INOVASI
Inovasi untuk Anak Sekolah Indonesia
Kemitraan Australia Indonesia

Kementerian PPN/ Bappenas

Australian Government

THE LEARNING GAP SERIES – ONE

Beyond letters and numbers: the COVID-19 pandemic and foundational literacy and numeracy in Indonesia

THE LEARNING GAP SERIES – ONE
Beyond letters and numbers: the COVID-19 pandemic and foundational literacy and numeracy in Indonesia

Jeaniene Spink
Dan Cloney
Amy Berry

Published by INOVASI in 2022

Disclaimer:

INOVASI – Innovation for Indonesia's School Children
Ratu Plaza Office Tower 19th Floor,
Jl. Jend. Sudirman Kav 9, Jakarta Pusat, 10270
Indonesia
Tel : (+6221) 720 6616
Fax : (+6221) 720 6616
http://www.inovasi.or.id

The governments of Australia and Indonesia are partnering through the Innovation for Indonesia's School Children (INOVASI) program.

INOVASI is an Australia–Indonesia Government Partnership – Managed by Palladium.

THE LEARNING GAP SERIES – ONE
Beyond letters and numbers: the COVID-19 pandemic and foundational literacy and numeracy in Indonesia

April 2022

Table of Contents

List of Tables

List of Figures

Abbreviations and acronyms

ACER	Australian Council for Educational Research
AKM	*Asesmen Kompetensi Minimum* (Indonesian Minimum Competency Assessment)
GPF	Global Proficiency Framework
IRT	Item Response Theory
INOVASI	Innovation for Indonesia's School Children
KK	*Kurikulum Kombinasi* (Combined Curriculum)
K–13	*Kurikulum 2013* (Curriculum 2013)
MELQO	Measuring Early Learning Quality and Outcomes
MoEC	Ministry of Education and Culture
MoECRT	Ministry of Education, Culture, Research and Technology
MoRA	Ministry of Religious Affairs
MI	*Madrasah Ibtidaiyah* (Elementary Madrasah)
MPL	Minimum Proficiency Level
OECD	Organisation for Economic Co-operation and Development
OLS	Ordinary Least Squares
PAUD	*Pendidikan Anak Usia Dini* (Early Childhood Education and Development)
PSKP	*Pusat Standar dan Kebijakan Pendidikan* (Education Standards and Policy Centre)
SDG	Sustainable Development Goal
SD	*Sekolah Dasar* (Elementary School)
SES	Socioeconomic Status
SLA	Student Learning Assessment

Acknowledgements

The research project on which this report draws is a collaboration between INOVASI, the Education Standards and Policy Centre (PSKP – *Pusat Standar dan Kebijakan Pendidikan*) of the Indonesian Ministry of Education, Culture, Research and Technology (MoECRT) and ACER. The project was co-designed and INOVASI oversaw the project implementation. ACER undertook the analysis and authored the report for this study, which relates to students' learning during the school closure period due to COVID-19, while INOVASI developed a separate analysis of learning and participation loss from a subgroup of the schools in the project for which the program had pre-COVID-19 baseline data.

ACER would like to acknowledge the support of INOVASI (Rasita Purba, Mary Fearnley-Sander, George Adam Sukoco, Wahyu Widhiarso, and Senza Arsendy) and the MoECRT PSKP team to this study; they contributed technical assistance to instrument development and psychometric analysis, and to the quality assurance of the study.

Executive summary

In 2020, with the COVID-19 pandemic spreading across Indonesia and around the world, INOVASI (the Innovation for Indonesia's School Children) [1] and the Indonesian Ministry of Education, Culture, Research and Technology (MoECRT) initiated a study of foundational literacy and numeracy learning by Indonesian students and the impact of the pandemic on this learning. The Australian Council for Educational Research (ACER)[2] provided technical support. The study developed descriptions of literacy and numeracy achievement levels, referencing both global proficiency frameworks and Indonesia's curriculum and assessment standards. Student results were benchmarked against descriptors of skills for the Global Proficiency Framework (GPF) Minimum Proficiency Levels (MPLs).[3]

The study aims to identify learning gaps between what the standards set for students to learn and actual student attainment. The findings of this study provide clear directions for taking necessary actions to optimise students' learning recovery, in addition to designing and implementing school curricula in Indonesia. This study is the first series of INOVASI and PSKP study on learning gap. The second series is a study of the effects of curriculum and its interaction with factors affecting delivery within different regions of Indonesia, including a case of learning loss in INOVASI partner schools, is presented in a separate report in the Learning Gap Study series. Specific gender-based analysis on learning loss, experienced by girls and boys, and how they differ, is described in the third Learning Gap series, exploring gender, disability and social inclusion lenses implicated by the pandemic. The last series describes teacher and school capacities in teaching during the pandemic.

The key data elements for the study were (1) data from literacy and numeracy tests administered to students in Grades 1, 2 and 3; and (2) survey responses from principals, teachers, and parents of the sampled students. The survey provided contextual data on the students' experience of learning the during the school closure.

The instrument used for measuring students' literacy and numeracy attainment was the Student Learning Assessment (SLA)[4], previously used by the INOVASI team and psychometrically reviewed by ACER. To compare the results of the SLA with the GPF MPLs, the skills measured by the SLA were mapped against the international framework. Student

[1] INOVASI is a partnership program between the Government of Australia and Indonesia, working directly with the Ministry of Education and Culture, Research and Technology. The program is seeking to understand how learning outcomes in foundational literacy and numeracy can be improved (https://www.inovasi.or.id/en/)
[2] ACER is an independent, a non-profit educational research organisation (https://www.acer.org/id/)
[3] The GPF MPL outlines minimum proficiency levels children are expected to obtain at the end of each of grades for both reading and mathematics (https://gaml.uis.unesco.org/wp-content/uploads/sites/2/2019/05/GAML6-REF-16-GLOBAL-PROFICIENCY-FRAMEWORK.pdf)
[4] The SLA is literacy and numeracy student learning assessment tools. The SLA covers math and Indonesian language tests for students in lower and upper grades, developed by INOVASI in collaboration with KIAT Guru (*Kinerja dan Akuntabilitas Guru* or Teachers' Performance and Accountability, a TNP2K - National Team to Accelerate Poverty and World Bank project). The literacy and numeracy SLA was constructed, following TIMSS Numeracy and PIRLS Literacy frameworks, Indonesia K-13 curriculum, using items from KIAT Guru Project, the Indonesian National Assessment Program (INAP) and some items adapted from MoEC's electronic textbooks. For the purpose of this learning gap study, SLA has been further adjusted to integrate several items from GPF and AKM Kelas (the Indonesia's Minimum Competencies Assessment for classroom use).

results from the SLA tests were reported as proportions of students attaining a particular level of proficiency in literacy and numeracy.

The study covered more than 18,000 students in Grades 1–3 in 19 districts across eight provinces, representing the western and eastern regions of Indonesia. The data were collected in mid-2021, which was after 12 months of school closures.

This study points out several **high-level findings:**

1. Most students did not have foundational literacy and numeracy skills. However, with additional support, a large number of students may be able to achieve these standards.

Only 16% and 32% of students in Grades 2 and 3 met the numeracy GPF MPLs, while for literacy the number was higher (39% and 55%, respectively). While there was no significant gender gap in numeracy, there was a gender gap in favour of girls in literacy; the proportion of girls meeting the GPF MPLs was significantly higher than that of boys.

The study discovered that approximately one out of 10 Grade 2 and 3 students were still performing at pre-primary levels in literacy. For numeracy, the figure was approximately two out of 10. These children have not yet acquired the essential foundation skills to learn at primary schools.

Conversely, results show the average performance of students increasing year on year, demonstrating the value-add of schools and teaching. This finding focuses attention on the possibility that, with additional remedial support targeted to areas of difficulty, another 40% of Grade 2 and Grade 3 students should be able to meet the GPF MPLs standard.

2. There is a mismatch in the sequencing of curricula and assessment frameworks against global frameworks and student abilities.

Attempts were made to map the GPF against the competencies specified in the 2013 Curriculum, Emergency Curriculum and Indonesia's Minimum Competencies Assessment (AKM) for classroom use. This mapping was possible in the mathematics curricula but not for literacy in the Bahasa Indonesia curricular frameworks, highlighting insufficient articulation of literacy competencies. For both literacy and numeracy, mapping was possible in the AKM.

In literacy, the 2013 curriculum framework for Bahasa Indonesia does not articulate the essential skills for reading literacy that students need to demonstrate to become good readers: listening, decoding, and reading comprehension, which comprises retrieving information, interpreting, and reflecting on text. While the AKM identifies skills related to reading comprehension, it does not include skills in the domains of decoding and listening comprehension. The emphasis on reading comprehension (rather than listening comprehension and decoding) in the AKM at Levels 1 and 2 means the expectations for minimum competency in the early grades may exceed the global expectations for what early readers can do. Students who are likely to become effective independent readers typically demonstrate listening comprehension skills in Grades 1–3 that are well in advance of their reading comprehension skills (typically two years ahead). If students are not yet able to comprehensively process a short piece of oral text, it is likely that in reading they focus on matching words in the text to the comprehension question, while paying little attention to the overall meaning of the text.

In numeracy, where it was possible to align the Indonesian curricula with the GPF MPLs, some of the expectations in the mathematics curriculum appear to be too high compared to global standards. For instance, (i) computing the addition of numbers up to 20 is a competency that must be mastered by Grade 1 students in the Indonesia Emergency Curriculum and the AKM, but by Grade 2 in the Sustainable Development Goal (SDG); and (ii) describing and determining the relationship between standardized units (e.g., kg, g, m, cm) is a competency that must be mastered by Grade 3 students based on Indonesia's Emergency Curriculum and the AKM, but it is allocated to Grade 6 level in the SDG.

3. The factors most correlated with learning outcomes, including family and school factors, indicate the role that inequality plays in outcomes.

Family factors strongly and positively associated with learning outcomes are mothers' fluency in Bahasa Indonesia, mothers' education level (secondary education or higher), household expenditure, households with connectivity and computers used for leaning activities, books in the home, children's preschool experience, living in a developed area (most strongly correlated with literacy outcomes), and a child feeling supported at home in study. **School dominant factors** for positive learning outcomes are the status of the school as a government school, teachers' internet and computer access, and teachers having a four-year teaching qualification.

Findings from this study provide important **insights and recommendations** for national, district, and school policy, in support of students' learning recovery and continued learning. These fall into two areas:

a. Equity and inclusion. The study revealed a significant spread of student proficiencies in literacy and numeracy associated with a range of disadvantages for a significant proportion of students. The disadvantages have been greatly extended by COVID-19 and school closure. As schools reopen, the achievement gap may become even more pronounced.

Strategies that enable a more targeted system of support to significantly improve learning outcomes for all students include the following:

(i) investing in underserved areas – those less developed areas that indicate gaps in infrastructures, facilities, and resources for public services, including education;

(ii) supporting schools to improve the extent and quality of engagement with families and local communities; and

(iii) identifying the more marginalised students in schools and developing and implementing learning assistance programs.

b. Systemic support for targeted teaching. The study found that teachers who adopted the numeracy and literacy modules, developed by MoECRT for the Emergency Curriculum, had better student results. This may be in part because the modules clearly defined skills and learning objectives against each level of learning and provided additional support for teachers on how to teach these skills.

Observations about the success of the literacy modules and observations about weaknesses in the existing curricular frameworks highlight the need for the following:

(i) greater attention to, and understanding of, what students can do (current levels of attainment);

(ii) clear identification of next steps in learning (ideally on an individual student basis); and

(iii) resource and professional development support to enable targeting teaching according to student needs.

Targeting teaching to the point of student needs requires a systematic approach that sets clear expectations for learning and measures progress and attainment through a system of continuous classroom assessment. Curricular frameworks should set out the knowledge, understanding, and skills that students should attain as they progress through school, particularly in the early years for literacy and numeracy. Of equal importance is the need to develop teachers' knowledge and capability of teaching and monitoring progress in these essential areas of learning.

1. Introduction

The INOVASI team undertook a situational analysis to better understand the status of students' learning and the potential impact of COVID-19 in Indonesia. A student test of literacy and numeracy levels was administered to students in Grades 1, 2 and 3. Teachers, principals, and parents of the sampled students were also surveyed to provide contextual data on factors relevant to students' experience of learning during school closure and to their association with the test performance. The study intends to inform timely curricular, assessment, and pedagogical recalibration as students return to school in the aftermath of the COVID-19 pandemic and to identify the most pressing areas for teacher professional development in delivering instruction and assessing at the right level so students can learn essential competencies they may have missed.

The study benchmarked the findings from the student assessment against national and global points of reference, including the following:

- *Global alignment:* The Global Proficiency Framework (GPF; USAID, 2019) can be used as a foundation for linking assessments to the MPLs of the Sustainable Development Goals (SDGs) against SDG 4.1.1. The GPF's main purpose is to provide a common set of descriptors for aligning and interpreting grade-level student performance in reading and mathematics. The proficiency levels are purposely set low so as to be achieved by most students in the Grade 2/3 band.

- *Local alignment:* Results were aligned across the various national frameworks in Indonesia including the Indonesian Minimum Competency Assessment (AKM) levels, which were specifically designed for classroom diagnostic use, and the Curriculum 2013 (K-13) and the Emergency Curriculum.

Benchmarking student results against these external frameworks made it possible to report gaps in learning achievements against expected curricular or global standards. Students' socio-emotional and executive function skills were also analysed as were students' background factors, including those related to the student, family, school, and community, to determine what effect these may have on student learning.

The study highlights several high-level findings that warrant further investigation:

1. gaps in student foundational competencies in literacy and numeracy;
2. spread of student abilities, highlighting issues of equity in learning outcomes;
3. mismatches in curricular sequencing against global frameworks and student abilities;
4. higher targeting of student standards in the Indonesian national curriculum, compared with international proficiency benchmarks; and
5. important student, family and school factors that affect student learning outcomes.

Findings from this study provide important insights into student learning for national policy and explain how these policies apply to districts, schools and teachers as they facilitate the transition back to school and support students' continued success in learning beyond COVID-19.

2. Scope and method

2.1. Study participants

The study assessed the learning proficiencies in literacy and numeracy of students in Grades 1, 2 and 3 in eight provinces across Indonesia. A total of 18,370 students participated in the study from across 612 randomly selected schools. Of the 18,370 students total, 6094 were children with disabilities who were evenly distributed across the grades.

Given the grade levels included in the study, students were assessed one-on-one.

Figure 1: number of study participants by grade, gender, and disability

Study participants by gender			
Grade levels	**Girl**	**Boy**	**Total**
Grade 1	3060	3068	6128
Grade 2	3066	3053	6119
Grade 3	3063	3060	6123
TOTAL	9189	9181	18 370
Study participants with disability[1]			
	Non-physical disability	**Physical disability**	
TOTAL	5119	974	6094

Approximately the same number of students participated in the study by grade, with an almost even distribution between girls and boys. The same number of students participated in the mathematics and the literacy assessment.[5]

The study includes a representative sample of students from the 11 INOVASI districts in the provinces of West Nusa Tenggara, East Java, North Kalimantan, and East Nusa Tenggara. To provide coverage and balance across aspects of Indonesia's education system, an additional eight non-INOVASI partner districts were added. Given the spread of the study participants, the results provide important insights into learning gaps of students in the early years.

[1] Student disability status was assessed through parent responses using adjusted Washington Group Child Functioning Module questions. Students who experienced 'some difficulty', 'a lot of difficulty', and/or 'cannot do at all' in one or more functional domains were categorised as student with disabilities and were further categorised into two sub-types: physical and non-physical disabilities. Physical disabilities included difficulties in the domains of: seeing, hearing, walking, and speaking. Non-physical disabilities included difficulties in the domains of: learning, remembering, and focusing attention. Some child functioning module domains that were not included are self-care, behaviour, accepting change, and making friends. Source: https://www.washingtongroup-disability.com/question-sets/.

[5] Ten students from each of the three target grades were randomly sampled from each participating school. The study allows for reporting at the district level.

Figure 2: number of students per province

Province	Students
Jambi	2340
East Java	3186
South Kalimantan	2524
North Kalimantan	1999
North Maluku	692
West Nusa Tenggara	4201
East Nusa Tenggara	1777
Southeast Sulawesi	1651
TOTAL	18 370

Figure 3: number of schools in the study (by type)

School Type	Private	Public
Elementary School (SD)	62	433
Elementary Madrasah (MI)	99	18
TOTAL	161	451

2.2. Test measures

The study used the Student Learning Assessment (SLA) instrument to measure student literacy and numeracy abilities. The instrument included a set of items that measured students' Bahasa Indonesia and mathematical ability, as well as a set of items to measure students' socio-emotional and executive function skills. The instrument was previously tested by the INOVASI team and psychometrically reviewed by ACER. With some adjustments, it was found to have sufficient psychometric properties for use in this learning gap study. The SLA was adapted to include additional items from the global item pool to enable the comparison of results against global proficiency benchmarks. The instrument also included the addition of executive function items that were drawn from the Measuring Early Learning Quality and Outcomes (MELQO) instruments. The assessment was administered orally, using one-on-one test administration procedures. A screening test was first administered to determine at which level of test the student should participate to ensure better targeting of test instruments. All assessments were conducted in Bahasa Indonesia.

2.3. Analytical approach

The learning gap study followed three stages of analysis for benchmarking (see Annex B for a summary of the research method). The first stage of the study focused on conceptually benchmarking the skills measured by the SLA test instrument to GPF grade level skills and to local proficiency expectations. These local proficiencies are those of the K-13, the Emergency Curriculum, a reduced curriculum scope introduced during school closure, and the national

assessment framework, the AKM. The purpose of this exercise was to report student results against the GPF and national curricula.

The second stage included constructing common scales of student performance in Bahasa Indonesia and mathematics, and mapping student results from the SLA onto these common scales. The development of these scales allowed for the ranking of student performance from the weakest to the most able of students. A set of common scales enabled the comparison of learning outcomes accurately and reliably between population subgroups. As part of this process, a series of cut points against these empirical scales was established to create levels of student proficiency, as tested by the SLA instrument. A set of proficiency descriptors was developed for each of the bands created against the scale to describe what students can do at that level.

The third stage included reporting the proportion of students in each of these proficiency levels. These actual proficiency levels were then benchmarked against the global and local proficiency levels to show where students in the study sat in relation to the performance expected in those frameworks, using the analysis from the first stage of the study.

> *It is now possible to report the proportion of students (included in this study) who are at or above minimum proficiency expectations in their first three years of school. The study provides a deeper appreciation of the equity of student learning outcomes across population groups and reveals important new insights into the locations of learning gaps and suggests how these can be addressed.*

3. SLA proficiency descriptors and alignment to external benchmarks

The SLA instrument was developed prior to the implementation of the learning gap study. While the instruments provide a good measure of student performance in literacy and numeracy, the assessment was not specifically designed to assess grade level expectations as determined by national curricular frameworks or international benchmarks of performance. While alignment to the mathematics GPF with the SLA was reasonable, and some indication of performance by grade level is possible, alignment to the literacy domain was less clear and not as well defined. An exact determination of the proportion of students not performing at grade level, therefore, is not possible using the existing SLA instrument. For a stronger alignment with the GPF, a wider range of core skills would need to have been assessed. Because the area of number and operations was the dominant focus of the SLA, this was selected as the focus for describing the different levels of proficiency and mapping them against the curricula frameworks. The one exception is at Level 4, where the measurement domain is mentioned in the proficiency level descriptor, so a connection to the measurement descriptors in the K-13 and the Emergency Curriculum frameworks was included. Noting SLA limitations and after exploring other possible alternatives, it was decided to use SLA for this study with following reasons: (1) there is a continuity of data collected prior and during the pandemic using SLA; (2) the current version of SLA has incorporated some of GPF and AKM items; (3) SLA does not require a lengthy test which is especially important element in conducting literacy and numeracy tests for early grade students during the pandemic.

One of the challenges of mapping the descriptors from the curricula frameworks (e.g., what students are expected to know or do in relation to number at Grade 1) against the descriptions of the different levels of proficiency in the SLA (e.g., what students at Proficiency Level 1 are able to do in relation to number) is that the skills demonstrated by students in a given SLA proficiency level often span across multiple grade levels in the curriculum. For example, at Level 3, the students can read numbers up to 1,000 (roughly aligned with Grade 3 of K-13), are able to identify common fractions represented as objects or pictures (Grade 2 of K-13) and can add and subtract within 100 (Grade 1 of K-13). It is also noted that students at Level 1 have only partially met the curriculum expectations for Grade 1 in both of the curricula frameworks because they have only demonstrated their knowledge of numbers up to 20 rather than 99 (K-13 Curriculum Grade 1) and 50 (Emergency Curriculum Grade 1).

It remains unclear with which level of the AKM Level 4 might align. The focus of the mapping was on the Grades 1–3 curriculum descriptors from the K-13 and Emergency Curriculum, and the descriptors from Levels 1 and 2 of the AKM. Given that Level 4 (from the SLA) is the highest level of proficiency and includes skills that would be considered above Grade 3 in the GPF, it is possible that Proficiency Level 4 might align with Level 3 or 4 of the AKM. Based on the documentation used in the mapping, it appears that Proficiency Level 4 sits beyond AKM Level 2.

3.1. Proficiency descriptors and alignment to external benchmarks for SLA mathematics

The learning gap study developed four levels of proficiency for mathematics, based on the results from the SLA. Level 1 is the easiest level and Level 4 the most difficult. Each of these levels was aligned (where possible) to the national curricular standards for Grades 1–3, the AKM Levels 1 and 2, and the GPF for Grades 1–3. Table 1 outlines these descriptors and presents the approximate national curricula alignment to each level. The Grade 2 or Grade 3 MPL for mathematics was assessed as being at Level 4 and above.

Table 1: SLA proficiency descriptors for Mathematics

Proficiency Level	Description of skills	K-13 Curriculum	Emergency Curriculum	AKM
Level 4 and above	Students at this level can identify and express proper fractions as equivalent fractions and use place value concepts for hundreds, tens, and ones. They can add and subtract within 1,000 and multiply and divide a two-digit number by a one-digit number. They can apply their understanding of multiplication to solve simple real-world problems involving whole numbers to 5. Students understand standard units of measurement for length and weight and can make conversions between adjacent units, such as metres and kilometres.	Grade 2 (Number) Grade 3 (Measurement)	Grade 2 (Number) Grade 3 (Measurement)	Unclear which level of the AKM this equates to*
Minimum Proficiency Level (MPL) for Grade 2/3 against SDG indicator 4.1.1a				
Level 3	Students at this level can read whole numbers up to 1,000 in numerals, identify everyday unit fractions such as ½ and ¼ when represented as objects or pictures, and use place value concepts for tens and ones. They can add and subtract within 100 and multiply and divide within 25.	Grade 1 (Number) Grade 2 (Number) Grade 3 (Number)	Grade 2 (Number) Grade 3 (Number)	Level 1 (Number) Level 2 (Number)
Level 2	Students at this level can read, compare, and order whole numbers up to 100 in numerals and compare whole quantities up to 30 represented as objects, pictures, and numerals. They can add and subtract within 20 and apply their understanding of addition and subtraction to solve simple real-world problems within 10.	Grade 1 (Number)	Grade 1 (Number)	Level 1 (Number)

Level 1 and below	Students at this level can read and compare whole numbers up to 20 in numerals and whole quantities up to 5 that are represented as objects, pictures, and numerals.	Below Grade 1** (Number)	Below Grade 1 ** (Number)	Level 1

*The skills at Proficiency Level 4 appear to be beyond those described at Levels 1 and 2 of the AKM.

** The K-13 Curriculum refers to knowing whole numbers up to 99 and the Emergency Curriculum expects whole numbers up to 50. Therefore, students at this level of proficiency have not yet achieved the curricular expectations for Grade 1.

3.2. Proficiency descriptors and alignment to external benchmarks for SLA Bahasa Indonesia

The learning gap study developed three levels of proficiency for literacy outlined in Table 2. Level 1 is the easiest level and Level 3 the most difficult. Each of these levels was aligned to the AKM and the GPF.

Table 2: SLA proficiency descriptors for Bahasa Indonesia

Proficiency Level	Description of skills	AKM
Level 3 and above	When reading independently, students can identify the main idea in texts of several paragraphs. When listening to texts, they can recall some directly stated information and make some simple inferences. When a short sentence is dictated to them, students can use correct spelling and basic punctuation.	Unclear which level of the AKM this equates to*
	Minimum Proficiency Level (MPL) for Grade 2/3 against SDG indicator 4.1.1a	
Level 2	When reading independently, students can, in texts of about one paragraph, make simple inferences across the text. They can locate information when synonymous matching is required or when there is some competing information to discount. When listening to very short texts, students can locate information using close or synonymous matching and make simple inferences. When a short sentence is dictated to them, students can spell a small number of words correctly.	Level 1 or Level 2 depending on text level
Level 1 and below	When reading independently, students can, in texts of about one paragraph, locate information using close matching when there is no competing information to discount. They can make inferences across adjacent sentences. When listening to texts, they can recall directly stated information from the start of the text. They use spaces between words when writing a sentence. They are developing their vocabulary and can name familiar objects when presented with a picture.	Level 1 or Level 2 depending on text level

*The skills at Proficiency Level 3 appear to be beyond those described at Levels 1 and 2 of the AKM.

It was not possible to align the literacy items in the SLA with the Emergency or K-13 Curriculum frameworks for three main reasons. In some instances, the descriptions from the curriculum did not clearly articulate what the student was expected to do (e.g., explaining the preparation

activities for early reading in the right way). In other cases, the descriptions from the curriculum related to very specific activities involving applying literacy skills to specific domains or content areas (e.g., dig up information about the sources and forms of energy presented in verbal, written, visual and/or exploration of the environment).

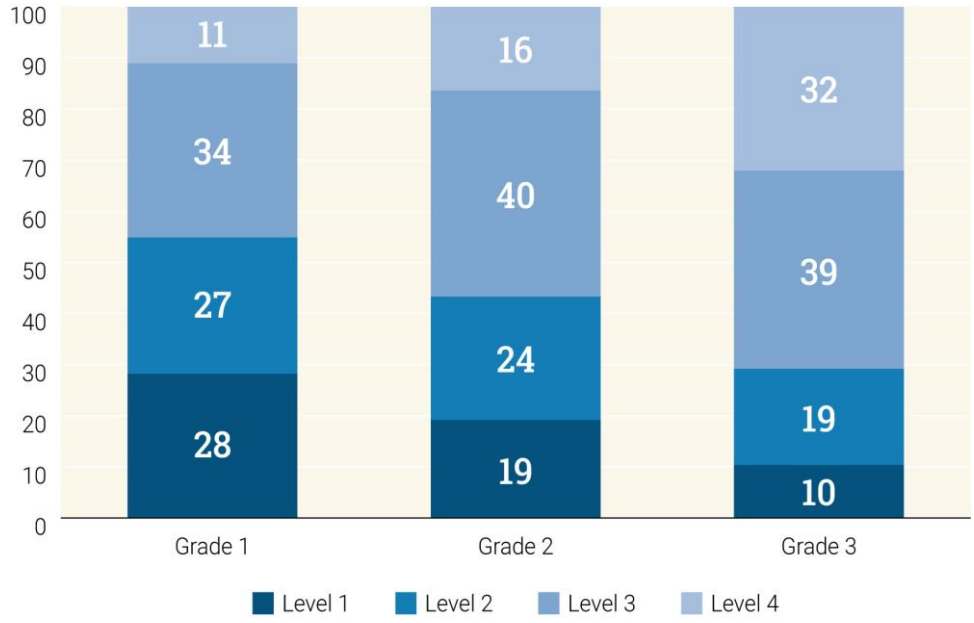
Finally, there were instances where the descriptions in the curriculum did not relate to the specific skills assessed in the SLA (e.g., practice how to use the expressions of gratitude, apologies, favours, and compliments using polite language to other people, both verbally and in writing). While it was possible to broadly map the Proficiency Levels against the AKM, without knowing how the level of texts is determined in the AKM (e.g., what is a Level 1 text, what is a Level 2 text), it is difficult to determine whether the Proficiency Level equates to Level 1 or Level 2 of the AKM (or what grade level) because the wording of the described skill is the same for both levels. Therefore, the SLA Proficiency Levels for Bahasa Indonesia are mapped only against the AKM and the MPL, and do not include connections to the curricula frameworks. The MPL for Bahasa Indonesia was assessed as being at Level 3 and above.

4. Student performance

4.1. Mathematics

A total of 18,370 students participated in the SLA mathematics assessment. While a direct and accurate alignment with the SDG MPLs was not possible, students performing at Level 4 were assessed as meeting or exceeding the MPL for SDG 4.1.1a. Students at Levels 1 and Level 2 were assessed as not meeting the minimum level of proficiency expected by the end of Grade 2/3, and those in Level 3 had met some but not all proficiencies expected for early primary. Results from the assessment showed on average students in Grade 3 were not meeting the SDG MPLs for early primary. By Grade 3, approximately two out of three students did not meet expected standards, and 84% of students in Grade 2 did not yet meet minimum mathematical proficiencies. A more promising result was that a significant proportion of students (those in Level 3) were working towards expected standards and with additional remedial support targeted to areas of difficulty, could reach expected standards by the end of Grade 3. Results also showed proportionate increases in student performance by grade level, with the average performance of students increasing year on year, demonstrating the value-added of schools and teaching.

Figure 4: proportion of students by level by grade—Mathematics



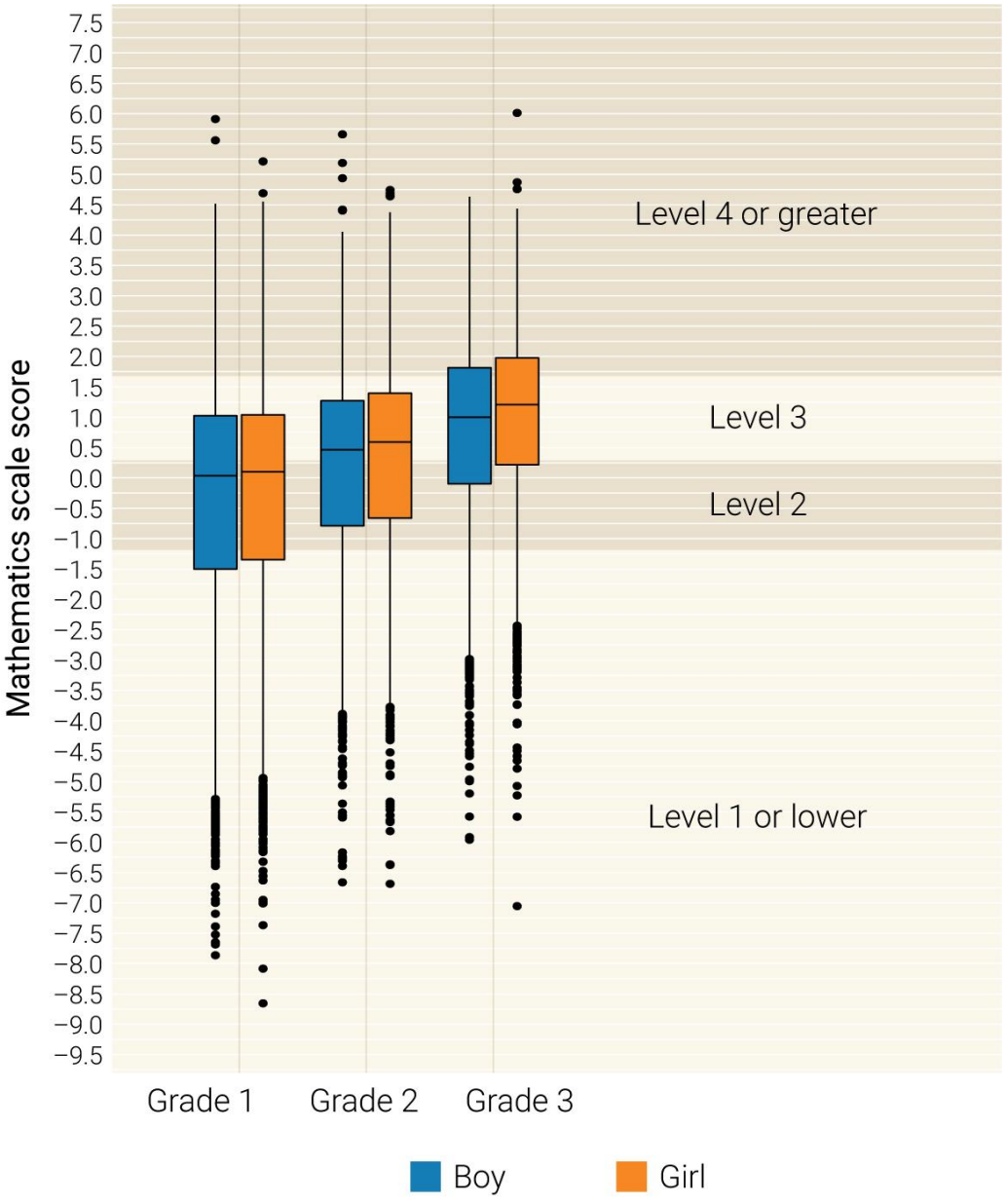
Note: due to rounding, totals may not add up to 100.

This study has revealed significant equity gaps in mathematics learning outcomes. The box plot below shows the spread of student results in proficiency level, by gender. The centre box shows 50% of the student population, the horizontal line through the centre of the box is the mean performance of those students, and the outer solid lines to the boxes show 75% of students. The small circles represent the remaining 25% of students. These graphs show that

there was some part of the student population that was performing significantly above grade level expectations; however, a significant number of students were falling far behind.

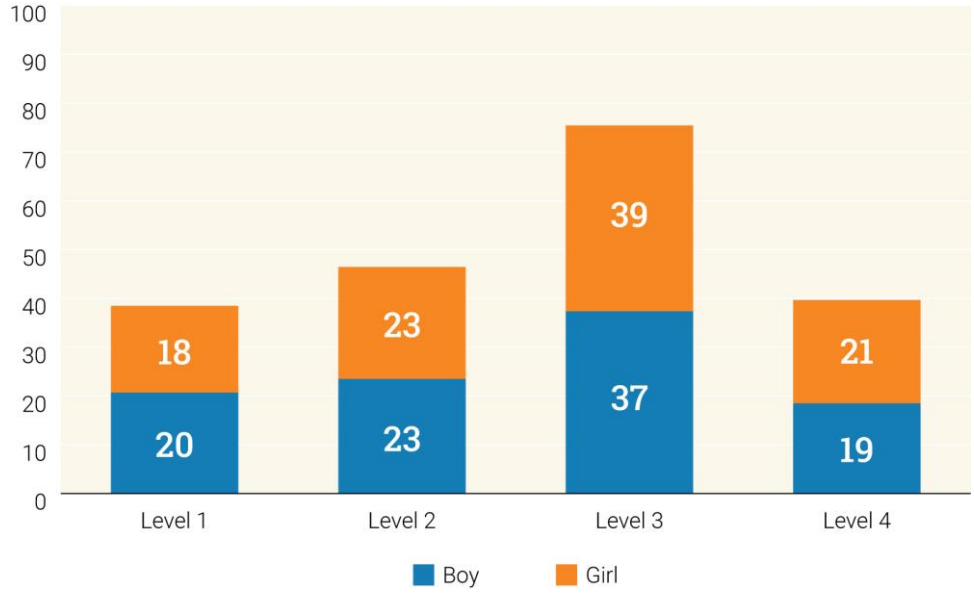
In this graph, considering that a difference in student results of approximately 1.4 on the graph equates approximately to one year level against the GPF, students were performing three and four grade levels below expectations. On a more positive note, the results confirm that students in the lower performing bands were showing growth in learning between the year levels. It is concerning, however, that the top performing students were showing little to no growth in mathematical proficiencies between the year levels. This suggests that while remedial support for poorer performing students appeared to be working, higher performing students were not being adequately extended in their learning.

Figure 5. box plot depicting distribution of abilities for the mathematics domain by grade and gender



In mathematics, the study shows there was a slight emerging gender gap. The box plots in Figure 5 show that students in Grade 1 were close to equal in ability, but by Grade 3 there was a distinct positive effect for girls who were consistently outperforming boys on average. Figure 6 shows the proportion of all girls and boys included in the study in each of the proficiency bands for mathematics. Across all grades, there were higher proportions of girls in Level 3 and Level 4 than there were of boys.

Figure 6: proportion of girls and boys by performance levels (all grades)—mathematics



4.1.1. Key findings in the mathematics domain

Each of the SLA tests contains a high proportion of items assessing the number and operations domain, with more limited coverage of the other four domains. This is particularly true in the Grade 1 and Grade 2 booklets. For example, the Grade 1 test has only one measurement item, one geometry item, and one algebra item. While there are eight items for measurement, most of which were in the Grade 3 test, five of those items are well beyond the Grade 3 level on the GPF.

In terms of reporting on learner progress towards meeting the global minimum proficiency level, it is important to consider their progress across all domains as together they represent the global expectations for what students should know and be able to do in mathematics. The coverage across the domain of the number and operations domain in the SLA means that partial alignment with the GPF is possible. However, the limited coverage of the other domains means students may not have had sufficient opportunity to demonstrate what they know and can do in Geometry, Measurement, Statistics and Probability, and Algebra. More information may be required to better understand where students are positioned in relation to meeting the grade level MPLs of the GPF and the end of lower-primary MPLs for SDG 4.1.1.

Looking at the scaling of the items, at least four things can be observed:

1. As expected, students proceeded from being able to identify, read, order, and compare whole numbers up to 20 to applying those same skills to larger whole numbers.
2. While students may have been able to use addition or subtraction and multiplication or division to solve problems presented as number sentences (e.g., 12 + 3), applying their understanding of those operations to solve real-world problems was more challenging (e.g., Fikri has 3 cakes. While visiting grandfather's house, Fikri was given 12 cakes. How many cakes does Fikri have?).[6]
3. The single item assessing Statistics and Probability, relating to the use of tally marks to represent data, proved to be very difficult for students, which may indicate this was not a concept they had been exposed to in Grades 1–3.
4. Items relating to recognising and extending a pattern (algebra), measuring an object using non-standard units (measurement), and recognising the basic shape of a square (geometry) proved to be more difficult than anticipated, given their location on the GPF.

4.1.2. Implications for teaching and learning —mathematics

Given the limited coverage of the mathematical domains other than Number (i.e., measurement, algebra, statistics and probability, and geometry), and the fact that items assessing those other domains proved to be more difficult than anticipated, it is worth considering how the current curriculum supports mathematical development across the five sub-constructs of the mathematics domain, especially in the early years of school (because this is the focus of the study), and how teachers are supported in teaching mathematical sub-constructs other than number. In relation to the teaching of number, it is also worth considering how teachers can be supported in extending students' skills beyond solving equations, and how they can take learning deeper, enabling students to apply their knowledge to more practical, real-world problems.

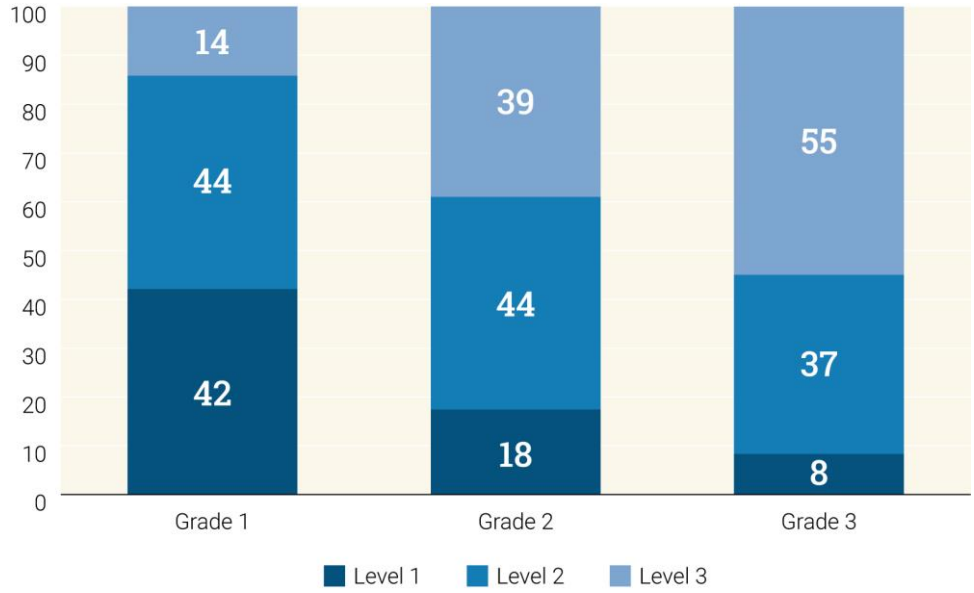
4.2. Bahasa Indonesia listening and reading comprehension

A total of 18,370 students participated in the SLA literacy assessment. Three performance levels were developed as part of the SLA assessment for Bahasa Indonesia. While a direct and acurate alignment with the SDG MPLs was not possible, students performing at Level 3 and above were assessed as meeting or exceeding the MPL for SDG 4.1.1a. Students at Level 1 were assessed as not meeting the mimimum level of proficiencies expected by the end of Grade 2/3 reading, and those in Level 2 met some but not all of the required proficiencies. Results from the assessment showed that approximately 55% of students in Grade 3 were meeting or exceeding the SDG MPLs for early primary reading; however, as shown in Figure 7, approximately three out of five students in Grade 2 still fell short of the mimimum reading proficiencies. A more encouraging result was that there was a signicant proportion of students (44% in Grade 2 and 37% in Grade 3) who, with additional support targeting a greater range of skills essential to listening and reading comprehension, could meet international benchmarks for performance by the end of Grade 3. Results also showed

[6] 65.2% of students could answer the question '12 + 4' (item cma111) while only 46.2% of students could answer the word problem 'Fikri has 3 cakes. … Fikri was given 12 cakes. How many cakes does Fikri have?' (item cma117). Item cma117 was more difficult than 70% of all the items.

a positive proportionate increase in student performance by grade level, with the average performance of students increasing year on year, demonstrating the value-added of schools and teaching.

Figure 7: proportion of students by performance level—Literacy

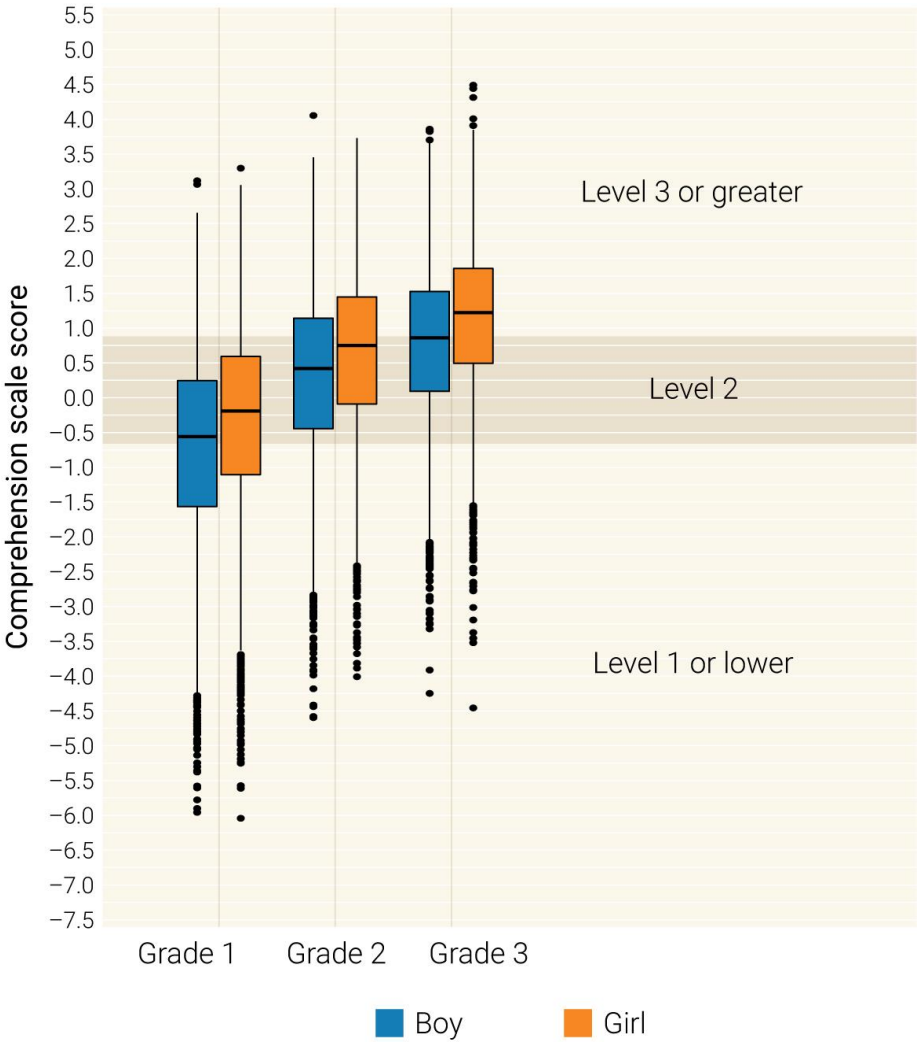


Note: because of rounding, totals may not add up to 100.

Figure 8 shows that a significant equity gap existed in student literacy skills. The box plot below shows the spread of student results in each grade level by gender. The centre box shows 50% of the student population, the horizontal line through the centre of the box is the mean performance of those students, and the solid vertical lines extending from the boxes show 75% of students. The small circles represent the final 25% of students. These graphs show that there were some portions of the student population that were performing significantly above grade level expectations; however, there were also a significant number of students falling far behind.

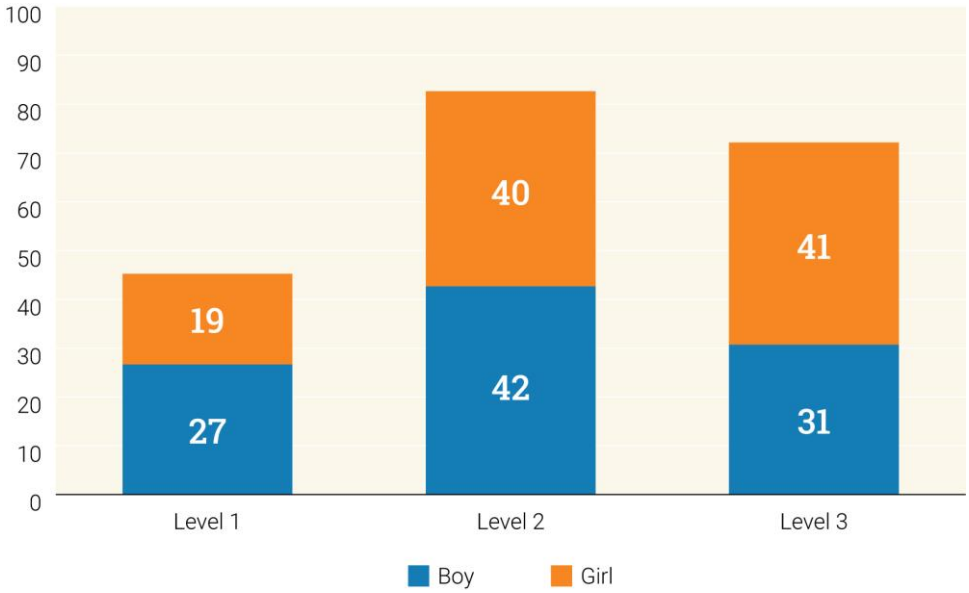
A difference in student results of approximately 1.4 on the literacy scale equates to approximately one year level of learning progress against the GPF. This learning gap study therefore shows that there was a proportion of students who were performing three- and four-year levels below expectations. This means that a proportion of the student population was still performing at a level expected of pre-primary students and had not yet formed the foundation skills necessary for progressing through primary school. More positively, the results indicated that students in the lower performing bands were showing growth in learning between the year levels. Unlike mathematics, the top performing students were also showing growth in learning between the year levels.

Figure 8: box plot depicting distribution of abilities for the literacy (comprehension) domain by grade and gender



The learning gap study shows there was a significant gender gap in student reading comprehension abilities. Figure 9 shows the proportion of all girls and boys included in the study in each of the proficiency bands for Bahasa Indonesia. Across all grades, there were significantly higher proportions of girls in Level 3 than there were of boys. A full set of interactions between gender and grade can seen in Annex C.

Figure 9: proportions of girls and boys (all grades) by performance levels—Literacy



4.2.1. Key findings in Bahasa Indonesia reading comprehension

Most of the reading texts in the tests for Grades 1, 2, and 3 were classified as Grade 3-level reading texts. This is an important factor in determining the level of alignment between the SLA and the GPF and in assessing reading development.

The mismatch between the level of skill being assessed and the level of the text means that it is likely that if some students could demonstrate the skill with the harder text, even more could have done so if the text was easier and at the appropriate level. For this reason, it is recommended that some simpler Grade 2 level reading texts be included in the tests for Grades 1 and 2 in the future.

The hardest text in the Grade 3 test was classified as Grade 6 level. It is suggested that this text was too hard and that the items failed to address comprehensive understanding of the text. The items were either too easy, or misleading because students were likely to be able to guess the answers. It is strongly recommended that the set of items addressing a text comprehensively cover the main meaning of the text and that a very hard text with easy or misleading items be avoided.

All the texts were simple narratives. A greater diversity of text types and narrative styles would support the assessment of a wider range of reading skills. It is strongly recommended that some informational texts be included in the future.

Most of the reading items entail retrieving information or making simple inferences. This means the reading sub-constructs are well covered for Grade 2, but not for Grade 3, which is why the estimated strong alignment for Grade 3 is problematic. It is strongly recommended that a greater diversity of reading skills be addressed in the items. This requires selecting texts where it is possible to ask questions about the meaning of unfamiliar words or expressions, as well as about the main idea.

The learning gap study noted that students found it very difficult to rearrange eight words to make a sentence and concluded that this task appears above the estimated point alignment with Grade 3 MPLs. This is not a task that aligns with the GPF, but nonetheless it illustrates students' limited comprehension skills. Students who can read and fully understand a 60-word text should be able to rearrange eight common words into a short sentence.

4.2.2. Key findings in Bahasa Indonesia listening comprehension

Key findings from the Bahasa Indonesia assessment show that all students typically demonstrated greater levels of skills in reading comprehension than in listening comprehension, which is a surprising outcome. One of the shortest, simplest listening comprehension texts (approximately 30 words) with simple, literal questions was only answered by Grade 3 students, which shows that students found this to be one of the most difficult listening comprehension items. Students demonstrated greater skill in answering questions that asked them to retrieve and interpret simple information from a written text when they could write the answer or select a multiple-choice answer than in answering oral questions about a short text that was read aloud to them. This suggests that many students were not yet able to comprehensively process a short piece of text. If they could only listen, they did not yet have the skills to process the text in meaningful blocks that would allow them to remember the main ideas and attach the details.

Students' reading strategies were likely to focus on matching words from the question to the text to locate nearby information while paying little attention to the overall meaning. This suggests these students may only have been capable of making very literal, superficial interpretations of texts. Students who were likely to become effective independent readers, typically demonstrated listening comprehension skills in Grades 1–3 that were well in advance of their reading comprehension skills (typically two years ahead).

It will be important to examine pedagogical practices in the classrooms where students are meant to be learning from oral texts. Administering more listening comprehension assessments with items that address a range of comprehension skills to students in Grades 1, 2, and 3 would also provide more information to help clarify what is happening in relation to students' listening comprehension skills.

4.2.3. Implications for teaching and learning—Bahasa Indonesia

While the curricula frameworks give some guidance about how literacy skills like reading, writing, or speaking might be applied across different learning contexts (e.g., learning about specific concepts in science), they offer limited guidance about how children learn to read and what reading skills students are expected to demonstrate at different levels of the curriculum.

Looking at the AKM, there is more guidance for skills related to reading comprehension, but no clear guidance for skills related to the domains of decoding and listening comprehension. From the perspective of the GPF, listening comprehension and decoding are also important aspects of early reading development in Grades 1–3. For this reason, it is worth considering whether the AKM incorporates the full range of domains associated with reading and provides sufficient information about the development of early reading.

Regarding the skills covered in the AKM, the emphasis on reading comprehension (rather than on listening comprehension and decoding) at Levels 1 and 2 means the expectations for minimum competency in the early grades may exceed the global expectations for what early readers can do. For example, whereas *making simple inferences from a grade-level text that is read by the student* is a skill that appears in the GPF at Grade 3 and beyond, this appears to be an expectation at Level 1 of the AKM. If so, this expected minimum competency in Indonesia might exceed the global minimum competency level. Likewise, *evaluating and reflecting on a grade-level text that is read by the student* is a skill that appears in the GPF at Grade 4 and beyond. In the AKM this skill is described at both Level 1 and Level 2, which may indicate another area where the expectations in the AKM exceed the global minimum competency described in the GPF.

These findings provide important insights into how the teaching of reading might be strengthened and supported within Indonesia. First, the results of the listening comprehension items and the absence of listening comprehension from the AKM suggest that this may be an area to focus on. Listening comprehension is an important aspect of learning how to be an effective reader in the early years of school. Reading is not just about decoding words, but about making meaning from a text. As students are improving their skills in decoding, they can be supported in the development of their comprehension skills by listening to texts that are read to them and engaging in rich discussions about the text. Students who are likely to become effective independent readers, typically demonstrate listening comprehension skills in Grades 1–3 that are well in advance of their reading comprehension skills (typically two years ahead). Thus, their ability to make sense of a text is not constrained by their ability to decode, and once they have developed these important skills, they can then apply them to texts that they are able to read once their decoding skills have developed.

Second, in relation to the variety of text types and text levels, it is worth considering the extent to which teachers have access to a range of different text types and levels to support the teaching of reading in the early years. Equally important is the support that teachers are given for selecting appropriately levelled texts and extending students' comprehension skills, from retrieving explicit information and making simple inferences to more complex tasks such as interpreting the meaning of unfamiliar words and communicating the main idea.

Finally, the lack of guidance in the curriculum for the teaching and learning of reading is an important area for consideration. Teachers need explicit information about how students develop their skills as readers, about how students progress in their reading development and about how to identify where individual students are placed on the pathway to learning. With this information, teachers will be better placed to target their teaching and selected texts to the right level and monitor individual progress in reading and listening.

5. Equity of learning outcomes

Box 1: Interpreting regression analysis

In the following section, a series of regression analyses is presented to provide evidence for the risks and protective factors that influence students' learning and development outcomes. To support interpretation of these analyses, the following note is provided:

The reader can explore the relative magnitude of the effects in the regression analysis in Figure 10 (where mathematics and literacy comprehension are combined on one plot). The independent variables in the regression are listed down the Y-axis. The dots represent the point estimates of the regression coefficients that have been converted into standard deviation units for each of the dependant variables,[7] and the horizontal lines passing through each point represent a 95% confidence interval.

Where the horizontal lines do not intersect the vertical dashed line at zero, the estimate is significantly different from zero. The direction of the effect matters: for example, a child who has a physical disability has a negative coefficient (they score lower on the literacy assessment relative to students with no physical disability) whereas girls have a positive coefficient indicating they score higher on average than boys.[8]

Coefficients for categorical independent variables are standard deviations away from the omitted level, and coefficients for continuous variables (executive function, social, and emotional skills) are standardised only in terms of the dependant variable, therefore representing the elasticity of the change in standard deviations on the dependant variable given a one-unit change in the independent variable.

In interpreting these results, it is important to consider that each effect is estimated by holding the other variables constant at zero. Therefore, students with multiple risk or protective factors will have multiple additive impacts on their learning. For example, in Figure 10, students who are boys, who have lower than average executive function capabilities, who did not attend preschool and who have a disability will be behind their peers proportional to the sum of the individual effects. These sum to large effects—children who experience multiple risk factors can be behind by the equivalent of a year of school education given sufficient risk or protective factors.

5.1. Student background

Student background and context are related to the attributes of the child. Students' age, gender, cultural background, cognitive capabilities, social and emotional skills, and physical abilities are all typically associated with leaning outcomes. Findings related to student background may indicate policy areas of interest (e.g., gender gaps) that can contribute to reducing inequities. This section describes how these traits link to learning outcomes.

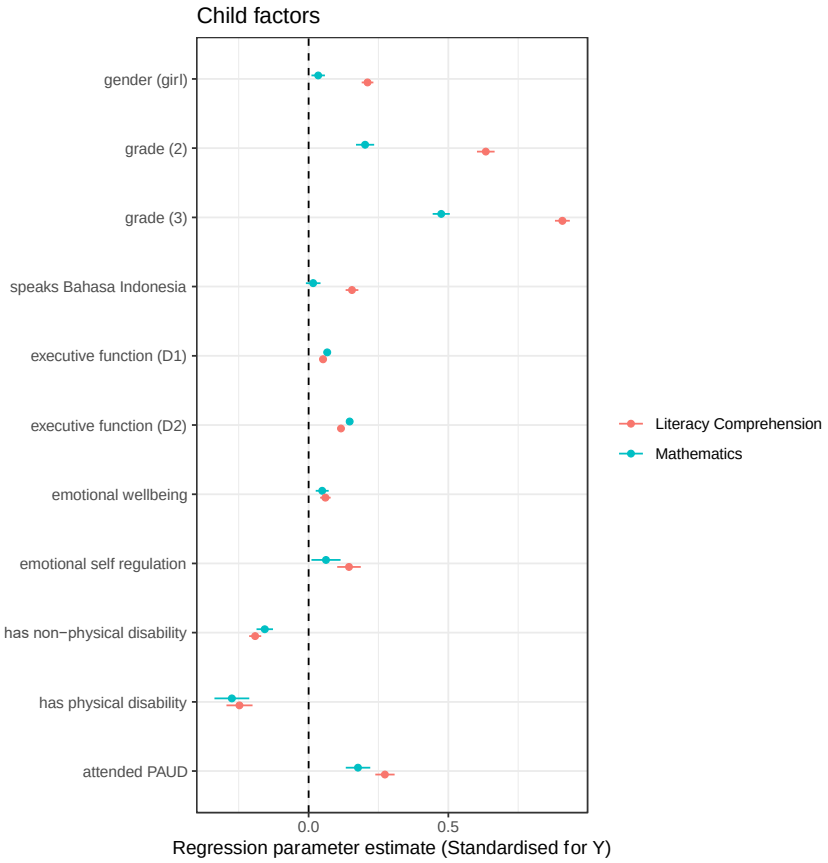
[7] So called, *standardised for Y*.

[8] All regression parameters are conditional on a fully specified regression model.

The results of regressing the mathematics domain on the background factors showed that there were several expected findings: students in Grade 2 scored higher than students in Grade 1, and students in Grade 3 scored higher again, on average. Students with disabilities scored lower than those who were typically developing. There was also a small gender effect—girls outperformed boys. Of note, students with stronger working memory,[9] inhibition control,[10] emotional regulation,[11] and well-being[12] also scored higher on numeracy. This is consistent with international evidence about the relationship of cognitive and social and emotional skill to academic learning (O'Connor et al., 2019; OECD, 2020). There was also a positive effect for students who attended preschool prior to school.

The results of regressing the comprehension domain on the background factors showed a similar pattern with the notable additional factor of language. Students who spoke Bahasa Indonesia, rather than a local language as their home language, scored higher on the (Bahasa Indonesia) comprehension domain.

Figure 10: plot of regression parameter estimates (and 95% confidence intervals) for the literacy comprehension and mathematics domains regressed on child factors



[9] Executive function (D1).

[10] Executive function (D2).

[11] The ability to control emotions and social behaviour in the interest of engagement and participation in both social interactions and independent work.

[12] The absence of anxiety, the presence of sadness, and the inability to control aggressive impulses.

5.2. Family and home

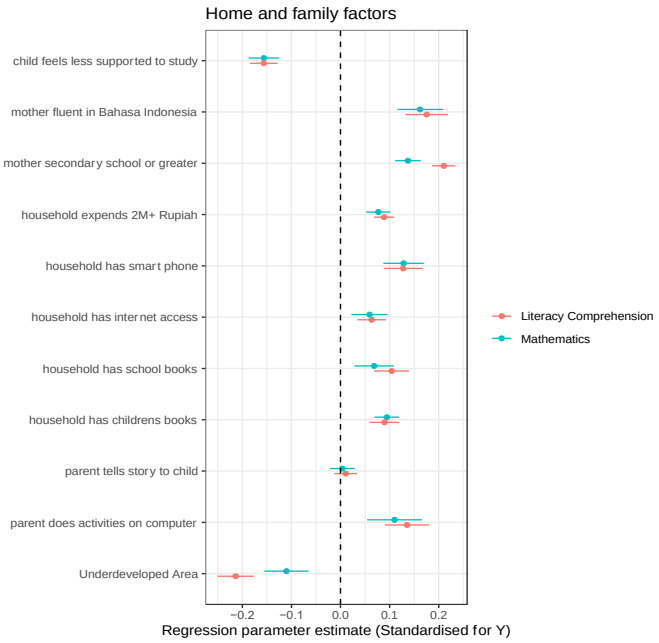
The family and home context involves the attributes of the child's family, the home environment they live in, and their local community. Families tend to vary not only in their resources and capacities—for example, in their socioeconomic status (SES)—but also in the kind of home-learning interactions and practices they engage in, and these can influence student learning either directly or indirectly. For example, although good quality interactions in the home such as reading to children may promote learning directly, families experiencing socioeconomic disadvantages may experience barriers to selecting high quality schools, so their socioeconomic status indirectly impacts learning. Findings relating to family background and home environment may not only indicate policy areas of interest (e.g., SES gradients) that can contribute to reducing inequities but also indicate where additional support to families may increase their capacity to act as first educators. This section describes how these traits relate to learning outcomes.

The results of regressing the mathematics domain (Figure 11) on the family and home background factors showed that there were several expected findings related to socioeconomic advantage; for example higher SES households (higher maternal education and household expenditure) were associated with higher achieving students, as were households with more possessions, including mobile phones and internet access.

There was a pronounced language effect—mothers who were fluent in Bahasa Indonesia were related to greater student achievement. Parents who interacted with their children by working together or doing an activity on a computer were also associated with higher student outcomes. The nonsignificant finding for telling a story to a child is perhaps related to the age of the focal children who were transitioning toward independent reading. Additionally, living in an underdeveloped location was strongly associated with poorer outcomes, especially literacy. It is importantly to note that students who reported feeling supported by their parents when studying at home did significantly better than their peers who felt unsupported. This last finding implies a psychosocial element to the home-learning environment and suggests that improved support at home operates through a sense of engagement, motivation, or well-being that, in turn, leads to improved learning outcomes.

Effects for the comprehension domain were nearly identical to those of the mathematics domain with matching direction and similar magnitudes. The reader is referred to Box 1 for a description of how to interpret the figures that summarise the regression analysis.

Figure 11: plot of regression parameter estimates (and 95% confidence intervals) for the literacy comprehension and mathematics domains regressed on household and family factors



The findings from this analysis reveal that the language of the parents was especially important in the students' learning outcomes—especially when the mother spoke Bahasa Indonesia. There were also strong SES effects: both mother's education and household expenditure were significant factors in learning outcomes. Parental involvement in student learning was also important and the availability of books in the home showed a strong effect on the literacy abilities of students. One of the most significant factors that influenced learning outcomes was the location of where a student lived. Students living in underdeveloped areas had significantly poorer outcomes.

5.3. School and learning environment

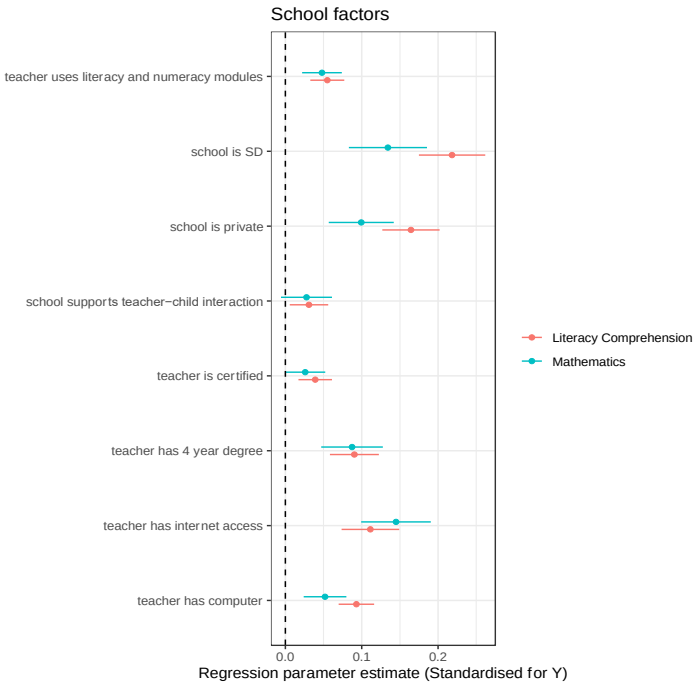
The school and learning context relates to the attributes of the child's school, the staff that work within them, and their resources and behaviours. Schools tend to vary in their characteristics (e.g., public and private auspices) and also in the behaviours of their staff. Findings about schools are a key aspect of the study and may lead to policy insights regarding the best way to support school education in recovering from the COVID-19 pandemic.

The results of regressing the mathematics domain (Figure 12) on the school factors showed that there were several expected findings: schools in the private sector (an index of aggregate SES) tended to include higher achieving schools, as did SD schools as opposed to madrasahs (which are also an index of aggregate SES—see Niaz Asadullah, 2018). Teachers who had completed a four-year degree and who had access to technology (a computer and internet access) also tended to be associated with better student learning outcomes. When teachers used the literacy and numeracy modules developed by MoECRT as a resource for the Emergency Curriculum, a significant positive effect was found. A study of the effects of the

curriculum, including interaction with factors affecting delivery within different regions of Indonesia, will be the subject of a separate report in the Learning Gap Study series.

The results for the comprehension domain were similar to the mathematics domain. In addition to the findings described above, in the comprehension domain there was a positive effect for teacher certification. The reader is referred to Box 1 for a description of how to interpret the following figures that summarise the regression analysis.

Figure 12: plot of regression parameter estimates (and 95% confidence intervals) for the literacy comprehension and mathematics domains regressed on school factors



6. Emerging themes

6.1. School readiness and preparedness

The learning gap study has revealed that students are coming to school without the level of preparedness required and that schools are providing insufficient support for students to build essential foundational skills in early literacy and numeracy. Most students in Grade 3 were not meeting the SDG mathematics MPLs for Grade 2/3 standards. By Grade 3 approximately two out of three students did not meet expected mathematical standards, and 84% of students in Grade 2 did not yet meet minimum mathematical proficiencies. For Bahasa Indonesia, approximately 55% of students in Grade 3 were meeting the SDG MPLs for early primary reading; however, approximately three out of five students in Grade 2 still fell short of the mimimum reading proficiencies.

A number of factors are associated with the gap in students' foundational literacy and numeracy skills. A significant proportion of students enter school without foundational language skills in Bahasa Indonesia. This means they are not prepared for the multitude of language exchanges required in the classroom and will need more time to build Bahasa language proficiency and expand on their vocabulary. Furthermore, many students are entering school with limited executive function skills, in other words, the ability to concentrate, collaborate, organise, and prioritise at a level which is required for the classroom. The study has revealed that students who attended preschool do better than those students who did not. Quality preschool education helps students build the essential language and executive function skills needed to suceed in school. Additionally, students who come from homes where the parents are more involved in their child's learning or who support children in their schooling also do better. These homes typically have a selection of reading materials for children including story books, magazines and school text books, with access to further reading material via smart phones or through the internet. Students who attend schools that have access to the internet also do better. Teachers also play an important role in setting up students for success. Where students are taught by fully qualified teachers, learning outcomes are significantly better.

Identifying the gaps in student skills as they enter school so that teachers can in turn address these learning gaps in the early years, while considering the multifaceted components that drive learning, is the first critical step in ensuring long-term improved learning outcomes for all students in Indonesia.

6.2. Equity, inclusion, and engagement

The learning gap study has revealed a significant spread in student proficiencies in literacy and numeracy. Some parts of the student population in Grades 1 to 3 were performing significantly above grade level expectations; however, a large number of students that were three- and four-year levels below expectations. These differences became more pronounced between sub-populations including gender. Girls consistently outperformed boys in literacy and numeracy. Language is another factor across sub-populations, with SES being one of the most significant social drivers that underpin learning outcomes.

As the study has revealed, students who have more learning resources in the home including acess to reading materials (such as books and the internet) and parental support do better than those who do not. The COVID-19 pandemic has likely disporportionately affected students from more marginalised families and communities, magnifying learning inequities across the Indonesian population. With students forced into home learning, parental support and student engagement become central in supporting the educational challenges of students during remote schooling periods. Students who are marginalised by language, neurodiversity, or by living in rural or remote areas are most at risk of experiencing learning losses over extended periods of time. Students who live in lower SES households, requiring parents to engage in longer work hours, will likely experience heightened intergenerational poverty issues. Marginalised students who do not benefit from sustained home support for their learning during the home learning periods are at a higher risk of disengagement and consequent school dropout. Building protective factors such as social connections and relationships around those students who are most at risk may help to lower attrition levels.

A one-size-fits-all approach therefore will not address learning inequities, especially in the light of the COVID-19 pandemic. As students begin to return to school, the gap in the learning outcomes will become even more pronounced. A strategy that enables a more targeted system of support, investing in underserved areas, engaging with communities and families, and identifying the more marginalised students within schools to build programs of learning intervention will significantly improve learning outcomes for all students.

6.3. Systemic support for targeted teaching

The return to school, after prolonged learning-from-home periods, is likely to reveal more pronounced learning gaps between learners. Where learning has been missed, or students have not progressed in their learning, content delivered in teaching programs will need to be adjusted. Learning is a progression and students must first master foundational skills in literacy and numeracy before they advance to more complex content. The learning gap study has mapped student performance against a scale of proficiency in literacy and numeracy. The results from the study clarify what students can do so that their next steps for learning can be identified in support of targeting teaching to the point of student need. Using a common empirical scale of student proficiency, the learning gap study showed significant variation in student abilities with some students four and five years behind their peers. Teaching fractions to students who are still unable to count beyond 10 will have little effect on learning outcomes.

Targeting teaching to the point of student ability requires a systematic approach that helps teachers define, measure, and understand learning. Defining what students are expected to do at different levels provides clear pathways of learning for teachers, students, and parents. Grade competencies are outlined in the K-13 national curriculum and the Emergency Curriculum. However, the study has shown an uneven representation of the different learning constructs across literacy and numeracy. It also highlights the need for a clear set of learning progressions to support teachers' endeavours to understand what is expected of their students. At the same time, defined learning pathways must go beyond grade level expectations. They need to include a range of different levels of development that teachers can use to monitor the progress of learning throughout the year while making decisions on how to target instruction to the child's point of need.

The learning gap study found that teachers who adopted the numeracy and literacy modules in their teaching attained better student results. This may be due in part to the continuous learning objectives that were clearly defined against each level of learning. The literacy and numeracy modules were developed to help teachers and learners focus on foundational skills during school closure. These modules also include end-of-lesson reflective and self-assessment tasks to determine the level of understanding the student has for each activity, and importantly the modules provide resources and materials for parents to help with home-based learning initiatives. Measuring learning, through a system of continuous classroom assessment, which is embedded in the teaching and learning process and used to inform decisions about next steps, helps teachers target their teaching and support remedial action for students as needed. Pedagogical practices supported by a deep understanding of learning at the foundational level is critical. However, this is a sophisticated process and requires a level of skill and knowledge on the part of the teacher, as well as systemic support to enable the effective monitoring of learning at the classroom level, benchmarked against a clear set of learning proficiencies.

Teaching at the right level, therefore, requires a clear alignment between curricula, teaching, and assessment. Addressing gaps in learning cannot be achieved through a single point of intervention, whether through a new teaching technique, a new curriculum or a new assessment platform. Rather, these three areas must be interlinked and systemically supported so they may be enacted at the school level, addressing the specific learning needs of each student.

A main finding of this study is that a large majority of students could meet international proficiency standards with additional support. The constituents of that support are summarised above. They begin with targeted curricular and pedagogical planning for children entering school from different background conditions—including the key factor of readiness for the school language—and continue with regular assessment to identify problems and support remedial teaching. These fundamental processes of tailored teaching themselves depend on the fundamental need of to be informed about the progression of learning, so they know what they are aiming for, and can recognise and remediate blockages to progress.

As students approach nearly two years of disrupted schooling, the data behind this investigation could be used to start recalibrating the curriculum of the first three years of learning, to arrest the widening gaps in mastery due to the pandemic, and to help update the more permanent development of curricular pathways to the proficiencies students should attain by the end of the early grades learning.

7. Conclusion

A rich dataset on early grade learning from the learning gap study now exists and this report reflects on some of the higher-level findings emerging from this data. Further analysis could focus on drawing plausible causal inferences to provide powerful insights into the factors that affect learning. Methods like propensity score matching (Zanutto, 2021) would allow deeper interrogation of the effect of school inputs while adjusting more comprehensively for local and family contexts. Specific policy questions could also be answered by leveraging the IRT analysis done in this study to develop learning and development scales, which might include developing models of the proportion of students who meet the MPLs and explain how background factors are related to meeting the MPLs. By approaching similar research questions, authors could also explore the robustness of the magnitude of effects by triangulation (for example comparing the multivariate regression results reported here to propensity score analysis and perhaps other methods like multi-group regression).

The learning gap study has revealed that a significant proportion of Indonesian students in Grades 1, 2 and 3 are below expected learning levels as defined by international standards. The study also shows the grade level expectations as per the Indonesian curricula are higher than those of the international MPLs. This is not necessarily an issue, provided Indonesian curricular expectations are clear and realistic. What the learning gap study highlights, however, is that while a large proportion of students are not meeting international learning standards, an even higher proportion of students do not meet national grade level expectations. COVID-19 is expected to widen the learning gap, with a disproportionate effect on the most marginalised students. Immediate system level support to develop clearly defined learning objectives for students in literacy and numeracy that engage students, teachers, and parents is a critical first step and may help mitigate further disadvantage and learning inequity.

8. References

Adams, R. J., Wu, M. L., Cloney, D., & Wilson, M. (2020). *ACER ConQuest: Generalised Item Response Modelling Software* (Version 5) [Computer software]. Australian Council for Educational Research.

Cloney, D., & Adams, R. J. (2020). *Conquestr* (0.8.4) [Computer software]. Australian Council for Educational Research. https://cran.r-project.org/package=conquestr

Lumley, T. (2004). Analysis of complex survey samples. *Journal of Statistical Software*, *9*(1), 1–19. https://doi.org/10.18637/jss.v009.i08

Niaz Asadullah, M. (2018). Madrasah for girls and private school for boys? The determinants of school type choice in rural and urban Indonesia, *International Journal of Educational Development*, 62, 96–111. https://doi.org/10.1016/j.ijedudev.2018.02.006

O'Connor, M., Cloney, D., Kvalsvig, A., & Goldfeld, S. (2019). Positive Mental Health and Academic Achievement in Elementary School: New Evidence From a Matching Analysis. *Educational Researcher*, *48*(4), 205–216. https://doi.org/10.3102/0013189X19848724

OECD. (2020). *Early Learning and Child Well-being: A Study of Five-year-Olds in England, Estonia, and the United States*. https://www.oecd-ilibrary.org/content/publication/3990407f-en

R Core Team. (2020). *R: A Language and Environment for Statistical Computing* (4.0.0) [Computer software]. R Foundation for Statistical Computing. https://www.R-project.org/

USAID. (2019). Global Proficiency Framework for Reading and Mathematics Grades 2 to 6. http://gaml.uis.unesco.org/wp-content/uploads/sites/2/2019/05/GAML6-REF-16-GLOBAL-PROFICIENCY-FRAMEWORK.pdf

Von Davier, M., Gonzalez, E., & Mislevy, R. (2009). What are plausible values and why are they useful. *IERI Monograph Series*, *2*, 9–36.

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. https://ggplot2.tidyverse.org

Wu, M. (2005). The role of plausible values in large-scale surveys. *Studies in Educational Evaluation*, *31*(2–3), 114–128. https://doi.org/10.1016/j.stueduc.2005.05.005

Zanutto, E. L. (2021). A Comparison of Propensity Score and Linear Regression Analysis of Complex Survey Data. Journal of Data Science, 4(1), 67-91. doi:10.6339/JDS.2006.04(1).2335

Annex A: alignment against external benchmarks for the SLA

Alignment of the SLA against mathematics curriculum standards

Proficiency level and key skills	AKM descriptors and levels	Curriculum descriptors and grade levels	
		Emergency curriculum	K-13 curriculum
PROFICIENCY LEVEL 1 Read and compare whole numbers up to 20 in numerals. Identify equivalence between whole quantities up to 5 represented as objects, pictures, and numerals.	Compare two whole numbers (maximum three digits; Level 1)		C. Compare two numbers up to two digits by using a set of concrete objects (Grade 1). B. Explain up to two-digit numbers using a collection of concrete objects, and explain how to read the numbers (Grade 1).
PROFICIENCY LEVEL 2 Read, compare, and order whole numbers up to 100 in numerals. Identify and represent the equivalence between whole quantities up to 30 represented as objects, pictures, and numerals. Add and subtract within 20. Solve simple real world problems using addition and subtraction facts within 10.	Understand whole numbers (maximum three digits; Level 1). Solve simple equations using only the addition or subtraction operation (in a child-friendly form; Level 1). Get to know prisms and tubes (Level 2).	A. Explain the meaning of whole numbers up to 50 and the place value that composed the numbers using a collection of concrete objects and explain how to read the numbers (Grade 1). B. Compare two whole numbers up to 50 using a group of concrete objects. (Grade 1). C. Explain and perform the addition and subtraction of numbers involving whole numbers up to 20 in a daily life context by counting (Grade 1). D. Explain a plane and solid based on their characteristics (Grade 1).	A. Explain the meaning of whole numbers up to 99 as a number of members of a set of objects (Grade 1). B. Explain up to two-digit numbers and the place value that composed the numbers using a collection of concrete objects, and explain how to read the numbers (Grade 1). C. Compare two numbers up to two-digit numbers by using a set of concrete objects (Grade 1). D. Explain a plane and solid based on their characteristics (Grade 2).

Proficiency level and key skills	AKM descriptors and levels	Curriculum descriptors and grade levels	
		Emergency curriculum	K-13 curriculum
PROFICIENCY LEIEL 3 Read whole numbers up to 1000 in numerals. Identify everyday unit fractions represented as objects or pictures. Use place value concepts for tens and ones. Add and subtract within 100. Multiply and divide within 25.	Calculate the addition or subtraction of two whole numbers (maximum three digits; Level 1). Solve simple equations using only the multiplication or division operation (in a child-friendly form; Level 2). Understand simple unit fractions (1/2, 1/3, 1/4, and 1/5; Level 1).	C. Explain and perform the addition and subtraction of numbers involving whole numbers up to 100 in a daily life context by grouping based on the number place value and linking the additions and subtractions (Grade 2). G. Explaining the fractions of 1/2, 1/3, and 1/4 using concrete objects in daily life (Grade 2). E. Identify and explain the patterns of numbers and patterns of planes and solid rows by using images or concrete objects (Grade 1).	D. Explain and perform addition and subtraction of whole numbers up to 99 in a daily life context and link the additions and subtractions (Grade 1). G. Explain the fractions of 1/2, 1/3, and 1/4 using concrete objects in everyday life (Grade 2). J. Explain the patterns of rows of planes and solids using images or concrete objects (Grade 2).
PROFICIENCY LEVEL 4 Add and subtract within 1,000. Multiply and divide a two-digit number by a one-digit number. Identify and express proper fractions as equivalent fractions. Solve simple real-world problems involving multiplication of two whole numbers to 5. Use place value concepts for hundreds, tens, and ones. Make conversions between adjacent units of length and weight within a standard system of measurement.		D. Explain multiplication and division involving whole numbers that result in up to 100 in a daily life and link the additions and subtractions (Grade 2).	D. Explain multiplication and division involving whole numbers that result in up to 100 in a daily life and link the additions and subtractions (Grade 2). E. Explain and perform the addition and subtraction of fractions that have the same denominator (Grade 3).

Spread of SLA listening items across the domains of the GPF

Test booklet	Number of items for each domain			Other items*
	Listening Comprehension	Decoding	Reading comprehension	
Grade 1	3	50 letters 15 syllables 15 words	13	Two items unclassified—dictation and fill in the missing word
Grade 2	6	50 letters 15 syllables 20 words	11	Four items unclassified dictation, fill in the missing word, put words in the correct order to make a sentence
Grade 3	7	50 letters 15 syllables 25 words	7	Four items unclassified dictation, fill in the missing word, put words in the correct order to make a sentence

*Some items could not be aligned with the GPF (shown in the other items column). This was due to several factors, including items that had no clear match to the descriptors found in the GPF and items for which more information would be needed to (i) whether the words were common grade-level words and (ii) at which grade level they would be considered common.

Text levels in the SLA for listening and reading comprehension

SLA Test	Estimated GPF level for reading texts (number of texts)	Estimated GPF level for listening texts (number of texts)
Grade 1 SLA test	GPF Grade 2–(1) GPF Grade 3–(3)	GPF Grade 2–(1)
Grade 2 SLA test	GPF Grade 3–(3)	GPF Grade 2–(1)
Grade 3 SLA test	GPF Grade 3–(1) GPF Grade 6–(1)	GPF Grade 1–(1) GPF Grade 2–(2)

Annex B: summary of Analytical Method of Benchmarking

Stage 1: Conceptual alignment of the SLA to global and local proficiency levels

Global alignment of the SLA: The Global Proficiency Framework (GPF; USAID, 2019) can be used as a foundation for linking assessments to the minimum proficiency levels of the Sustainable Development Goals (SDGs) and it supports reporting against SDG 4.1.1. The GPF's main purpose is to provide a common set of descriptors for aligning and interpreting grade level student performance in reading and mathematics.

Local alignment of the SLA: Connections between these descriptions of the SLA items and the descriptions of student competencies from the AKM, the Emergency Curriculum, and the K-13 Curriculum were then explored qualitatively.

Stage 2: Scaling of student response data and setting proficiency 'cut points'

Student data was calibrated and scaled using Item Response Theory (IRT). IRT is a psychometric method that is commonly used in educational research (see e.g., Wu, 2005) and has significant strengths, including yielding linear measures of latent constructs and being robust to missing data. In this study, the domains of literacy and numeracy were scaled.

Initial analysis focused on calibration of the latent measures and assessment of item functioning. Items that were biased (e.g., unfair to some subgroups), poorly discriminant, or otherwise a misfit to the model were omitted from the scaling process. The final data set was scaled, which provided information on the range of items in the SLA assessment from the easiest to the most difficult along with the corresponding student proficiencies against each of these items. Items that clustered together on the scale provided the basis for developing a set of proficiency bands (or levels). A series of 'cut points' were set on the scale to enable reporting student performance by proficiency level, rather than only a population average.

A subset of the contextual data, selected to ensure important predictors of learning, was included in the population model (a latent regression model). Together the item response model and population model were fit to yield ability estimates that could be analysed with respect to the background data collected; these estimates included measures of child, family, home, school, and community characteristics.

Secondary analysis was conducted using a fully conditional OLS regression model. Analysis is reported based on pooled analysis of the ability estimates (plausible values; for information on calculating pooled estimates and justification [see e.g., Von Davier et al., 2009]) and robust standard errors were calculated to account for the two-stage sampling design (students within schools; Lumley, 2004).

All analyses were conducted in ACER ConQuest (Adams et al., 2020) and Conquest (Cloney & Adams, 2020). Analysis in Conquest relied on the R statistical environment (R Core Team, 2020) and visualisation was conducted using the ggplot2 library (Wickham, 2016).

Stage 3: Development of proficiency descriptors and reporting student performance

To support the translation of findings from this study into policy and practice, a set of proficiency descriptors for each of the proficiency levels was developed. Rather than only reporting student performance by levels, a set of descriptors, which outlined the proficiencies students are expected to know in each level, were developed. These descriptors were aligned to the GPF and to the various national curricular frameworks where possible.

Annex C: statistical output data tables

Table 3: fully conditional OLS regression for mathematics domain (robust standard errors for clustering within school).

term	estimate	std.error	statistic	df	p.value
Intercept	-2.27	0.10	-22.30	107.14	0.00
gender (girl)	0.06	0.02	2.81	205.30	0.01
grade (2)	0.35	0.03	12.25	87.92	0.00
grade (3)	0.83	0.03	30.32	300.09	0.00
speaks Bahasa Indonesia	0.03	0.02	1.23	2943.93	0.22
executive function (D1)	0.12	0.01	16.33	336.49	0.00
executive function (D2)	0.26	0.01	34.39	799.41	0.00
emotional wellbeing	0.09	0.02	4.13	145.41	0.00
emotional self regulation	0.11	0.05	2.35	1311.98	0.02
has non-physical disability	-0.27	0.03	-10.47	83.84	0.00
has physical disability	-0.48	0.06	-8.63	243.40	0.00
attended PAUD	0.31	0.04	7.90	207.66	0.00
child feels less supported to study	-0.27	0.03	-9.86	17418.71	0.00
mother fluent in Bahasa Indonesia	0.28	0.04	6.87	114.67	0.00
mother secondary school or greater	0.24	0.02	10.34	1073.44	0.00
household expends 2M+ Rupiah	0.13	0.02	6.07	870.32	0.00
household has smartphone	0.22	0.04	6.11	252.61	0.00
household has internet access	0.10	0.03	3.14	60.77	0.00
household has school books	0.12	0.04	3.36	10434.40	0.00
household has childrens books	0.16	0.02	7.29	1075.80	0.00
parent tells story to child	0.01	0.02	0.28	597.75	0.78
parent does activities on computer	0.19	0.05	3.86	38.72	0.00
teacher uses literacy and numeracy modules	0.08	0.02	3.61	973.06	0.00
school is SD	0.23	0.05	5.14	54.96	0.00
school is private	0.17	0.04	4.57	52.36	0.00
school supports teacher-child interaction	0.05	0.03	1.63	38.73	0.11
teacher is certified	0.05	0.02	1.95	225.91	0.05
teacher has 4 year degree	0.15	0.04	4.24	373.96	0.00
teacher has internet access	0.25	0.04	6.23	355.28	0.00
teacher has computer	0.09	0.02	3.63	439.97	0.00
Underdeveloped Area	-0.19	0.04	-4.79	115.66	0.00
id_kab2	0.15	0.05	2.86	2018.85	0.00
id_kab3	1.06	0.06	18.19	775.17	0.00
id_kab4	0.12	0.08	1.51	223.50	0.13
id_kab5	0.44	0.06	7.55	16324.14	0.00
id_kab6	-0.01	0.06	-0.12	306.29	0.90
id_kab8	0.30	0.09	3.23	153.03	0.00
id_kab9	0.89	0.07	12.44	52.30	0.00
id_kab13	0.76	0.06	13.25	302.18	0.00
id_kab15	0.75	0.09	7.95	89.97	0.00
id_kab17	-0.18	0.08	-2.18	118.44	0.03
id_kab18	0.76	0.08	9.28	282.41	0.00
id_kab29	0.91	0.06	15.31	197.78	0.00
id_kab71	0.26	0.05	5.13	557.35	0.00

Model R^2 0.411 (95% CI 0.399 0.423)

Table 4. fully conditional OLS regression for comprehension domain (robust standard errors for clustering within school)

term	estimate	std.error	statistic	df	p.value
Intercept	-2.76	0.08	-33.14	16.44	0.00
gender (girl)	0.29	0.01	19.73	65.96	0.00
grade (2)	0.88	0.02	39.49	20.16	0.00
grade (3)	1.26	0.02	66.53	72.64	0.00
speaks Bahasa Indonesia	0.22	0.02	13.32	187.68	0.00
executive function (D1)	0.07	0.00	15.71	306.91	0.00
executive function (D2)	0.16	0.01	25.99	20.75	0.00
emotional wellbeing	0.08	0.01	6.25	154.19	0.00
emotional self regulation	0.20	0.03	6.73	2421.10	0.00
has non-physical disability	-0.27	0.02	-17.37	2243.30	0.00
has physical disability	-0.34	0.03	-10.42	2593.65	0.00
attended PAUD	0.38	0.02	15.51	359.58	0.00
child feels less supported to study	-0.22	0.02	-10.97	120.32	0.00
mother fluent in Bahasa Indonesia	0.24	0.03	7.92	23.36	0.00
mother secondary school or greater	0.29	0.02	17.48	82.80	0.00
household expends 2M+ Rupiah	0.12	0.01	8.49	569.63	0.00
household has smart phone	0.18	0.03	6.26	26.13	0.00
household has internet access	0.09	0.02	4.26	85.42	0.00
household has school books	0.14	0.03	5.71	91.87	0.00
household has childrens books	0.12	0.02	5.86	13.07	0.00
parent tells story to child	0.01	0.02	0.89	50.73	0.38
parent does activities on computer	0.19	0.03	5.90	42.82	0.00
teacher uses literacy and numeracy modules	0.08	0.02	4.84	231.85	0.00
school is SD	0.30	0.03	9.86	35.19	0.00
school is private	0.23	0.03	8.58	25.31	0.00
school supports teacher-child interaction	0.04	0.02	2.40	100.37	0.02
teacher is certified	0.05	0.02	3.49	131.59	0.00
teacher has 4 year degree	0.13	0.02	5.52	424.29	0.00
teacher has internet access	0.15	0.03	5.78	111.72	0.00
teacher has computer	0.13	0.02	7.82	209.84	0.00
Underdeveloped Area	-0.30	0.03	-11.34	113.59	0.00
id_kab2	0.20	0.04	4.62	29.68	0.00
id_kab3	0.59	0.05	12.59	26.83	0.00
id_kab4	-0.08	0.06	-1.38	69.97	0.17
id_kab5	0.32	0.05	6.09	17.33	0.00
id_kab6	0.08	0.05	1.75	21.12	0.10
id_kab8	0.77	0.07	10.66	18.75	0.00
id_kab9	0.73	0.05	13.79	30.01	0.00
id_kab13	0.63	0.05	13.15	19.40	0.00
id_kab15	0.27	0.07	3.95	68.99	0.00
id_kab17	-0.25	0.06	-3.92	21.65	0.00
id_kab18	1.00	0.06	16.60	44.37	0.00
id_kab29	0.71	0.05	15.70	30.39	0.00
id_kab71	0.57	0.05	11.76	12.22	0.00

Model R^2 0.609 (95% CI 0.599 0.619)

INOVASI
Innovation for Indonesia's School Children
Australia Indonesia Partnership

Ratu Plaza Office Tower - 19th Floor
Jl. Jend. Sudirman Kav 9
Central Jakarta, 10270
Indonesia
Phone : (+6221) 720 6616
Fax : (+6221) 720 6616

info@inovasi.or.id
Inovasi untuk Anak Sekolah Indonesia
Inovasi Pendidikan
www.inovasi.or.id

INOVASI is managed by Palladium
on behalf of the Australian Government

Palladium
MAKE IT POSSIBLE