

# Comprehensive Reading and Mathematics Assessment Tool (CERMAT)



Niken Rarasati

Goldy Dharmawan

Arya Swarnata

Anisah H. Zulfa

Delbert Lim

**SMERU TECHNICAL REPORT**

**Comprehensive Reading and Mathematics  
Assessment Tool (CERMAT)**

Niken Rarasati

Goldy Dharmawan

Arya Swarnata

Anisah H. Zulfa

Delbert Lim

Reviewed by  
Thomas Coen and Daniel Suryadarma

**Editor**

Alia An Nadhiva

**The SMERU Research Institute**

**March 2020**

# Comprehensive Reading and Mathematics Assessment Tool (CERMAT)

Authors: Niken Rarasati, Goldy Dharmawan, Arya Swarnata, Anisah H. Zulfa, Delbert Lim

Reviewed by Thomas Coen and Daniel Suryadarma

Editor: Alia An Nadhiva

Cover photo: Goldy Dharmawan

The SMERU Research Institute Cataloguing in Publication Data

Niken Rarasati

Comprehensive Reading and Mathematics Assessment Tool (CERMAT) / Niken Rarasati, et al; Editor Alia An Nadhiva.

--Jakarta: Smeru Research Institute, 2020,

--67 p.; 27 cm.

ISBN 9786237492283

ISBN 9786237492276 [PDF]

1. Literacy 2. Numeracy 3. Learning assessment  
I. Title

370.7-ddc 23

Published by:  
The SMERU Research Institute  
Jl. Cikini Raya No.10A  
Jakarta 10330  
Indonesia

First published March 2020



This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

SMERU's content may be copied or distributed for noncommercial use provided that it is appropriately attributed to The SMERU Research Institute. In the absence of institutional arrangements, online content may be published on a website a link to SMERU's

The findings, views, and interpretations published in this report are those of the authors and should not be attributed to an agencies providing financial support to The SMERU Research Institute.

A significant part of the research in this publication uses interviews and focus group discussions. All recorded information is stored at the SMERU office.

For further information on SMERU (23-31936336 (phone), 620-31930850 (fax), or contact smeru@smeru.or.id (e-mail); or visit [www.smeru.or.id](http://www.smeru.or.id).

## ACKNOWLEDGEMENTS

The development of the Comprehensive Reading and Mathematics Assessment Tool (CERMAT) is part of RISE Programme in Indonesia's research project funded by the UK Department for International Development, Australian Department of Trade, and Bill and Melinda Gates Foundation. We would like to thank our partner, KIAT Guru, especially Sharon Lumbanraja and Indah Ayu Prameswari for their generosity in providing us the first stage pilot data and items so that we did not have to start this project from scratch. We also like to express our appreciation to INOVASI for providing us their resources and expertise regarding numeracy and literacy. Our special thanks go to Ariyadi Wijaya from Yogyakarta State University, Budi Poniam, Dhitta Puti Sarasvati, Desyarti Safarini Deshinta Argaswari, and Namirah Fatmanissa from Faculty of Education, Sampoerna University for their valuable inputs in our numeracy framework and items, and Dwi Firli Azhari from Singapore, Sakota as our literacy specialist.

# ABSTRACT

## Comprehensive Reading and Mathematics Assessment Tool (CERMAT)

Niken Rarasati, Goldy Dharmawan, Arya Swarnata, Anisah H. Zulfa, and Delbert Lim

RISE Programme in Indonesia seeks to learn whether the education policy reforms carried out in Indonesia over the past twenty years have been successful in improving learning outcomes. To evaluate how well particular education policies help schools conduct better learning, we developed Comprehensive Reading and Mathematics Assessment Tool (CERMAT) student learning assessment (SLA) tool that can assess specifically the mathematics skills of students in Grades 3-4. CERMAT uses a framework based on the research for cognitive domains; stages of numeracy development that has been adapted into Indonesian literacy context; and Indonesian curricula. We use the Rasch model to evaluate the quality of the psychometric properties of CERMAT. After the instrument was piloted and underwent two revision cycles, it finally reached a sufficient reliability score and contained items with a wide range of difficulty levels. Thus, CERMAT is sensitive enough to detect an increase in student abilities. There are two methods for administering CERMAT: (i) individual and oral tests for students in Grades 3 and (ii) classical and written tests for students in Grades 4.

Keywords: literacy, numeracy, learning assessment

# TABLE OF CONTENTS

|                                                            |     |
|------------------------------------------------------------|-----|
| ACKNOWLEDGEMENTS                                           | i   |
| ABSTRACT                                                   | ii  |
| TABLE OF CONTENTS                                          | iii |
| LIST OF TABLES                                             | iv  |
| LIST OF FIGURES                                            | iv  |
| LIST OF APPENDICES                                         | iv  |
| LIST OF ABBREVIATIONS                                      | v   |
| I. INTRODUCTION                                            | 1   |
| 1.1 Background                                             | 1   |
| 1.2 Theoretical Background                                 | 2   |
| II. TEST DESIGN                                            | 10  |
| 2.1 C E R M A T ' s Cognitive and Content Domain Framework | 10  |
| 2.2 Test Format and Scope                                  | 14  |
| 2.3 Test Scoring                                           | 15  |
| III. DEVELOPMENT OF THE TEST                               | 16  |
| 3.1 School Index for Sample Selection                      | 16  |
| 3.2 Item Development                                       | 17  |
| IV. PSYCHOMETRIC PROPERTIES OF THE TEST                    | 20  |
| 4.1 Test Validity                                          | 20  |
| 4.2 Test Reliability                                       | 21  |
| LIST OF REFERENCES                                         | 23  |
| APPENDICES                                                 | 24  |
| GLOSSARY                                                   | 55  |

## LIST OF TABLES

|                                                                          |    |
|--------------------------------------------------------------------------|----|
| Table 1. Bloom's Dimensions and Its Equivalent to TIMSS and PIRLS        | 3  |
| Table 2. Text Level Presented in the SLA Test Booklets                   | 6  |
| Table 3. Cognitive Domain Framework for Numeracy Test                    | 10 |
| Table 4. Content Domain Framework for Numeracy Test                      | 11 |
| Table 5. Cognitive Framework for Literacy Test                           | 12 |
| Table 6. Allocation of Item Presentation in Numeracy Assessment by Grade | 13 |
| Table 7. Allocation of Item Presentation in Literacy Assessment by Grade | 14 |
| Table 8. Number of Items by Grade and Instrument Type                    | 14 |
| Table 9. Mean of the Index in Each Quality Group                         | 16 |
| Table 10. Outfit and Infit Indices                                       | 21 |
| Table 11. Item Person Reliability Scores and Separation Indices          | 22 |

## LIST OF FIGURES

|                                                                           |    |
|---------------------------------------------------------------------------|----|
| Figure 1. Number domain staircase                                         | 4  |
| Figure 2. Levels in measurement                                           | 5  |
| Figure 3. Item person map for Indonesian Grade 1                          | 8  |
| Figure 4. SLA scores by school quality category                           | 17 |
| Figure 5. Wright Map of Indonesian Grade 3 mathematics items and students | 18 |
| Figure 6. Wright Map of Indonesian Grade 4 mathematics items and students | 19 |

## LIST OF APPENDICES

|                                              |    |
|----------------------------------------------|----|
| Appendix1. Test Administration Manual        | 25 |
| Appendix2. Fountas and Pinnel Text Levelling | 31 |
| Appendix3. Numeracy Staircase                | 35 |
| Appendix4. Wright Map                        | 38 |

## LIST OF ABBREVIATIONS

|           |                                            |                                                                                                                    |
|-----------|--------------------------------------------|--------------------------------------------------------------------------------------------------------------------|
| CERMAT    |                                            | Comprehensive Reading and Mathematics Assessment Tool                                                              |
| COMPASS   |                                            | Computer Adaptive Placement Assessment and Support System                                                          |
| DAPODIK   | <i>Data Pokok Pendidikan</i>               | The Ministry of Education Data Centre                                                                              |
| EGMA      |                                            | Early Grade Mathematics Assessment                                                                                 |
| EGRA      |                                            | Early Grade Reading Assessment                                                                                     |
| GRM       |                                            | Graded Response Model                                                                                              |
| IB        |                                            | International Baccalaureate                                                                                        |
| IEA       |                                            | International Association for the Evaluation of Educational Achievement                                            |
| INAP      |                                            | Indonesian National Assessment Programme (the Indonesian version is or <i>Asesmen Kompetensi Siswa Indonesia</i> ) |
| INOVASI   |                                            | Innovation for Indonesia's School Child                                                                            |
| KIAT Guru | <i>Kinerja dan Akuntabilitas Guru</i>      | Teacher Performance and Accountability                                                                             |
| KTSP      | <i>Kurikulum Tingkat Satuan Pendidikan</i> | 2006 Curriculum                                                                                                    |
| MoEC      |                                            | Ministry of Education and Culture                                                                                  |
| NAEP      |                                            | National Assessment of Educational Progress                                                                        |
| PCA       |                                            | Principal Component Analysis                                                                                       |
| PCM       |                                            | Partial Credit Model                                                                                               |
| PIRLS     |                                            | Progress in International Reading Literacy Study                                                                   |
| PISA      |                                            | Programme for International Student Assessment                                                                     |
| PRIORITAS |                                            | Prioritizing Reform, Innovation and Opportunities for R Teachers, Administrators, and Student                      |
| RISE      |                                            | Research on Improving Systems of Education                                                                         |
| SLA       |                                            | student learning assessment                                                                                        |
| TIMSS     |                                            | Trends in International Mathematics and Science Study                                                              |
| TNP2K     |                                            | Team for the Acceleration of Poverty Reduction                                                                     |

|       |                                          |                                                        |
|-------|------------------------------------------|--------------------------------------------------------|
| UN    | <i>Ujian Nasional</i>                    | National Examination                                   |
| USAID |                                          | The United States Agency for International Development |
| USBN  | <i>Ujian Sekolah Berstandar Nasional</i> | National-Based School Examination                      |

# I. INTRODUCTION

## 1.1 Background

Research on Improving Systems of Education Programme in Indonesia focuses on (i) evaluating the impact of education policy reform from the national level, especially those relating to the improvement of teacher and teaching quality; and (ii) understanding how education policy reform initiated in innovative districts improve learning in respective districts and spread to other districts. With these two focus areas, we aim to learn whether policies in question successfully improve the country's student learning outcomes.

In Indonesia, the student learning outcomes are assessed using both high-stakes tests. The high-stakes test, known as the National Examination (UN), is carried out nationally at the end of junior and senior high school year (Grades 9 and 12, respectively). The UN contains questions related to topics covered in the curriculum, focusing on the knowledge taught in school rather than their ability to apply the concept in everyday life or think of it critically. Such examination is a "test of knowledge" rather than focus on the learning process. Only students who have passed the minimum standard of the curriculum. Consequently, the test is dominated by items with low and medium difficulty levels. With only a few difficult items, this test is ineffective to capture a variety of abilities to detect score increase sensitively.

The low-stakes test is the Indonesian National Assessment Programme (INAP); it is used to map, diagnose, and evaluate education progress in each province in Indonesia by looking at the performance of sampled Grade 4 students. The INAP, administered by the Indonesian Ministry of Education and Culture (MoEC) and launched in 2016, is the only nationally standardized assessment tool that assesses middle and higher order thinking skills. However, the tool only contains items for Grade 4 students. A test designed for Grade 8 and 10 students is still under development.

In addition to the assessment tools administered by the MoEC, there are other student learning assessment (SLA) tools adapted or developed by several development projects. In early 2016, K-12 Guru<sup>2</sup> and INOVASI<sup>3</sup> developed a numeracy and literacy assessment tool that covers materials for students in Grades 1 to 5. The tool includes cognitive domain TIMSS (Trends in International Mathematics and Science Study) and PIRLS (Progress in International Reading Study). The development of the two projects' outputs in 256 schools is being implemented in the provinces of West Java, Special Region of Yogyakarta, Banten, and West Nusa Tenggara. The tool was followed by expert review on psychometric properties and content of the items. Since RISE Programme in Indonesia focuses its research on basic education (Grades 1 to 9), we developed Comprehensive Reading and Mathematics Assessment (CERMA) by adopting the KIAT Guru and INOVASI's SLA for students in Grades 6 to 9.

---

<sup>1</sup>RISE Programme in Indonesia defines innovative districts as cities or regencies with innovative policies specifically aimed at improving their student learning outcomes. The districts are selected based on their positive trends in the National Examination (UN) score.

<sup>2</sup>KIAT Guru (an Indonesian abbreviation of Teacher Performance and Accountability) is a World Bank and the National Team for the Acceleration of Poverty Reduction's (TNP2K's) initiative.

<sup>3</sup>INOVASI (Innovation for Indonesia's School Governments of Indonesia) is a partnership between the Australian Government and the Indonesian Government to improve student learning outcomes in literacy and numeracy in primary schools and districts across Indonesia.

CERMAT aims to evaluate how well the particular education reforms or policies help schools in providing their students with skills that could prepare them to live in this fast world.

Accordingly, the development of CERMAT considers the following aspects:

- a) It assesses not only procedural knowledge, but also the abilities to apply, analyze and evaluate.
- b) It has items that cover a wide range of student
- c) It is sufficiently sensitive to capture student
- d) The test has vertical continuity across grade levels, meaning that the ability of other students in his/her grade.

This technical report outlines the process of developing CERMAT, including the manual on how to implement and score the tests.

## 1.2 Theoretical Background

As stated in the previous section, there are four aspects to consider in the development process of CERMAT. An overview of the theory that we use as the basis of our instrument adaptation to meet the four aspects is presented in this section.

### 1.2.1 Cognitive Domain

In assessing the cognitive dimension of student literacy and literacy skills, CERMAT refers to the revised Bloom's Taxonomy (Krathwohl, 2002). This new classification makes the taxonomy easier to operationalize into a learning outcome assessment. The taxonomy provides six levels of cognitive skills, which starts from retrieving the relevant knowledge that a student has remembered to synthesizing the elements together to create a new original idea or product. Cognitive skills from the taxonomy that are relevant for elementary and junior high school students have been adapted by the International Association for the Evaluation of Educational Achievement (TIMSS) and PIRLS framework (Mullis and Martin, 2013). The cognitive domain framework of CERMAT also adopted the TIMSS and PIRLS framework. For comparability and practical reasons, we use the TIMSS and PIRLS cognitive domain framework in CERMAT respectively. The cognitive process dimension of Bloom's Taxonomy and PIRLS framework (Mullis and Martin, 2013) is presented in the following table.

**Table 1. Bloom's Cognitive Process Dimension and Its Equivalent to TIMSS and PIRLS**

| Bloom's Cognitive Process Taxonomy                                                                                                                                                                                                        | Definition                                                                                                                            | TIMSS Framework for Numeracy  | PIRLS Framework for Literacy                        |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------|-------------------------------|-----------------------------------------------------|
| Remember<br><ul style="list-style-type: none"> <li>&lt; Recognizing</li> <li>&lt; Recalling</li> </ul>                                                                                                                                    | Retrieving relevant knowledge from long-term memory                                                                                   | Knowing                       | Focus on and retrieve explicitly stated information |
| Understand<br><ul style="list-style-type: none"> <li>&lt; Interpreting</li> <li>&lt; Exemplifying</li> <li>&lt; Classifying</li> <li>&lt; Summarizing</li> <li>&lt; Inferring</li> <li>&lt; Comparing</li> <li>&lt; Explaining</li> </ul> | Determining the meaning of instructional messages, including oral, written, and graphic communication                                 | Knowing                       | Make straightforward inferences                     |
| Apply<br><ul style="list-style-type: none"> <li>&lt; Executing</li> <li>&lt; Implementing</li> </ul>                                                                                                                                      | Carrying out or using a procedure in a given situation                                                                                | Applying                      | Make straightforward inferences                     |
| Analyze<br><ul style="list-style-type: none"> <li>&lt; Differentiating</li> <li>&lt; Organizing</li> <li>&lt; Attributing</li> </ul>                                                                                                      | Breaking material into its constituent parts and detecting how the parts relate to one another and to an overall structure or purpose | Reasoning                     | Interpret and integrated ideas and information      |
| Evaluate<br><ul style="list-style-type: none"> <li>&lt; Checking</li> <li>&lt; Critiquing</li> </ul>                                                                                                                                      | Making judgments based on criteria and standards                                                                                      | Reasoning                     | Evaluate and criticise content and textual elements |
| Create<br><ul style="list-style-type: none"> <li>&lt; Generating</li> <li>&lt; Planning</li> <li>&lt; Producing</li> </ul>                                                                                                                | Putting elements together to form a novel, coherent whole, or make an original product                                                | Not assessed in TIMSS and SLA | Not assessed in PIRLS and SLA                       |

### 1.2.2 Content Domain

#### a) Numeracy

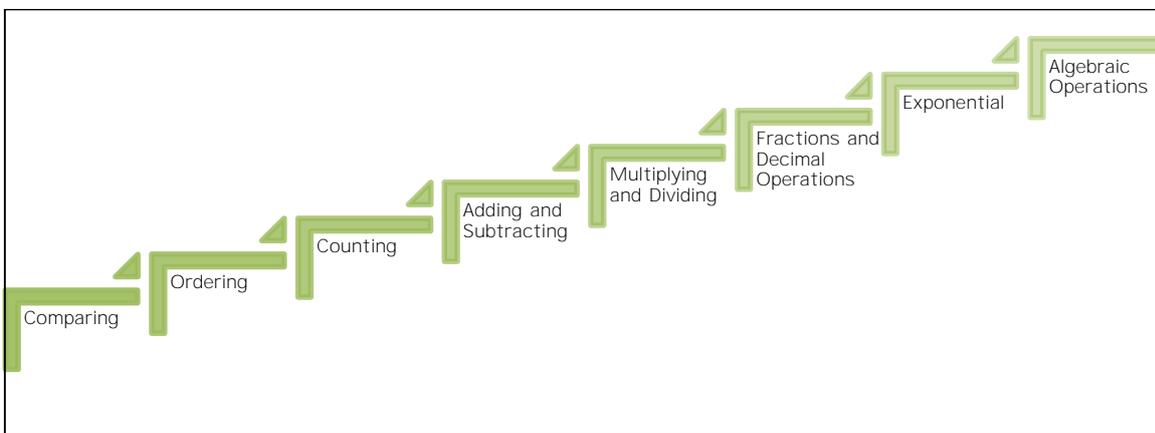
The numeracy assessment tool contains three primary mathematics content domains for elementary and junior high schools: numbers, geometry, and measurement, as well as data and statistics. The tool framework covers both the scope of mathematical content and also the complexity of mathematical skills. Referring to the Indonesian curriculum<sup>4</sup> and theories on numeracy development (Education Scotland, 2017; Booker et al., 2014; Van de Walle, Karp and Bay Williams, 2013), we created a numeracy staircase that implies skills that a student needs to acquire before acquiring other skills with broader scope or higher complexity. In general, it is a bit in dealing with numbers categorized into two broad levels.

<sup>4</sup>We specifically refer to two Minister of Education and Culture Regulation (No. 17/2013 on Basic Framework and Curriculum Structure of Elementary School *Madrasah Ibtidaiyah* and No. 68/2013 on Basic Framework and Curriculum Structure of Junior High School *Madrasah Sanawiah*).

- (1) Level 1: The number concepts. At this stage, children are required to have the sense of quantities and sizes and how they are represented by numbers, identify the place value of each digit in a number and identify sequences or patterns.
- (2) Level 2: The number operations. This level requires skills of using numbers to solve mathematical problems, such as addition, subtraction, multiplication, and division.

Apart from the above categories, the type of number also determines the complexity of the number concepts. We introduce numbers starting from whole numbers, followed by fractions, decimals, and negative integers. In each type of number, we always start with items asking about the concept (level 1), then followed by mathematical problems that include those types of numbers (level 2).

With respect to skills, the framework includes comparing, ordering, counting, adding, subtracting numbers, multiplying and dividing numbers, and solving word problems. The levels of number domain are presented in the following staircase.

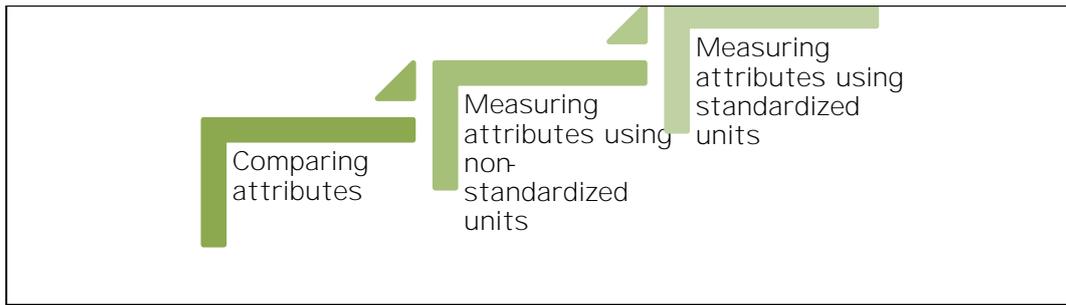


**Figure 1. Number domain staircase**

The geometric material (SERMAT) are put into levels based on five levels of geometric thinking (Van de Walle, Karp and Williams, 2013). The levels are labeled as visualization, analysis, informal deduction, formal deduction, and rigor.

- (1) Level 0: recognition or visualization. In this stage, children recognize shapes in terms of what they resemble (e.g., triangle resembles mountain).
- (2) Level 1: analysis. In this level, children start analyzing and naming the geometric figures.
- (3) Level 2: abstraction (informal deduction). Children start to see the relationships between properties and figures, categorize the shapes based on their characteristics, and make meaningful definitions (e.g., all squares are rectangles, but not all rectangles are squares).

The third and fourth levels, formal deduction and rigorous, are not included in the test. Meanwhile, the measurement domain is categorized into three levels as illustrated in the following staircase.



**Figure 2. Levels in measurement**

Data and statistics domain covers on data representation in the form of pictograms and simple tables for elementary school level and central tendency as well as the basic concept of probability for junior high school level. The third and fourth levels for each content domain guide us in deciding materials to be assessed in each booklet of the test. Table for each domain can be seen in Appendix 3.

b) Literacy

This literacy assessment tool is designed to cover various forms of texts. During our review of the National Curriculum<sup>5</sup>, we found that they focus on two aspects: reading skills and text types. Since the function of the SLA is to assess students' ability to construct a meaning from the text, we have integrated both aspects into the framework. The reading skills in the national curriculum for the first semester of Grade 1 focus on early literacy skills, such as syllable recognition and word recognition. From the second semester of Grade 1 up to Grade 3, the students learn to comprehend the content of short narrations and be able to get the main idea from them. From Grades 4 to 6, students learn how to create linkages between ideas, organize information, evaluate text, and give opinions. Domains related to reading skills for junior high school level (Grades 7 to 9) are similar to those for higher grades of elementary level. It is the complexity that increases in terms of the length of the passages and the complexity of the text genre (e.g., experiment report, popular science report, and classic literature).

The domain in the national curriculum that is related to text types only focuses on the genre, but not specifically on the complexity level of the text. Without the explanation of the complexity of text, the text genre on the curriculum draws questions on what type of text is appropriate for a particular level. Take the Grade 1 curriculum as an example. It is written in the curriculum that one of the texts to be given to students is a poem. Comprehending typical poems requires a skill to understand the implicit message or interpreting the text. An implicit meaning of a text is the second stage of the pyramid. Reflecting on this case, we refer to the international benchmark of text levelling, which has been adapted into an Indonesian context.

The Fountas and Pinnell Text Leveling Guide is a widely used framework to determine the complexity of the text appropriate to a particular grade level (Fountas & Pinnell, 2016). The Literacy levels are constructed based on several factors:

- (1) Genre. Each genre has characteristics and features. Understanding the features can help us to determine in which level the text belongs to.

<sup>5</sup>The National Curriculum in this document refers to domains included in the 2006 and 2013 curriculum.

- (2) Text structure. The more complex the information is structured in the text, the higher the level of the text.
- (3) Content. The content factor refers to how complex the topics presented in the text.
- (4) Themes and ideas. The number of ideas and themes presented in the text is one of the factors determining the difficulty level of the text.
- (5) Language and literary features. This aspect refers to the style of the written language used in the text. The children's familiarity with the style appropriate to the subject to be the factor that determines the difficulty level of the text.
- (6) Sentence complexity. Texts with simpler and more natural sentences are considered to be easier to process.
- (7) Vocabulary. The difficulty level of the vocabulary is determined by the familiarity with the words. The more commonly used words in the text are, the easier the text will be.
- (8) Words. This factor refers to the number of difficult words or commonly used words are repeated in the text.
- (9) Illustrations. For young readers, illustrations provide information related to the text can make the text easier to understand.
- (10) Book and print features. This factor refers to the physical aspects of the text, such as layout, font size, and length.

PRIORITAS Prioritizing Reform, Innovation and Opportunity for Indonesia. The project is funded by the United States Agency for International Development (USAID), has adapted the Fountas and Pinnell Text Levelling by taking into account the context relating to the above ten factors of text difficulty level (PRIORITAS 2015). The level ranges from A for texts with only one to two words per page to Z for complex texts, such as classic literature or popular science (see Appendix 2 for a more detailed level). Based on the review by literacy experts and findings of SLA pilot study, we decided to use the range of reading level that would be suitable for assessing students in Indonesia. The following table shows a range of text level of the SLA booklets.

**Table 2. Text Level Presented in the SLA Test Booklets**

| Grade Level | Level | Appear in the Test Booklet for |         |         |         |         |         |           |
|-------------|-------|--------------------------------|---------|---------|---------|---------|---------|-----------|
|             |       | Grade 1                        | Grade 2 | Grade 3 | Grade 4 | Grade 5 | Grade 6 | >=Grade 7 |
| Pre-school  | A. D  |                                |         |         |         |         |         |           |
| Grade 1     | E. J  |                                |         |         |         |         |         |           |
| Grade 2     | K. M  |                                |         |         |         |         |         |           |
| Grade 3     | N. P  |                                |         |         |         |         |         |           |
| Grade 4     | Q. S  |                                |         |         |         |         |         |           |
| Grade 5     | T. V  |                                |         |         |         |         |         |           |
| Grade 6     | W. Y  |                                |         |         |         |         |         |           |
| Grades 7. 9 | Z     |                                |         |         |         |         |         |           |

### 1.2.3 The Use of the Rasch Model in Designing the Test

We used the Rasch model to evaluate the quality of psychometric properties of each item included in the test as well as the overall set of instruments (Staver and Yale, 2014). The model predicts the probability of a person with a particular ability to give a correct answer to an item at a particular difficulty level. The model can be written as follows

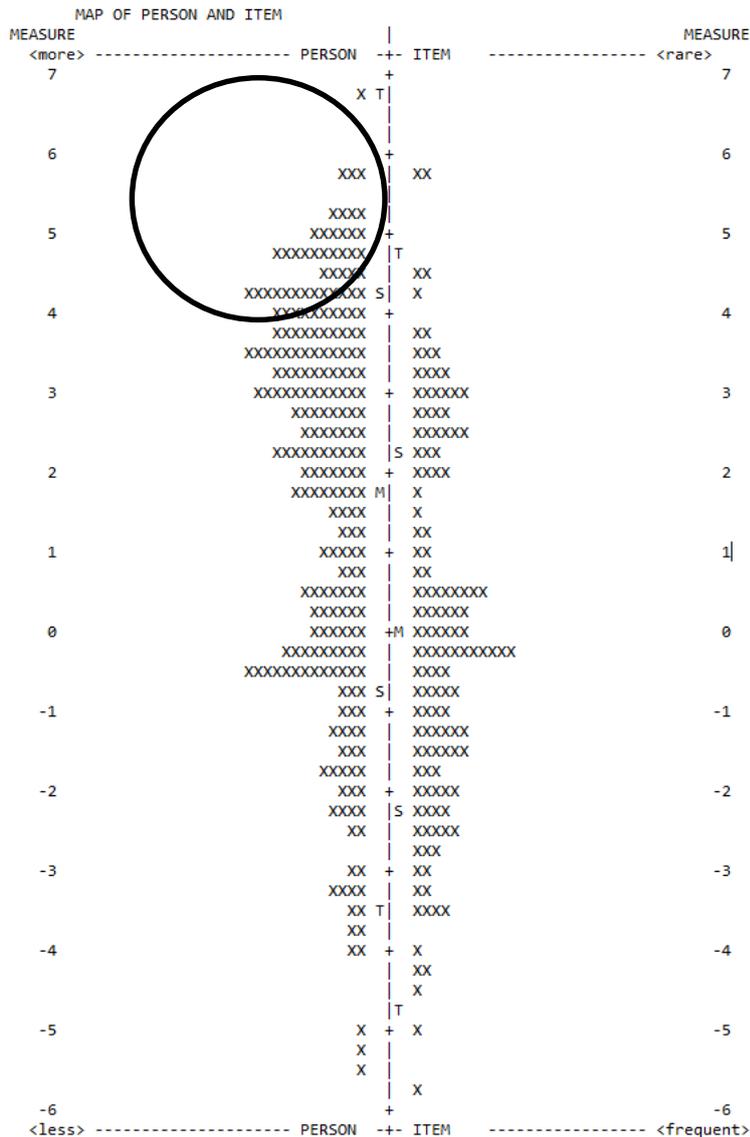
where  $p_{ij}$  is the probability of the response of person  $i$  to item  $j$ ,  $\theta_i$  is the person's ability, and  $\delta_j$  is the item's difficulty level. Based on the attributes that can describe the quality of a test: item and person measure. The person measure is considered as a latent ability of the person, which is the closest to the true score of person. This latent ability cannot be predicted if we estimate the score using a raw score calculation. This model assumes that the difficulty level of items is the most relevant parameter to predict the person's latent ability as well as to evaluate the quality of the test. Therefore, the difficulty level should be the focus of the Rasch analysis.

A good test is a test that has a distribution of difficulty level that is relevant to the purpose of the test. If a test is designed for selection purposes, it should contain difficult items. On the other hand, if a test is designed for measurement purposes, it should contain items with a wide range of difficulty levels to be able to assess both low and high intelligent people.

Increasing the precision of a measurement is the first step in test development using the Rasch approach. Using the Rasch model, we wanted to create a test that is adaptive to our targeted students' ability. To see how to fit the distribution of item difficulty to the ability of the students, we created an item-person map (known as the Wright Map) using the Rasch model. In the following figure, we can see the distribution of item difficulty in the Indonesian test for Grade 1 (on the right side) mapped against the

---

<sup>4</sup>Hereinafter referred to as Indonesian.



**Figure 3. Item-person map for Indonesian - Grade 1**

Figure 3 shows that the test items have a difficulty level that can sensitively assess students with various abilities, except a small group of students who have high ability (see the circle on the above figure). This figure suggests adding more difficult items to make the instrument more sensitive in assessing the high ability students.

### 1.2.4 The Link of Booklet, Grade, and Other Assessment Instruments

To be able to compare the score of this SLA with the score obtained from another instrument, the two instruments need to have items in common, known as anchor items. The function of these anchor items is to equate the psychometric properties of the test. The proportion of anchor items in each test booklet is approximately 30% of the total number of items in the booklet. The anchor items allow us to equate the results among different grades using one metric. When choosing the anchor items, we must ensure that there are no subdomains omitted from the booklet.

Apart from the anchor items between the modules, initially, we designed our instrument to be comparable with other general assessments, such as Programme for International Student Assessment (PISA), TIMSS, and INAP. Therefore, we adopted some items from other tests as anchor to compare our instruments with the results of other assessments. However, there was insufficient information on the other assessments for high school level. Consequently, our high school level instruments did not adopt the items from other assessment instruments.

#### a) Horizontal Equating Test Strategy

In CERMA, we have several identical items to INAP. We can equate our test with INAP by putting the anchor items' psychometric properties as a fixed parameter in the Rasch analysis. In CERMA, the psychometric properties of the rest of the items will be adjusted to these properties.

#### b) Vertical Equating Test Strategy

Similar to the horizontal equating test strategy, this strategy is a procedure to compare a test with a higher or lower grade. In CERMA, we have several items that anchor the Grade 1 test to the literacy or numeracy test, Grade 2 test to the Grade 3 test, and so on. The aim of employing the vertical equating test strategy is to track progress across grades to see whether or not a participant is performing at or above his/her age group.

## II. TEST DESIGN

### 2.1 CERMAT's Cognitive and Content Domain Framework

#### 2.1.1 Numeracy

To design a valid and efficient instrument, we studied instruments used by other institutions to know the composition of knowing, applying, and reasoning domains. Some instruments, such as the UN and the National Based School Examination (USBN), emphasize more on procedural knowledge than middle or higher order thinking skills. On the other hand, international assessments such as PISA and TIMSS try to balance the composition between knowing, applying, and reasoning.

We have conducted multiple pilots to find the combination of knowing, applying, and reasoning in each grade ability. The instrument for the low level elementary school is designed to include the domains of knowing and applying. Moreover, the instrument also emphasizes more on the knowing domain than the applying domain. Although the instrument for Grades 1, 2, and 3 items is similar, we increased the proportion of items to assess the applying skill in the Grade 3 instrument. This increase is based on cognitive child development theory.

For the higher level elementary school, we designed half of the items to cover the knowing domain, while the rest are divided into the applying and reasoning. The small proportion of higher order thinking items in the test was decided based on findings from a study where there are so many higher order thinking items that the students could not answer. At the same time, the findings also suggested that we add more variation of easy items. In the Grade 4 instrument, we introduced the reasoning. The difficulty of the item increases gradually following the topic learned at a specific curriculum level. Thus, although the domain compositions for the higher elementary school are rather similar, the difficulty levels should be adjusted to be at the right level.

The junior high level school numeracy instrument emphasizes more on the applying domain following the Indonesian curricular properties. This is aligned with what is stated in the curriculum that the junior high school students are required to be able to apply mathematical concepts that have been learned during elementary schools in more complex problems. The cognitive domain framework of our numeracy instrument can be the following table.

**Table 3. Cognitive Domain Framework for Numeracy Test**

| Level              | Grade | Knowing (%) | Applying (%) | Reasoning (%) |
|--------------------|-------|-------------|--------------|---------------|
| Elementary School  | 1     | 76          | 24           |               |
|                    | 2     | 76          | 24           |               |
|                    | 3     | 67          | 33           |               |
|                    | 4     | 52          | 30           | 18            |
|                    | 5     | 42          | 39           | 18            |
|                    | 6     | 52          | 36           | 18            |
| Junior High School | 7     | 35          | 40           | 25            |
|                    | 8     | 35          | 40           | 25            |
|                    | 9     | 35          | 40           | 25            |

In constructing the framework for content domain for our numeracy test, we mapped the Indonesian curricula that are currently being used in schools, the 2006 Curriculum (KTSP) and the 2013 Curriculum. We only picked content domains that appear in both curricula and mapped those domains into grade level according to numeracy staircase and literacy level (see section 1.2.2). By only including contents that are supposed to be taught in the classroom, the test will have content validity. The following table explains the content domain framework of our numeracy test.

**Table 4. Content Domain Framework for Numeracy Test**

| Level              | Grade | Numbers (%) | Geometry and Measurement (%) | Data and Statistics (%) |
|--------------------|-------|-------------|------------------------------|-------------------------|
| Elementary School  | 1     | 85          | 15                           |                         |
|                    | 2     | 85          | 15                           |                         |
|                    | 3     | 60          | 30                           | 10                      |
|                    | 4     | 75          | 15                           | 10                      |
|                    | 5     | 70          | 20                           | 10                      |
|                    | 6     | 70          | 25                           | 5                       |
| Junior High School | 7     | 70          | 15                           | 15                      |
|                    | 8     | 60          | 20                           | 20                      |
|                    | 9     | 60          | 25                           | 15                      |

### 2.1.2 Literacy

In designing the literacy assessment instrument, we tried to adopt the framework used in PIRLS. We consider the compatibility and comparability of the framework in adopting PIRLS.

To develop valid and efficient instrument, we conducted multiple pilots. From our pilots, we learned that there were many students from Grades 1 to 3 in our sample group who could not recognize letters and words. Therefore, we decided to insert the early literacy components not only in the module for Grade 1 but also in modules for Grades 2 and 3. To assess the reading comprehension, we included three cognitive process domains:

- a) Focus on and retrieve explicitly stated information
- b) Making straightforward inferences
- c) Interpret and integrate ideas and information

In addition to the three cognitive process domains, we added another domain for the higher-level elementary school instrument, which designated students in Grades 4, 5, and 6 cognitive process domain is ability to evaluate and criticize content textual elements. We also added more proportion of items (Table 5) to assess the ability to interpret and integrate ideas and information. Following the Indonesian curricula, the length of the passages for the higher grades is similar, but the vocabulary richness we used in passages are gradually increasing in difficulty.

The last group of the instrument is the module for Grades 7, 8, and 9. Passages belonging to this group are relatively longer and have a more advanced vocabulary. Most of the items in the booklets assess students' ability in making inferences from a passage (40% to 50%) and only contain a small proportion of items that ask students to retrieve explicitly stated information (10%). We picked text passages according to the Indonesian curricula.

introduce various genres for each grade. We put more advanced vocabulary and more difficult items for the later grades.

**Table 5. Cognitive Framework for Literacy Test**

| Level              | Grade | Early Literacy Components (%) | Focus on and Retrieve Explicitly Stated Information (%) | Making Straightforward Inferences (%) | Interpret and Integrate Ideas and Information (%) | Evaluate and Criticize Content and Textual Elements (%) |
|--------------------|-------|-------------------------------|---------------------------------------------------------|---------------------------------------|---------------------------------------------------|---------------------------------------------------------|
| Elementary School  | 1     | 10                            | 40.50                                                   | 25.30                                 | 15.20                                             |                                                         |
|                    | 2     | 10                            | 40.50                                                   | 25.30                                 | 15.20                                             |                                                         |
|                    | 3     | 10                            | 40.50                                                   | 25.30                                 | 15.20                                             |                                                         |
|                    | 4     |                               | 10.20                                                   | 15.20                                 | 30.60                                             | 10.30                                                   |
|                    | 5     |                               | 10.20                                                   | 15.20                                 | 30.60                                             | 10.30                                                   |
|                    | 6     |                               | 10.20                                                   | 15.20                                 | 30.60                                             | 10.30                                                   |
| Junior High School | 7     |                               | 10.20                                                   | 40.50                                 | 30.40                                             | 10.20                                                   |
|                    | 8     |                               | 10.20                                                   | 40.50                                 | 30.40                                             | 10.20                                                   |
|                    | 9     |                               | 10.20                                                   | 40.50                                 | 30.40                                             | 10.20                                                   |

### 2.1.3 Item Presentation

#### a) Numeracy

In designing our test items, we chose two types of item presentation for the numeracy section: multiple-choice and closed constructed-response items. The proportion of the two types of item presentations is presented in Table 6.

##### (1) Multiple-Choice Items

The multiple-choice problems require the students to choose the most correct and relevant answer from the choices provided. The multiple-choice problems in elementary and junior high school modules have four choices. Most of the elementary school modules use multiple-choice problem type, with a few constructed-response items. Learning from the experience in constructing items for Grades 1 to 6, we would need considerable resources to score the answers of the closed constructed-response item. Hence, when creating the items for Grades 7 to 9, we tried to minimize the number of constructed-response items. We then designed the mathematics to only consist of multiple-choice items.

##### (2) Closed Constructed-Response Items

In the numeracy instrument, the closed constructed-response problems require the students to write their answers in the provided field. This type of item presentation may reduce guessing bias by the students. In our pilots, we found that many students tend to repeatedly answer A or D on a different item based on their guesses. Therefore, to minimize correct answers made by guessing, we introduced the constructed-response item type.

**Table 6. Allocation of Item Presentation in Numeracy Assessment by Grade**

| Level              | Grade | Multiple Choice (%) | Closed Constructed-Response (%) |
|--------------------|-------|---------------------|---------------------------------|
| Elementary School  | 1     | 50                  | 50                              |
|                    | 2     | 50                  | 50                              |
|                    | 3     | 50                  | 50                              |
|                    | 4     | 50                  | 50                              |
|                    | 5     | 70                  | 30                              |
|                    | 6     | 70                  | 30                              |
| Junior High School | 7     | 100                 |                                 |
|                    | 8     | 100                 |                                 |
|                    | 9     | 100                 |                                 |

b) Literacy

In the literacy assessment instrument, we presented three types of item presentation: multiple choice, closed constructed response, and open constructed response. The proportion of the three types of item presentation is presented in Table 7.

(1) Multiple-Choice Items

The multiple-choice problems require the student to choose the most correct and relevant answer from the choices provided. The multiple-choice problems in both elementary school and junior high school modules have four choices.

(2) Closed Constructed-Response Items

The closed constructed-response problems require the student to give short and straightforward answers. For Grades 1 to 3, the students are required to answer verbally or in a written form. For Grades 7 to 9, the students are required to provide short and concise answers in a written form. Questions with this type of presentation usually have specific correct and explicit answers. In scoring this type of answers, we compare correctness and relevance of the answers to the rubric. Different students may choose different words or terminology to answer the questions; however, if the meaning of the answer is relevant and correct, it can get full score.

(3) Open Constructed-Response Items

The open constructed-response problems require the student to provide concise and elaborative answers in a written form. To answer the questions that fall into this type of presentation, the student would have to explain or describe a process, situation, or object. Moreover, the questions may also require the student to mention items questioned in the problem set.

**Table 7. Allocation of Item Presentation in Literacy Assessment by Grade**

| Level              | Grade | Multiple-Choice (%) | Closed Constructed-Response (%) | Open Constructed-Response (%) |
|--------------------|-------|---------------------|---------------------------------|-------------------------------|
| Elementary School  | 1     | 10.20               | 75.85                           | 10.20                         |
|                    | 2     | 10.20               | 75.85                           | 10.20                         |
|                    | 3     | 10.20               | 75.85                           | 10.20                         |
|                    | 4     | 80.85               |                                 | 15.20                         |
|                    | 5     | 80.85               |                                 | 15.20                         |
|                    | 6     | 80.85               |                                 | 15.20                         |
| Junior High School | 7     | 80.90               | 10.15                           | 10.15                         |
|                    | 8     | 40.60               | 10.15                           | 30.45                         |
|                    | 9     | 40.60               | 10.15                           | 30.45                         |

## 2.2 Test Format and Scope

In general, our instrument comprises two types of assessment: literacy and numeracy. Both assessments cover modules of all grades in elementary and junior high school levels. We conducted several pilots to ensure that our instruments are valid and efficient. In practice, our instruments are designed to be concise and able to explain the learning outcome of the specific grade using as few items as possible. The table below shows the number of items in each type and specific grade.

**Table 8. Number of Items by Grade and Instrument Type**

| Instrument Type | Level              | Grade | Number of Items |
|-----------------|--------------------|-------|-----------------|
| Numeracy        | Elementary School  | 1     | 20              |
|                 |                    | 2     | 25              |
|                 |                    | 3     | 25              |
|                 |                    | 4     | 23              |
|                 |                    | 5     | 23              |
|                 |                    | 6     | 25              |
|                 | Junior High School | 7     | 12              |
|                 |                    | 8     | 12              |
|                 |                    | 9     | 12              |
| Literacy        | Elementary School  | 1     | 17              |
|                 |                    | 2     | 23              |
|                 |                    | 3     | 23              |
|                 |                    | 4     | 19.21           |
|                 |                    | 5     | 19.21           |
|                 |                    | 6     | 19.21           |
|                 | Junior High School | 7     | 12              |
|                 |                    | 8     | 12              |
|                 |                    | 9     | 12              |

Table 8 shows the number of items in each booklet. All booklets were designed to be completed in less than one hour. For example, a Grade 7 student will have two hours to finish both numeracy and literacy modules. Since more time is needed to complete the high school problem set than the elementary school problem set, we assigned fewer items per module for the junior school module.

## 2.3 Test Scoring

### 2.3.1 Item Scoring

For multiple choice and closed constructed response items, the score is as simple as 1 for a correct answer and 0 for an incorrect one. In the Rasch model, we name this type of scoring the Graded Response Model (GRM) where the variables are dichotomous (either 0 or 1). In the open constructed response items, some questions require lengthy or multiple responses to have a full answer. In this case, sometimes the students can provide a partially correct response. To accommodate this condition, we use the Partial Credit Model (PCM) where the variables can be polytomous (for our SLA, it is either 0, 1, or 2).

### 2.3.2 Estimating Students' Score

As discussed in the previous section, the Rasch model enabled us to estimate the true score or latent ability of a student. In estimating the score, the model predicted the maximum likelihood a person to get a full score in one particular item by looking at the pattern of the responses to items in the whole test. Hence, this model could still predict the likelihood of someone getting a full score for a missing item. To create a test that is comprehensive yet still has an efficient number of items in each module, we divided items for each grade in several test booklets with anchor items that link one booklet to another.

### III. DEVELOPMENT OF THE TEST

#### 3.1 School Index for Sample Selection

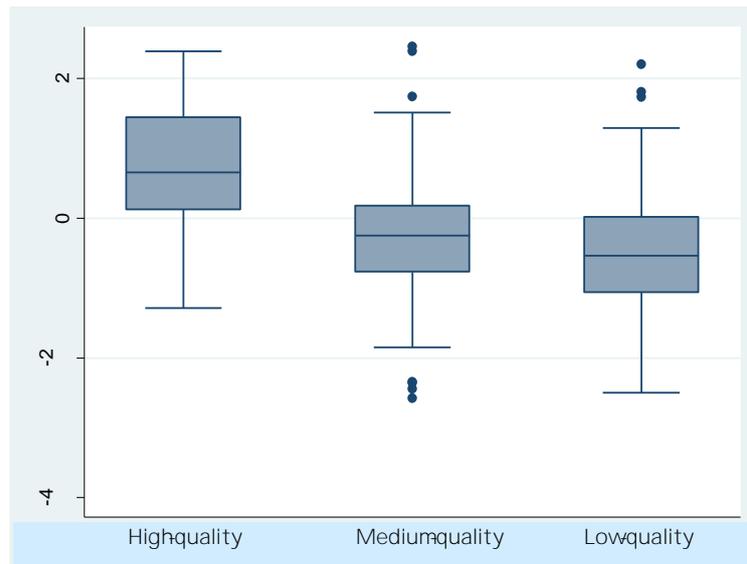
As mentioned before, we aim at creating a test that is sufficient for literacy and numeracy improvement in RISE's target population. The test booklet to have items with a wide range of difficulty levels that can capture the performance of students with low and high ability.

To obtain a sample of students with a high variation of abilities that could represent the range of abilities of RISE studies' targeted population, we conducted a Principal Component Analysis (PCA) based on the available school characteristics data that potentially have linkage to the learning outcome. Given the limited national data availability, we used data from DAPODIK (Data Pokok Pendidikan, The Ministry of Education and Culture) as variables relating to learning outcome. The variables used were: teacher quality, student ratio, classroom quality, internet access, the proportion of civil servant and contract teachers, library quality, access to clean water, accreditation score, and number of special education teachers. We excluded schools in two regions, Maluku and Papua, that would not be selected as RISE Indonesia study area. From the PCA, we took the first component as the school quality index. We grouped schools into three categories: high, medium, and low quality schools. The high quality schools are those with the 20% highest index. It was followed by the middle schools, whose indexes are in the fourth and third quintiles, and low quality schools that are in the first and second quintiles. The mean of the quality index in each category is presented in the table below.

**Table 9. Mean of the Index in Each Quality Group**

| School Category | Mean of Quality Index | Min     | Max    |
|-----------------|-----------------------|---------|--------|
| High quality    | 0.937                 | 0.496   | 2.835  |
| Medium quality  | 0.003                 | -0.526  | 0.496  |
| Low quality     | -1.655                | -27.230 | -0.527 |

For each of our pilot, we picked schools that had the closest index to the mean of each of the category. Our data shows that the means of schools' quality index of the schools (Figure 4). It can be implied that the chosen variables from DAPODIK can describe the quality of the schools.



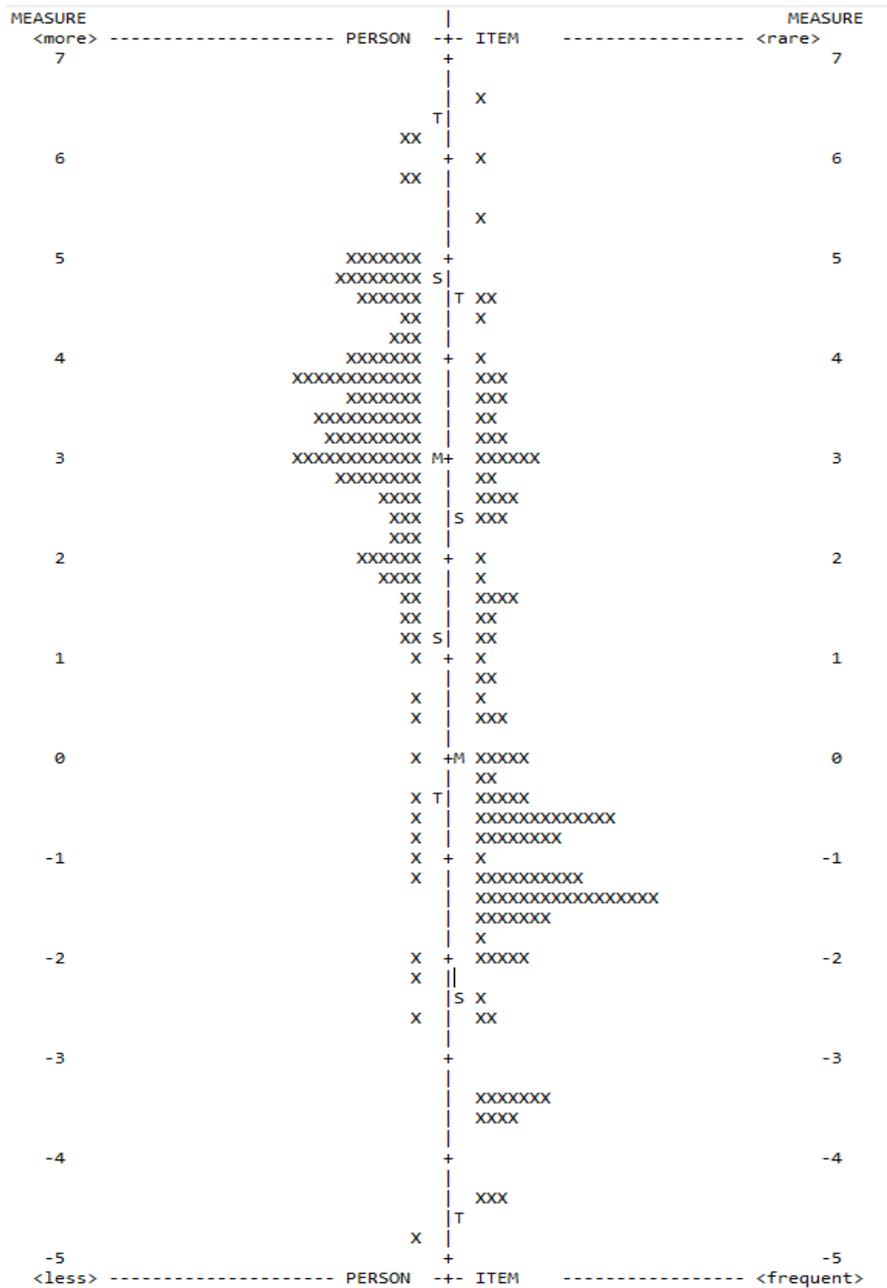
**Figure 4. SLA scores by school quality category**

## 3.2 Item Development

### Revisiting and Revising the SLA Framework

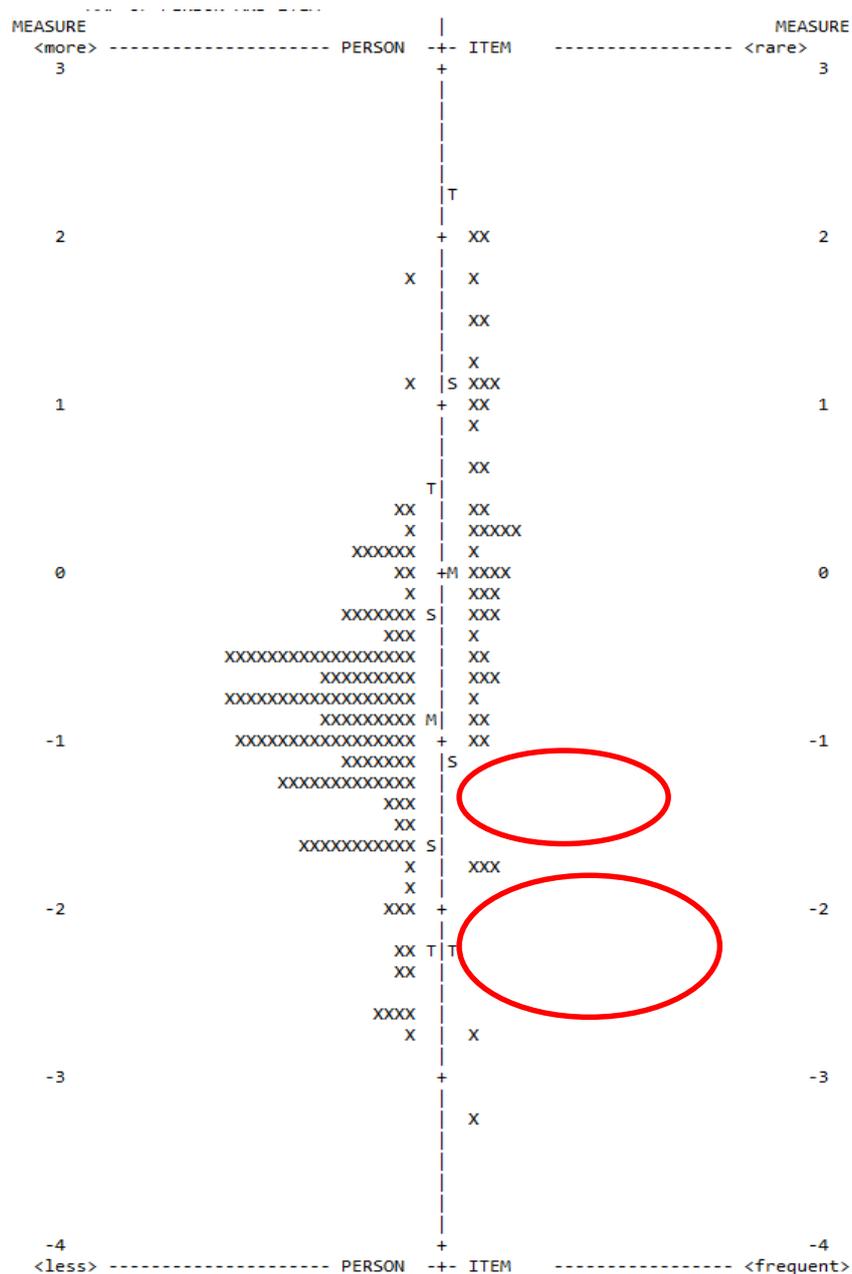
The development of GERMAT for Grades 1 to 9 was an adaptation process from the framework developed by INOVASI and KIAT Guru. We started by mapping the 2006 and 2013 National Curriculum as the currently used curriculum in Indonesian public schools. To enrich our reference on content domain, we also reviewed content domain from other standardized tests, e.g., National Assessment of Educational Progress (NAEP), the Computer Adaptive Placement Assessment and Support System (COMPASS), PIRLS, and INAP. With the help of literacy and numeracy experts, we decided the proportion of content and cognitive domain for each grade level. Then, we mapped the good quality items from INOVASI and KIAT Guru to our framework and created new items as needed. INOVASI and KIAT Guru consist of items designed for Grades 1 to 5 and draft items designed for Grade 6; all were piloted in six schools in West Lombok in April 2018. The schools represent the three categories of quality.

Findings from the first pilot show that the Grades 1 to 3 tests are good quality. The difficulty level of the items is widely distributed and with the ability of the targeted sample. Hence, we only did minor revision to the tests for Grades 1 to 3 and the items were not included in the subsequent pilot. We used the Wright Map (Figure 5) to help us determine which items needed revising what kind of items we needed to add. The left side of the graph shows the distribution of students. The higher the position of a student, the higher his/her ability is. The right side of the graph shows the distribution of items based on their difficulty. The items placed at the top of the graph are easier than those placed at the bottom. The following graph exemplifies the distribution of Indonesian items for Grade 3 students that captures all levels of students' ability, even items that are too easy and to be removed from the test.



**Figure 5. Wright Map of Indonesian items and students' abilities - Grade 3**

In contrast, we found that the distribution of item difficulty in tests for Grade 4 to 6 had yet capture the range of students' ability. Take a l



**Figure 6. Wright Map of mathematics items and students' abilities - Grade 4**

Figure 6 shows that the overall test was too difficult for the students. There are also blank (marked in red circle) where we can see a group of students whose abilities could not be captured by items with the appropriate level of difficulty. After looking at these findings, the numeracy literacy expert revised the items to make difficulty level appropriate to the ability of the targeted students.

In July 2018, we conducted the second cycle of the pilot study. This cycle also included booklet students in Grades 7 to 9 with mathematics items constructed by a mathematics expert Sampoerna University. From the second cycle, we found that booklets for Grades 4 and 5 need minor revisions while booklets for Grades 6 to 9 required more. After the booklets were revised, we tested them again in September 2018. The last cycle of the pilot study was followed by the finalization of the whole SLA instrument.

## IV. PSYCHOMETRIC PROPERTIES OF THE TEST

This chapter describes the quality of the test in terms of its psychometric properties. The following subsections present data that shows how valid and reliable the test is.

### 4.1 Test Validity

#### 4.1.1 Content Validity

Validity tells us how accurate a measurement is to assess a particular aspect. The validity of a test relies more on a judgment of whether the content of the test is in line with the purpose of the test rather than the outputs of the statistical test. To make sure that our SLA validly assess what students are supposed to learn at school, we referred to the content domain of the test to the national curricula.

All of the items that we created went through a content review process by literacy and numeracy experts. The experts are teachers and lecturers of Indonesian and mathematics education who have experience in teaching Grades 1 to 3, 4 to 6, or 7 to 9. They also have at least a formal training certificate or experience in creating higher order thinking for internationally standard assessments curricula e.g., Early Grade Reading Assessment (EGRA), Early Grade Mathematics Assessment (EGMA), and International Baccalaureate Assessment. During the development process, they were involved in finalizing the content and cognitive domain framework, creating items, and reviewing items based on findings from the psychometric analysis.

#### 4.1.2 Outfit and Infit Test

To support the above judgment, we also checked the validity of the SLA tool by estimating the outfit and infit mean squares of the items. The outfit test shows how much noise the items measure. The expected value is 1.0. The value that is below 1.0 indicates that the observations are too predictable, while if it is too far above 1.0 indicates unpredictability. Meanwhile, the purpose of estimating the infit index is to see whether a predictable pattern captured in the test (e.g., there is a group of people with low ability who have higher probability to answer difficult items). The good infit index ranges from 0.5 to 1.5. The following table shows that the outfit and infit test of our SLA tool indicate good validity.

**Table 10. Outfit and Infit Indices**

| Test            | Outfit            |                 | Infit |      |
|-----------------|-------------------|-----------------|-------|------|
|                 | MNSQ <sup>a</sup> | SD <sup>b</sup> | MNSQ  | SD   |
| <b>Numeracy</b> |                   |                 |       |      |
| Grade 1         | 1.15              | 0.64            | 1.01  | 0.12 |
| Grade 2         | 1.15              | 0.85            | 0.99  | 0.14 |
| Grade 3         | 1                 | 0.46            | 0.98  | 0.15 |
| Grade 4         | 1.09              | 0.48            | 1     | 0.2  |
| Grade 5         | 1.03              | 0.32            | 0.98  | 0.17 |
| Grade 6         | 1.11              | 0.75            | 1.02  | 0.24 |
| Grade 7         | 1                 | 0.27            | 1.01  | 0.14 |
| Grade 8         | 1.06              | 0.45            | 1     | 0.16 |
| Grade 9         | 1.08              | 0.39            | 1.04  | 0.2  |
| <b>Literacy</b> |                   |                 |       |      |
| Grade 1         | 1.12              | 0.91            | 0.99  | 0.26 |
| Grade 2         | 1.06              | 0.96            | 1     | 0.24 |
| Grade 3         | 0.97              | 0.76            | 0.99  | 0.26 |
| Grade 4         | 1.22              | 1.09            | 1.01  | 0.26 |
| Grade 5         | 1.01              | 0.53            | 1     | 0.45 |
| Grade 6         | 0.94              | 0.24            | 1     | 0.1  |
| Grade 7         | 1.11              | 0.85            | 1     | 0.19 |
| Grade 8         | 1.02              | 0.42            | 0.98  | 0.14 |
| Grade 9         | 1.04              | 0.5             | 0.99  | 0.16 |

<sup>a</sup>MNSQ mean square

<sup>b</sup>SD: standard deviation

## 4.2 Test Reliability

We used the Rasch model to estimate the reliability of the test. In the classical approach, reliability is only seen as how consistent the sample could answer the items correctly by performing correlation between an item with the rest of the items. In the Rasch model, the person ability and item difficulty were taken into account in estimating the consistency of pattern. The item reliability index indicates the consistency of the difficulty level of the items after they were tested in different levels of ability. The expected reliability score is above 0.75.

In addition to reliability score, we also have the separation index. Person separation index shows how well the test can differentiate students into groups of ability. For example, the separation index of 2.9 could divide students into three groups of high, medium, and low. Likewise, the item

separation index shows how well the items can be separated into number of levels (e.g., index of 5.47 shows five levels of difficulty)

The following table shows that the items in our tests are reliable (0.8-0.98). However, a low person reliability score can still be found in numeracy tests for Grades 1, 3, 7, 8, and 9; and literacy test Grade 8. By looking at the person separation index, we can see that this may be due to the small variation of the sampled students' ability (separate person reliability score and separation index is that the test cannot be used to classify students who attend the same grade into many levels of competence). However, the wide range of item difficulty with a high item separation index makes the test sufficiently sensitive for evaluative purposes.

**Table 11. Item-Person Reliability Scores and Separation Indices**

| Test            | Person      |                  | Item        |                  |
|-----------------|-------------|------------------|-------------|------------------|
|                 | Reliability | Separation Index | Reliability | Separation Index |
| <b>Numeracy</b> |             |                  |             |                  |
| Grade 1         | 0.68        | 1.45             | 0.98        | 6.55             |
| Grade 2         | 0.75        | 1.72             | 0.98        | 6.51             |
| Grade 3         | 0.66        | 1.4              | 0.97        | 5.47             |
| Grade 4         | 0.71        | 1.56             | 0.95        | 4.47             |
| Grade 5         | 0.73        | 1.66             | 0.95        | 4.15             |
| Grade 6         | 0.75        | 1.71             | 0.96        | 4.83             |
| Grade 7         | 0.68        | 1.44             | 0.92        | 3.32             |
| Grade 8         | 0.72        | 1.59             | 0.92        | 3.3              |
| Grade 9         | 0.65        | 1.35             | 0.8         | 2.02             |
| <b>Literacy</b> |             |                  |             |                  |
| Grade 1         | 0.98        | 6.39             | 0.98        | 8.02             |
| Grade 2         | 0.97        | 5.5              | 0.98        | 6.91             |
| Grade 3         | 0.94        | 3.92             | 0.96        | 5.13             |
| Grade 4         | 0.8         | 1.98             | 0.95        | 4.59             |
| Grade 5         | 0.79        | 1.93             | 0.82        | 2.12             |
| Grade 6         | 0.57        | 1.16             | 0.94        | 3.8              |
| Grade 7         | 0.78        | 1.89             | 0.93        | 3.6              |
| Grade 8         | 0.7         | 1.52             | 0.92        | 3.29             |
| Grade 9         | 0.71        | 1.56             | 0.89        | 2.92             |

## LIST OF REFERENCES

- Booker, George Denise Bond, Len Sparrow and Paul Swain (2014) *Teaching Primary Mathematics* 5<sup>th</sup> ed. Frenchs Forest: Pearson Australia.
- Boone, Wiam J John R Staver and Melissa Sale (2014) *Rasch Analysis in the Human Sciences* 1<sup>st</sup> ed. Dordrecht: Springer Netherlands.
- Education Scotland (2017)  *Benchmarks: Numeracy and Mathematics* Livingston: Education Scotland
- Fountas & Pinnell (2016). *Fountas & Pinnell* [online] <[https://www.fountasandpinnell.com/Authenticated/ResourceDocuments/FP\\_FPL\\_Characteristics\\_Difficulty.pdf](https://www.fountasandpinnell.com/Authenticated/ResourceDocuments/FP_FPL_Characteristics_Difficulty.pdf)> [4 December 2018].
- Krathwohl, David R (2002) A Revision of Bloom's Taxonomy: An Overview *Theory Into Practice* 41(4) 212-218.
- Mullis, Ina V. Sand Michael O Martin (eds.) (2013) *TIMSS 2015 Assessment Frameworks* Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College and International Association for the Evaluation of Achievement.
- USAID PRIORITAS (2015) *Buku Sumber untuk Dosen LPTK: Pembelajaran Literasi di Sekolah Dasar/ Madrasah Ibtidaiyah* [online] <[https://prioritaspendidikan.org/Pembelajaran\\_Literasi\\_SD\\_di\\_LPTK1.pdf](https://prioritaspendidikan.org/Pembelajaran_Literasi_SD_di_LPTK1.pdf)> [4 December 2018]
- Van de Walle, John A., Karen S. Karp, Jennifer M. Bay Williams (2013) *Elementary and Middle School Mathematics: Teaching Developmentally* ed. Boston: Pearson.
- Government Laws and Regulations
- Regulation of the Minister for Education and Culture No. 67/2013 on Basic Framework and Curriculum Structure of Elementary School/Madrasah Ibtidaiyah
- Regulation of the Minister for Education and Culture No. 68/2013 on Basic Framework and Curriculum Structure of Junior High School/Madrasah Sanawiah

# APPENDICES

# APPENDIX 1

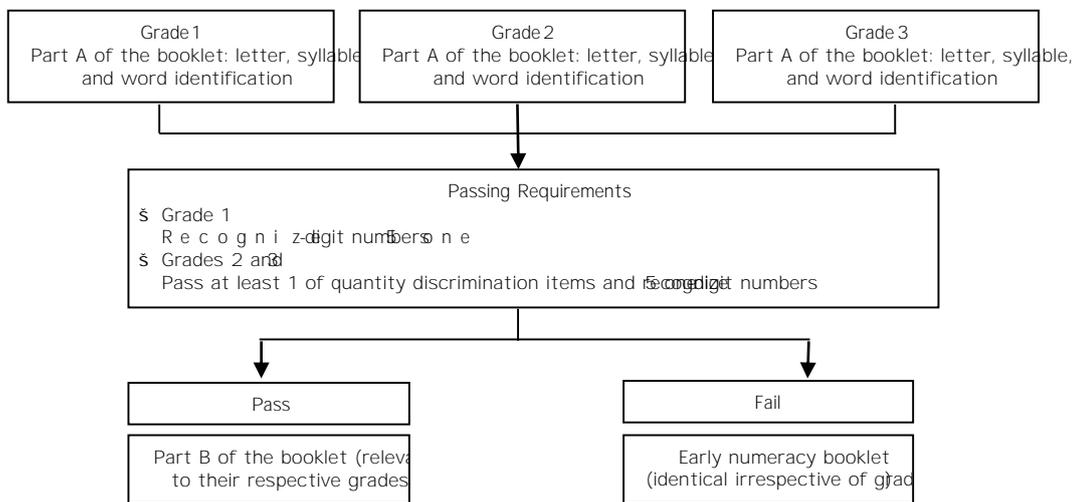
## Test Administration Manual

### 1. Test Flow

We used three booklets when we administered the test of Grades 1 to 3; each had a different purpose. Booklet A serves as a screening to see if the students have the required skills to attempt Booklet B, which will then measure their skills respective to the grade. Failure to pass Booklet A indicates that the student does not possess the required literacy skills to be able to understand the questions in Booklet B. To be able to capture the possible literacy skills of the students, we gave them a distinct booklet applicable to their capability; the early numeracy and early literacy booklets serve the purpose.

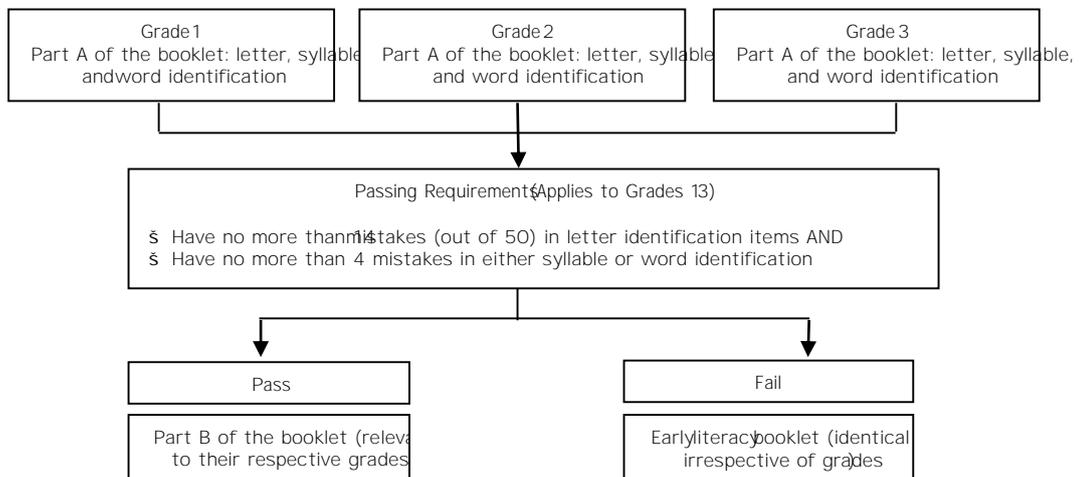
#### a) Numeracy for Grades 1 to 3

The following chart outlines the flow of the mathematics tests for Grades 1 to 3.



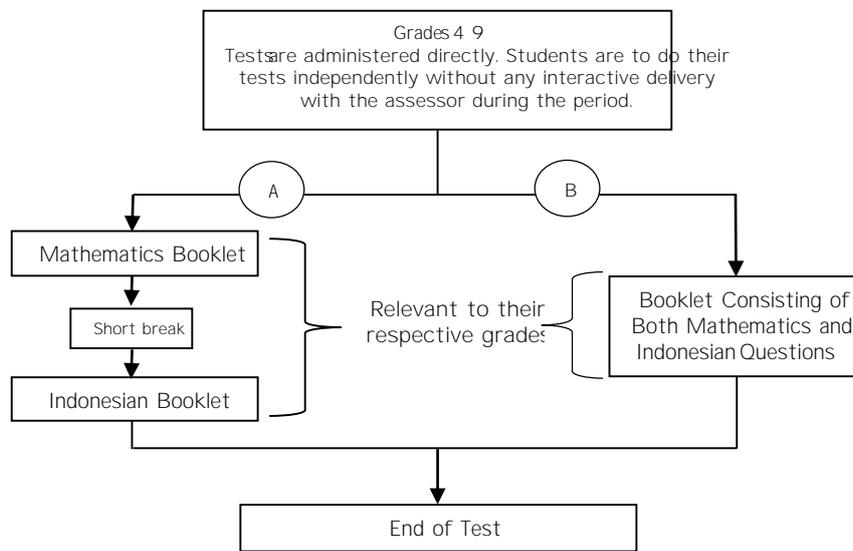
#### b) Literacy for Grades 1 to 3

The following chart outlines the flow of the Indonesian tests for Grades 1 to 3.



c) Grades 4 to 9

The following chart outlines the flow of the mathematics and Indonesian tests for Grades 4 to 9



Flow A is used when the total number of questions may be overwhelming for the student impacting their capacity to answer the questions. Mathematics precedes Indonesian due to the likelihood of Indonesian being more taxing for the student due to the higher word count. Hence, the order prescribed is aimed at minimizing any adverse effects between the two tests.

Flow B is used when the number of questions is appropriate to be bundled in one single test. Bundling together the two disciplines allow flexibility for the students to decide which subject to do first according to their individual capabilities.

## 2. Test Procedure

### a) Test Preparation

#### (1) General

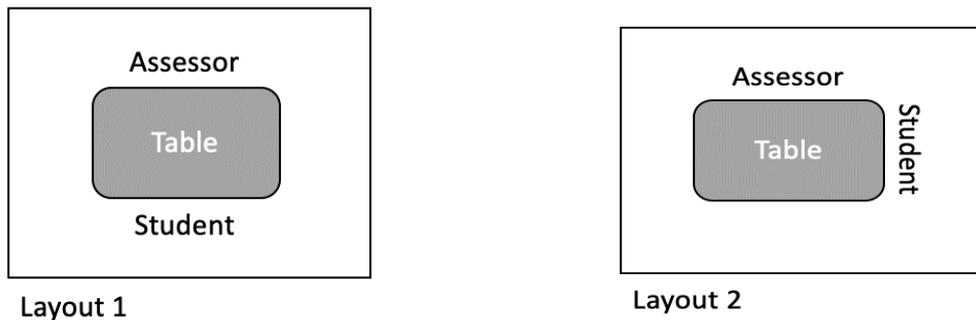
In conducting the exam, a conducive environment should be arranged to the best of the circumstances. An ideal condition includes

- (a) an uncrowded room or a clear area (if a room is unavailable),
- (b) suitable lighting,
- (c) good ventilation, and
- (d) a venue free of noise and interference.

The assessor should also make sure that there are no materials, such as posters, relevant to the exam visible in class; coverage or removal may be necessary.

## (2) Grades 1 to 3

As early grade tests are conducted one-on-one, the assessor should arrange their seating appropriate for such type of assessment, as outlined by the following layout options. Also should ensure that the distance between any two students is not intrusive of each other, as shown in the following layout



The following equipment should be prepared and ensured to be correct:

- (a) Tablet with test materials
- (b) Tests' show card
- (c) Required materials for each subject
  - i) Mathematics: Blank paper and pencil
  - ii) Indonesian Dictation sheet

Considering the interactive nature of the tests, assessors should build rapport with the students prior to the test; this can be achieved by asking several simple questions and making warm contact. Finally, they should ensure that the students are fully prepared to start the test.

## (3) Grades 4 to 9

The assessor should coordinate directly with the relevant field observer for scheduling purposes (assessment to be done prior after the classroom observation). Permit should also be obtained prior to the assessment from the teachers whose teaching period is to be replaced by assessment.

The following equipment should be prepared and ensured to be correct:

- (a) Question sheets. Several versions may be employed to prevent cheating.
- (b) Answer sheet
- (c) Consent form to be filled by the homeroom teacher
- (d) Stationery

### b) During the Test

#### (1) Grades 1 to 3

Upon starting the assessors should consult with the students of their preferences on which exam they would like to take first. They should ensure that the show card is faced directly to the student.

## During the Mathematics Assessment

### Part A: Reading Numbers

Instruction for students: Clearly read aloud the following numbers from left to right.

Instructions for assessors:

(a) Always begin with the assumption that students are unable to read the text.

The students are to be shown figures or numbers on the whole page. Questions are being read out

(b) The sample page should be shown to the student to starting the actual test and the following procedures should be followed

i) Ask the student to read the sample questions.

ii) Emphasize correct answer with phrase "Yes, that is correct. This is the number ..."

iii) Correct their mistakes when wrong with phrase "emphasize the pronunciation 'puluh' and 'ndeh'." and "hundred"

(c) Give the following instruction: "I will keep counting the numbers. So please attempt to read as fast and clearly as possible. You are allowed to point at the numbers as you are reading them".

Early Numeracy Test: there are two parts involved in the assessment.

(a) Reading numbers with illustrations.

i) The students are requested to read the numbers pointed by assessors.

ii) The assessors are allowed to express the number with their fingers.

(b) Pointing numbers read out by assessors

Assessors are allowed to provide an example of the instruction by saying phrases such as "Can you point at a pencil? Good. Please do read out".

Part B: The students should attempt to answer the questions independently. Assessors may only assist in reading out the questions.

## During the Indonesian Assessment

### Part A: Reading Letters, Words, and Syllables

Instruction for students: Clearly read aloud the following letters from left to right.

Instructions for assessors:

(a) The sample page should be shown to the student to starting the actual test and the following procedures should be followed

i) Ask the student to read the sample questions.

ii) Emphasize correct answers with phrase "This is correct. This is the letter ..."

iii) Correct their mistakes when wrong with phrases such as "This letter is ..."

(b) Give the following instruction: "I will keep counting the letters. So please attempt to read as fast and clearly as possible. You are allowed to point at the letters as you are reading them".

Note: The students are allowed to pronounce each alphabet differently from the Indonesian pronunciation rules EXCEPT for F, P, and

Early Literacy Test: there are three parts involved in the assessment.

(a) Reading letters with illustrations.

- i) The students are requested to read the numbers pointed by the assessors.
- ii) The assessors are allowed to express the number with their fingers.

(b) Pointing letters read out by the assessors.

The assessors are allowed to provide an example of the instruction by saying phrases such as "Can you point at a pencil? Good. Please do read out".

(c) Listening

- i) The assessors should read out the passage with a clear articulation to be fast.
- ii) The assessors should ensure that the student is ready to pay attention before reading.
- iii) After the passage has been read, assessors should immediately instruct: "answer the following question..."

The assessors should avoid giving out too much instruction between the reading activity and the question asked.

## Part B

Dictation instructions to be read to students and followed by assessors:

- (a) "I will read out a sentence THREE times."
- (b) "You should listen carefully. I will provide you with a pencil, ruler, and paper."
- (c) "Then, I will read out the sentence what you heard on the paper. I will provide you with a writing time."
- (d) "Finally, I will read out the same sentence your answer."

(2) Grades 4 to 9

Instruction for assessors:

- (a) Introduce yourself
- (b) Explain clearly that the assessment is meant to be a form of practice for mathematics in Indonesian, and WILL NOT influence their school grades in any way.
- (c) Distribute the answer sheets.
- (d) Guide students in filling their identity and other personal information on the answer sheet. Make sure that all of the students have filled in the required fields before distributing the question booklets.
- (e) Explain that they are about to be given a booklet consisting of both mathematics and Indonesian questions, of which they are free to choose particular questions to do first according to their preference and capabilities. They are allowed to immediately move on to the next set of questions after they are finished on a subject.

- (f) Clearly explain how to fill the answer sheet and emphasize that the question booklet MUST NOT be used for any writing or rough work.
- (g) Distribute question booklets. If different versions are used, make sure that no two students sitting side by side receive the same version of questions.

c) Rules

- (1) The questions are to be attempted individually without any form of cooperation between students.
- (2) The students are allowed to ask clarificatory questions by raising their hand towards the assessor
- (3) The students are allowed to go to the bathroom.
- (4) The students cannot leave the room before the whole duration of the test is over. They should quietly remain in their seats even if they finish early.

d) Test Duration

The SLA is designed as a power test rather than a speed test. In this case, the test is designed to assess intelligence, not speed. Therefore, the test has no time limit. However, for practical matter in a classical test (Grades 4 to 9), we set a time limit that is minutes more than the time that would actually be needed to complete the whole test booklet. If a booklet consists of fifteen reading questions, the average time of students to complete the test is 30 to 35 minutes. Hence, the given time limit is 45 minutes. The same amount of time is also applied to a booklet consisting of 20 mathematics questions. It takes around 30 minutes for students to finish the 20 item mathematics test.

## APPENDIX 2

### The Fountas and Pinnel Text Levelling

| Text Level | Indicators                                                                                                                                                                                                                                                                                                                                                                                                                    |
|------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| A          | <ul style="list-style-type: none"> <li>◁ Recognize letters and their sounds</li> <li>◁ Point to words while reading</li> <li>◁ Use picture to support understanding</li> <li>◁ Know the difference between words and pictures</li> <li>◁ One sentence per page with simple words</li> <li>◁ Read easy, high frequency words</li> </ul>                                                                                        |
| B          | <ul style="list-style-type: none"> <li>◁ Follow a sentence over 2 lines of text</li> <li>◁ Continue to point to words while reading</li> <li>◁ Recognize pattern throughout story</li> <li>◁ Reread to fix reading mistakes</li> <li>◁ Read, easy high frequency words</li> </ul>                                                                                                                                             |
| C          | <ul style="list-style-type: none"> <li>◁ Read simple stories with 2. 6 lines of text on page</li> <li>◁ Notice repeated lines and phrases</li> <li>◁ Begin to follow text with eyes, rather than pointing</li> <li>◁ Use strategies to help understanding</li> <li>◁ Begin to correct reading mistakes</li> <li>◁ Read easy, high frequency words</li> </ul>                                                                  |
| D          | <ul style="list-style-type: none"> <li>◁ Read fiction and simple nonfiction</li> <li>◁ Continue to follow text with eyes, rather than pointing</li> <li>◁ Read text with fewer lines of repeated words</li> <li>◁ Read compound words (e.g., newspaper, sandbox)</li> <li>◁ Continue to correct reading mistakes</li> <li>◁ Read easy, high frequency words</li> </ul>                                                        |
| E          | <ul style="list-style-type: none"> <li>◁ Read books with 3. 8 lines of text</li> <li>◁ Follow text with eyes, rather than pointing</li> <li>◁ Read texts that require more attention for understanding</li> <li>◁ Follow punctuation correctly</li> <li>◁ Take apart long words</li> <li>◁ Rely on meaning from the text, rather than pictures</li> <li>◁ Read fluently</li> <li>◁ Read easy, high frequency words</li> </ul> |
| F          | <ul style="list-style-type: none"> <li>◁ Begin to understand genres (fiction, nonfiction, etc.)</li> <li>◁ Read and understand dialogue in text</li> <li>◁ Read words with multiple syllables</li> <li>◁ Automatically read high frequency words</li> </ul>                                                                                                                                                                   |
| G          | <ul style="list-style-type: none"> <li>◁ Continue to understand different genres</li> <li>◁ Read 3. 8 lines of text per page; text is smaller</li> <li>◁ Read difficult words by using letter/sound information, thinking of familiar words, taking apart words</li> <li>◁ Read text with a few challenging vocabulary words</li> <li>◁ Automatically read high frequency words</li> </ul>                                    |

| Text Level | Indicators                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
|------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| H          | <ul style="list-style-type: none"> <li>Read longer text with more challenging vocabulary</li> <li>Read difficult words by using letter/sound information, thinking of familiar words, taking apart words</li> <li>Automatically read high frequency words</li> </ul>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
| I          | <ul style="list-style-type: none"> <li>Read short texts (8. 16 pages) and easy chapter books (40. 60 pages)</li> <li>Understand longer sentences of more than 10 words</li> <li>Read many texts silently, without pointing to words</li> <li>Automatically read a large number of high frequency words (from all previous levels and more)</li> </ul>                                                                                                                                                                                                                                                                                                                                                                                                                   |
| J          | <ul style="list-style-type: none"> <li>Read many types of texts (informational texts, short chapter books, simple biographies)</li> <li>Understand a large number of longer sentences</li> <li>Use strategies to figure out hard words (go back and reread, use picture clues, find smaller words inside the bigger word, etc.)</li> <li>Automatically read a large number of high frequency words (from all previous levels and more)</li> </ul>                                                                                                                                                                                                                                                                                                                       |
| K          | <ul style="list-style-type: none"> <li>Read many types of texts (biographies, informational texts, realistic fiction stories, fantasy stories, traditional literature, simple texts)</li> <li>Read many illustrated chapter books</li> <li>Must remember many details</li> <li>W} â ^ ! • c æ} â Á â ã æ  [ * ~ ^ Á æ} â Á c @^ Á ~ • ^ Á [ ~ Á ~ ~ [ c æ c</li> <li>Books have many characters that change a little in the story</li> <li>Read stories with diverse cultures</li> <li>Use strategies to figure out hard words (go back and reread, use picture clues, find smaller words inside the bigger word, use word parts like prefixes/suffixes, etc.)</li> <li>Automatically read a large number of high frequency words (from all previous levels)</li> </ul> |
| L          | <ul style="list-style-type: none"> <li>Read easy chapter books with less pictures</li> <li>Read short informational and fiction books</li> <li>Read slower or faster depending on the book</li> <li>Learn new concepts through reading</li> <li>Use what they already know to help their reading</li> <li>Use pictures and text to help understand</li> <li>Connect known facts to new information</li> <li>Understand difficult ideas</li> <li>Understand a large number of words (plurals, contractions, possessives, multi-syllable words, content-specific words, technical words)</li> <li>Understand difficult sentences</li> </ul>                                                                                                                               |
| M          | <ul style="list-style-type: none"> <li>Know the characteristics of different genres (realistic fiction stories, fantasy stories, informational text, traditional literature, biographies, etc.)</li> <li>Read fiction chapter books, such as series books (e.g., Junie B. Jones) or mysteries</li> <li>Read fiction texts that have many characters that change in the story</li> <li>Read shorter nonfiction texts on one topic</li> <li>Understand difficult sentences</li> </ul>                                                                                                                                                                                                                                                                                     |
| N          | <ul style="list-style-type: none"> <li>Process short fiction stories, chapter books, short informational texts, series books (e.g., most Amber Brown books) or mysteries</li> <li>Read fiction texts that have many characters that change in the story</li> <li>Read nonfiction texts on many related topics</li> <li>Automatically use strategies (find smaller words inside the bigger word, use word parts like prefixes/suffixes, etc.)</li> <li>Read and understand descriptive words</li> </ul>                                                                                                                                                                                                                                                                  |

| Text Level | Indicators                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
|------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|            | <ul style="list-style-type: none"> <li>Understand difficult sentences</li> </ul>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
| O          | <ul style="list-style-type: none"> <li>Know the characteristics of most genres</li> <li>Read chapter books, shorter informational texts, mysteries, series books, books with sequels, or short stories</li> <li>Read fiction stories with many characters that change throughout the story</li> <li>Read nonfiction texts that give information on many related topics</li> <li>Understand difficult sentences and words</li> <li>Figure out new vocabulary words by using clues</li> </ul>                                                                                                                                                                       |
| P          | <ul style="list-style-type: none"> <li>Know the characteristics of most genres</li> <li>Read chapters books, shorter informational texts, mysteries, series books, books with sequels, short stories</li> <li>Read fiction stories with many characters that change throughout the story</li> <li>Read nonfiction texts that give information on new topics</li> <li>Understand mature themes (race, language, culture, etc.)</li> <li>Make sense of new vocabulary words</li> </ul>                                                                                                                                                                              |
| Q          | <ul style="list-style-type: none"> <li>Automatically read and understand characteristics of most genres, including biographies on new topics, chapter books, shorter informational texts, mysteries, series, books with sequels, short stories.</li> <li>Read fiction stories with many characters that change throughout the story</li> <li>Make sense of new vocabulary words</li> <li>Look for information in pictures, photographs, maps, charts, etc.</li> <li>Can break words into syllables</li> <li>Understand texts with different layouts</li> <li>Look for information in pictures, photographs, maps, charts, etc.</li> </ul>                         |
| R          | <ul style="list-style-type: none"> <li>Automatically read and understand characteristics of most genres, including biographies on new topics, fantasies, chapter books, shorter informational texts, mysteries, series, books with sequels, short stories, diaries, and logs.</li> <li>Read fiction stories with many characters that change throughout the story</li> <li>Make sense of new vocabulary words</li> <li>Look for information in pictures, photographs, maps, charts, etc.</li> <li>Can break words into syllables</li> <li>Use strategies to figure out difficult words</li> <li>Understand texts with different layouts</li> </ul>                |
| S          | <ul style="list-style-type: none"> <li>Automatically read and understand characteristics of most genres, including biographies on new topics, fantasies, chapter books, shorter informational texts, mysteries, series, books with sequels, short stories, diaries, and logs.</li> <li>Read fiction stories with many characters that change in the story</li> <li>Understand hard sentences and words</li> <li>Can break words into syllables</li> <li>Understand texts with different layouts</li> <li>Look for information in pictures, photographs, maps, charts, etc.</li> </ul>                                                                             |
| T          | <ul style="list-style-type: none"> <li>Automatically read and understand characteristics of most genres, including biographies on new topics, fantasies, chapter books, shorter informational texts, mysteries, series, books with sequels, short stories, diaries, logs, fantasies, myths, and legends.</li> <li>Read longer texts with many lines of print that require the reader to remember lots of information</li> <li>Can break words into syllables</li> <li>Use strategies to figure out difficult words</li> <li>Look for information in pictures, photographs, maps, charts, etc.</li> <li>Use what they already know to understand a text</li> </ul> |

| Text Level  | Indicators                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
|-------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| U           | <ul style="list-style-type: none"> <li>◁ Automatically read and understand characteristics of most genres, including biographies on new topics, fantasies, chapter books, shorter informational texts, mysteries, series, books with sequels, short stories, diaries, logs, fantasies, myths, and legends.</li> <li>◁ Read longer texts with many lines of print that require the reader to remember lots of information</li> <li>◁ Can break words into syllables</li> <li>◁ Use strategies to figure out difficult words</li> <li>◁ Search for and use information in a text</li> <li>◁ Look for information in pictures, photographs, maps, charts, etc.</li> </ul>                                                                                                                                    |
| V           | <ul style="list-style-type: none"> <li>◁ Read and understand characteristics of most genres, including biographies on new topics, fantasies, chapter books, shorter informational texts, mysteries, series, books with sequels, short stories, diaries, logs, fantasies, myths, and legends</li> <li>◁ Read texts that are longer and involve remembering information</li> <li>◁ Can break words into syllables</li> <li>◁ Search for and use information in a text</li> <li>◁ Look for information in pictures, photographs, maps, charts, etc.</li> </ul>                                                                                                                                                                                                                                               |
| X, Y, and Z | <ul style="list-style-type: none"> <li>◁ Read and understand characteristics of all genres, including biographies on new topics, fantasies, chapter books, shorter informational texts, mysteries, series, books with sequels, short stories, diaries, logs, fantasies, myths, and legends.</li> <li>◁ Use critical thinking skills</li> <li>◁ Read long texts with long sentences and paragraphs</li> <li>◁ Understand mature themes (abuse, poverty, war, etc.)</li> <li>◁ Read texts with many characters that change in the story</li> <li>◁ Use what they know to understand a text</li> <li>◁ Search for and use information in a text</li> <li>◁ Look for information in pictures, photographs, maps, charts, etc.</li> <li>◁ Read texts that require knowing about history and science</li> </ul> |

Source: Fountasandpinnell.com, 2014; Schem.edu, 2019.

## APPENDIX 3

### Numeracy Staircase

| Domain                  | Level 1                                                                 | Level 2                                                            | Level 3                                                                                                         | Level 4                                                                          | Level 5                                                                                                           | Level 6                                                        | Level 7 | Level 8 | Level 9 |
|-------------------------|-------------------------------------------------------------------------|--------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------|---------|---------|---------|
| <b>Number Concept</b>   | Determining quantity: presented in iconical or pictorial representation | Recognizing simple fractions represented in concrete form          | Representing fraction as a part of whole and a part of a collection by using pictorial form (e.g., shaded area) | Making equivalent fractions                                                      | Representing power by two as repeated multiplication and understanding square root as the inverse of power by two | Recognising negative integers                                  |         |         |         |
|                         | Recognizing numbers up to two digits                                    | Recognizing numbers up to three digits                             |                                                                                                                 | Recognizing prime numbers                                                        |                                                                                                                   |                                                                |         |         |         |
|                         | Comparing the quantity of objects                                       | Comparing and ordering whole numbers                               | Comparing and ordering like fractions                                                                           | Comparing and ordering unlike fractions, decimals, and percentages               |                                                                                                                   | Comparing and ordering positive and negative integers          |         |         |         |
|                         | Identifying place values up to two digits                               | Identifying place values up to three digits                        |                                                                                                                 |                                                                                  |                                                                                                                   |                                                                |         |         |         |
|                         | Determining a missing number in an ordered pattern on a number line     | Determining a missing number in a skipped pattern on a number line | Determining two missing numbers in a skipped pattern on a number line                                           | Determining two sequential missing numbers in a skipped pattern on a number line |                                                                                                                   |                                                                |         |         |         |
| <b>Number Operation</b> | Adding and subtracting whole numbers up to 99                           | Adding and subtracting whole numbers up to 999                     | Performing mixed math operations of whole numbers: addition and subtraction                                     | Performing mixed math operations of whole numbers: multiplication and division   |                                                                                                                   | Performing number operations of positive and negative integers |         |         |         |

| Domain                          | Level 1                                                                                                | Level 2                                                                                                 | Level 3                                                                                                                                  | Level 4                                                                                                                                        | Level 5                                 | Level 6                                                                 | Level 7 | Level 8 | Level 9 |
|---------------------------------|--------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------|-------------------------------------------------------------------------|---------|---------|---------|
|                                 |                                                                                                        | Multiplying and dividing whole numbers up to 100                                                        |                                                                                                                                          | Estimating and rounding                                                                                                                        |                                         |                                                                         |         |         |         |
|                                 |                                                                                                        |                                                                                                         |                                                                                                                                          | Least common multiples and greatest common divisors                                                                                            | Multiplying and dividing fractions      |                                                                         |         |         |         |
|                                 |                                                                                                        |                                                                                                         | Adding and subtracting like fractions                                                                                                    | Converting fractions from and into decimals and percentages                                                                                    | Adding and subtracting unlike fractions | Performing mixed math operations of whole numbers and fractions         |         |         |         |
| <b>Word Problem</b>             | Simple word problem supported by illustration                                                          | Simple word problem <b>without</b> illustration: two available numbers and one missing number           | Simple word problem without illustration: more than two available numbers and one missing number; requiring single-step solution process | Simple word problem without illustration: more than two available numbers and one missing number; requiring simple multi-step solution process |                                         |                                                                         |         |         |         |
| <b>Geometry and Measurement</b> | Recognizing simple plane figures (triangle, rectangle, square) and simple solid figures (cube, cuboid) | Categorizing simple plane figures (triangle, rectangle, square) and simple solid figures (cube, cuboid) | Describing the properties and characteristics of simple plane figures (triangle, rectangle, square)                                      | Recognizing and categorizing polygon                                                                                                           |                                         | Describing the properties and characteristics of a circle               |         |         |         |
|                                 |                                                                                                        |                                                                                                         | Categorizing simple solid shapes (cube, cuboid)                                                                                          | Describing the characteristics of simple solid shapes (cube, cuboid)                                                                           | Making the nets of cube and cuboid      | Recognizing and categorizing prism, cylinder, pyramid, cone, and sphere |         |         |         |

| Domain                     | Level 1                     | Level 2                                                | Level 3                                                       | Level 4                                                   | Level 5                                                                 | Level 6                                            | Level 7 | Level 8 | Level 9 |
|----------------------------|-----------------------------|--------------------------------------------------------|---------------------------------------------------------------|-----------------------------------------------------------|-------------------------------------------------------------------------|----------------------------------------------------|---------|---------|---------|
|                            | Comparing length and weight | Measuring length and weight using nonstandardized unit | Measuring length, weight, and time using standardized unit    | Unit conversion for length, weight, and time              | Recognising and comparing distance, time, and speed                     |                                                    |         |         |         |
|                            |                             | Comparing time and size                                | Recognizing and comparing angles                              | Measuring angles                                          |                                                                         |                                                    |         |         |         |
|                            |                             |                                                        | Measuring area using nonstandardized unit such as unit square | Determining the area of square, rectangle, triangle       |                                                                         | Measuring the area and perimeter of circle         |         |         |         |
|                            |                             |                                                        |                                                               | Measuring volume using nonstandardized unit: cube, cuboid | Determining the volume of cube and cuboid                               | Determining the volume of a combination of cuboids |         |         |         |
|                            |                             |                                                        | Determining the perimeter of square, rectangle, triangle      |                                                           | Coordinate point, map, and scale                                        |                                                    |         |         |         |
|                            |                             |                                                        | Point and linear symmetry                                     |                                                           | Point of compass                                                        |                                                    |         |         |         |
| <b>Data and Statistics</b> |                             |                                                        | Identifying information from pictogram and 2 x n table        | Identifying information from table, bar, and pie chart    | Using information from charts and tables to solve mathematical problems | mean, median, mode                                 |         |         |         |



Math-Grade 2

TABLE 1.2 MTK\_2.xlsx ZOU424WS.TXT Dec 31 7:11 2018  
 INPUT: 233 PERSON 304 ITEM MEASURED: 233 PERSON 42 ITEM 84 CATS 3.69.1

```

PERSON - MAP - ITEM
<more>|<rare>
5 .##### +
  .
  . | T ma125_sc rm205_sc
4 +
  |
  | rm215_sc
  |
  | rm203_sc
3 # T+
  # | ma326_sc
  .
  #
  .# | rm105_sc rm206_sc
  ### | S rm208_sc
2 ## +
  # |
  ## | rm202_sc rm209_sc
  .#### | S | rm204_sc rm207_sc rm214_sc
  ###
  .####
1 ## +
  .##### | rm217_sc
  .##### | m126_sc ma117_sc rm218_sc
  .##### | ma320_sc
  .##### | m120_sc ma322_sc rm106_sc
  .#### | M | ma112_sc ma321_sc
0 .##### +M | rm212_sc
  .##### | rm201_sc
  .## | ma323_sc rm211_sc
  .#####
  .##
  .#####
-1 .#### +
  .## | S | m110_sc
  .# | ma123_sc rm210_sc rm216_sc
  # | m18_sc
  .
  # +
  | S
  | ma203_sc
  # T | rm213_sc
  . | ma002_sc ma106_sc ma107_sc ma121_sc rm104_sc
-3 . + | m15_sc
  .#####
  |
  | ma001_sc ma105_sc
-4 ## +
  <less>|<frequ>
EACH "#" IS 2. EACH "." IS 1.
    
```





Math-Grade 5

```

PERSON - MAP - ITEM
<more>|<rare>
4      . +
      # |
      | |
3      . +
      # |
      T |
2      T+
      ## | sc_ma404
      | | sc_rm525
      .# | sc_m2016inapbt141
      . |
      ## | S sc_m418          sc_rm514
      | | S sc_ma506
      + | sc_rm524
1      .## | sc_ma503          sc_rm423
      . S | sc_rm407a
      #####
      | | sc_rm404
      .### | sc_rm425          sc_rm523
      . |
      #### | sc_m2016inapbt934 sc_m425          sc_ma512
0      .## +M sc_m320
      ## | sc_rm401
      ####
      ### M | sc_ma510          sc_rm505
      ## | sc_rm510
      ##### | sc_ma508
      ## | sc_ma505          sc_rm418
      .###
-1     ##### +
      # | S
      #####
      . | sc_rm521
      ## S | sc_m2016inapbt64 sc_rm520
      .##### | sc_rm417
      . |
      | | sc_rm519
-2     .#### +
      | |
      . | T
      # |
      . |
      | |
      T | sc_rm516
      # |
-3     +
      <less>|<frequ>
EACH "#" IS 2. EACH "." IS 1.

```

Math-Grade 6

```

PERSON - MAP - ITEM
<more>|<rare>
4      .## +
      |
      ##
      #
      .
3      T+
      |
      .
      .### sc_rm514
      |T
      |
#####
2      +
      . S sc_rm617
      .###
      ##### sc_rm602
      . S
      ##### sc_rm619
1      .##### + sc_ma506
      .## | sc_rm423
      #### | sc_ma605
      .##### M | sc_m418
      .# | sc_ma503 sc_rm618 sc_rm623
      ##### | sc_rm603 sc_rm626
      ##
0      .##### +M sc_m52 sc_rm505 sc_rm625
      .##### | sc_rm605
      ##### | sc_ma512 sc_rm621
      .##
      ## | sc_rm510 sc_rm622
      .##### S | sc_rm620
      # | sc_ma611
-1     .##### +
      |
      . S
      ## | sc_ma510
      . | sc_ma505
      .
      T
-2     +
      #
      |
      |T
      |
      .
      | sc_rm516
-3     + sc_rm624
      |
      <less>|<frequ>
EACH "#" IS 2. EACH "." IS 1.

```

Math-Grade 7

```

PERSON - MAP - ITEM
<more>|<rare>
4      # +
      |
      | sc_rm73
3      . + sc_rm72
      |
      | T
      |
      | sc_rm72
2      # + sc_rm74
      ## T | sc_rm72
      |   | sc_rm74
      #   |
      .#   | sc_rm71
      .#   | S sc_rm74
      ##   | sc_rm70 sc_rm70 sc_rm71
1      #### + sc_rm71 sc_rm71 sc_rm73
      . S |
      .### | sc_rm70 sc_rm73
      #   |
      ### | sc_rm73 sc_rm74
      .#   | sc_rm61 sc_rm74
      ##### | sc_m418 sc_rm72 sc_rm73 sc_rm73 sc_rm74
0      .## +M sc_rm72
      .##### | sc_rm70 sc_rm71 sc_rm72 sc_rm73 sc_rm75
      ##### M | sc_rm72 sc_rm74
      ##### | sc_rm70 sc_rm71 sc_rm72
      .##### | sc_ma61 sc_rm70 sc_rm71 sc_rm71 sc_rm72 sc_rm73
      .#### | sc_rm71
      .#### | sc_rm70 sc_rm74
-1     ### + sc_rm62 sc_rm74
      .## | sc_rm74
      ##### S S | sc_rm61 sc_rm70
      ### | sc_rm73
      ### |
      .# | sc_rm73
      .## | sc_rm70
-2     # +
      # |
      | sc_ma50
      | sc_rm51
      | T
      | T
      |
      |
-3     # +
      |
      | sc_rm71
-4     +
      |
      | <less>|<frequ>
EACH "#" IS 2. EACH "." IS 1.

```



```

PERSON - MAP - ITEM
<more>|<rare>
7      +
      X T|
      |
6      +
      XXX | ba1 rb1
      |
      XXXX|
5      +
      XXXXXX|
      XXXXXXXXXX| T
      XXXXXX| ba1 rb1
      XXXXXXXXXXXXXX| S| rb1
4      +
      XXXXXXXXXXXXXX|
      XXXXXXXXXXXXXX| ba1 rb1
      XXXXXXXXXXXXXX| b21 ba1 rb1
      XXXXXXXXXXXXXX| b11 ba1 rb1 rb1
3      +
      XXXXXXXXXXXXXX| b11 ba1 ba1 ba1 men rb1
      XXXXXXXXXX| ba1 ba1 ba1 ba1
      XXXXXXXXXX| ba1 ba1 rb1 rb1 rb1 rb1
      XXXXXXXXXXXXXX| S| b11 ba1 rb1
2      +
      XXXXXXXXXX| ba1 ba1 men rb1
      XXXXXXXXXX M| ba1
      XXXX| ba1
      XXX| ba1 ba1
1      +
      XXXX| ba1 ba1
      XXX| ba1 ba1
      XXXXXXXXXX| ba1 ba1 ba1 ba1 ba1 ba1 ba1 ba1
      XXXXXXXXXX| ba1 ba1 ba1 ba1 ba1 ba1
0      +M
      XXXXXXXXXX| ba1 ba1 ba1 ba1 ba1 ba1 ba1 ba1
      XXXXXXXXXXXXXX| ba1 ba1 ba1 ba1
      XXX S| ba0 ba1 ba1 ba1 ba1
-1     +
      XXX| ba1 ba1 ba1 ba1
      XXXX| ba1 ba1 ba1 ba1 ba1 ba1
      XXX| ba0 ba1 ba1 ba1 ba1 ba1
      XXXXX| ba0 ba1 ba1
-2     +
      XXX| ba0 ba1 ba1 ba1 ba1
      XXXX| S| ba0 ba0 ba1 men
      XX| ba0 ba0 ba1 ba1 ba1
      | ba0 ba1 ba1
-3     +
      XX| ba0 ba1
      XXXX| ba0 ba0
      XX T| ba0 ba0 ba1 ba1
      XX
-4     +
      XX| ba0
      | ba0 ba0
      | ba0
      | T
-5     +
      X| ba0
      X|
      X|
      | ba0
-6     +
      <less>|<frequ>

```

```

PERSON - MAP - ITEM
<more>|<rare>
7      +
      T|
      .|
      .|
6      . + rb2
      # |
      # | rb2 rb2
      # |
5      .# +
      ### | T b21
      .### S|
      .#### |
      .##### |
4      .##### + rb2 rb2
      ##### | ba2 ba2 rb2
      ##### | ba1 ba1 ba1 ba1 ba2 rb1
      .##### | ba1 ba1 rb1 rb2 rb2
3      ##### + b11 ba1 rb1
      .##### | b21 b22 ba1
      .### M| rb2 rb2 rb2 rb2
      ##### S| S ba1 ba1 ba2 rb2
2      ##### + rb2
      ## | ba1 ba1 ba1 rb2 rb2
      .### | b11 ba1 ba1 ba1
      # | ba1 men rb2
1      . + ba1 rb2 rb2
      ## | ba1
      .## | ba1 ba1
      ## S|
0      .## +M ba1 ba1 ba1 ba1 ba1 ba1 ba1 ba1 ba1 ba2
      . | ba1 ba1 ba1 ba1 ba1 ba1 ba1 ba1 ba1
      . | ba1 ba1
-1     # + ba1 ba1 ba1 ba1 ba1 ba1 ba1
      | ba0 ba1 ba1 ba1 ba1 ba1
      . | ba1 ba1 ba1 ba1 ba1 ba1 ba1 ba1 ba1
      .# | ba1 ba1 ba1
-2     # T+ ba1
      # | S ba0 ba1 ba1 ba1
      # | ba0 ba0 ba0 ba0 ba0 ba1 ba1 men
      | ba0 ba1
-3     .# + ba0 ba0
      # | ba0 ba0 ba1
      # | ba0 ba0 ba0 ba1
      | ba0 ba0
-4     # + ba0
      # | ba0 ba1
      |
-5     | T ba0
      +
      #
      |
-6     +
      <less>|<frequ>
EACH "#" IS 2. EACH "." IS 1.

```



| PERSON - MAP - ITEM |              | <more> <rare> |          |
|---------------------|--------------|---------------|----------|
| 4                   |              | +             |          |
|                     | X            |               |          |
| 3                   |              | +             |          |
|                     | T            |               |          |
|                     | XXXXX        | T             | sc_b430  |
|                     |              |               | sc_ba416 |
|                     | XX           |               | sc_rb409 |
| 2                   | XXXXXXXX     | +             | sc_ba415 |
|                     | XXXXXXXX     |               |          |
|                     | XXXXXXXX     |               |          |
|                     | XXXXXXXX     | S             | sc_ba321 |
|                     | XXXXXXXX     |               |          |
|                     | XXXX         |               |          |
| 1                   | XXXXXXXX     | +             | sc_rb410 |
|                     | XXXXXXXXXXXX |               | sc_ba513 |
|                     | XXX          |               |          |
|                     | XXXXXXXX     |               | sc_ba512 |
|                     | XXXXXX       |               | sc_b420  |
|                     | XXXXXX       | M             | sc_rb309 |
| 0                   | XXXX         | +M            | sc_b313  |
|                     | XXXXX        |               | sc_rb503 |
|                     | XXXXXXXXXXXX |               | sc_b431  |
|                     | XX           |               | sc_rb404 |
|                     | XXXXXXXX     |               | sc_rb402 |
|                     | XXXX         |               | sc_b43   |
|                     | XXXX         |               | sc_rb504 |
| -1                  | XXXXXXXXXXXX | +             |          |
|                     | XXXXXXXXXXXX | S             |          |
|                     | XXXXX        |               | sc_rb422 |
|                     | XXXXXXXXXXXX |               | sc_rb301 |
|                     |              |               | sc_rb424 |
|                     | XXXXXX       |               | sc_b312  |
| -2                  |              | +             | sc_rb302 |
|                     | X            |               | sc_ba418 |
|                     |              | T             |          |
|                     | XX           |               |          |
|                     | X            |               |          |
| -3                  |              | +             |          |
|                     | X            |               |          |
| -4                  |              | +             |          |
|                     | <less>       |               | <frequ>  |

```

PERSON - MAP - ITEM
<more>|<rare>
4
+
.
3
+
sc_rb60 sc_rb60 sc_rb61 sc_rb61 sc_rb61
T
2
+
.
#
.# T
# sc_rb52
.# S
1
+
#
### S
. sc_rb52
#####
##### sc_ba41
.##### sc_rb41 sc_rb52
.##### sc_rb51
##### sc_b430 sc_rb51
### M sc_rb40 sc_rb42 sc_rb50
0
##### +M sc_ba51 sc_ba51 sc_rb51
## sc_rb40 sc_rb61
#####
.#
#### sc_b431 sc_b48 sc_rb40 sc_rb50 sc_rb51 sc_rb51
#### sc_b420 sc_rb40 sc_rb50
### S sc_ba41
### sc_b421 sc_rb51 sc_rb61
.###
## sc_rb41 sc_rb51 sc_rb51
-1
+
.# S sc_ba41
.# sc_b410 sc_b43 sc_rb50 sc_rb50
T
.
.
sc_rb50
-2
+
<less>|<frequ>
EACH "#" IS 2. EACH "." IS 1.

```





```

PERSON - MAP - ITEM
<more>|<rare>
3
+
#
.
.
2
.# +
.
.
T
.# sc_rb8
.###
.
## | T
.###
## S | sc_rb8
.#### sc_rb8
1
.##### +
##### | sc_rb8 sc_rb8 sc_rb8 sc_rb8
##### | sc_rb8
.#### | sc_rb8
.##### S | sc_rb8 sc_rb8 sc_rb8
.##### M
##### | sc_rb8
.##### | sc_rb7 sc_rb8
.### | sc_rb7 sc_rb8 sc_rb8 sc_rb8
#### | sc_rb7 sc_rb8 sc_rb8 sc_rb8 sc_rb8 sc_rb8 sc_rb8
.##### | sc_rb8 sc_rb8 sc_rb8 sc_rb8 sc_rb8
##### | sc_rb8 sc_rb8 sc_rb8
0
##### +M | sc_rb8 sc_rb8 sc_rb8
.### S | sc_rb7 sc_rb8 sc_rb8 sc_rb8 sc_rb8 sc_rb8 sc_rb8
.## | sc_rb7 sc_rb8
.#
##
.# | sc_rb8 sc_rb8
# | sc_rb7
. | sc_rb7
. T S | sc_rb8
.# | sc_rb8 sc_rb8 sc_rb8 sc_rb8
# | sc_rb8 sc_rb8
-1
+ | sc_rb8
| sc_rb8 sc_rb8 sc_rb8
.
. | sc_rb8 sc_rb8
. T
|
| sc_rb8
-2
+
<less>|<frequ>
EACH "#" IS 2. EACH "." IS 1.

```

| PERSON - MAP - ITEM |              |         |          |          |          |
|---------------------|--------------|---------|----------|----------|----------|
|                     | <more>       | <rare>  |          |          |          |
| 3                   |              |         |          |          |          |
|                     | X            |         |          |          |          |
| 2                   |              |         |          |          |          |
|                     | XXX          |         |          |          |          |
|                     | X            | T       |          |          |          |
|                     | XXXXX        |         |          |          |          |
|                     | XXXXX        |         |          |          |          |
|                     | X            | T       | sc_rb842 | sc_rb936 |          |
|                     | XXXXX        |         |          |          |          |
|                     | XXX          | S       |          |          |          |
|                     | XXXXX        |         | sc_rb938 |          |          |
|                     | XXXXXX       |         |          |          |          |
| 1                   | XXXXXX       |         | sc_rb939 |          |          |
|                     | XXXXXXXX     |         | sc_rb922 | sc_rb943 |          |
|                     | XXXXXXXXXXXX |         | sc_rb945 |          |          |
|                     | XXXXXXXX     |         | sc_rb919 |          |          |
|                     | XXXXXXXXXXXX | S       | sc_rb912 | sc_rb935 | sc_rb944 |
|                     | XXXXXXXXXXXX | M       | sc_rb937 |          | sc_rb946 |
|                     | XXXXXXXXXXXX |         | sc_rb910 | sc_rb914 | sc_rb920 |
|                     | XXXXXXXX     |         | sc_rb841 | sc_rb907 | sc_rb921 |
|                     | XXXXXX       |         | sc_rb918 | sc_rb933 |          |
|                     | XXXXXXXX     |         | sc_rb845 | sc_rb901 |          |
|                     | XXXXXXXX     |         | sc_rb847 | sc_rb942 |          |
|                     | XXXXX        |         | sc_rb911 |          |          |
| 0                   | XXX          | S+M     | sc_rb905 | sc_rb913 | sc_rb928 |
|                     | XXX          |         | sc_rb904 | sc_rb908 | sc_rb909 |
|                     | X            |         | sc_rb846 | sc_rb848 | sc_rb915 |
|                     | XX           |         | sc_rb814 | sc_rb903 | sc_rb906 |
|                     | XX           |         | sc_rb934 |          | sc_rb932 |
|                     | XXX          |         | sc_rb849 | sc_rb902 | sc_rb926 |
|                     | XX           |         | sc_rb850 |          | sc_rb930 |
|                     | X            | T       | sc_rb813 |          |          |
|                     |              | S       |          |          |          |
|                     | XXX          |         | sc_rb812 |          |          |
|                     | XX           |         | sc_rb843 | sc_rb844 | sc_rb916 |
|                     | X            |         | sc_rb931 |          |          |
| -1                  |              |         | sc_rb924 |          |          |
|                     |              |         | sc_rb929 |          |          |
|                     |              |         | sc_rb815 |          |          |
|                     |              |         |          | sc_rb927 |          |
|                     |              | T       |          | sc_rb917 |          |
|                     |              |         |          | sc_rb925 |          |
| -2                  |              |         |          |          |          |
|                     | <less>       | <frequ> |          |          |          |

# GLOSSARY

|                                     |                                                                                                                                                                   |
|-------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Anchor items                        | Test items administered in combination with two or more problem sets to connect the tests so that their scores are comparable                                     |
| Assessment framework                | A framework that explains the detailed content of the assessment                                                                                                  |
| Bloom's Taxonomy                    | A set of hierarchical models used to classify educational learning objectives into levels of complexity and specificity                                           |
| Classical test                      | A test administered in groups, commonly involving all students in one classroom                                                                                   |
| Closed constructed-response problem | A question similar to the more traditional multiple-choice type of questions and having only one right answer                                                     |
| Cognitive domain                    | A domain that involves the cognitive skills and the acquisition of knowledge linked to the subject                                                                |
| Content domain                      | The body of knowledge, skills, or abilities being measured by the student learning assessment tool                                                                |
| Graded response model (GRM)         | A model to analyse whether a dichotomous response is correct                                                                                                      |
| Highstakes test                     | A test that has major consequences for the test taker                                                                                                             |
| Horizontal equating strategy        | A strategy that refers to the equating of test scores administered to groups with similar abilities (e.g., different tests for students in the same grade or age) |
| Infit                               | Inlier-sensitive fit that shows how sensitive the pattern of the response to the item that detects the ability of the test taker                                  |
| Latent ability                      | Individual's actual ability                                                                                                                                       |
| Lowstakes test                      | A test with less important consequences for the test taker                                                                                                        |
| Lowerorder thinking skills          | A list of thinking skills which include remembering and understanding, and are generally attained by rote memorization                                            |
| Middleorder thinking skills         | A list of thinking skills which include applying and analyzing, and are generally attained after accomplishing lower-order thinking skills                        |
| Numeracy staircase                  | A set of numeracy abilities divided into several different levels                                                                                                 |
| Open constructed-response problem   | A test problem that requires the test taker to answer a constructed response                                                                                      |
| Outfit                              | Outlier-sensitive fit where more difficult items are more sensitive to people with lower ability, and vice versa                                                  |

|                                    |                                                                                                                                              |
|------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------|
| Partial credit model (PCM)         | A model used to analyse polytomous response, where item has its own scale                                                                    |
| Principal component analysis (PCA) | A statistical procedure that converts a set of observations of possibly correlated variables into a set of values of uncorrelated variables  |
| Procedural knowledge               | The explicit knowledge that is required in performing straightforward tasks                                                                  |
| Psychometric properties            | Properties that refer to the reliability and validity of instrument (e.g., difficulty level, discrimination power, pseudoguessing parameter) |
| Rasch model                        | A family of psychometric models for creating measures from categorical data that has difficulty level as the main parameter                  |
| Reliability                        | Quality of the test that shows how consistent it is for a particular ability                                                                 |
| Teaching to test                   | A practice of teaching the purpose of which is answer test problems correctly                                                                |
| Text level gradient                | A gradual change of difficulties of the text, including number of difficult words, length, and genre                                         |
| Validity                           | The quality of a test that shows how accurately it assesses the intended ability                                                             |
| Vertical continuity                | The continuous properties of a scale owned by different levels of tests to make them comparable                                              |
| Vertical equating strategy         | A strategy that refers to the equating of test scores administered to groups with similar abilities                                          |
| Wright map (item-person map)       | A map that juxtaposes the item's difficulty level and the person's ability level                                                             |

## **The SMERU Research Institute**

Telephone : +62 21 3193 6336

Fax : +62 21 3193 0850

E-mail : [smeru@smeru.or.id](mailto:smeru@smeru.or.id)

Website : [www.smeru.or.id](http://www.smeru.or.id)

Facebook : [@SMERUInstitute](https://www.facebook.com/SMERUInstitute)

Twitter : [@SMERUInstitute](https://twitter.com/SMERUInstitute)

YouTube : [The SMERU Research Institute](https://www.youtube.com/TheSMERUResearchInstitute)

