Tanoto Foundation



Cohort 1 District and TTI Partner Schools Baseline Monitoring Report Volume 2:

Student Assessments in Bahasa Indonesia, Mathematics and Science

December 2018



Promoting Improvement to INnovate, Teach And Reach (PINTAR)

Cohort 1 District and TTI Partner Schools Baseline Monitoring Report Volume 2: Student Assessments in Bahasa Indonesia, Mathematics and Science

Cover Photo: Grade 5 students in Central Java completing the science assessment

This report was prepared by staff of the PINTAR program

INTRODUCTIO	ON	1
THE SAMPLIN	IG OF THE SCHOOLS	2
PROGRAM N	ONITORING AND EVALUATION	3
AN OUTLINE	OF THE ASSESSMENT PROGRAM	3
		~
1 SUMMARY	OF THE RESULTS OF THE TESTS AND RECOMMENDATIONS	6
1.1 IMPLEM	ENTATION OF THE TESTS	6
1.2 How тн	e Results are Presented	6
1.3 Summa	ry of Results in Primary Schools (SD and MI)	7
1.4 Summa	ry of Results in Junior Secondary Schools (SMP and MTs)	8
1.5 IMPLICA	TIONS AND RECOMMENDATIONS FOR PINTAR PROGRAM	10
2 FIRST ROUN	ID ASSESSMENT OF STUDENTS IN PRIMARY SCHOOLS	13
2 1 Βαμαςα	INDONESIA GRADE A	13
2.1 0411434	Introduction	13
212	The Results	13
213	Readina	14
2.1.4	Writing	1.5
215	Implications and Recommendations for PINTAR program	17
2 2 Mather	MATICS TEST GRADE 4	18
221	Introduction	18
222	The Results	18
223	Implications and Recommendations for PINTAR program	20
2.3 SCIENCE	TEST GRADE 5	21
2.3.1	Introduction	21
2.3.2	The Results	21
2.3.3	Implications and Recommendations for PINTAR Program	23
3 FIRST ROUN	ID TESTING OF JUNIOR SECONDARY SCHOOLS	25
3.1 BAHASA	INDONESIA GRADE 8	25
3.1.1	Introduction	25
3.1.2	The Results	25
3.1.3	Keaaing	26
3.1.4	Writing	27
3.1.5	Implications and Recommendations for PINTAR Program	28
5.2 IVIATHEN	NATICS TEST GRADE 0	29
3.2.1	Introduction	29
.∠.∠ 2 7 2	Inc. NESULS	23 21
3.2.3		27
2 2 1	I ESI URADE 0	32 27
227	The Results	22
222	Implications and Recommendations for PINTAP Drogram	25
3.3.3		20
ANNEY 2. CO		20
		39
ANNEX 3: CRI		45
ANNEX 4: SU	VIMARY OF THE TESTS AND THEIR DEVELOPMENT	46

Table of Contents

List of Tables

Table 1:	The Number of Partner Schools in Partner Districts	1
Table 2:	The Number of Partner Schools in 10 TTIs	1
Table 3:	The Number of Sample Schools by Province	2
Table 4:	Summary of Test Result for all Test in Primary Schools	7
Table 5:	Highest and Lowest School's Average Scores in Primary School Tests	8
Table 6:	Highest and Lowest Student's Scores in Primary School Tests	8
Table 7:	Summary of Test Results for all Tests in Junior Secondary Schools	9
Table 8:	Highest and Lowest School's Average Scores in Junior Secondary School Tests	10
Table 9:	Highest and Lowest Student's Scores in Junior Secondary School Tests	10
Table 10	Participant Data and Average Scores in Grade 4 Reading and Writing Tests	13
Table 11	Scores by Section in Primary School Reading Test	15
Table 12	Percentage Scores for Elements of Written Work in Primary School Writing Test	17
Table 13	Participant Data and Average Scores in Primary School Mathematics Test	18
Table 14	Participant Data and Average Scores in Primary School Science Test	21
Table 15	Average Scores by Section in the Primary School Science Test	23
Table 16	Summary of Test Result for all Test in Junior Secondary Schools	25
Table 17	Participant Data and Average Scores in Grade 8 Reading and Writing Tests	26
Table 18	Scores by Section in the Grade 8 Reading Comprehension Test	26
Table 19	Percentage Score for Components of Written Work in Grade 8 Writing Test	28
Table 20	Participant Data and Average Scores in Grade 8 Mathematics Test	29
Table 21	Participant Data and Average Scores in Grade 8 Science Test	32
Table 22	Average Scores by Section in the Grade 8 Science Test	34

List of Charts

Chart 1: Primary Reading Test - Comparison between Different Groups	14
Chart 2: Average Score by Quartile in Primary School Reading Comprehension Test	15
Chart 3: Primary Writing Test - Comparison between Different Groups	16
Chart 4: Average Score by Quartile in Primary School Writing Test	17
Chart 5: The Grade 4 Mathematic Test Results Disaggregated by Six Grouping	19
Chart 6: Distribution of Scores by Quartile in Primary School Mathematics Test	19
Chart 7: Results of Mathematic Test Sorted by Their Level of Difficulties	20
Chart 8: Primary Science Test - Comparison between Different Groups	22
Chart 9: Distribution of Scores by Quartile in Primary School Science Test	22
Chart 10: Analysis of Scores by Question in Primary School Science Test	23
Chart 11: Junior Secondary B. Indonesia Test in Reading- Comparison between Different Groups.	27
Chart 12: Average Student Scores by Quartile in Grade 8 Reading Comprehension Test	27
Chart 13: Average Student Scores by Quartile in Grade 8 Writing Test	28
Chart 14: Junior Secondary Mathematics Test - Comparison between Different Groups	30
Chart 15: Distribution of Scores by Quartile in Grade 8 Mathematics Test	30
Chart 16: Analysis of Scores by Questions in Grade 8 Mathematics Test	31
Chart 17: Comparison between Different Groups	32
Chart 18: Average Percentages of Student Scores by Quartiles in Science Test	33
Chart 19: Analysis of Scores by Questions in Grade 8 Science Test	34

Acronyms, Abbreviations, Terms

EGR	Early Grade Reading
EGRA	Early Grade Reading Assessment
EGMA	Early Grade Mathematics Assessment
Training facilitators	Fasilitator Daerah (District Facilitator)
GOI	Government of Indonesia
Gugus	Cluster of primary schools
IRR	Inter-Rater Reliability
JSS	Junior Secondary School
KKG	Kelompok Kerja Guru (Teacher Working Group- Primary Schools)
KKS	Kelompok Kerja Kepala Sekolah (School Principal Working Group)
Kota	City or Municipality
M&E	Monitoring and Evaluation
MGMP	Musyawarah Guru Mata Pelajaran (Local Association of Secondary Teachers, grouped by subject matter)
MI	Madrasah Ibtidaiyah (Islamic Primary School)
MOEC	Ministry of Education and Culture
MORA	Ministry of Religious Affairs
MORTHE	Ministry of Research, Technology and Higher Education
MTs	Madrasah Tsanawiyah (Junior Secondary Islamic School)
PAKEM	<i>Pembelajaran yang Aktif, Kreatif, Efektif dan Menyenangkan</i> (Active, Creative, Effective, and Joyful Learning)
PDF	Portable document format
PMP	Performance Management Plan
PRIORITAS	USAID Prioritizing Reform, Innovation, and Opportunities for Reaching Indonesia's Teachers, Administrators, and Students Project
PS	Primary School
SD	Sekolah Dasar (Primary School)
SMP	Sekolah Menengah Pertama (Junior Secondary School)
ТТІ	Teacher Training Institute
UIN	Universitas Islam Negeri (State Islamic University)
UN	Universitas Negeri (State University)
WSD	Whole School Development

INTRODUCTION¹

The Tanoto Foundation PINTAR program is working in 440 partner schools in 14 districts in five provinces and 10 Teacher Training Institutes (TTIs). In each of the 14 districts, the Program has at least 24 partner schools comprising 16 primary and eight junior secondary schools. Four districts have slightly higher numbers: Wonogiri 29, Siak 28, Bengkalis 27, and Batanghari 26 (Table 1). In addition, the program is also working with 90 partner schools of the ten TTIs. Each TTI has nine partner schools (Table 2).

Drevines	District	Primary		Junior Secondary		Total
Province	DISTLICT	SD	MI	SMP	MTs	Total
North Sumatra	Batubara	12	4	6	2	24
	Karo	13	3	6	2	24
	Kota Pematangsiantar	12	4	6	2	24
Riau	Bengkalis	16	3	6	2	27
	Kota Dumai	12	4	6	2	24
	Kota Pekanbaru	12	4	6	2	24
	Siak	16	4	6	2	28
Jambi	Batang Hari	15	3	6	2	26
	Tanjung Jabung Barat	12	4	6	2	24
	Tanjung Jabung Timur	13	3	6	2	24
Central Java	Kendal	12	4	6	2	24
	Wonogiri	18	3	6	2	29
East Kalimantan	Kota Balikpapan	12	4	6	2	24
	Kutai Kartanegara	12	4	6	2	24
Total		187	51	84	28	350

Table 1: The Number of Partner Schools in Partner Districts

Table 2: The Number of Partner Schools in 10 TTIs

Drovinco / LDTV	Primary		Junior Secondary		Grand
	SD	MI	SMP	MTs	Total
Sumatera Utara	8	4	3	3	18
UIN Sumatera Utara	2	4		3	9
Universitas Muhammadiyah	6		3		9
Riau	6	6	3	3	18
UIN Sultan Syarif Kasim		6		3	9
Univeritas Riau	6		3		9
Jambi	6	6	3	3	18
UIN Sulthan Thaha Saifuddin		6		3	9
Universitas Jambi	6		3		9
Jawa Tengah	6	6	4	2	18
LPTK UIN Walisongo		6	1	2	9

¹ PINTAR program adopted the entire strategy and methodology of student assessment that had been successfully implemented by a number of projects in the past, including the USAID PRIORITAS. The adoption included the test instruments, the arrangement of data collection, and the data analysis. To facilitate the comparison of test results in the past with the current results of PINTAR program, it was considered efficient to use the same structure of reports and the methods (formats) for presenting the results (tables and charts).

Drovinco / LDTK	Primary		Junior Secondary		Grand
	SD	MI	SMP	MTs	Total
LPTK Universitas Sebelas Maret	6		3		9
Kalimantan Timur	6	6	3	3	18
LPTK IAIN Samarinda		6		3	9
LPTK Universitas Mulawarman	6		3		9
Total	32	28	16	14	90

The program activities at school level are focusing on instruction (teaching and learning), leadership and management of the schools.

The sampling of the schools

The baseline survey which **included the student tests** was conducted in the same sample schools in all 14 districts in five provinces and 10 TTIs. The sample was selected non-randomly from the partner schools. Since the partner schools comprise primary and secondary schools, regular and madrasah (Islamic schools), and in rural and urban locations, the sample was selected so that all of the groups were proportionally represented in the sample.

The following are the steps in selecting the samples from the 24 or more partner schools in each district.

- 1. The 16 or more partner schools at primary level comprise 75% regular (SD) and 25% Islamic schools (MI). As samples, three regular (SD) and one Islamic school (MI) were selected.
- 2. The eight partner schools at secondary level comprise six regular schools (SMP) and two Islamic schools (MTs). As samples, two regular (SMP) and one Islamic (MTs) were selected. Altogether, a district has seven school sample.

The selection of samples at the two TTIs in each province is different because of two reasons. First, each TTI has smaller number of partner schools: six primary and three junior secondary schools. Secondly, the non-religious Universities have only regular schools (SD and SMP) as partners. The State Islamic Universities (UIN/IAIN) has only Islamic schools (MI and MTs) as partners. Because of this, for then non-religious universities, two SD and one SMP were selected as a sample. For the Islamic universities, two MI and one MTs were selected as a sample. For details of the sample see Table 3.

Drovince	N of district	Primary School		Junior Secondary		Total
Province	IN OF district	SD	МІ	SMP	MTs	rotar
North Sumatra	3	9	3	6	3	21
Riau	4	12	4	8	4	28
Jambi	3	9	3	6	3	21
Central Java	2	6	2	4	2	14
East Kalimantan	2	6	2	4	2	14
Total	14	42	14	28	14	98

Table 3: The Number of Sample Schools by Province

Drovinco		Primary School		Junior Secondary		Tatal
Province	NOTII	SD	МІ	SMP	MTs	TOLAT
North Sumatra	2	2	2	1	1	6
Riau	2	2	2	1	1	6
Jambi	2	2	2	1	1	6
Central Java	2	2	2	1	1	6
East Kalimantan	2	2	2	1	1	6
Total	10	10	10	5	5	30
Grand Total (14 district + 10 TTI)	24	52	24	33	19	128

Program Monitoring and Evaluation

The Program has undertaken baseline monitoring activities in a sample of schools in the districts and TTIs listed above in order to assess their needs at the start of the project and to provide a baseline against which, in subsequent years, to assess the impact of the project.

Three activities have been undertaken as follows:

- 1. Monitoring of Teaching and Learning, School Management, and Community Participation.
- 2. Student Assessments in Bahasa Indonesia, Mathematics and Science (for primary and junior secondary schools)
- 3. An Early Grades Reading Assessment (EGRA) and Early Grade Mathematics Assessment (EGMA) (for grade 3)

These activities are reported in separate volumes. This volume concerns the Student Assessments in Bahasa Indonesia, Mathematics and Science.

This monitoring and evaluation activities will be repeated at approximately the same time of year after two years in the same set of schools to assess the project impact in these schools. Analysis will take place after that monitoring to investigate correlations between changes in school management, community participation, teaching and learning and student performance. This will assess whether changes in the way schools are managed and teachers teach are reflected in improved student performance.

An Outline of the Assessment Program

The ultimate success of the PINTAR program must be assessed in terms of the impact on students through the improved quality of teaching and learning. However student performance and its assessment are complex, since they encompasses knowledge and understanding, skills and attitudes. The national school examination and half-yearly tests are limited in their nature mainly to factual recall of knowledge and in many cases are not comparable from year to year or between different geographic areas. The program has, therefore, undertaken its own student performance assessment. The assessment was matched to the objectives of the teacher training program and the government's competency-based curriculum.

The tests have been conducted in a total of four partner primary schools and three partner junior secondary schools in each of the 14 districts. In each TTI, the tests have been conducted in two primary and one junior secondary schools. The tests are as follows:

Primary Schools (SD and MI)	Junior Secondary School (SMP and MTs)
Grade 4: Bahasa Indonesia (Reading and	Grade 8: Bahasa Indonesia (Reading and
Writing)	Writing)
Grade 4: Mathematics	Grade 8: Mathematics
Grade 5: Science	Grade 8: Science

The tests used in primary schools are based on those developed under the World Bank PEQIP² and Basic Education Programs and subsequently also used in the CLCC³, MBE⁴, MGP-BE⁵ and USAID PRIORITAS programs. They have been used over a period of 20 years by these and other programs and undergone revisions based on experience in using them. As a result, the tests can be considered valid and reliable. Tests for Bahasa Indonesia and Mathematics for junior secondary schools were developed by the MBE program and used in the DBE3⁶ program. The science test for junior secondary schools was developed under the PRIORITAS project. Personnel from the Curriculum Development Centre and a number of Teacher Training Universities were involved in the development and subsequent revision of the tests.

The tests will be implemented with the current cohort of students in the above classes in the same schools every two year and at the same time of the school year in order to ensure comparability. For example, the Bahasa Indonesia and Mathematics tests for primary schools will be conducted every two years in the same schools with the current cohort of grade 4 children at the time of testing. This report concerns the first round assessment of students in a sample of schools in PINTAR program partner districts and is intended to establish a baseline against which to assess progress in subsequent years.

The tests have been designed to measure key aspects of the impact of the PINTAR program's teacher training program, as reflected in the development of student competencies. They measure a range of competencies and use a number of different techniques to measure these, including traditional multiple-choice questions, open ended questions and essay questions in the language tests. All the tests are believed to be compatible with the current curriculum. More details of each of the tests is shown in a matrix in Annex 4.

The written tests were developed to take not more than an hour each. The Bahasa Indonesia and Mathematics tests in both primary and junior secondary schools were conducted with half of the relevant class, while the Science tests were conducted with a maximum of 25 randomly selected students per class. The first round of assessment took place in September and November 2018.

When these tests have been used in previous projects, they have included word recognition and reading comprehension tests for grade 1. For PINTAR Program these tests have been replaced by a more comprehensive Early Grades Reading Assessment (EGRA) in grade 3, which has been reported separately.

² Primary Education Quality Improvement Program (1992 – 98)

³ Creating Learning Communities for Children (UNESCO-UNICEF, 1999 – 2010)

⁴ Managing Basic Education (USAID, 2003 – 7)

⁵ Mainstreaming Good Practices in Basic Education (UNICEF-EC, 2007 – 2010)

⁶ Decentralized Basic Education 3 Program (USAID, 2005 – 11)

The report of results of the assessment is set out in three separate parts:

- 1. Summary of the results and recommendations
- 2. First round assessment of students in primary schools

3. First round assessment of students in junior secondary schools

Some implications and recommendations for the implementation of the PINTAR program based on the assessment are included in the report. These have drawn on the extensive experience of the author in working with Indonesian schools and districts as well as reports from those who implemented the testing in the field. It is intended that the report will be discussed with project staff and consultants, trainers and district personnel to make them aware of the results and assess the implications for future PINTAR program activities.

The total possible number of marks in each test varies (e.g. 20 for grade 1 reading, 28 for grade 4 writing, 24 for grade 4 mathematics). However, in order to avoid confusion **all marks have been converted to percentages**.

- In calculating the scores, two types of questions should be taken into account. The first type is a multiple-choice question, for which the answer has only two values: either 1 for the correct answer, or 0 for the wrong answer. If five students in a class of 20 can answer a question correctly, it will be reported as "25% of students could answer the question."
- The second type of question has multiple answers, where each answer can have a different score, depending on how complete an answer is provided. For example, the first question of the Grade 5 Science Test Section B asks students to find three signs in a provided picture that a boat is traveling in a certain direction. The student who can identify at least three signs scores 3, two signs score 2, one sign scores 1, and no signs scores zero. In a class of 10 students, the highest possible score is 10 x 3 = 30. If the actual total score of the students is 12, the average percentage of the (correct) students' answer for this question is (12 ÷ 30) x 100 = 40%. This does not mean that 40% of the students answer correctly, rather it means that the students could achieve 40% of the highest possible score for the question. In this report, such a result is called "the percentage of correct answers."

This method of scoring for the second type of question can be applied in the same way to multiple-choice questions as described above. If five students in a class of 20correctly answer a question, it is reported that "the question has 25% correct answers."

Copies of the tests have not been included with this report in order to avoid their inadvertent dissemination to schools which would make their further use invalid. It is intended that they will be used again in the repeat testing.

1 SUMMARY OF THE RESULTS OF THE TESTS AND RECOMMENDATIONS

1.1 Implementation of the Tests

The first round of tests was administered between September and November 2018 in primary and junior secondary schools in each of the 14 partner districts and 10 TTIs. The samples of school monitored included four partner primary and three junior secondary schools in each district. At the TTI level, the tests were administered in samples of two primary schools and one junior secondary level.

This assessment covered about 25% of the project partner primary schools in those districts. The schools tested included regular schools (SD) and Islamic Schools (MI). The partner schools were chosen from each of two sub-districts targeted by the program. At the Islamic TTIs, all of the sampled schools are Madrasah and at the non-Islamic TTIs, all of the samples are regular schools.

A list of schools and districts tested with the average mark per student in each test is attached in 0.

1.2 How the Results are Presented

The results of these tests are discussed in part 2 of the report (primary schools) and part 3 (junior secondary schools) for each subject separately. The overall average score is given and comparative scores disaggregated for boys and girls. The average scores of higher and lower achieving groups of students are also presented by quartile.

The primary schools scores are also disaggregated between (i) those students who have attended pre-school education (TK) and those who have not, and (ii) regular primary schools (SD) and Madrasah Ibtidaiyah (MI); (iii) public and private schools; (iv) urban and rural, and (VI) district and TTI partner schools. A breakdown of the scores on individual questions is presented on the mathematics and science tests and for each section of the science test.

The junior secondary school scores are also disaggregated between (i) boys and girls,(ii) regular junior secondary schools (SMP) and Madrasah Tsanawiyah (MTs),(iii) urban and rural schools, (iv) public and private, and (v) districts vs TTI schools. A breakdown of the scores on individual questions is presented on the mathematics and science tests and for each section of the science test.

Future rounds of testing will track whether any improvements in scores are spread over the different groups, for example: Do SD and MI improve equally? Are low achieving and higher achieving groups improving equally?

It needs to be stressed that the samples in districts, only four primary schools and three junior secondary schools were included in the test. Results of the tests from individual schools are included in Annex 1, but should not be viewed as being a representative sample of the districts' schools. The results of the primary and junior secondary school tests have been aggregated also by district, but the sample of the schools were too small and cannot be used to represent the results of each district as a whole. Because of this, comparisons of individual school or district performance are kept to a minimum in the report.

1.3 Summary of Results in Primary Schools (SD and MI)

The partner schools tested in the 14 districts and ten TTIs included 52 partner primary schools (SD) and 24 partner religious primary schools (MI). A total of 973 students were involved in reading and writing test, 983 in mathematic test and 1,144 in science test. Table 4 gives a summary of the results of each test.

			Grade 5		
		Reading	Writing	Math	Science
		(%)	(%)	(%)	(%)
N Student Tested		973	973	983	1,144
Attended Pre-School (TK)	85.5%	85.5%	88.8%	87.3%
Test Results (% of Correc	ct Answers)				
Gender	Boys	42.4%	40.8%	37.5%	32.8%
	Girls	50.5%	38.6%	41.4%	34.6%
	Attend	47.3%	39.8%	40.4%	34.6%
Attend Pre School (TK)	Not Attend	41.1%	39.3%	32.6%	27.5%
Cohool Turo	Regular	46.1%	39.4%	40.4%	33.4%
School Type	Madrasah	47.2%	40.4%	37.4%	34.6%
School Status	Public	46.5%	38.7%	39.8%	33.2%
	Private	46.1%	42.7%	38.7%	35.5%
Leastien	Urban	50.4%	39.8%	41.8%	36.4%
LOCATION	Rural	40.6%	39.6%	36.1%	30.1%
Cohool Comula	District School	45.7%	41.0%	37.4%	32.0%
School Sample	TTI School	48.4%	36.0%	45.5%	38.8%
Average Score		46.4%	39.7%	39.5%	33.7%

Table 4: Summary of Test Result for all Test in Primary Schools

Grade 4 Bahasa Indonesia Test: In the grade 4 Bahasa Indonesia test, the average score of all schools tested was 46.4% for reading and 39.7% for writing. 9% of children in partner schools wrote nothing. It is evident that many grade 4 children in PINTAR program schools have difficulty in comprehending meaning in what they read and in communicating ideas in a coherent and legible manner.

Grade 4 Mathematics Test: In the grade 4 Mathematics test, the overall average score was 39.5%. Areas in which students had particular difficulties included recognising the value of both decimal and simple fractions and operations with decimal fractions. Students also scored very low on questions which required problem solving and creativity in their answers.

Grade 5 Science Test: In the grade 5 science test, the overall average score on the test was 33.7%. Children found the traditional format of questioning (with multiple choice answers) in Section A easier than in Section B, which required them to make deductions and apply concepts which they have learned.

Comparisons between Different Groups: On all tests (except on writing) girls scored higher than boys. Scores of children who attended TK (pre-school) were substantially higher than those who had not. From observations at school level it appears that many children who have

attended TK enter primary school already having mastered some of the basics of literacy and numeracy, which gives them a significant advantage over the length of their school career.

Average scores at MI were higher than at SD except in Mathematics. Historically madrasahs have tended to underperform secular schools on these tests, and the reason for the apparent over-performance on these tests is not evident from the data collected. The majority of these private schools were madrasahs, which suggest that the MI selected to take part in the assessment were among the better endowed madrasahs, possibly run by well-funded foundations. The number of private schools was in any case small, 33 schools or 26% of the sample, which may reduce the significance of the figures for private schools.

Differences between Schools and Districts: The schools chosen to take part in the PINTAR program and then chosen as samples from among these schools to take part in the student assessment were not intended as a representative sample of the schools in each district. However, the average school and district scores for primary schools and the average school scores for junior secondary schools have been included in Annex 1. There were large differences in scores between schools. The school's and student's highest and lowest scores on each of the tests were as set out in Table 5 and 6 below.

Test	Highest School's Average Score	Lowest School's Average Score	Number of Schools with Average Score below 10%
Reading Grade 4	74.5%	7.9%	1
Writing Grade 4	65.0%	15.0%	0
Mathematics Grade 4	60.6%	2.8%	1
Science Grade 5	54%	17%	0

 Table 5:
 Highest and Lowest School's Average Scores in Primary School Tests

Table 6:	Highest and Lowest Student's Scores in Primary School Tests

Test	Highest Student's Score	Lowest Student's Score	Number of Students with Score below 10%
Reading Grade 4	100%	0%	60
Writing Grade 4	100%	0%	78
Mathematics Grade 4	92%	0%	54
Science Grade 5	83%	0%	65

While some differences can be explained by different student intakes, the largest reason for the differences must lie with the quality of teaching.

A table comparing the results from the USAID PRIORITAS, MGMP-BE, MBE and CLCC programs is set out in Annex 2.

1.4 Summary of Results in Junior Secondary Schools (SMP and MTs)

The student assessments took place between October and November 2018 in 52 partner schools (28 SMP and 14 MTs). At least 950 students were tested overall in each group for each subject. The results are summarized in Table 7 and discussed briefly below. They are then

analyzed and discussed in more detail in part 2 of this report. The results for each school can be found in Annex 1.

Descriptions		Grade 8				
		Reading Test	Writing Test	Mathematics Test	Science Test	
		(%)	(%)	(%)	(%)	
N Student Tested	l	956	956	963	944	
Test results						
Condor	Boys	60.8%	40.9%	30.8%	33.9%	
Gender	Girls	68.1%	51.7%	31.3%	35.2%	
	Regular	63.7%	45.7%	31.5%	34.2%	
School Type	Madrasah	66.0%	47.7%	30.4%	35.2%	
School Status	Public	64.8%	46.9%	31.3%	35.1%	
School Status	Private	64.0%	45.0%	30.4%	33.0%	
Location	Urban	65.6%	47.4%	33.3%	36.9%	
Location	Rural	62.7%	44.7%	26.8%	30.2%	
Comula	District School	63.7%	46.9%	29.8%	34.3%	
Sample	TTI School	67.9%	44.7%	36.2%	35.6%	
Overall		64.6%	46.5%	31.1%	34.6%	

 Table 7:
 Summary of Test Results for all Tests in Junior Secondary Schools

Bahasa Indonesia Test: The average scores in Bahasa Indonesia Reading and Writing tests were 64.6% and 46.5% respectively. The highest quartile of students in sampled schools scored 82% on the reading test and the lowest quartile 41%.

Between 24% and 39% of students scored poorly or very poorly in the writing test in terms of the ability to write in sentences, the quality of ideas, spelling and punctuation and handwriting. Almost 25% had difficulty writing in paragraphs.

Mathematics Test: The scores on the mathematics test were relatively low (on average only 31.1% of the possible score), which reflects the difficulty students had answering the questions with an emphasis on understanding and needing the ability to apply concepts. Students found considerable difficulty with questions which involved problem solving and had to be worked in two or more stages (i.e. solving one part of the problem first and then using the answer from that part of the problem to solve the whole problem).

Science Test: The average overall score in the test was 34.6%, with partner schools. Students were relatively weak in all areas, but especially where they had to reason or make deductions from data. They also seem not to have acquired measuring skills through practical work. For example, they had difficulty in reading measurements of a ruler and reading weighing scales and measuring cylinders. They also had a weak knowledge of technical terms and difficulty in applying concepts to everyday situations.

Comparisons between Different Groups: Girls performed considerably better than boys in the Bahasa Indonesia reading and writing tests. There was little difference between boys and girls in the mathematics and science test. MTs students performed better than SMP students on three tests (except in Mathematics).

Differences between Schools: There were wide differences in average scores between schools in every subject, indicating that students are learning much better in some schools

than in others. In some cases there will be mitigating social and economic circumstances. However, it is noticeable that many schools rate relatively well in one subject and poorly or very poorly in another (see Annex 1 for a complete list of school scores). This indicates variable quality in the teaching within the same school. Tables 8 and 9 show the highest and lowest scores on each test.

Test	Highest School's Average Score	Lowest School's Average Score	Number of Schools with Average Score below 10%
Reading Grade 8	80%	47%	0
Writing Grade 8	66%	22%	0
Mathematics Grade 8	70%	17%	0
Science Grade 8	54%	17%	0

Table 8: Highest and Lowest School's Average Scores in Junior Secondary School Tests

Table 9:	Highest and Lowest Student's Scores in Junior Secondary School Tests
----------	--

Test	Highest Student's Score	Lowest Student's Score	Number of Students with Score below 10%
Reading Grade 8	100%	0%	3
Writing Grade 8	100%	0%	81
Mathematics Grade 8	95%	0%	21
Science Grade 8	80%	0%	40

1.5 Implications and Recommendations for PINTAR program

A. General

- The better scores achieved by children who have attended kindergarten (TK) suggest that district should prioritize the provision of pre-school education, but also make sure that teachers are well trained to help children make the best of their opportunity.
- Based on general experience in Indonesian schools, improvements both in the areas of teachers' subject knowledge and the approaches and methodology of teaching are needed to improve student performance. However, the PINTAR program should build on lessons learned from previous programs concerning the need to link improved teaching to other system improvements at school and district level, including improving school management, increasing community support for schools and improved support and management from district governments.
- Government policy stresses the use of Active, Joyful, Creative and Effective Learning (better known by its Indonesian acronym – PAKEM) as its preferred teaching approach. However, teacher training under PINTAR program needs to identify *specific weaknesses in the teaching of the various subjects* and help teachers develop strategies and methodologies to address these weaknesses. Some more specific suggestions are set out below. Training should pay attention to the special needs of specific districts and schools.

B. Bahasa Indonesia

- A problem which was reported from a number of primary schools was a lack of mastery of Bahasa Indonesia, with schools which appear to have similar backgrounds showing different levels of success in helping their students master the language. Previous experience has shown that this is often dependent on the will and commitment of teachers and that local government and especially school supervisors and principals can do much to promote the use of Bahasa Indonesia in their schools.
- It is evident that many grade 4 children in PINTAR program schools have difficulty in comprehending meaning in what they read, especially when it is not explicit in the text and in communicating ideas in a coherent and legible manner. Mastery of language is the key to success across the curriculum and, in many cases, in later life. This highlights the importance of training in the teaching of Bahasa Indonesia. From observations in many schools around the country language teaching focuses too narrowly on the mechanics of reading (often barking at print) and writing is confined largely to copying words and sentences or filling in words in sentences from the text book or the teacher.
- In line with the competency based curriculum, Bahasa Indonesia training should focus on developing students' language skills. Teachers should be trained to give their students opportunities to write for a variety of purposes including reporting facts and events, writing instructions and expressing their feelings and opinions. Children also need to be given the opportunity and to be taught to read for different purposes including for enjoyment and finding information and to reflect on and report back on what they have read.
- Teachers need to give their students the opportunity to develop their speaking and listening skills by giving them the opportunity to discuss a variety of issues and problems. Speaking and listening can and should often be linked to reading and writing activities with students being invited to discuss and comment on what they read and discuss ideas before they begin to write. They should also be given the opportunity to read and give feedback on each other's work.
- Teaching should pay attention to handwriting, spelling and punctuation, which need to be taught regularly and systematically and appear to have been neglected in many schools. While punctuation and spelling should be introduced through special lessons, they need to be reinforced through the children's own writing. Children need to be encouraged to get into the habit of re-reading their own writing and correcting spelling, punctuation and other errors.

C. Mathematics

- Experience in Indonesia has shown that mathematics is generally poorly taught. Many teachers have a poor understanding of the concepts they are teaching and tend to teach rules and procedures for doing mathematical operations rather than cultivating an understanding of the concepts. As a result students have difficulty applying the concepts and using mathematics as a tool for solving problems.
- Training for teachers should focus on helping both teachers and students to gain an understanding of mathematical concepts, especially by relating them to real situations in areas such as number, measurement, geometry and graphical representation.

- Poor language skills also inhibit children's ability to read and understand information and instructions and to explain how they have solved problems and what these solutions are.
- Teachers should be encouraged to adopt 'problem solving' approaches to teaching mathematics, which also encourage creativity and develop understanding. This can include children being asked to think of a variety of answers to open ended problems, being asked to make up their own questions for other children to answer and being asked to make up a variety of questions which will result in the same answer (e.g. how many questions can you make with the answer '20', how many different shapes can you make with an area of 24cm²?).

D. Science

- Science teaching focuses too much on the memorisation of rules and concepts and too little on developing understanding of and applying concepts. Too little practical work takes place to support science teaching. Students spend much of their time memorising information from books rather than developing scientific skills such as measuring, observing real phenomena, data analysis, making hypotheses and drawing conclusions.
- Teacher training should focus on developing students' scientific skills based on the
 observation of the real environment and doing experiments to investigate natural
 phenomena. Training should include helping students to make systematic reports on the
 experimental and observational work they undertake. Simple technology activities should
 be promoted to encourage students to apply scientific concepts in real situations.

2 FIRST ROUND ASSESSMENT OF STUDENTS IN PRIMARY SCHOOLS

The student assessment took place between September and November 2018. During September-October, the test was carried out in 56 primary schools (42 SD and 14 MI) in the Cohort 1 District. During October-November the second phase, the test was carried out in a sample of 20 TTI partner schools. The total sample in primary level schools was 76 primary schools. The results are reported below by subject, together with the implications and recommendations for PINTAR program.

2.1 Bahasa Indonesia Grade 4

2.1.1 Introduction

Traditional Bahasa Indonesia tests assess knowledge of the Indonesian language rather than children's functional language skills although the new curriculum emphasizes the development of all four language skills. This particular test focused on skills and was divided into two parts. The first part, reading comprehension, tests children's ability to read an extended piece of writing with understanding. The second part, story writing, tests children's ability to extract ideas from a picture and, using their imagination, to produce a story based on that picture. The final score for writing was a composite of five scores for the different skills of handwriting, spelling, punctuation, length of the written piece and the quality of language used. A matrix showing how the scores for the writing test were compiled in included as Annex 3.

2.1.2 The Results

Table 10 shows the average scores obtained in the two tests. The average score was 46% for reading and 40.2% for writing. Comparison across groupings shows that girls have higher scores than boys, students attending kindergarten have higher scores than who did not; students in urban schools have higher scores that in rural, and regular schools have higher than Madrasah schools in writing but the other way around in reading.

		Student Test		Reading	Writing
		Ν	%	Test (%)	Test (%)
Condor	Boys	493	51%	42%	41%
Gender	Girls	480	49%	51%	40%
Attend Pre School	Attend	832	86%	47%	41%
(ТК)	Not Attend	141	14%	41%	40%
	Regular	684	70%	46%	41%
School Type	Madrasah	289	30%	47%	40%
	Public	743	76%	47%	40%
School Status	Private	230	24%	46%	43%
Location	Urban	582	60%	50%	41%
LOCATION	Rural	391	40%	41%	40%
	District School	714	73%	46%	42%
School Sample	TTI School	259	27%	48%	36%
Overall		973	100%	46%	41%

Table 10: Participant Data and Average Scores in Grade 4 Reading and Writing Tests

There were large differences between individual schools with the highest having an average student score of 74.5% and the lowest 7.9% on the reading test and the highest 65.0% on the writing test compared to 15.0% for the lowest. One school had average scores below 10% on the reading test and no school has the writing test below 10% (Table 5).

2.1.3 Reading

The results of reading test disaggregated by various grouping are shown in Chart 1. Three groupings have considerable differences: girls, pre-school attendance and urban schools have higher percentages than boys, no pre-school attendance and rural schools. In three other groupings such as between regular vs madrasah, public vs private, and district vs TTI schools the percentage differences are relatively small.



Chart I: Primary Reading Test - Comparison between Different Groups

Chart 2 shows the average score per quartile in district and TTI schools. In the first three quartiles, there are no differences between district and TTI schools. There was a small difference in the lowest quartile. These scores will be monitored in the future to see whether improvement is spread evenly over all abilities (i.e. Does each quartile show improvement in both district and TTI school or, for example, just the highest or lowest quartiles?)



Chart 2: Average Score by Quartile in Primary School Reading Comprehension Test

The test was divided into three sections. Section A gave multiple choices of words to complete sentences about a reading passage. Section B required the students to evaluate whether statements about the passage were true or false, while Section C required students to deduce information from or attempt to explain what they had read. As can be seen from Table 11 below, the students found Section C most difficult with an average of 32.2% of questions answered correctly compared to around well over 55% for the other sections.

Section	% Correct
Section A	59.6%
Section B	64.9%
Section C	32.2%
Total	46.4%

Table 11:	Scores by Section	in Primary School	Reading Test
-----------	--------------------------	-------------------	---------------------

2.1.4 Writing

The results disaggregated by various grouping are shown in Chart 3. In five groupings (gender, preschool attendance, school type and status, geographic location) the scores only ranges between 40% to 43%. Only in the last group the difference is slightly bigger: district schools have higher percentage (42%) than TTI schools (36%).



Chart 3: Primary Writing Test - Comparison between Different Groups

The writing test was assessed according to five elements: handwriting, spelling, punctuation, length and the quality of the writing. The weighting in the overall score was handwriting (15%), spelling (15%), punctuation (15%), length (20%), quality of the writing (35%).

The results are summarized in Table 12..

- **Handwriting**: only 10% of children could write neat, cursive writing, while another 54% wrote neatly but without joining their writing. 24% of the children's writing was classified as poor and 12% of children did not score.
- **Spelling**: only 7% of children wrote without spelling mistakes, while in 32% more cases the spelling was good (largely correct). In 45% of cases the children's spelling was considered poor while 16% obtained no score, which implies that their writing was not comprehensible.
- **Punctuation**: only 5% of children punctuated their work well (classified as perfect or good), while in 19% of cases children's punctuation was classified as good. The punctuation of 49% of children was considered as poor and 27% did not use punctuation.
- Length: 25% of children wrote half a page or more, 46% wrote more than two sentences while 29% wrote less than two sentences (including 8% who wrote nothing).
- **Quality**: The quality of the content of the 19% children's writing was classified as very good or good, i.e. their ideas were clear and well expressed and in logical order. 42% of children's writing was considered as poor, 30% was ranked as poor in quality, while 9% received no score, because they either wrote nothing or what they wrote was incomprehensible. In general, children were weak in organising their thoughts in a systematic way.

Handwriting	Good Joined	Good Printed	Poor		No Score
	10%	54%	24%		12%
Spelling	Perfect	Good	Poor		No Score
	7%	32%	45%		16%
Punctuation	Perfect	Good	Poor		No Score
	5%	19%	49%		27%
Length	> 1 Page	Half Page	>2 Sentences	<2 Sentences	No Writing
	2%	23%	46%	21%	8%
Quality	Very Good	Good	Fair	Poor	No Writing
	4%	15%	42%	30%	9%

Table 12: Percentage Scores for Elements of Written Work in Primary School Writing Test





Chart 4 shows the average score per quartile. The top 25% of students in partner schools scored on average 66%, whereas the lowest quartile scored on average of 13%. The differences between district and TTI are small except in the lowest quartile. These scores will be monitored to see whether any improvement is spread evenly over all abilities.

2.1.5 Implications and Recommendations for PINTAR program

- It is evident that many grade 4 children in the schools tested have difficulty in comprehending meaning in what they read and in communicating ideas in a coherent and legible manner. They have particular difficulty in 'reading between the lines' to extract meaning that is not explicit. Mastery of language is the key to success across the curriculum and, in many cases, in later life. This highlights the importance of training in the teaching of Bahasa Indonesia. From observations in many schools around the country, language teaching focuses too narrowly on the mechanics of reading (often barking at print) and writing is confined largely to copying words and sentences.
- The emphasis in future teacher training should be on improving students' communication skills including the ability to get meaning from what they hear and read and communicate their own ideas better in both spoken and written form. The ability to communicate for

different purposes to different audiences should be taught by the introduction of appropriate text types, as outlined in Curriculum 2013. Reading should be embedded in a more complete literacy program which develops these skills. One activity which should be encouraged is children reading their peers' work and giving feedback. The benefits are two way as the reader gains new skills and insights and the writer is given new ideas and feedback.

 Language teaching should pay attention to handwriting, spelling and punctuation, which need to be taught regularly and systematically and appear to have been neglected in many schools. While punctuation and spelling should be introduced through special lessons, they need to be reinforced through the children's own writing. Children need to be encouraged to get into the habit of re-reading their own writing and correct spelling, punctuation and other errors.

2.2 Mathematics Test Grade 4

2.2.1 Introduction

The mathematics test was revised substantially in 2004 compared to the test used in PEQIP and the World Basic Education Projects in order to give a greater emphasis on testing children's understanding and their problem solving capabilities.

2.2.2 The Results

Table 13 and Chart 5 show that the overall average score on the test was 40%. In four groupings, the differences between TTI and district schools and between preschool attendance and no attendance are the highest (7%), followed by difference between urban and rural (6%), and between boys and girls (4%). The differences between Madrasah and regular school and between public and private are rather small (3% and 2% respectively).

		N of Stude	N of Student Tested		
		Ν	%	Test (%)	
Condor	Boys	469	48%	37%	
Gender	Girls	514	52%	41%	
Attend Pre School	Attend	873	89%	40%	
(ТК)	Not Attend	110	11%	33%	
School Turno	Regular	691	70%	40%	
School Type	Madrasah	292	30%	37%	
School Status	Public	748	76%	40%	
	Private	235	24%	39%	
Location	Urban	593	60%	42%	
LOCATION	Rural	390	40%	36%	
School Somolo	District School	728	74%	37%	
School Sample	TTI School	255	26%	45%	
Overall		983	100%	40%	

 Table 13: Participant Data and Average Scores in Primary School Mathematics Test



Chart 5: The Grade 4 Mathematic Test Results Disaggregated by Six Grouping

There were large differences between individual schools with the highest having an average student score of 60.6% and the lowest 2.8% (one school). The next lowest score is 17.3% (Table 5).

Chart 6 shows that the average scores per quartile in district and TTI sample schools are very similar. These scores will be monitored to see whether any improvement is spread evenly over all abilities (i.e. Does each quartile show improvement or, for example, just the highest or lowest quartiles?).



Chart 6: Distribution of Scores by Quartile in Primary School Mathematics Test

Chart 7 presents the result of the mathematic test and rank the test item according their level of difficulties measured by the percentage of students who could answer them correctly. The number in front of the item is the order of the item in the test.



Chart 7: Results of Mathematic Test Sorted by Their Level of Difficulties

As shown in the chart, four of the most difficult questions are as follows. The first question asks the students to buy four different items with a Rp.10.000 bill. The second question asks the student to order four decimal number from the smallest to the largest. The third question is about addition of four decimals numbers, and the fourth is completing a number series.

2.2.3 Implications and Recommendations for PINTAR program

- Experience in Indonesia has shown that mathematics is poorly taught in many classes. Many teachers have a poor understanding of the concepts they are teaching and tend to teach rules and procedures for doing mathematical operations rather than cultivating an understanding of the concepts. As a result students have difficulty applying the concepts in real life and using mathematics as a tool for solving problems.
- Training for teachers should focus on the development of students' conceptual thinking and the systematic teaching of number concepts from the physical to the verbal to the symbolic. It should focus on helping both teachers and students to gain an understanding of mathematical concepts by relating them to real situations in areas such as number, money, measurement, geometry and graphical representation.

 Teachers should be encouraged to adopt 'problem solving' approaches to teaching mathematics, which also encourage creativity and develop understanding. This can include children being asked to think of a variety of answers to an more open ended problem, being asked to make up their own questions for other children to answer and being asked to make up a variety of questions which will result in the same answer (e.g. how many questions can you make with the answer '20', how many different shapes can you make with an area of 24cm²?).

2.3 Science Test Grade 5

2.3.1 Introduction

This test was divided into two sections. Section A used the familiar format of multiple choice questioning to assess students' understanding of concepts they have already learnt. Section B assessed their process skills such as the ability to observe, interpret and hypothesize (i.e. providing tentative answers based on previous knowledge and experience). Some of the test items also assessed the ability to apply basic science concepts to everyday situations.

2.3.2 The Results

Table 14 and Chart 8 show that the overall average score on the test was 34%. In six groupings, the differences between TTI and district schools and between preschool attendance and no attendance are the highest (7%), followed by difference between urban and rural (6%), and between boys and girls (4%). The differences between Madrasah and regular school and between public and private are rather small (2% and 3% respectively).

		Stude	Score	
		Ν	%	Test (%)
Condor	Boys	545	48%	33%
Genuer	Girls	599	52%	35%
Attend Pre School	Attend	999	87%	35%
(ТК)	Not Attend	145	13%	28%
School Turno	Regular	806	70%	33%
School Type	Madrasah	338	30%	35%
	Public	867	76%	33%
School Status	Private	277	24%	36%
Location	Urban	658	58%	36%
LOCATION	Rural	486	42%	30%
	District School	849	74%	32%
school sample	TTI School	295	26%	39%
Overall		1,144	100%	34%

 Table 14: Participant Data and Average Scores in Primary School Science Test



Chart 8: Primary Science Test - Comparison between Different Groups

There were large differences between individual schools with the highest having an average student score of 56.6% and the lowest 15.9%. No school had average scores below 10% (See Table 5).

Chart 9: Distribution of Scores by Quartile in Primary School Science Test



Chart 9 shows the average score per quartile. The top 25% of students averaged 54.3%. The lowest 25% scored on average 14%. The differences between district and TTI schools are rather small in all four quartiles.

As can be seen from Table 15 below children found the traditional format of questioning (with multiple choice answers) in Section A is easier than Section B, answering an average of 41.3

in section A. Section B, where they answered an average of 30.7% require them to make deductions and apply concepts which they have learned.

Section	Average Score
Section A	41.3%
Section B	30.7%
Total	33.7%

 Table 15: Average Scores by Section in the Primary School Science Test

Chart 10 shows the percentage of correct answers to individual questions. The questions where students had most difficulty were those where they had to interpret data and where they had to give open ended answers, i.e. there were no multiple-choice answers to select from. This suggests that students are more confident selecting right answers, when they are given a choice, but lack confidence or skills to construct an answer by themselves.



Chart 10: Analysis of Scores by Question in Primary School Science Test

2.3.3 Implications and Recommendations for PINTAR Program

 Science teaching currently focuses too much on the memorisation of rules and concepts and too little on developing understanding of and applying concepts. Too little practical work takes place to support student learning. Students spend much of their time memorising information from books rather than developing scientific skills such as observation of real phenomena, data analysis, making hypotheses and drawing conclusions.

• Teacher training should focus on developing students' scientific skills based on the observation of the real environment and doing experiments to investigate natural phenomena. Training should include helping students to make systematic reports and draw their own conclusions on the experimental and observational work they undertake. Simple technology activities should be promoted to encourage students to apply scientific concepts in real situations.

3 FIRST ROUND TESTING OF JUNIOR SECONDARY SCHOOLS

The student assessment took place between September and November 2018 in 52 sample of PINTAR program schools (42 in district schools and 10 in TTI schools) as shown in Table 16.

		Grade 8								
Desc	rintions	Reading	Writing	Mathematics	Science					
Dese		Test	Test	Test	Test					
		(%)	(%)	(%)	(%)					
N Student Tes	ted	956	956	963	944					
Test results										
Condor	Boys	60.8%	40.9%	30.8%	33.9%					
Gender	Girls	68.1%	51.7%	31.3%	35.2%					
Location	Urban	65.6%	47.4%	33.3%	36.9%					
Location	Rural	62.7%	44.7%	26.8%	30.2%					
	Regular	63.7%	45.7%	31.5%	34.2%					
School Type	Madrasah	66.0%	47.7%	30.4%	35.2%					
School	Public	64.8%	46.9%	31.3%	35.1%					
Status	Private	64.0%	45.0%	30.4%	33.0%					
Commis	District School	63.7%	46.9%	29.8%	34.3%					
Sample	TTI School	67.9%	44.7%	36.2%	35.6%					
Overall		64.6%	46.5%	31.1%	34.6%					

Table 16: Summary of Test Result for all Test in Junior Secondary Schools

3.1 Bahasa Indonesia Grade 8

3.1.1 Introduction

Traditional Bahasa Indonesia tests assess knowledge of the Indonesian language rather than children's functional language skills although the new curriculum emphasizes the development of all four language skills. This particular test focused on skills and was divided into two parts. The first part, reading comprehension, tests children's ability to read an extended piece of writing with understanding, including the ability to deduce meaning from a text. The second part, the writing test, assesses children's ability to extract ideas from a picture and, using their imagination, to produce a logical and well-ordered piece of writing based on the picture. The final score for writing was a composite of five scores for the different components of (i) paragraphing and (ii) sentencing, (iii) the quality of the ideas expressed, (iv) spelling and punctuation and (v) handwriting.

3.1.2 The Results

Table 17 shows the average scores obtained in the two tests. The average score was 63.7% for reading and 46.9% for writing. Girls scored substantially higher than boys in both reading and writing. Madrasah students scored higher than SMP students on both tests. The scores are almost the same between urban and rural schools.

Dece	intions	Studen	t Tested	Grade 8			
Desci	npuons	N	%	Reading	Writing		
Condox	Boys	463	48%	61%	41%		
Gender	Girls	493	52%	68%	52%		
Location	Urban	624	65%	66%	47%		
Location	Rural	332	35%	63%	45%		
	Regular	592	62%	64%	46%		
School Type	Madrasah	364	38%	66%	48%		
Sahaal Status	Public	724	76%	65%	47%		
School Status	Private	232	24%	64%	45%		
Sample	District School	760	79%	64%	47%		
	TTI School	196	21%	68%	45%		
Average		956	100%	65%	46%		

Table 17: Participant Data and Average Scores in Grade 8 Reading and Writing Tests

3.1.3 Reading

The test was divided into three sections. Section A gave multiple choices of words to complete sentences about a reading passage. Section B required the students to evaluate whether statements about the passage were true or false, while Section C required students to deduce information from or attempt to explain what they had read. As can be seen from Table 18, the students found section A the easiest with an average score of 71%. However, they did not find much greater difficulty with the other sections. This appears to show that many had reasonable facility in understanding both overt and hidden meaning in the passage of reading.

Table 18: Scores by Section in the Grade 8 Reading Comprehension Test

Section	Score
Section A	71.2%
Section B	60.3%
Section C	64.6%
Total	66.5%

Chart 11 shows that the overall average score on the test was 64%. In six groupings, the differences between TTI and district schools and between preschool attendance and no attendance are the highest (7%), followed by difference between urban and rural (6%), and between boys and girls (4%). The differences between Madrasah and regular school and between public and private are rather small (2% and 3% respectively).



Chart II: Junior Secondary B. Indonesia Test in Reading- Comparison between Different Groups

Chart 13 shows that the differences between district and TTI in the first three quartiles are small but in substantial in the lowest quartiles (42% in district and 38 in TTI schools).





3.1.4 Writing

Table 19 shows data for each of the components of the writing test: (i) paragraphing and (ii) sentencing, (iii) the quality of the ideas expressed, (iv) spelling and punctuation and (v) handwriting. About 18% to 37% of students scored perfectly in these components (excellent and very good lumped together) in the five component. However, 28% to 52% of students scored 'poor" and "very poor".

Paragraph								
Excellent (%)	Very Good (%)	Good (%)	Poor (%)	Very Poor (%)				
2%	16%	30%	43%	9%				
	Se	entences						
Excellent (%)	Very Good (%)	Good (%)	Poor (%)	Very Poor (%)				
2%	24%	44%	20%	10%				
Quality of Ideas								
Excellent (%)	Very Good (%)	Good (%)	Poor (%)	Very Poor (%)				
3%	23%	45%	20%	8%				
	Spelling a	nd Punctuat	tion					
Excellent (%)	Very Good (%)	Good (%)	Poor (%)	Very Poor (%)				
2%	22%	41%	26%	9%				
Handwriting								
Excellent (%)	Very Good (%)	Good (%)	Poor (%)	Very Poor (%)				
6%	31%	35%	20%	9%				

 Table 19:
 Percentage Score for Components of Written Work in Grade 8 Writing Test

Chart 13 shows the scores per quartile of students in the writing test from the highest to the lowest in district and TTI schools. TTI has the higher score than the district schools in the first (highest) quartile. In the last quartile, the district schools have higher percentages.

The scores of each quartile will be reported next year to assess whether progress is evenly (or not evenly) spread in district and TTI schools across all learners, from the best to the poorest.



Chart 13: Average Student Scores by Quartile in Grade 8 Writing Test

3.1.5 Implications and Recommendations for PINTAR Program

• As in primary schools, much of the emphasis in language teaching has been on teaching about language rather than developing students' skills in using language. Where students get to write, it is often only by inserting words in sentences provided by the teacher or

the textbook. There are few opportunities for students to express their own thoughts by, for example, making reports or expressing their feelings or opinions. Reading comprehension also tends to be confined to repeating facts set out in the text. There are few opportunities research information or to read 'behind the text'.

- To conform with the competency-based curriculum, Bahasa Indonesia training should focus on developing students' skills in reading and writing. Teachers should be trained to give their students opportunities to write for a variety of purposes including reporting facts and events, write instructions and expressing their feelings and opinions. Children also need to be given the opportunity and taught to read for different purposes, including for enjoyment and finding information, as well as to reflect on and report back on what they have read.
- Teachers also need to give their students the opportunity to develop their speaking and listening skills by giving them the opportunity to discuss a variety of issues and problems. Speaking and listening can and should often be linked to reading and writing activities with students being invited to discuss what they read and discuss ideas before they begin to write. They should also be given the opportunity to read and give feedback on each other's work.

3.2 Mathematics Test Grade 8

3.2.1 Introduction

The mathematics test was designed to lay emphasis on testing children's understanding of mathematical concepts and their ability to apply these concepts in solving problems. The test was revised and some of the questions simplified, following their use between 2005 and 2007 in the assessment of the MBE program in Central and East Java.

3.2.2 The Results

Table 20 shows that the overall average score on the test was 31.1%. Boys and girls had the same average scores. Urban schools scored higher than rural schools and Madrasah scored slightly higher than SMP.

		Stud	ent Test	Score		
		Ν	%	Test (%)		
Condor	Boys	443	46%	31%		
Gender	Girls	520	54%	31%		
School Type	Regular	596	62%	31%		
School Type	Madrasah	367	38%	30%		
School Status	Public	731	76%	31%		
School Status	Private	232	24%	30%		
Location	Urban	627	65%	33%		
Location	Rural	336	35%	27%		
Samplo	District School	769	80%	30%		
Sample	TTI School	194	20%	36%		
Overall		963	100%	31%		

Table 20: Participant Data and Average Scores in Grade 8 Mathematics Test

There were large differences between individual schools with the highest having an average student score of 70% and the lowest 17% (Table 8). The highest student's score is 95% and the lowest is 0%. (Table 9). A total of 44 students (5.8%) scored over 80% and 12 students (1.1%) failed to score (Table 9).

Table 20 and Chart 14 disaggregate the results of the test. In five background variables ((gender, school type, school status), the differences between the categories are not more than two percent. Only in location (rural-urban) and sample (district-TTI), the differences are six percent.



Chart 14: Junior Secondary Mathematics Test - Comparison between Different Groups

Chart 15 shows the average scores per quartile of students. Only in the highest quartile, the difference between district and TTI sample schools is significant. The other three quartiles, the averages are exactly the same. The scores of each quartile will be reported next year to assess whether progress is evenly spread across all learners, from the best to the poorest.





Chart 16 shows the percentage of children scoring correct in each of the 15 questions in the test. The chart shows four questions which the students found most difficult (less than 20% of students able to answer correctly). The questions are dealing finding the number of squares (14%), open ended number problems (16%), ordering decimals and fractions (18%), and openended area problem.

Many of the questions which involved problem solving had to be worked in two or more stages (i.e. solving one part of the problem first and then using the answer from that part of the problem to solve the whole problem). Students found this especially difficult.





3.2.3 Implications and Recommendations for PINTAR program

- As in primary schools, mathematics is poorly taught in many secondary school classes. Many teachers have a poor understanding on the concepts they are teaching and tend to teach rules and procedures for doing mathematical operations rather than cultivating an understanding of the concepts. As a result, students have difficulty applying the concepts in real life and using mathematics as a tool for solving problems.
- Training for teachers should focus on helping both teachers and students to gain an understanding of mathematical concepts, especially by relating them to real situations in areas such as number, money, measurement, geometry and graphical representation.
- Teachers should be encouraged to adopt 'problem solving' approaches to teaching mathematics, which also encourage creativity and develop understanding. This can include children being asked to think of a variety of answers to a more open-ended problem, being asked to make up their own questions for other children to answer and being asked to make up a variety of questions which will result in the same answer.

3.3 Science Test Grade 8

3.3.1 Introduction

This science test is divided into two sections. Section A has ten questions using the familiar format of multiple choice questioning to assess children's understanding of concepts they have already learnt. Section B consisted of six questions and assessed children's process skills such as the ability to observe, interpret and hypothesize (i.e. providing tentative answers based on previous knowledge and experience). Some of the test items also assessed the ability to apply basic science concepts to everyday situations. A number of the test items were adapted from TIMSS⁷ test items.

3.3.2 The Results

		Stude	Score	
		N	%	Test (%)
Condor	Boys	464	49.2%	34%
Gender	Girls	480	50.8%	35%
School Turno	Regular	589	62.4%	34%
School Type	Madrasah	355	37.6%	35%
Cabaal Chatura	Public	717	76.0%	35%
School Status	Private	227	24.0%	33%
Lesstien	Urban	617	65.4%	37%
Location	Rural	327	34.6%	30%
School Sampla	District School	750	79.4%	34%
School Sample	TTI School	194	20.6%	36%
Overall		944	100.0%	35%

Table 21: Participant Data and Average Scores in Grade 8 Science Test

Table 21 and Chart 17 show that the overall average score on the test was 35%. There was little difference in scores between boys and girls. Students attending SMP scored slightly lower than those attending MTs. Public schools scored higher than private schools.

⁷ The Trends in International Mathematics and Science Study which is implemented in many countries with 4th and 8th grade students every four years



Chart 17: Comparison between Different Groups

There were large differences between individual schools with the highest having an average student score of 54% and the lowest 17%. (Table 8). The highest student's score is 80% and the lowest is 0% (Table 9). A total of 39 students (1.9%) scored over 70%, while only two students failed to score (Table 9).

Chart 18 shows the average scores per quartile of students from the highest to the lowest 25%. The highest and second quartiles scored the same: 55% and 39% for district in TTI schools. In the third and fourth (lowest) quartiles, the differences are only one percent. The scores of each quartile will be reported in the next two years to assess whether progress is evenly spread across all learners, from the best to the poorest





As can be seen from Table 22 below students were able to answer questions the traditional format of questioning (with multiple choice answers) in Section A just as easily those in Section B, which required written answers.

Section	Score
Section A	38.7%
Section B	40.3%
Total	39.3%

Table 22: Average Scores by Section in the Grade 8 Science Test

Chart 19 shows the percentage of children scoring correct in each of the 16 questions in the test. The chart shows the questions which students had the most difficulty answering correctly (the scores are <30%). The students were relatively weak in all areas, but especially where they had to reason or make deductions from data. They also seem not to have acquired measuring skills through practical work. For example, they had difficulty in reading measurements of a ruler and reading weighing scales and measuring cylinders. They also had a weak knowledge of technical terms and difficulty in applying concepts to everyday situations.





3.3.3 Implications and Recommendations for PINTAR Program

- The results of the junior secondary school science test reinforce those of the primary schools test. Science teaching focuses too much on the memorisation of rules and concepts and too little on developing understanding of and applying concepts. Too little practical work takes place to support science teaching. Students spend much of their time memorising information from books rather than developing scientific skills such as observation of real phenomena, data analysis, making hypotheses and drawing conclusions.
- Teacher training should focus on developing students' scientific skills based on the
 observation of the real environment and doing experiments to investigate natural
 phenomena. Students need to be trained in measuring and other observational skills.
 Training should include helping students to make systematic reports and draw their own
 conclusions on the experimental and observational work they undertake. Simple
 technology activities should be promoted to encourage students to apply scientific
 concepts in real situations.

ANNEX 2: Comparison among previous projects on these tests

The table on the next page and the charts on the following pages summarize the results of the tests used by PINTAR, when they were used under other previous projects, compared with the results of the PINTAR tests. The results of three other tests are also included, which were not used by PINTAR. These are a reading word recognition test and a reading comprehension test for grade one students, which has been replaced by the EGRA assessment, and an English language test for grade eight.

The projects, which have used these tests and for which results are available include:

- Creating Learning Communities for Children (CLCC) managed by UNICEF and UNESCO and funded by NZAID and others from 1999-2010
- Managing Basic Education (MBE) managed by RTI International funded by USAID from 2003-2007
- Mainstreaming Good Practices in Basic Education (MGP-BE) managed by UNICEF and funded by the EU from 2007-2010
- Decentralized Basic Education 3 (DBE3) managed by Save the Children and funded by USAID from 2005-2011
- USAID PRIORITAS managed by RTI International and funded by USAID from 2012 to 2017.

Here are some general remarks about the results:

- The number of schools surveyed include only project partner schools, not comparison or control group schools
- Where projects worked mainly or wholly in provinces in Java (such as MBE), the results are considerably higher than projects that worked mainly outside Java (CLCC and MGMP-BE).
- Students' results in primary school across all subjects are considerably higher where large proportions of students attended pre-school (TK). It is also significant that pre-school participation is higher in Java than elsewhere, which may explain some or much of the better results from project working on Java. Students who have attended TK appear to have largely mastered word recognition by the time they enter grade 1.

There are a number of further cautionary factors in these comparisons.

- The primary school mathematics test was partially revised in 2004 after experience of using it on CLCC.
- The grade 8 Mathematics test was somewhat simplified for the USAID PRIORITAS and MGMP-BE districts, based on experience of its use in MBE.

COMPARATIVE TEST SCORES FROM VARIOUS PROJECTS

PROJECT NAME	CLO	CC			М	BE				MGP-B	E		DBE3	USAID PRIORITAS							PINTAR		
			1	Phase 1		Pha	se 2	Aceh								Cohort 2	1	1	Cohort 2		Coho	rt 3	Cohort 1
Round of Testing	1	2	1	2	3	1	2	1	1	2	. 3	1	2	3	1	2	3	1	2	3	1	2	1
Year of Testing	2003	2010	2004	2005	2006	2005	2006	2006	2008	2009	2010	2009	2010	2011	2012	2014	2016	2013	2015	2016	2014	2016	2018
#ofprovinces (of which on Java)	6 (2)	6 (2)	2 (2)	2 (2)	2 (2)	2 (2)	2 (2)	1 (0)	6 (1)	6 (1)	6 (1)	6 (4)	6 (4)	6 (4)	7 (4)	7 (4)	7 (4)	7 (4)	7 (4)	7 (4)	2 (1)	2 (1)	5 1)
PRIMARY SCHOOLS ASSESSMENTS																							
# of districts (of which on Java)	15 (5)	15 (5)	9 (9)	9 (9)	9 (9)	11 (11)	11 (11)	2 (0)	12 (2)	12 (2)	12 (2)	25 (15)	25 (15)	25 (15)	23 (15)	23 (15)	23 (15)	20 (10)	20 (10)	20 (10)	7 (4)	7 (4)	14 (2)
# of schools surveyed	45	45	54	54	54	66	66	20	72	72	72				92	92	92	80	80	80	28	28	72
% of Children with pre-school	42.4	66.4	90.7	92.7	92.5	91.3	95.7	81.7	55.2	57.9	71.0				78.1	84.0	87.5	80.0	85.8	85.6	86.1	92.0	81.7
Reading Word Recognition, Grade 1	47.1	71.3	87.3	91.4	94.6	87.9	91.9	50.4	56.4	61.9	70.6							ļ					
Reading Comprehension Grade 1	20.5	59.4	60.8	61.8	67.6	56.6	63.8	23.8	19.9	20.2	30.4							ļ					
Reading Comprehension Grade 4	40.1	46.9	53.0	62.8	64.8	59.9	61.4	38.8	35.7	35.9	39.6				43.2	47.3	56.1	37.1	53.4	55.2	42.1	55.8	38.8
Writing Grade 4	34.1	40.4	58.1	54.5	58.5	51.0	58.2	40.2	38.9	43.0	45.6				41.8	44.6	52.2	38.7	47.1	45.3	35.6	50.4	40.2
Mathematics Grade 4	47.0	47.0	61.1	65.5	65.0	64.7	65.0	41.3	39.4	38.1	43.7				40.9	44.5	52.4	39.2	47.0	53.1	47.8	56.1	41.3
Science Grade 5	28.8	39.8	44.3	50.4	53.4	48.8	54.5	29.0	28.1	28.9	31.9				35.8	43.0	49.5	33.8	42.0	48.3	38.2	49.1	29.0
JUNIOR SECONDARY SCHOOL ASESSEM	ENTS																						
#of districts (of which on Java)						20 (20)	20 (20)		12 (2)	12 (2)	12 (2)	25 (15)	25 (15)	25 (15)	23 (15)	23 (15)	23 (15)	20 (10)	20 (10)	20 (10)	7 (4)	7 (4)	14 (2)
#ofschools surveyed						60	60		36	36	36	54	54	54	69	69	69	56	56	56	21	21	56
Reading Comprehension Grade 8						78.3	78.5		58.7	64.9	66.2	66.6	73.0	75.1	63.9	70.0	72.7	66.0	69.6	74.1	70.6	72.1	72.1
Writing Grade 8						54.1	62.1		46.6	50.6	46.4	51.6	60.4	64.7	50.0	52.3	59.0	49.6	49.8	57.9	47.3	57.9	57.9
Mathematics Grade 8						36.7	35.2		23.3	26.7	27.4	32.0	41.7	47.4	28.9	36.8	54.4	33.9	38.1	43.9	35.8	43.4	43.4
English Grade 8						41.4	45.7		26.0	26.4	27.4	38.4	49.7	46.8				ļ					
Science Grade 8															41.1	43.9	46.0	39.6	42.1	42.7	46.0	51.3	51.3
PROVINCES	Central Java, S Sulawes NTB &	& East South Si, NTT, Papua		Centr	al & Eas	t Java		Aceh	Riau, La NTB, Go	ampung, prontalo,	Banten, , Maluku	N. Su West Java,	matra, Ba , Central a South Sul	anten, & East lawesi	Aceh,	N. Suma	tra, Banto	en, West,	Central 8	e East Java	a, South Sul	awesi	N. Sumatra, Riau, Jambi, C. Java, E. Kalimantan



PRIMARY SCHOOL ASSESSMENTS







JUNIOR SECONDARY SCHOOL ASSESSMENTS









ANNEX 3: Criteria for Marking the Grade 4 Writing Test

Tulisan (3)
3. Rapih, teratur, bersambung
2. Kerapihan kurang tetapi mudah dibaca
1. Kurang rapih dan sulit dibaca
0. Sangat kurang, kurang dapat dibaca
Ejaan (3)
3. Sempurna
2. Sedikit kesalahan
1. Banyak kesalahan tetapi masih dapat dimengerti
0. Hampir semua salah sehingga kurang dapat dimengerti
Tanda Baca (3)
3. Lengkap (titik, hurup besar dan tanda baca lain)
2. Titik dan hurup besar lengkap, lain-lain belum
1. Tanda baca kurang lengkap
0. Belum ada tanda baca
Panjang (4)
4. Lebih dari satu halaman
3. Lebih dari 1/2 halaman
2. Lebih dari dua kalimat
1. Satu atau dua kalimat
Mutu bahasa (7)
7-6. Gagasan menarik, kreatif dan diuraikan jelas dan berurut
5-4. Gagasan baik tetapi kurang original, penjelasan cukup baik.
3-2. Gagasan kurang menarik tetapi masih dapat dimengerti
1. Gagasan kurang menarik dan kurang dapat dimengerti
Jumlah (20)

Test	Development History	Broad Competencies assessed	Notes on the tests
Reading Grade 1 Test 1 Test 2	Developed by Muhlisoh (Puskur), Elizabeth Sweeting and Stuart Weston in 1996	Word recognition Simple comprehension	The tests are administered orally to 12 grade 1 children in each class, chosen at random Words in the word recognition test are taken from the grade 1 reading book. Only students able to complete test 1 are asked to do test 2
Bahasa Indonesia Grade 4 Reading Writing	Developed by Muhlisoh (Puskur) and Elizabeth Sweeting and Stuart Weston in 1996.	Finding information in a passage Inferring information Predicting future events Handwriting Spelling Punctuation Ability to express ideas logically Length of writing	The reading test is based around comprehension of a story. The writing test is based on an essay about a picture. The test is administered to half the class, while the other half takes part in the mathematics test (max. 20 per school)
<u>Mathematics Grade 4</u>	Revised substantially in 2004 by Ujang Sukandi (Puskur) and Ar. Asari (UM)	Various of operations of whole numbers and fractions Number series Shape Length Solving problems (money, shape, number series)	The questions have a mixture of multiple choice, closed ended calculation, problem solving and open-ended problems requiring creativity The test is administered to half the class, while the other half takes part in the B. Indonesia test (max. 20 per school)
<u>Science Grade 5</u>	Designed in 1996 by Gunadi (Puskur) Minor revisions in 2002 and 2004 by Masjudi (Puskur), Sup. Koes (UM) and Andreas Priyono (UNES)	Air Water Plants and animals Food chain Force and energy Resources etc Process skills including observing, interpreting data and hypothesizing	This test is divided into two sections. Section A used the format familiar to students of multiple choice questioning to assess children's understanding of concepts they have already learnt. Section B assesses children's active learning or process skills such as the ability to observe, interpret and hypothesize and requires the children to apply basic science concepts to everyday situations.
Bahasa Indonesia Grade 8 Reading Writing	Developed in 2004 by Wahyudi (ex- Puskur), Moh. Najid (UNESA) and Lynne Hill (MBE)	Finding information in a passage Inferring information Predicting future events Paragraphs Sentencing Quality of ideas	The reading test is based around comprehension of a story. It includes multiple choice, right and wrong and essay style answers. The writing test is based on an essay about a picture.

ANNEX 4: Summary of the Tests and their Development

Test	Development History	Broad Competencies assessed	Notes on the tests
		Spelling and punctuation Handwriting	The test is administered to half the class, while the other half takes part in the mathematics test (max. 20 per school)
<u>Mathematics Grade 8</u>	Developed in 2004 by Ujang Sukandi (Puskur) and Ar. Asari (UM). Revised 2008 by Ujang Sukandi and Eddy Budiono (UM)	Number operations Graphs and maps Geometry and angles Measurement Problems solving using a variety of concepts	The test is divided into a multiple choice answer section and an open ended answer section based around problem solving. The questions have a mixture of multiple choice, closed ended calculation, problem solving and open-ended problems requiring creativity The test is administered to half the class, while the other half takes part in the B. Indonesia test (max. 20 per school)
<u>Science Grade 8</u>	Developed in 2012 by Ferdy Rondonuwu (Universitas Satya Wacana, Salatiga) and Hadi Suwono (Universitas Negeri, Malang)	Classifying animals and plants Buoyancy Expansion and contraction Evaporation and condensation Process skills including measurement of length, weight and volume, observing, interpreting data and hypothesizing	This test is divided into two sections. Section A used the format familiar to students of multiple choice questioning to assess children's understanding of concepts they have already learnt. Section B assesses children's active learning or process skills such as the ability to observe, interpret and hypothesize and requires the children to apply basic science concepts to everyday