems by applying the basic addition and subtraction facts assessed in the level I subtask. Pupils were allowed to use any strategy that they wanted, including using the paper and pencil supplied by the assessor. The problems required adding or subtracting two-digit numbers involving bridging. Five items per addition and subtraction subtask. Not timed.
Word problems	The ability to interpret a situation (presented orally to the pupil), make a plan and solve the problem.	Solve problems presented orally using any strategy, including using the paper and pencil and/or counters supplied by the assessor. The numerical values involved in the problem were deliberately small to allow for the targeted skills to be assessed without confounding problems with calculation skills that might otherwise impede performance. Six items. Not timed.

The data collection was fully conducted with tablets, using the Tangerine software developed by RTI specifically for the administration of EGRA and EGMA. In addition to the tablets, several stimuli that should be exposed to early grade students were also used with a one-onone student assessment.



Chapter 3. EGRA Results

Chapter 3. EGRA Results

3.1 EGRA Student Profile

As planned, the students who were assessed with EGRA were balanced according to gender. Almost all of them had an average age range of 3rd graders. In terms of language used for communication, almost all the students used Bahasa Indonesia in daily life, except in Central Java. In Central Java, half of them used the local language. Moreover, the pre-school attendance rate was high, especially in Central Java.

Category	Indicator	All	Jambi	Central Java	East Kalimantan	Riau	North Sumatra
		n = 815	n = 180	n = 128	n = 132	n = 190	n = 185
Gender	Воу	52%	50%	51%	51%	55%	51%
	Girls	48%	50%	49%	49%	45%	49%
Age	Less than 8 years old	6%	2%	8%	3%	12%	8%
	8 – 9 years old (ref)	90%	90%	90%	93%	86%	90%
	More than 9 years old	4%	8%	2%	4%	2%	3%
Language	Indonesian	72%	71%	52%	78%	67%	89%
	Local language	28%	29%	48%	22%	33%	11%
Pre school attendance	Νο	9%	10%	2%	9%	12%	9%
	Yes	91%	90%	98%	91%	88%	91%

Table 3.1: EGRA Student Profile

3.2 Students' Reading Abilities

To understand the students' reading abilities, four reader categories classified by RTI International were applied. The categories were based on the composite measure of reading ability which was calculated as a combination of oral reading fluency (i.e. correct words per minute) and reading comprehension (correct answers out of those attempted). The four categories were:

- 1. Reading fluently with comprehension: obtained 80% of correct answers for reading comprehension questions, with a note that all of the texts were completely read.
- 2. Reading with comprehension: received 60% of correct answers for reading comprehension questions from all of the questions asked.

- 3. Reading with limited comprehension: had an understanding of less than 60% of the texts, with a note that the oral reading fluency was higher than 0.
- 4. Non-reader: possessed an oral reading fluency of 0.



Figure 3.1: Overall Students' Reading Abilities

Based on the results of the study, on average 60% of the students could mostly read fluently with comprehension, meaning that more than half of the students had good oral reading ability accompanied by a good understanding. This was approximately 5% higher than the national average conducted by RTI in 2014. However, the percentage of non-readers and student reading with limited comprehension was relatively similar with the national data.

Looking at the results in each province, it can be seen that high performers were found in East Kalimantan and Central Java, with student reading fluency with comprehension around 67-65%. In addition, no non-readers were found in East Kalimantan. Meanwhile, the other provinces had similar accomplishments with each other and slightly performed above the national data results. The national EGRA Reading ability was segregated by 4 regions, Sumatra (all province in Sumatra Island, Java-Bali (all province in Java island combined with Bali), Kalimantan-Sulawesi and MNP (Maluku, Nusa, and Papua which represented the eastern Indonesia). Figure 3.2 shows the reading ability among three of these region (MNP excluded), compared to the results of the TF Baseline segregated by Province.

Comparing with the results with National EGRA by region, each district was performing higher compared to each region that the province belonged to. For example, Central Java, compared to Java-Bali from the national EGRA show that



Figure 3.2: Overall Students' Reading Abilities

3.3 EGRA Zero Scores

With the current level of students' reading abilities, it was not surprising that zero scores in EGRA were low in all the subtasks presented. A zero score is the percentage of students who were unable to provide a single correct response on a given subtask. Table 3.2 shows the zero score percentage of each subtask across the provinces.

Table 3.2: EGRA Zero Scor

Area	Letter-sound identification (letters/min)	Word reading (words/min)	Non word reading (words/min)	Oral reading fluency (words/min)	Reading comprehension (%Correct)	Listening comprehension (%Correct)	Dictation (%Correct)
Overall	0.1%	1.2%	3.1%	1.6%	4.9%	3.9%	0.7%
Jambi	0.6%	1.1%	2.2%	2.2%	4.4%	5.6%	1.7%
Central Java	0.0%	2.3%	3.1%	1.6%	6.3%	2.3%	0.8%
East Kalimantan	0.0%	0.0%	0.8%	0.0%	2.3%	5.3%	0.0%
Riau	0.0%	1.1%	4.7%	2.1%	5.3%	5.3%	1.1%
North Sumatra	0.0%	1.6%	3.8%	1.6%	6.0%	1.1%	0.0%

3.4 Overall EGRA Scores by Subtask

The results of each subtask are shown in Table 3.3. The timed subtasks are shown in scores of correct items/minute, while un-timed subtasks are shown in the percentage of correct answers per total number of items asked. The overall scores are the combined scores of the five provinces, while the national scores are the results of the EGRA conducted by RTI in 2014. It is to be noted that the national EGRA did not include the word reading subtask, and the items asked in the dictation subtask differed slightly.

An overview of the results indicate that students were able to read and identify 81 letters per minute, which is around 5 letters more than the national average. Meanwhile, for word reading and oral fluency, students were reading at around 66 words/minute and 71 words/minute respectively. Comparing students' fluency with the national score, students were reading around 19 words/minute faster. Furthermore, fluency in reading non-words was also higher than the national level; in this case, students were reading 38 non-words/minute, 8 words/minute faster than the national level.

Moving on to the untimed subtasks, for reading comprehension, on average students were able to answer 75% of the questions asked to them, while for listening comprehension the average was 68% correct responses. Both reading and listening comprehension scores were above the national scores. Lastly for dictation, the average number of correct responses was at 69%, slightly lower than the national level.

Area	Letter-sound identification (letters/min)	Word reading (words/min)	Non word reading (non-words/min)	Oral reading fluency (words/min)	Reading comprehension (%Correct)	Listening comprehension (%Correct)	Dictation (%Correct)
Overall	80.99	66.87	37.79	71.03	75.14%	68.42%	69.09%
Jambi	88.48	68.42	38.45	70.03	74.8%	63.5%	69.0%
Central Java	81.65	66.55	38.75	72.19	78.0%	73.7%	74.8%
East Kalimantan	79.71	72.24	40.91	79.67	79.2%	64.1%	68.1%
Riau	73.65	61.31	35.04	66.48	73.8%	68.7%	67.4%
North Sumatra	81.69	67.45	37.07	69.69	71.9%	72.3%	67.6%
National	75.00	-	29.90	52.10	62.80%	53.70%	72.%

Table 3.3: Overall EGRA Scores by Subtask

3.5 Letter Sound Identification

Letter sound identification was the first most basic EGRA subtask. Conducting a deeper analysis in this subtask provided a clearer understanding of which letters were considered difficult for students in this study. All sampled students were shown 100 letters on paper, in both upper and lower case form, arranged in a random sequence. Then they were asked to identify as many of the letters as they could within one minute.

As seen in Figure 3.2, the results show that amongst the various letters read in this subtask, it was seen that there was a tendency for the letters v, q, and f to be more difficult to read compared with the other letters for students who attempted to read them. For example, between all of the students who attempted to read the letter v, about 30-35% of them read it incorrectly. The red line on the graph also shows that the number of students who were still reading within the timed one minute started to decline at around the 28th letter, and eventually around 26% of the total students were able to read all the 100 letters.



Figure 3.3: Letter Identification Results Per Item

3.6 Word Reading

The second subtask, was assessing student's ability to read randomly sequenced familiar words. As shown on Figure 3.2, the majority of students were able to compete reading all the 50 words within the timed minute. Those that did also were able to accurately read the words. it can be seen that some words tend to be more difficult to read such as "merah" and "bunga"





Readers who answered correctly - The average number of students who tried to read the word

3.7 Non-word Reading

Known to be one of the more challenging subtasks in the EGRA, the non-word reading subtask measured students' ability to decode words and blend letter sounds together to pronounce words. Each student was shown 50 non-words, mostly 4-5 letters in length. The students were given one minute to read as many non-words as possible. Although the non-words were not real words in the language of the assessment, they did follow the structure and orthography of the language. Using non-word decoding is helpful for noting a student's ability to apply phonetic knowledge to unfamiliar words.

Using the national EGRA as a benchmark, the overall results as previously shown suggest that each province performed higher than the national score, which was overall 37 non-words/minute compared to 29 non-words/minute. In addition, accuracy in decoding non-words showed that on average more than 80% of the students were able to decode correctly any non-word they attempted to read. The results shown in Figure 3.3 also reveal that around 33% of them were able to attempt all 50 non-words. Pupils' performance on the non-word decoding subtask was positively correlated with their performance on the letter sound knowledge subtask, as well as with the oral passage reading subtask.



Figure 3.5: Non-Word Results Per Item

3.8 Oral Passage Reading & Comprehension

From the results of the oral passage reading subtask, or also known as oral reading fluency (ORF), without considering students in the non-reader category, it can be seen that 90% of the students were able to accurately read the words that they attempted to read in the passage. However, a few words dipped slightly below 90%, namely the words "kebingungan" and "uangnya" as well as the name "Intan" as depicted in Figure 3.4. Also in that figure, it was found that around 73% of the students were able to read until the end of the paragraph. The percentage of reading mistakes decreased as the number of students who continued to read decreased. This reveals that students who were able to read until the end of the paragraph were those who were more fluent in reading and had higher accuracy.



Figure 3.6: Oral Passage Reading

After reading the passage, students were asked a maximum of five questions, depending on where they stopped reading by the end of the timed minute. Among those that attempted, on average, 82% answered correctly. The results suggest, slower readers who only asked the first question could not answer correctly.



Figure 3.7: Oral Passage Reading

Furthermore, by combining the ORF score with the reading comprehension, Figure 3.5 shows that the more fluent a student was in reading, the greater the possibility was that the student would understand what was being read. This figure depicts that a student who had an ORF score of 21-30 on average answered 50% correct from the questions asked related to the paragraph read. This percentage also increased as the ORF score increased.





3.9 Listening Comprehension (% Correct)

Looking at the results of individual items in the listening comprehension subtask, it appears that the first question regarding the location of Maya's new school was the hardest question for students to remember. Moreover, the question related with the reason why Maya was eager to learn, could only be answered by less than 70% of the students. On the other hand, more than 90% of the students were able to answer the question 'What are the teachers like in Maya's new school?'





3.I0Dictation (% Correct)

Reading and writing are complementary literacy skills – children learn to read by writing; they learn to write by reading. In this sub-test, students were asked to write a brief sentence (5 words) which was read to them three times in total. The subtask not only demanded the students to write the words correctly, but also took into account the proper usage of capital letters, a full stop, and spaces between the letters. For example, the use of a capital letter anywhere that is not at the start of the first word would be considered incorrect, even though the student managed to spell the word correctly.

The overall results convey that only 5% of the students were able to write the sentence without any mistakes. To further see where the students mostly failed in this subtask, Figure 3.7 shows the percentage of correct responses per item assessed. From this figure, it can be seen that using a capital letter and a full stop at the end of a sentence were still less likely to be applied by early grade students. However, almost all the students correctly wrote the sentence from left to right and were also accustomed to leaving a space between words.



Figure 3.10: Dictation Results Per Item

3.11 Students' Backgrounds vs. ORF & Comprehension

After discussing the EGRA results, it can be viewed how several student profiles influenced their reading abilities, by using ORF and reading comprehension as an overall indicator of students' reading abilities. As observed in Table 3.4, students who were attending preschool were able to read at a faster rate of around 12 more words per minute, and they also averaged higher comprehension. Similarly, students who mostly spoke bahasa Indonesia at home also read at a faster rate.

Category	Indicator	% students	Oral Reading Fluency (mean score)	Reading Comprehension (mean score)
Gender	Воу	52%	66.87*	71.92%*
	Girls	48%	76.46	77.92%
Age	Less than 8 years old	7%	54.54*	66.67%*
	8 – 9 years old (ref)	90%	73.66	76.34%
	More than 9 years old	3%	48.56*	51.72%*
Language	Indonesian	72%	73.26	74.94%
	Local language	28%	66.93*	74.51%
Pre school attendance	Νο	9%	60.40	66.39%
	Yes	91%	72.58*	75.64%*

Table 3.4: Students' Backgrounds vs. ORF & Comprehension



Chapter 4. EGMA Results

Chapter 4. EGMA Results

4.1 EGMA Student Profile

A total of 815 students across the five provinces were evaluated with the EGMA assessment. The ratio of boys and girls was kept as even as possible throughout the provinces. Their ages were mostly in the 8 to 9 year old range. A further look into their background revealed that around 90% of the students claimed to have attended pre-school. Similarly, with the EGRA student profile, around a third of the students said that they mostly used their local language at home, except for Central Java, which was more than half.

Category	Indicator	All	Jambi	Central Java	East Kalimantan	Riau	North Sumatra
		n = 815	n = 180	n = 128	n = 132	n = 190	n = 185
Gender	Воу	50%	46%	51%	50%	53%	51%
	Girls	50%	54%	49%	50%	47%	49%
Age	Less than 8 years old	7%	3%	6%	3%	13%	11%
	8 – 9 years old (ref)	89%	90%	91%	96%	84%	84%
	More than 9 years old	4%	7%	3%	1%	4%	5%
Language	Indonesian	70%	67%	44%	77%	67%	87%
	Local language	30%	33%	56%	23%	33%	13%
Pre school	Νο	10%	9%	4%	10%	15%	9%
attendance	Yes	90%	91%	96%	90%	85%	91%

Table 4.1: EGMA Respondent Profile

4.2 Overall EGMA Scores by Subtask

The average scores of each subtask, in total and by province, can be seen in Table 4.2. Among the timed sub-tasks (number identification, level I addition and subtraction) two types of scores were calculated, first the % of correct/attempted answers, and second the average number of correct answers/minute. Meanwhile, all the other subtasks are shown by % of correct answers.

In general, an overall trend across the subtasks is evident. This can also be surmised across the 5 provinces. Students performed better on number identification, quantity discrimination, and the level I addition and subtraction. Their performance in these subtasks reveals that they learned the basic mathematics skills quite well. However, when students were faced with

more conceptual subtasks, namely missing numbers, level 2 addition and subtraction, and word problems, a decrease was evident in their performance as they seemed to struggle more in these subtasks. For example, looking at the overall province scores, it is depicted that even though students were able to answer the level I addition and subtraction well (84.3% for addition and 80.4% for subtraction), the level 2 addition and subtraction scores dropped by 20% and 35.5% respectively.

Comparing the average scores of the five provinces, it is displayed that students from Central Java seemed to perform higher than the other provinces on most of the subtasks. Their higher performance can clearly be seen in number identification, missing numbers, and level 2 addition and subtraction. On the other hand, Jambi was found to have the lowest performance on most of the subtasks, for example, 70.5% for number identification, 56% for level 2 addition, and 38.9% for level 2 subtraction. Meanwhile, the other provinces performed similarly.

Province	Number identification		Quantity discrimination	Missing Number	Addition Level 1		Addition Level 2	Subtraction Level 1		Subtraction Level 2	Word Problem
	% Correct/ Attempted	# Correct/ minute	% Correct	% Correct	% Correct/ Attempted	# Correct/ minute	% Correct	% Correct/ Attempted	# Correct/ minute	% Correct	% Correct/ Attempted
Overall	76.4%	28.11	91.5%	54.6%	84.3%	11.18	64%	80.4%	8.68	44.9%	38.6%
Jambi	70.5%	25.83	90.9%	49.9%	84.5%	11.12	56%	80.3%	8.88	38.9%	37.2%
Sumatera Utara	74.1%	29.98	91.5%	51.3%	84.6%	11.84	67%	78.8%	8.58	45.0%	43.6%
Riau	73.5%	25.08	89.1%	51.6%	84.3%	10.10	60%	80.5%	8.07	43.6%	38.3%
Jawa Tengah	92.7%	33.21	93.4%	65.7%	82.5%	11.75	73%	80.6%	9.38	56.9%	39.2%
Kalimantan Timur	75.9%	28.09	93.7%	59.4%	85.6%	11.34	67%	82.5%	8.78	43.6%	33.3%

Table 4.2: Overall EGMA Performance

4.3 EGMA Zero Scores

Like EGRA, another way to gauge the performance of students in the assessments was to calculate the percentage of students who could not answer a single item on a specific subtask, also known as the percentage of zero scores. Looking at the zero score results in Table 4.3, it is not surprising that the addition & subtraction of level 2 had the highest incidence of zero scores.

Area	Number identification	Quantity discrimination	Missing number	Addition (Level I)	Addition (Level 2)	Subtraction (Level I)	Subtraction (Level 2)	Word problem
Overall	0%	0%	1%	١%	10%	3%	21%	5%
Jambi	0.0%	0.0%	0.6%	0.0%	12.7%	1.1%	23.2%	2.8%
Central Java	0.0%	0.8%	1.6%	1.6%	7.8%	4.7%	13.3%	6.3%
East Kalimantan	0.0%	0.0%	0.0%	0.8%	3.0%	1.5%	18.2%	9.1%
Riau	0.0%	1.1%	1.1%	1.6%	14.7%	3.7%	25.3%	3.7%
North Sumatra	0.0%	0.0%	1.1%	0.0%	7.6%	5.9%	22.2%	3.8%

Table 4.3: Zero Score Percentage

4.4 Number Identification

The Number Identification subtest was timed (60 seconds) with no stop rules; and it consisted of 20 items that increased in difficulty. The first three items of the subtest included singledigit numbers. The next 12 items consisted of two-digit numbers from 10 to 99, and the last five items were three-digit numbers from 100 to 999. Students were asked to say each number aloud.

Because the positions of these numerals made a difference in the interpretation of the numerosity, understanding the place value (e.g., the value of a "5" in the ones and tens place means 5 and 50, respectively) was essential in the conceptual understanding of the number values. Therefore, the Number Identification subtest consisted of both single- and multi-digit items, and correct answers had to reflect the place value in multi-digit numerals (e.g., two-two is not a correct answer to 22, but twenty-two is). With this in mind, the results of the number identification can be seen in Figure 4.1.

Almost all the students had no problems with single digit numbers. However, when they were confronted with two or three digit numbers, roughly around a third of the students were not able to answer correctly. This most likely meant that they were pronouncing multi-digit numbers not reflecting the place value. For example, the number 22 was pronounced as *dua-dua*, instead of *dua puluh dua*. Nevertheless, it was interesting to see that when the second digit was 0, their automaticity to pronounce the number as thirty (tiga puluh) was high. This

might suggest that a student's ability to identify the multi-digit numbers was also a matter of habit in pronouncing multi-digit numbers.



Figure 4.1: Number Identification Results Per Item

4.5 Addition and Subtraction

The level I addition and subtraction questions involved numbers in a low number range, and it was expected that children could perform these calculations mentally. Learners were not given access to counters and/or paper and pencils for the questions at this level. The test administrators were asked to record whether or not the learners used their fingers. Both the level I addition & subtraction subtask consisted of 20 items, and it was timed for I minute.

Figure 4.2 shows the overall results of each individual item of the level I addition subtask as percentages of correct/attempted answers. This figure also shows how many students attempted the item, shown by the red dotted line. Students performed quite well in terms of correct/attempted answers, meaning that the majority of students who attempted the item got it correct. However, looking at the decline in the red dotted line, it seems that their capability in solving these single digit addition problems can still be improved.

Comparing the results of level I addition to level I subtraction, it is quite evident that students struggled more with subtraction, especially subtraction that involved bridging to ten. Their average % correct/attempted can be seen above 80% on the first seven items that does not involve bridging to ten, and then it started to decline onwards as the items got more complex.

A steeper decline on the red dotted line also signified that students required extra effort in solving these subtraction items, as they needed more time to calculate each item.



Figure 4.2: Level | Addition Results Per Item





The level 2 addition and subtraction questions involved numbers in a higher number range, and it was expected that some of these children would have to use tools other than mental arithmetic to solve them. Learners were provided with paper and a pencil, which they were allowed to use to solve these problems. It was made clear to the learners that they were allowed to use these tools but that they did not have to. The results are shown in Figure 4.4, as the average percentage of correct items out of the total items in each subtask.

In the level 2 addition, as the items involved higher numbers, the number of students who were able to answer correctly also was not as high as the level 1 items. It started with 77.3% on the first item, and eventually only 46.3% on the last item. Meanwhile, looking at the subtraction level 2, again it was obvious that more than half of the students still needed to improve their subtraction ability, especially when subtraction involved bridging to ten. When bridging to ten was not involved, around half and slightly more could answer the questions, but once it did involve bridging to ten, less than half could answer correctly, which can be viewed in item 2 (39.8%) and item 5 (21.2%).



Figure 4.4: Level 2 Addition Results Per Item

4.6 Quantity Discrimination

The quantity discrimination subtask in EGMA measured pupils' ability to make judgements about differences by comparing quantities, as represented by numbers. The subtask measured the pupils' sense of magnitude. Being able to compare numbers/quantities is a foundational mathematical skill that is critical to effective and efficient problem-solving strategies. For example, being able to compare numbers/quantities is important when estimating the reasonableness of answers to problems. In the early school years, this means developing an awareness that addition results in a larger number, subtraction produces an answer that is smaller than at least one of the original numbers, multiplication can result in answers that are larger than the addition of the same numbers, and so on.

The results of this subtask, as depicted in Figure 4.4, reveal that more than 80% of the students were able to comprehend how big a number/quantity was, and could compare numbers/quantities. Even so, the results showed that there was a 10% decline from the first six items which involved single and double digit numbers, to the last 4 items which consisted of three digit items.



Figure 4.5: Quantity Discrimination Results Per Item

4.7 Missing Numbers

One of the most challenging subtasks was the missing number subtask. Solving the missing number problems in the EGMA subtask involved studying the evidence available and using this to determine the step size of the pattern, as well as whether the pattern was increasing or decreasing, and then determining the missing number by extending the existing pattern.

Looking at the items presented in the subtask, the first 4 items were the easiest for students to complete, with more than 70% able to answer correctly. But accuracy dropped drastically for the next couple of items. The lowest percentages of students responding correctly can be seen in item 7 (17.1%), which involved a missing number in the middle of a descending

pattern, and the last item (7.4%) which was a missing number in the middle of an ascending pattern of 5 additions.



Figure 4.6: Missing Number Results Per Item

4.8 Word Problems

A word problem was the final subtask which consisted of 5 items. Each item only involved a very small numerical value, as this subtask was designed to assess more on students' critical and conceptual problem solving skills and less on their arithmetic skills. Figure 4.6 displays the results of this subtask as the percentage of correct responses per item. Students were provided with counters (manipulatives) as well as paper and pencil, which they were allowed to use to solve these problems. It was made clear to the learners that they were allowed to use these tools but that they did not have to. The test administrators were asked to record whether or not the learners used their fingers, the counters, and/or the paper and pencil.

The first item was the easiest and was also correctly answered by 90% of the students. The second item was also quite as straight forward as the first, though involving subtraction, and was answered correctly by 60% of the students. These first two items were presented in a more traditional format of $\sqrt{+\sqrt{=}}$ [?] and/or $\sqrt{-\sqrt{=}}$ [?]. When presented not in the traditional format, such as the third and fourth items, the percentages declined drastically. On average, only around 20% were able to answer the last four items.

Figure 4.7: Word Problem Results Per Item



4.9 Observed Student Calculation Methods

Observations of how students calculated the items in the level 2 additions and subtractions can be seen in Figure 4.7. The level 2 additions and subtractions show that most of the students utilized a pencil and paper to write and solve the problems. Meanwhile, it can be viewed that for the word problems, students did not utilize the counter as much, and most of them were found to just use their fingers.

Observed student's calculation method	Level 2 Additions	Level 2 Subtractions	Word Problems
Using fingers	63.29%	63.90%	51.53%
Using paper & pencil	74.78% 74.43%		30.43%
Using both finger and Paper & pencil	46.85%	47.39%	18.16%
Use counter	-	-	23.93%

Table 4.4: Student Calculation Methods

4.10 Students' Backgrounds vs. Basic Addition and Subtraction

Using the addition & subtraction level I to see how math performance differed based on the respondents' profiles, as described in Table 4.5, age and experience in attending pre-school were two indicators that were significantly different. Younger students completed an average of 2 less items for addition and I less item for subtraction per minute. Students who did not attend preschool also completed less subtractions per minute.

Category	Indicator	% students	Addition level I # Correct/ minute (mean score)	Subtraction Level I # Correct/ minute (mean score)
Gender	Воу	50%	11.14	8.55
	Girls	50%	11.22	8.82
Age	Less than 8 years old	7%	9.41*	7.49*
	8 – 9 years old (ref)	89%	11.31	8.82
	More than 9 years old	4%	11.57	7.81
Language	Indonesian	70%	11.31	8.82
	Local language	30%	9.41	7.49
Pre school attendance	Νο	10%	10.59	7.55
	Yes	90%	11.24	8.81*

Table 4.5: Students' Backgrounds vs. Basic Addition and Subtraction



Chapter 5. Conclusion

Chapter 5. Conclusion

The results of the EGRA in the five provinces convey that most of these students were quite fluent in reading. Across the five provinces, students read an average of 71 words/minute, and 60% could read with a high level of fluency and comprehension, achieving 80% or more in their comprehension scores. The students were equipped with the foundation of pre-reading and early reading skills on which to build fluency and comprehension. The EGRA study also showed that when children had a grasp of some of the basic 'building blocks' in learning to read, such as an understanding of letter sounds and the ability to decode or 'sound out' new words, they were more likely to be able to read fluently (e.g., quickly) and to understand what they read. Most students demonstrated these basic skills in the EGRA. Therefore, it was not surprising that most of the children in the EGRA study were able to read fluently and with comprehension. Even so, there is certainly still room for improvement in reading fluency and comprehension.

In addition, even though the data depicts that all the provinces performed above the national average, there were still some performance differences amongst the five provinces. Comparing the EGRA results by province, the students with the highest reading performance were from East Kalimantan and Central Java. Meanwhile, students in North Sumatra had the lowest performance.

The EGRA results also showed that students who mentioned that they had attended preschool could read more fluently compared to those who did not enroll in preschool. Interestingly, even though the overall results indicate that the use of a local language daily at home lowers the reading performance, Central Java, where half of the sampled students claimed to use a local language daily at home, still performed better than the other provinces.

The overall EGMA results reveal that while students performed well with some of the procedural mathematics skills, their conceptual understanding needs to be strengthened. Students performed best on single-digit items that required little critical or conceptual thinking to establish and complete the problems or patterns, which can be seen in the results

for the level I addition and subtraction. But when faced with problems that involved a more conceptual understanding, namely a missing number, level 2 addition and subtraction, and the word problem tasks, the students' performance indicates that they still struggle with such problems.

Comparing the mathematic performance across the five provinces, Central Java was found to have the highest score in almost all the subtasks, outperforming the other provinces more than 10% in the more challenging subtasks, such as level 2 addition and subtraction. On the other hand, students in Jambi seemed to perform lower compared to the other provinces.



Tanoto Foundation