THEMATIC CASE STUDY

Literacy

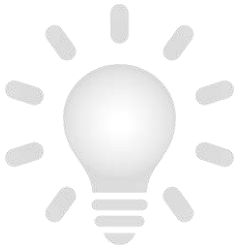
What works and why: Emerging evidence from INOVASI on effective practice in early grades literacy

June 2020

INOVASI
Innovation for Indonesia's School Children
Australia Indonesia Partnership

KEMENTERIAN PENDIDIKAN DAN KEBUDAYAAN

Australian Government



Thematic Case Study

Literacy

What works and why: Emerging evidence from INOVASI on effective practice in early grades literacy

June 2020

THEMATIC CASE STUDY

Literacy

What works and why: Emerging evidence from on effective practice in early grades literacy

The author of this study is Mary Fearnley-Sander, a member of the INOVASI team.

June 2020

Suggested citation: Fearnley-Sander, Mary. 2020. *Literacy. What works and why: Emerging evidence from on effective practice in early grades literacy*. Jakarta: INOVASI. https://www.inovasi.or.id/en/publication/literacy-what-works-and-why-emerging-evidence-from-inovasi-on-effective-practice-in-early-grades

Acknowledgements: The author would like to acknowledge and thank the INOVASI MERL team for the analysis of the program's survey data relating to literacy – particularly Rasita Purba, George Sukoco, Robertus Sodo and Senza Arsendy; the Education Program Development (EPD) team: Lynne Hill, as the lead education advisor in INOVASI for key advice and support throughout the study development; and Juprianto, Kania Dewi, Evi Rospita for their central role in implementing the video study of teaching and in the analysis of the teaching episodes; Stephanie Carter and Annisa Rachmawati for organising the videoing and development of the transcripts; INOVASI program managers for supporting the study in many different ways, and colleagues involved in the development of the other thematic studies for their contributions to the study methodologies. Mostly importantly the study acknowledges and thanks the three teachers who generously allowed the recording and analysis of the teaching and their reflection on it for the case studies section of the study, so that others could benefit from the exploration of their practice.

INOVASI - Innovation for Indonesia's School Children
Ratu Plaza Office Tower 19th Floor
Jl. Jend. Sudirman Kav 9, Jakarta Pusat, 10270
Indonesia
Phone : (+6221) 7206616
Fax : (+6221) 7206616
Email : info@inovasi.or.id

The governments of Australia and Indonesia are partnering through the Innovation for Indonesia's School Children (INOVASI) program. INOVASI seeks to understand how to improve student learning outcomes in literacy and numeracy in diverse schools and districts across Indonesia. The first phase of the program (AUD49 million) began in January 2016 and will continue until June 2020. Working with Indonesia's Ministry of Education and Culture, INOVASI has formed partnerships with 17 districts in four provinces namely West Nusa Tenggara, East Nusa Tenggara, North Kalimantan, and East Java.

INOVASI is an Australia-Indonesia Government partnership, managed by Palladium

Table of Contents

List of Figures

List of Tables

List of Acronyms, Abbreviations and Bahasa Indonesia Terms

ACER	Australian Council for Educational Research
ADD	*village funds (anggaran dasar desa*) ADD
AKSI	national survey of student competencies (*Asesmen Kompetensi Siswa Indonesia*)
ASEAN	Association of Southeast Asian Nations
ASER	Annual Status of Education Report
Bappenas	National Development Planning Agenc
BERSAMA	INOVASI pilot on community engagement pilot (Belajar di Sekolah dan Masyarakat
BOS	schools operational grants (*bantuan operasional sekolah)*
BOSDA	district supplementary operational funds for schools (*bantuan operasional sekolah daerah)*
EGRA	early grade reading assessment survey
EPD	education program development
GEMBIRA	INOVASI pilot on transition from mother-tongue to Bahasa Indonesia *(Gerakan Menggunakan Bahasa Indonesia yang Baik dan Benar)*
HOTS	higher-order thinking skills
IDR	Indonesian rupiah
INAP	Indonesian National Assessment Program
IRT	Item Response Theory
KEQ	key evaluation question
KKG	teachers' working group *(kelompok kerja guru)*
LPMP	Provincial quality assurance bodies (*lembaga penjaminan mutu pendidikan*)
madrasah	Islamic primary school
MERL	monitoring, evaluation, research and learning
MoEC	Ministry of Education and Culture
MoHA	Ministry for Home Affairs
MoRA	Ministry of Religious Affairs
NRP	United States National Reading Panel
OECD	Organisation for Economic Co-peration and Development
OPOB	One person, one book program
OPOB	One Person, One Book program
PAUD	early childhood centres (*Pendidikan Anak Usia Dini*)
PDIA	problem-driven iterative adaptation

PELITA	INOVASI pilot on specific issues in literacy
PIRLS	Progress in International Reading Literacy Study
PISA	Programme for International Student Assessment
PRIORITAS	Prioritizing Reform, Innovation and Opportunities for Reaching Indonesia's Teachers, Administrators and Students
Puspendik	National assessment centre, Ministry of Education and Culture (Pusat Asesmen *dan Pembelajaran)*
RPJMN	National mid-term development plan (*rencana pembangunan jangka menengah nasional*)
SDG	Sustainable Development Goal
SES	socioeconomic status
SETARA	INOVASI pilot on inclusive education
SIL	*Suluh Insan Lestari* foundation
SIPPI	INOVASI's baseline survey *(Survei Inovasi Pendidikan dan Pembelajaran Indonesia*)
TASS	Technical Assistance for Education System Strengthening program
TIMSS	Trends in International Mathematics and Science Study
UBT	University of Borneo in Tarakan
UINSA	Sunan Ampel Islamic University
UIS	UNESCO Institute of Statistics
UK	United Kingdom
UNESCO	United Nations Eductional Scientific and Cultural Organisation
UNM	State University of Makassar
US	United States
YLAI	Indonesian Children's Literature Foundation (*Yayasan Literasi Anak Indonesia*)

About the study

This study is a compilation of what we have learned about improving literacy outcomes in the course of INOVASI Phase 1. It provides emerging evidence of what can work to bring about improvement in the program's regional contexts. INOVASI's development experience of seeking local ownership of problems and solutions is a key component of the evidence the program has produced.

The orientation to literacy in INOVASI derives from its critical importance as the foundation of learning, and Indonesia's own ambitions for the literacy capabilities of its youth. These ambitions have two sources. One is the country's own research establishing the existing distance between Indonesian students' performance and proficiency in higher order comprehension as measured globally by international literacy assessments. The other is the current nation-building vision of the Nawa Cita, to which literacy is intended to contribute by widening horizons and capacity for self-development.[1]

INOVASI's support has been designed to meet these capability objectives; and what the program has worked for in teacher, school and district support for literacy is best encompassed by the definition of the Programme for International Student Assessment (PISA): *Reading literacy is understanding, using, reflecting on and engaging with written texts, in order to achieve one's goals, to develop one's knowledge and potential, and to participate in society.*[2]

INOVASI's main activity in literacy has been through piloting approaches to strengthening the teaching of early grades literacy, specifically Grades 1-3. Supporting access to written texts, as well as systemic, partnership and policy developments for effective literacy teaching and learning, have been integral to the piloting process. INOVASI has supported 38 literacy pilots, in partnership with four Indonesian Universities, and international and non-government organisation literacy foundations in Indonesia; and has provided technical support for the implementation of its literacy pilot model in East and Central Java by Muhammadiyah and Ma'arif Nahdlatul Ulama.

In keeping with the program's theory of development, the literacy piloting was an iterative process, seeking local ownership of problems and solutions that work in the context. A central dynamic in this approach is the relationship between evidently effective practice in global literacy research and the differences, choices, mindsets and capacities found in local contexts. Out of the continuous negotiation between the particularities and universals of early literacy learning, we developed a model for teacher development in literacy teaching and several significant variations on this main theme.

The model prioritises teachers' pedagogical knowledge of literacy: how to help children decipher the codes of written language; and how to help them access the literal and implied meaning in texts. In a context where *know-how* for the teaching of reading is often absent, it emerged as the priority. The program's main literacy pilots — Literacy 1 and 2 — are professional development pilots. An integral objective in these pilots is to strengthen Indonesia's established professional development system and develop the personnel to ensure it works.

In this study to explore what worked to improve literacy outcomes in the first phase, we focus on the Literacy 1 and Literacy 2 pilots. Our first inquiry is whether students whose teachers participated in these pilots had better literacy scores in the endline test than in the baseline test.

[1] Permen 23/2015 Tentang Penumbuhan Budi Pekerti Lampiran Peraturan Menteri. A.p.4. This Ministerial regulation derives from the Nawa Cita as does the parallel development of Indonesia's national and school literacy movements. (Gerakan Literasi Nasional; Gerakan Literasi Sekolah)
[2] OECD, 2015. PISA Assessment and Analytical Framework.

The second inquiry is whether teaching practice improved through the pilots. The third is what evidence do we have that INOVASI's interventions on teaching practices and children's access to books are associated with students' increased scores; and what aspects among the variant pilots on teacher development had the most effect?

Program wide, student results on a beginning reading skills test (component skills: letters, word construction and word recognition) show modest gains over the baseline in the Literacy 1 pilot at Grade 1 and Grade 2 level: seven and three percentage points respectively, after allowing for natural growth.[3] For reading comprehension and on higher order thinking skills (HOTS) at Grade 2 level, the gains were appreciably higher: 17 percentage point increases in both. (At Grade 1 the results were much more modest-unsurprisingly with many in Grade 1 still struggling with letter knowledge.) The students whose schools continued into Literacy 2 had 14 point gains in word recognition at the Literacy 2 endline over the Literacy 1 endline.

The student results established the pattern which runs through all the results in the study: of large provincial differences in outcomes. The lowest baseline provinces had the highest gains. At Grade 1, from baselines of 19% and 3% per cent of students passing the component skills test respectively, North Kalimantan and Sumba gained over 10 percentage points, on the component skills test, with the same level of gain at Grade 2. By contrast Java Timur with a high Grade 1 baseline of 58% of students passing, had no endline gain; and on a baseline of 85% passing at Grade 2 showed a slight loss in endline performance. The pattern is even more marked for reading comprehension. At Grade 2, North Kalimantan gained 37 percentage points, and Sumba 23, to Jawa Timur's 11 percentage point gain.

Teacher practice was the main construct on which teachers' development through the Literacy 1 and 2 pilots was measured. To develop the construct the study undertook a literacy review of effective literacy pedagogy to identify practices most associated with such pedagogy. For Literacy 1, findings from classroom observations of 100% of the pilot schools were that nearly 60% of teachers observed included in literacy lessons shared reading and questioning to build comprehension skills. However, only around 30% were implementing beginning reading skills. For Literacy 2, findings from classroom observation and teacher interview were that 92% of schools were implementing formative assessment and nearly 60% were able to identify different levels of reading proficiency among the students in their class; and organise students into group to teach to their level. General classroom practice skills highly relevant to literacy teaching were also measured. Large gains over baseline practice were found in attention paid to all students in the class (31% gain over baseline); and in teachers' use of appropriate media to teach a concept (24% gain).

Statistically, what worked in the pilots — that is, what statistically significant associations were established between student scores and INOVASI's interventions on beginning skills, comprehension and HOTS—was established through regression analyses. Support of literacy through reading corners with books that engaged students' interest, was the variable with the strongest correlation for all of the student outcomes. The strength of the regression coefficient was small to medium, but higher than any other variable measured, including student background measures, except for SES on some subskills.

Pilot teachers' own reading literacy proficiency had been tested in INOVASI in a baseline test of comprehension and HOTS constructs. Teachers' scores on this test turned out to be the only teacher variable associated with student scores in reading comprehension and HOTS. Across

[3] The methodology for estimating whether improvement may be due to the pilot did not allow for the retrieval of gains at Grade 3 level. The methodology was to use the baseline of the succeeding grade as a "control" and to identify whether gains at endline of the preceding grade exceeded that base-line of the succeeding grade indicating more at work than natural growth. INVASI did not take a Grade 4 baseline.

these skills it was consistently the third highest performer of all variables (after reading corners and teacher certification). While INOVASI had not explicitly targeted teachers' own reading literacy proficiency, growth in this capacity may have been an outcome of teachers learning how to support text comprehension in children. The variable Classroom Practice showed as being only very weakly associated with beginning reading skills; and not at all with reading comprehension or HOTS. Because constructs relating to literacy subject pedagogy were not included in the program baseline for teachers, teachers' performance in this area could not be included in the regression analysis.

Comparative analysis of the endline gains of students in variant pilot models of teacher development is another way of assessing what worked. An important finding from the overall experience is that a pilot focusing teachers on student problems—namely the Guru BAIK pilot—in combination with teachers' participation in Literacy 1, produced the highest gains. This leading performance was closely followed by pilots using a language transition approach to literacy teaching where children's home language was different from the language of instruction; and also by pilots where the Literacy 1 model had been supplemented by partnerships supplying early grades readers and storybooks; and additional teacher training in their use for balanced literacy instruction.

The study also sought to understand what the INOVASI model looked like in practice in three case studies of teachers delivering a literacy lesson. Video-recordings of the lesson were analysed by the teachers themselves and by the INOVASI education team to interpret how the teachers construed the strategies they had learnt in Literacy 1 and 2. Much was learnt about "what worked" through these thick descriptions, including ways of interpreting some of the quantitative results. A nutshell encapsulation of what the case studies showed is that strong progress has been made at the technical level in comprehension teaching—highly integrated skills in this displayed in one instance. Nevertheless, these literacy classrooms are still sites of teacher dominance, which limit opportunities for children to form language, infer for themselves and initiate responses to text, prarctices that are essential for the development of independent expressiveness and reasoning.

The case studies also throw light on many of the patterns in both the student and the teacher data.

They fit with the trend of findings in the student and teacher outcomes—that teachers' have taken up some of the key comprehension strategies in effective teaching of literacy and these are working. This success also implicates the success of the focus on access to books and levelled readers in the program.

The picture on teaching beginning skills reading is less clear. Great gains have been made in places where grade level reading lags greatly behind other provinces—in Sumba and North Kalimantan. The gains over the grades in these regions show how this success in beginning reading can reduce the learning gap with other provinces by the end of early grades; and therefore the importance of skills of beginning teaching in particular, remote locations. Elsewhere the lower effect of pilots on teaching beginning reading — and the struggles the case study teachers had with decoding—may indicate teachers' existing phonological strategies work better with Bahasa Indonesia than imported phonemic ones.

"Problem-based" has been a great teacher in INOVASI addressing literacy problems. The pilots that have been most successful are ones that emphasise the importance of students' problems as a point of departure for teaching. That includes problems created by distinguishing features of context such as mother tongue or remoteness from books.

Students' problems as a point of departure in teaching is another way of saying that student-centred teaching is effective practice. In its diagnostic approach to the teaching reading, INOVASI has introduced pedagogies which are logically student-centred. The insight of the original pilot Guru BAIK on *identifying* student problems aome to fruition in later pilots that built up expertise for *solving* problems in literacy. Fuller understanding of how to develop a student-centred mindset on the part of teachers is a remaining challenge: how to open up spaces in teaching and learning interactions for literacy to fulfil the promise that Indonesia's literacy policy holds out for literacy: of widening students' horizons and developing their full potential.

1 Introduction

Purpose of the study

This study is a compilation of what we learned about improving literacy outcomes during the first phase of INOVASI. It provides emerging evidence of what works to improve learning outcomes in the contexts where we worked.

Emerging evidence means the evidence-base of promising local solutions in classrooms, schools and clusters and in supporting policies and programs at district and national levels.

At this stage of the INOVASI program, emerging *evidence* means credible evidence that is convincing to policy makers, plausible and persuasive. The evidence derives from the following sources: baseline–endline comparisons of quantitative data on student achievement levels, teachers' knowledge and teachers' beliefs (mindset data); classroom observation data; qualitative case study data from classrooms and schools; and district management data on teaching and learning in literacy. This evidence is yet to be tested with more robust methodologies in the next phase to reach the standards of certainty of random-controlled trials or experimental studies.

The outcomes of the study provides evidence on the effectiveness of different contextualised strategies and broader policy-related considerations for national and sub-national governments in Indonesia and for the Australian government. This meets INOVASI's third program outcome: *national and sub-national stakeholders have access to emerging evidence of what does and does not work to improve student learning outcomes.*

A secondary audience is national and regional educational institutions, think tanks, supporting development partners and non-governmental organisations. The study is composed for this readership, with an emphasis on succinctness and the usability of the findings.

A. Focus of the study

The literacy study focuses on the overarching evaluation question:

What works to improve literacy outcomes in INOVASI's partner districts?

This inquiry has four sub-inquiries, each with their own key evaluation question (KEQ), as shown in box 1.

Box 1: Key evaluation questions for the literacy study

KEQ 1: To what extent does training teachers to teach reading result in children's improved reading outcomes?

KEQ 2: To what extent does providing appropriate books improve children's reading outcomes?

KEQ 3: To what extent does training teachers in mother tongue transition improve children's reading outcomes?

KEQ 4: Is there any evidence that improved literacy outcomes resulting from the pilots will lead to better learning outcomes at higher levels or across the curriculum? Or better higher-order thinking skills (HOTS)?

B. Outline of the study

After this introduction, the study commences in Chapter 2 by analysing the policy and situational context of literacy teaching and learning at the national level and in INOVASI's targeted provinces. Chapter 3 then outlines the scope of INOVASI's intervention in literacy – in pilots, policies and partnerships. In preparation for the analysis of INOVASI's effectiveness in literacy improvement, Chapter 4 follows with a contextualised review of the literature on effective approaches to early grades literacy, presenting global evidence on the "science of reading" and its relevance to contexts such as those that INOVASI works in. Chapter 5 describes the analytical methodology of the study: the development of the analytical framework, drawing on the literature review, for addressing the evaluation questions; the data sources used by the study; and the analytical pathways through the different pilot types. The findings from INOVASI's pilots then follow in Chapters 6-9 on whether student learning outcomes improved; whether teaching practices improved; and on what worked: in terms of statistical correlation with student outcomes; and comparatively, in relation to the different levels of success of different pilot approaches. This leads to the culminating section of the study, which reviews the implications of the findings for understanding the interactions between the 'science of literacy', contextual issues and what actually works in context. Successfully adjusting evidenced effectiveness to context is the contribution INOVASI's literacy pilots set out to make.

2 The context

This chapter presents two aspects of the context framing INOVASI's literacy intervention. Part one covers policy developments and issues relevant to literacy at the national and district level. This includes the outcomes of INOVASI's strategies to influence policy and regulation relevant to literacy at both levels.

Part two analyses the literacy attainment of Indonesian students to understand how the policy relates to the realities and to provide a point of reference in understanding the choices and outcomes of INOVASI's interventions in the chapters that follow.

Part one: National policy relevant to literacy 2015–2020

President Joko Widodo's vision of how Indonesian society should develop has a close affinity with definitions of literacy as enabling individual and community capabilities to develop to their full potential.[4] (UNESCO, 2006:137). The Programme for International Student Assessment (PISA) also uses an empowering definition of literacy as allowing one to 'achieve one's goals, to develop one's knowledge and potential and to participate in society' (OECD, 2017). Acquiring that kind of national significance for Indonesia is a turning point in the role literacy has so far played in schooling.

INOVASI's approach to literacy also aims to develop individual and social capabilities and, in keeping with the government's own strategy, we focus on proficiency in reading literacy as a means to these ends.

Jokowi's imperative for Indonesia at the beginning of his first term of office (2015) was 'revolutionalising the character of the nation', the fourth of the nine principles in the *Nawa Cita* underpinning development during his administration. The vision was driven by two imperatives. The first was enabling Indonesian youth to compete in the context of economic globalisation, knowledge based futures, and rapid change — all sharpened by the advent of the 2015 ASEAN Economic Community. The second was to strengthen the cohesiveness of Indonesian society by building up national identity and local Indonesian cultures.

Both these imperatives resulted in literacy becoming a priority in Indonesia's mid-term development plan (*rencana pembangunan jangka menengah nasional* – RPJMN). This mid-term plan and the sector plans that derive from it integrated literacy into their strategies to operationalise the national revolution.

The first Jokowi administration: the discourse and objectives of the mid-term development plan 2015–2019

The mid-term development plan's analysis of performance in the education sector is shaped by Indonesia's results in the 2012 PISA and particularly in comparison to its Southeast Asian neighbours of similar low-middle income status (Bappenas, 2015; MoEC, 2015).[5] Government aspirations for improvement are not directed towards the input-driven 'quality improvement' typical of past plans but towards specific learning outcomes. These are the 21st century skills and 'literacies' of the new vision for education of the World Economic Forum that widely influenced the education discourse in the mid-term development plan – the ability to apply

[4] UNESCO Institute of Statistics, 2006. Education For All *Global Monitoring Report on Literacy,* p.137.
5 RPJMN, II pages 2–34; Strategic plan for education (Renstra), pages 18-19, 41. In this assessment it was 100 points behind the OECD average.

knowledge and understanding to the different contexts and problems in life and work (ACER, 2017).[6]

The mid-term development plan 2015–19 targeted improvement at the next PISA through policy priorities for both curriculum and learning assessment. Its policy directions are listed in Box 2.

Box 2: Priorities for curriculum and assessment in the national mid-term development plan 2015–2019

Priorities for curriculum and assessment in the national mid-term development plan (RPJMN) 2015–2019

1. *Strengthening the curriculum* to deliver 21st century skills (policy 2.3.3/4a) – specifically**,** increasing the quality of literacy, mathematics and science learning as the foundational competencies that are needed in everyday life and in the community (2.3.3/4i) (RPJMN:105-06);
2. *Diversifying the curriculum so that students can develop their individual potential to the maximum* (RPJMN 2.3.4b; Target 2.3.3/3m). This policy recommendation is extended in the strategic plan for education (Renstra) to support learning up to grade three by using local languages for instruction in remote areas (Renstra, 2015: 24);
3. *Evaluating the implementation of the curriculum* closely, comprehensively and continuously (longitudinally) (2.3.3/4d);
4. *Increasing the culture of reading* in the community by providing library services and socialising the reading culture **(**2.3.4**) (**RPJMN: 116**);**
5. Increasing the quality of character education to foster and build character, and develop the selfhood of students (2.3.3 3m).

Revolutionising the character of the nation and promoting literacy

The policy objective for character education in the mid-term development plan has an unexpected significance for literacy development. Linking the plan with the aim to revolutionise the character of the nation, the ministerial regulation No 23 of 2015 on developing character includes schools' obligation to develop the full potential of each student. To achieve this, the regulation elevates *reading* to a critical role and gives teachers and schools the responsibility to:

'…develop the unique potential of every student through encouraging in learners ***a love of reading,*** *developing their interests and talents and extending their horizons and capacity for self-development'* (Ministerial regulation 23/2015, Appendix:4)*.* [7]

This link between reading and broadening students' potential for self-development encapsulates the human capabilities definition of literacy in PISA and similar assessments, referred to earlier. This is a pivotal moment in Indonesia because it elevates literacy from its transactional reading and writing ('*baca-tulis*') function in primary education to being a means of personal empowerment.

[6] Reading and writing literacy was one among six nominated in the new education agenda. The others are: numerical literacy, science literacy, digital literacy, financial literacy, and cultural and civic literacy.
[7] The original is in Bahasa Indonesian: '*penghargaan terhadap keunikan potensi peserta didik untuk dikembangkan, yaitu mendorong peserta didik gemar membaca dan mengembangkan minat yang sesuai dengan potensibakatnya untuk memperluas cakrawala kehidupan di dalam mengembangkan dirinya sendiri*.'

There are implementation guidelines for reading in the ministerial instruction *On Developing Character*—15 minutes mandated reading a day before the start of lessons. Reinforcing the idea of instilling love of reading, the material for reading is specifically *not* the school text book.

The Ministry of Education and Culture (MoEC) responded to this new national vision for the role of literacy by launching the national literacy movement (*Gerakan Literasi Nasional* – GLN) in 2016. Under the agency for the development of the Indonesian language, the movement promoted both national culture and a culture of reading by preserving and publishing quality local stories and developing public facilities for literacy.[8]

Communities had the opportunity to enculturate reading through the village budget (*anggaran dasar desa* — ADD) of IDR200 million for community empowerment, another strategy in the administration's nation-building mission. In ministerial regulation No 11 of 2019 on priorities for the use of village funds in 2020, building and resourcing community libraries, study centres and community reading facilities are all considered eligible. While books for schools are excluded, the regulation specifies supplying early childhood centres (PAUD) with books and creating a story resource that children in early grades can access with their parents (article E2b, page 25). INOVASI used this opportunity to support MoEC in also including these early childhood books in the eligible book list for schools so they can be used in the early grades.

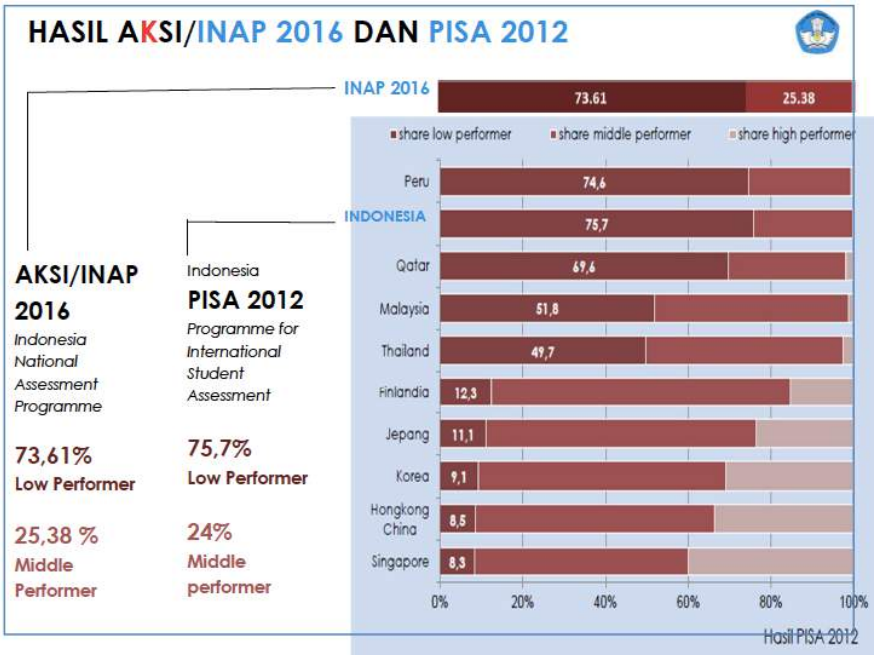
Implementing the education policy priorities of the mid-term development plan

In implementing the key policy directions of the mid-term development plan targeting students' performance on PISA, reforming student assessment has taken the lead. From 2016 the assessment unit (Puspendik) in MoEC with assistance from INOVASI developed the *Asesmen Kompetensi Siswa Indonesia* (AKSI) a recurring national survey of student competencies in reading literacy, mathematics and science.

The test items for reading literacy are based on literacy competencies drawn from PISA and the Progress in International Reading Literacy Study (PIRLS) assessment, and reflect their hierarchical order: retrieving direct information, interpreting or understanding texts, and evaluating and reflecting. The test was initially implemented at grade four, which was useful for showing the outcome of early grades education. In presenting its analysis of the results the assessment centre emphasised the predictive power of students' performance on these constructs in the early grades, for their performance as 15-year olds on PISA. This comparison is shown in figure 1.

8 Interview with the head of the agency for the development of the Indonesian language (*Badan Pengembangan dan Pembinaan Bahasa* – BPPB) and the team, 26 April 2018

Figure 1: Sampled results on Indonesia's national survey of student competencies at grade four and results on the Programme for International Student Assessment for Indonesia and nine other countries



Source: Assessment centre, MoEC (2017)[9]

Since its launch at the grade four level, AKSI has been implemented at grade eight level and plans are underway to use it as a benchmark of achievements at different assessment points throughout schooling.

Assessment is driving curriculum change through this comparative assessment on performance benchmarks. With benchmarks articulating what students should be able to do by the end of a given interval of schooling they will make the disconnection conspicuous between the national criteria for learning performance and the different objectives of the current Curriculum 13.

At the root of this disconnection are competing interpretations of how to produce students capable of higher-order thinking skills. The developers of Curriculum 2013 used the model of scientific inquiry. In early grades this is the integrative principle of a thematic curriculum. The outcome competencies and the grade level competencies require cognitive strategies, topics of study, priority vocabulary and text types that mainly derive from the scientific paradigm. By contrast, in most early grades curricula learning is dominated by students' need to acquire the foundational skills of early literacy and numeracy. Particularly for literacy, the thematic curriculum leaves little scope for sequencing and consolidating the early skills for reading acquisition. The units in the teachers' guides seem to assume that students know sounds and letters by the time they start school and so expect them to be able to read text early in the first semester. This situation is exacerbated by the extensive learning areas included in the unit themes, further reducing time for reading acquisition.

One important reform that Curriculum 2013 and its teachers' guides reflect is the use of local languages to deliver the curriculum where Bahasa Indonesia is not the students' mother tongue.

[9] Presented by the head of the assessment centre at a North Kalimantan provincial education meeting convened by INOVASI in March 2017

The guides suggest teachers accept students using Bahasa Indonesia or their local language to present their ideas.

This may derive from the more definite stance than previously in the strategic plan for education (Renstra) on using mother tongue in early grades: *Teachers are expected to use mother tongue as the language of instruction for primary students until grade three so that it is easier for them to understand the subject matter*. Also for the first time the plan acknowledges that developing second language speakers' competence in Bahasa Indonesia needs to be a graduated, systematic and ongoing process *(*Renstra: page 24).

Nevertheless, there is no methodological outline in the curriculum framework or in the teachers' guides to show teachers how to transition learning in the first language to learning in Bahasa Indonesia. Teachers depend on strategies that may hinder children's progress in a new language – for example, opportunistic code switching or teaching exclusively in the local language, leaving children unable to make the transition by upper primary.

Over the President's first term, the logic of including the PISA goals in the mid-term development plan has emerged. Developing learning progressions and using sub-sectoral benchmarks in key learning areas are likely to frame the process of redeveloping the curriculum.

In fact a start was made in 2019 to revise the 2003 Education law itself to "reconstruct" the education system. The objective is equip—inclusively —"Generation 45", one hundred years on from Indonesian Independence—with the skills needed to support Indonesia's entry into the ranks of higher income countries (MoEC, 2019).

Once Indonesia resolves the curriculum issue it can make progressive reforms in how it is delivered, particularly by bringing together the momentum for literacy in districts, village communities and schools, and the line ministries that support them at the national and sub-national levels.

Futhermore, this curriculum reform is critical for sustaining the teaching gains from INOVASI. Without it , teachers do not have the mandate to change their practices. They cannot systematically apply what they learned in their monthly planning and daily teaching while they are obliged to teach to the existing curriculum units and report monthly on students' performance against them.

Aware of the importance of teacher development and curriculum proceeding hand in hand, INOVASI along with its sister program *Technical Assistance for Education System Strengthening* (TASS) has been supporting MoEC's curriculum reform in early grades literacy and numeracy on the basis of evidenced progressions of learning in these domains, and helping to keep a focus on the diversity of contexts and learners in Indonesia that national curricula frameworks need to accommodate.

The second term: Jokowi's administration, 2020–2024

The link between changing the character of the community and promoting literacy continues in repeated statements in the technical plan developed for the mid-term development plan during the second Jokowi administration, for example:

'…the mental revolution is strengthened through efforts to conserve and promote local culture, religious moderation … and a culture of literacy, innovation and creativity to create a community that is knowledgable, innovative, creative and of high character' (Technical plan, RPJMN 2020–2024:120; Bappenas, 2019).

There is increased precision about what literacy can contribute in terms of higher order skills. These are the skills of the PISA constructs: identifying understanding and interpreting information to transform it into productive activities that bring social and economic benefits and wellbeing. The plan is clearer on how to accomplish this revolution too. In the policy objectives for teaching and learning, the plan singles out the skills of literacy, numeracy and science at every level of school education – specifically naming early grades. For the community it is about developing a culture of reading and this is recognised as a priority if Indonesia is to meet current challenges *(*Technical plan*,* RPJMN 2020–2024:127). Policies for these ambitious plans include providing more libraries, promoting book production and supporting civil society organisations that promote books and reading.

Progress in these social goals is quantified through performance on the PISA indicators. The 2024 targets for improving service delivery that increases productivity and competitiveness are: *to increase reading literacy scores from 397 to 412* and *to increase the proportion of Indonesian students above the minimum competency level from 44 per cent to 49 per cent* (Technical plan*,* RPJMN 2020–2024:104).

The Ministry of National Planning and Development's (Bappenas) mid-term development plan for the next administration was informed by a policy paper that draws together INOVASI's data on conditions affecting learning in Indonesia (INOVASI, 2018). These data nuance the guidance in the plan for continuing to prioritise diversifying the curriculum so that students can develop their individual potential to the maximum. Via the plan, they also influence the new strategic plans for MoEC and the Ministry of Religious Affairs (MoRA).

The most transformative action from the new administration to date has been to appoint a minister whose entrepreneurial skills and experience exemplify the kind of creative prowess that Indonesia seeks from its mental revolution. The emblematic command of the new education minister, Nadiem Makarim has been 'to free' (*merdekakan)*. By this he means freeing the talent and creativity of teachers and students from the stifling 'bureacratisation' of learning. He has so far announced three iconic policy directions to spur this on:

1. Replacing the national primary school examination with a school assessment, trusting teachers and principals to make judgments on how to assess their students in accordance with their contexts.

2. Abolishing the national examinations and replacing them with an assessment of minimum competency, based on the PISA constructs and character. This is to be recurrent and take place at mid-levels of schooling (grades four, eight and eleven) to avoid the exam being used as a basis of selection for the next level of schooling. The assessment will be used diagnostically by the ministry to improve the quality of education provided.

3. Simplifying the teaching plans required of teachers – replacing the 20-page (unimplementable) plan with one-page coherent statements of lesson objectives, activities and assessment.

These promising reforms for recovery focus on what matters in teaching and learning in schools, and are highly conducive for the change that INOVASI has been working on in teaching and assessment.

The most radical reform from government has been to restructure the national Ministry of Education. The concept is to professionalise the education service, advising a thinner band of decision makers at the top levels of the structure on MoEC policy and management. This

process is ongoing and consequently activity on the specific policies that INOVASI supported based on its pilot results, is currently awaiting the final restructuring.

District reform initiatives

The two policy and system developments at district level that are most relevant to INOVASI's pilots relate to book availability and systems for teacher development.

Reading support

To support the reading imperative in ministerial regulation No 23 of 2015 *On Developing Character*, MoEC launched the national literacy movement. Its national task force produced guidelines on developing reading in schools (MoEC, 2016). Its suggested 'balanced literacy' model for early grades literacy integrates the component skills of reading with a focus on comprehension. The model is a potent resource for instructional reform and meets the national agenda for higher-order thinking skills. The guidelines focus on narrative text in the early grades because of the well-established value of stories and literature for 'extending horizons and capacity for self-development' from the start of schooling.

The national literacy movement also proposed that schools organise a literacy task force to raise the profile of reading in the school and the community, and meet the challenge of supplying engaging, grade-appropriate books for children.

It was the GLS program that the Ministry for Home Affairs (MoHA) advocated in its circular (No 420/9240 of 2018) to all provincial governors and district regents in Indonesia, ordering them to implement early grades literacy education and include provision for it in the regional development plans and budgets of which that Ministry has oversight. This circular was issued as part of MoHA's responsibility for Indonesia's progress on Sustainable Development Goal 4 on inclusive, quality and lifelong learning. MoHA's intervention is potential support for early grades literacy reform in the districts. Nevertheless it is perplexing for schools that also have to implement Curriculum 2013 that has different competency requirements for teachers and schools to report on.

Many of INOVASI's partner districts responded positively to the literacy movement. Some declared themselves as literacy districts or cities and developed literacy roadmaps and local regulations to encourage reading.[10] INOVASI helped Batu city, for example, in developing a mayor's regulation (No 93 of 2018) on implementing literacy activities to support families, community library resources and the schools' focus on literacy, including in extracurricular activities (INOVASI, 2019).[11] Other districts issued regulations on allocating village funds to support community libraries (West Sumba, Bulungan and Malinau). In Bima, the regional development planning agency is coordinating relevant local authorities, including the district education office, to guide village administrations in effectively using the village funds to promote literacy.

As part of the districts' literacy movement, Batu city, Bulungan, West Sumba, East Sumba Probolinggo, Sidoarjo, Central Lombok and Bima all specified support for book purchase for schools in their supplementary operational funds for schools (BOSDA). Bulungan and East Sumba also issued regulations requiring schools to purchase books with their national schools operational grants (BOS).

[10] For example, among INOVASI's public districts, Bima, Batu city, Malinau, Bulungan.
[11] Mayoral regulation No 93 of 2018 on Batu as a literacy city, page 14

Bulungan closely followed the policy guidelines for the literacy movement agenda. By regulating the BOSDA support for schools almost 10,000 non-textbooks, made up of more than 1,500 titles, were purchased for the district's primary schools. The district also issued regulations on classroom reading corners and enlisted school supervisors to act as the local literacy movement task force. Their job descriptions were adjusted and they were trained to support schools in buying appropriate books. Bulungan credits its outstanding increase in the percentage of grade one children passing the basic literacy test (from a baseline of 17 per cent to 98 per cent in one year) to its integrated literacy movement campaign that engaged government and non-government sectors as well as communities. The non-government sector was represented by Rainbow Reading Gardens (*Taman Baca Pelangi*) and a distinguishing feature of its support to literacy was help for struggling readers, delivered in partnership with INOVASI.

Support for teachers' professional development

The main institution that INOVASI used in piloting teachers' professional development was the primary teachers' working group (*kelompok kerja guru* – KKG). This longstanding institution is the only professional development mechanism in Indonesia with sufficient reach to service the whole workforce affordably. For untrained or undertrained teachers these groups are their only means of acquiring some know-how. Consequently these working groups are essential in taking the teacher development pilots to scale. However the institution has been beset with apparently intractable issues of quality: lack of skilled facilitators to resource learning; limited support for teachers' participation from principals; lack of a monitoring system; and most determining of all, the lack of accountability to any authority for its operation.

INOVASI targeted the development of the teachers' working groups on various fronts – in its political work and policy dialogue with districts – considering their strategic significance to the piloting project. This led to three changes to the KKG in a number of districts. The most widespread change is that local authorities now seek official recognition from the provincial quality assurance bodies (*Lembaga Penjaminan Mutu Pendidikan* – LPMP) for professional development delivered through the KKG. Another change is that Sumba, Bulungan and Sidoarjo districts acknowledge that schools' funds should also cover teachers' participation in the working groups; and supplement these funds from the district budget. Nine of INOVASI's partner districts are preparing for increases in budget allocations to teacher quality as a result of the collaborative financial analyses INOVASI conducted with local governments (INOVASI, December 2019). The other relatively easy change to accomplish was greater flexibility in how the working groups are organised. This sometimes involved re-zoning the school catchments for the KKG to give better access to remote schools. For very remote schoois in Malinau it included using digital solutions (supervisory and technical support through mobile connectivity).

A significant way of addressing the issue of quality would be districts' institutionalisation in some form of KKG facilitators, starting with the group that INOVASI has already trained. While several districts have extended their facilitator numbers and funded and authorised their ongoing availability, they are not yet part of the formal operation of the local education systems.

Institutionalised reform in general has eluded the KKG so far. However the problem-based approach to improving these groups that INOVASI developed with local authorities is gradually resulting in more accountability for the quality of the professional development they offer. The most promising indications of this are institutionalised monitoring of learning outcomes in two districts (Bulungan and East Sumba). The main lesson learned from watching stakeholders apply problem-based approaches to dysfunctionalities in local education systems, is that solutions are likely to be different and not comprehensive at the outset. What districts working

successfully in KKG improvement have done is pick out a problem that they saw possible to resolve, and which, like a pulled thread, might lead to the unravelling of others.

Part two: Literacy attainment in Indonesia 2015–2018

This brief overview of Indonesia's performance in literacy provides the context that INOVASI was working in during the first phase and explains the program's motivation in developing and adapting the pilots; as well as the background to results achieved.

Student performance

A performance indicator adopted in the Strategic Plan for education 2015–2019 was to raise the literacy rates for Indonesian students taking the PISA test from 396 (2012) to 414 in the 2018 test. However, Indonesia's reading literacy performance in 2018 fell back to its 2001 level after a peak in 2009.[12] The mean score was among the lowest in the PISA test with Indonesia ranked 71st out of 76 participating countries.

With 70 per cent of Indonesian students performing below the minimum level of proficiency (level 2) in reading in 2018 the country was lagging way behind the OECD average of 23 per cent of low performers (OECD, 2019). At level 2, students can identify the main idea in a text of moderate length, find information based on explicit though sometimes complex criteria, and can reflect on the purpose and form of texts when explicitly directed to do so. As its descriptor suggests these are minimum performance levels on the three constructs of information retrieval, interpretation and evaluation or reflection on what is read.

In 2016 Indonesia conducted its first national sampling — the Assessment of Indonesian Student Competence (AKSI) at grade four level — a level that reflects on their performance in the early grades. The reading literacy assessment was based on the Grade 4 PIRLS assessment, with similar to PISA: accessing and retrieving information; interpreting and integrating information or ideas across texts; and evaluation (PIRLS, 2011: 13; OECD, 2017; PISA 2015: 56).[13]

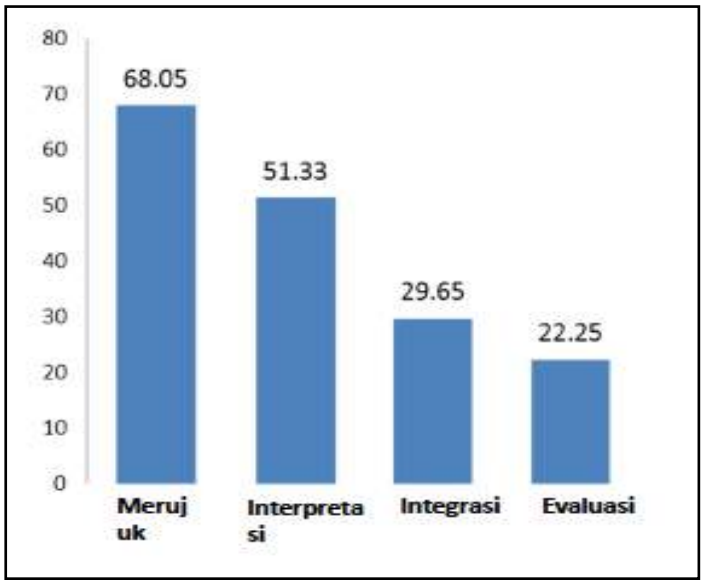
Table 1: Indonesia's student learning assessment 2016 national results for literacy: three performance bands

Low performance	47% (= < 41 points):
Satisfactory performance	47% (=< 59 points)
Good	6%

Figure 2 shows the performance of students on the three constructs.

12 The OECD report on Indonesia's 2018 PISA performance notes these results must be seen in the context of the vast strides that Indonesia has made in increasing enrolment. In 2001, the PISA sample covered only 46 per cent of 15-year-olds in Indonesia while in 2018, 85 per cent of 15-year-olds were covered (OECD, 2019).
13 The PIRLS 2011 framework developed for that year of the survey is used here as it is the most recent PIRLS survey that Indonesia participated in.

Figure 2: Proportion of correct answers on the three skills constructs in the grade four Indonesian students' performance assessment survey (AKSI), 2016 (percentages)



The proportion of correct answers on questions relating to retrieving information (*merujuk*) was relatively high at 68.05 per cent. By contrast, the proportions of correct answers on the higher-order items of integrating ideas and evaluating text were 29.65 per cent and 22.25 per cent respectively. These results show that students struggle with the higher skills of seeing relationships between ideas or information and integrating them across extended texts.

INOVASI supported AKSI sampling all ten districts in West Nusa Tenggara, with a sufficiently large sample to compare between districts. This finer-grained study found large disparities across districts. Average scores in reading and science in West Nusa Tenggara were 70 to 80 points (0.7 to 0.8 standard deviations) lower than the national average (with the exception of Mataram city) (INOVASI, 2017:6).

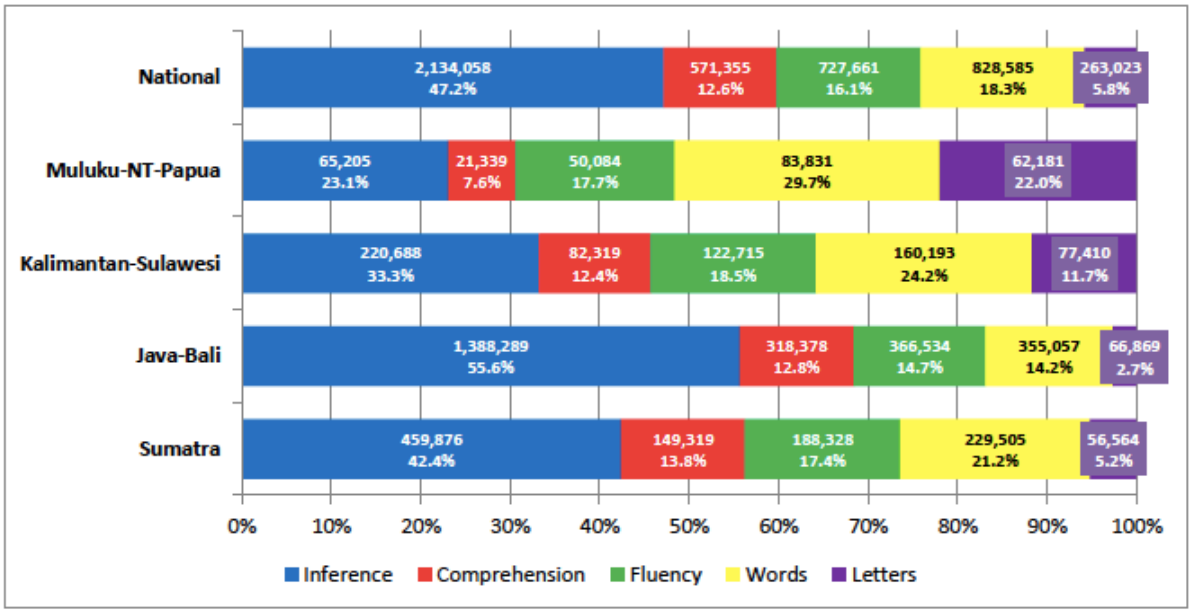
The only national early grades literacy results available are from the grade two Early Grade Reading Assessment (EGRA) survey conducted in 2014 (USAID, 2014).[14] The Indonesian government incorporated two of its key findings in the strategic plan for education 2015–2019.

Taking the average national performance on the main outcome measure of oral reading fluency (ORF) the survey found that 47 per cent of grade two students read fluently with comprehension (80 per cent correct answers on reading comprehension) (USAID, 2014:17). A second key finding was the disparity between the national average and results in the eastern regions. The performance of the eastern region in the survey (Maluku, East Nusa Tenggara, West Nusa Tenggara and Papua islands – shortened to MNP) showed 23 per cent of students in the top performing group – half the national percentage and nearly as many (22 per cent) in the non-readers group (USAID, 2014:31).

[14] The study sampled 4,812 grade two students across four 'regions' in Indonesia towards the end of the school year. The regions were (1) Sumatra and its adjacent islands; (2) Java and Bali; (3) Kalimantan, Sulawesi and its adjacent islands; and (4) the MNP region, consisting of Maluku, East Nusa Tenggara, West Nusa Tenggara and Papua islands.

The USAID study is useful for its diagnosis of the kinds of problems students have in literacy and their relative proportions in different locations.

Figure 3: Distribution of children by instructional need and region, national early grade reading assessment, 2014



Source: USAID (2014).[15]

Nationally, the most competent students — the 47 per cent reading 80 words per minute and understanding 80 per cent of the text – still have problems in inferencing (see figure 3). Aggregating the percentages of all those not reading with 80 per cent fluency in the USAID study, 53 per cent of students nationally have problems in comprehension. These are caused by inadequate skills ranging from mastering the written code and decoding meaning quickly, right down to the level of recognising letters. These basic word skill problems are more pronounced for students in the disadvantaged regions of Maluku, East Nusa Tenggara, West Nusa Tenggara and Papua islands (MPN) in the USAID study — 50 per cent of students are in this category.

These early grade findings on students' low performance in decoding skills and comprehension are repeated in other project results that fall within the mid-term development plan period. Over the period 2012–2017 period, the USAID project, PRIORITAS implemented the EGRA in 50

[15] The USAID EGRA report explicates the levels of capability in this graph: "The blue category represents children who are fluent and can read grade 2 level text with understanding (scoring at least 80% on the reading comprehension sub-task); and who would benefit from instruction that builds their inferential skills. The red category represents children who are nearly fluent (reading more than 50wpm) but have lower understanding of the text (scoring less than or equal to 60% on the comprehension sub-task). These children are nearly Grade 3 ready but need support to improve their understanding of the text, mainly through improved vocabulary to support comprehension. The green category represents children who are reading more slowly on average (between 26 and 50 words per minute). They understand some of what they read, but are reading too slowly to be functional in Grade 3. These children would benefit from practice to decode words as well as to improve their comprehension. The yellow category represents children who are beginner readers. They read between 1 and 25 words per minute, have higher levels of inaccuracy in word identification and limited comprehension. These children would benefit from word identification instruction. The purple category represents non-readers. These are children who were unable to correctly read a single word from the passage." P. 71.

districts. The study found that students beginning grade three on average understood about half of what they read (USAID PRIORITAS, 2017:46).

With much smaller samples, Save the Children implemented the EGRA tests in the Belu district of East Nusa Tenggara province in 2015 and 2016. The 2015 baseline findings from 20 schools in Belu showed that at the start of grade two, 19 per cent of students were 'readers' – they could read at least five words of the reading passage correctly in 30 seconds (Save the Children, 2015).

Teacher performance in literacy

This section is based on results from the reading literacy test that MoEC's assessment centre (Puspendik) set for teachers. The test is based on a grade four PIRLS reading literacy test. We included this element so we can assess the role and impact of teachers' own level of subject understanding on their learners' achievements. While teachers' proficiency in reading comprehension is different from their knowledge of the subject, it nevertheless reveals their understanding of how texts work and whether they can communicate this to their students. Teachers who lack skills in inferring and pursuing ideas across extended texts are unlikely to be able to teach these skills to others. Chapter 8 looks at the effect of this variable on students' scores.

Table 2: Program and province baseline profiles on teachers' literacy proficiency test

Locations (number of teachers in brackets)	Baseline mean scores
Program level	**53.79**
East Java (155): test mean	64.90
Higher-order thinking skills mean	54.60
North Kalimantan (54): test mean	40.37
Higher-order thinking skills mean	24.77
West Nusa Tenggara (139) test mean	50.83
Higher-order thinking skills mean	33.72
East Nusa Tenggara (Sumba) (132): test mean	49.36
Higher-order thinking skills mean	36.65

Across the provinces, there is less range in teachers' literacy proficiency compared to their students. East Java is the exception. Although scores are still not high on this grade 4 test, they are around 15 points higher than the nearest province. North Kalimantan has the lowest scores. Sumba has the lowest mean for student comprehension but not for teacher comprehension; and its higher-order thinking skills mean is the second highest after East Java. Overall, teachers have limited ability in higher-order thinking skills and, given this is the aggregated mean of three hierarchised skills, it indicates how low it is even on the lowest level of directly retrieving information, lower still on the higher skills.

Conclusion

Indonesia's strategy for improving the literacy capabilities of its youth by galvanising communities and schools around the love of reading provides the most conducive environment for developing the broader skills that literacy can bring. The direction of the reforms in the curriculum and approach to assessment, the power of the literacy movement and the enthusiasm in many districts for implementing it, will radically change the conditions of learning. The government's emphasis on literacy as development aligns with INOVASI's own approach to literacy that extends beyond functional literacy.

Students and teachers' current capacity needs to grow to meet these ambitions. The brief overview in this chapter sets out some of the challenges Indonesia faces in achieving its desired revolution. However the idea of a mental revolution creates exactly the kind of radical situation needed to legitimise and encourage other important departures from the prevailing low levels of performance in literacy.

3 INOVASI's approach to improving literacy

Piloting in INOVASI has been extensive and necessarily complicated. In the context of INOVASI's theory of development there could not be a planned roll out. INOVASI developed its models and partnerships with local stakeholders to meet problems identified by the districts and the districts ultimately chose their own pilots. This led to a tapestry of models and adaptations, continuities and discontinuities that means most districts have had distinct literacy experiences under INOVASI. (See annex 1 for profiles of INOVASI's literacy pilots.)

This chapter provides an overview of the type and scope of the literacy pilots that INOVASI supported. The two main types of literacy pilot were teachers' continuing professional development and books supply for engaging students in reading. These are presented in sequence and by provider: INOVASI-trialled pilots first, and then pilots developed or run by non-governmental organisations and university partnerships.

The coverage does not include scale-out adaptations of pilots by local government or by the large civil society providers of schooling, *Muhammadiyah* and *Ma'arif Nahdlatul Ulama*.

Teacher pilots directly supported by INOVASI

This category includes pilots that INOVASI technically developed, directly funded, managed and monitored. In this study they are referred to as the INOVASI pilots. They are the experimental pilots in the sense of seeking to prove concepts and incubate ideas.

The INOVASI pilots trialled two distinct ideas. The first idea was that an effective entry point to improving teaching was developing teachers' capacity to identify students' problems. The second idea was that proven literacy teaching methods still need to be trialled to assess their effectiveness in the particular cultural, work and capability environment of teachers and schools. If the evidence shows that these models improve the literacy outcomes of both students and teachers, then governments and Indonesian civil society can invest in approaches that work in analogous teacher and learner contexts.

These two ideas were trialled and developed in classroom action research in two different series.

The *Guru BAIK* pilots, January–June 2017: INOVASI's first pilots[16] were carried out in 25 schools each in the North Lombok and Sumbawa districts of West Nusa Tenggara, INOVASI's first partner province. The pilots applied problem-driven iterative adaptation (PDIA) in the classroom context to improve students' learning outcomes. Teachers did classroom action research using the established sequence of steps to identify the problem, develop and implement action plans, analyse the results; and reflect on what they learned from the process. The problems teachers identified had to be resolvable within a month, the time frame for trying out solutions. INOVASI trained local facilitators in the PDIA approach and its use in classroom action research. The program also held workshops and mentored teachers in probing problems and framing and reflecting on their solutions (INOVASI, 2017:6).

The *Guru BAIK* initiative in Sumbawa was independently evaluated a year after it was implemented, using a first iteration of INOVASI's education and learning survey, the *Survei INOVASI Pendidikan dan Pembelajaran Indonesia* (SIPPI).[17] The results show an average score increase of two percentage points in students' literacy scores over the baseline and an 11 per cent increase in teachers using active instruction, including learning media, over a low baseline of 15 per cent. Most interesting was the 10 per cent increase in teachers' scores on the literacy

[16] Guru BAIK= Aspirational, innovative and contextual learners.
[17] Evaluation in North Lombok was disrupted because of the Lombok earthquake.

proficiency test, considering that the pilot did not include teachers' own development in literacy. The evaluators attributed this result to teachers' deepened understanding of literacy teaching produced by the collaborative reflection process (REDI and INOVASI, 2019).

INOVASI learned some lessons from the *Guru BAIK* initiative. The first was that many teachers were held back from improving children's literacy and numeracy by not having the relevant subject knowledge – they could not correctly identify the nature of the problem students were facing or what the solution would be. The approach also had limited capacity to improve literacy and numeracy teaching. The kinds of problems teachers identified related to discrete items in the curriculum. Teachers were not gaining a sequenced understanding of foundation skills in literacy or numeracy and how to teach them.

Nevertheless there were lasting gains from the experiment. The main breakthrough was through teachers realising that students' difficulties could arise from their own teaching. The second was the pragmatic focus on two foundational teaching skills: the use of teaching media to help students understand; and lesson planning that addresses the learning problem. These gains lasted through all subsequent trials and, according to stakeholders, transformed classrooms.

Sumba district requested the *Guru BAIK* model in a later phase of INOVASI, alongside successor pilots, affording the opportunity to compare two different approaches to teacher development (see chapter 7).

The Literacy 1 pilot, January–May 2019: INOVASI reached 23,733 students through this pilot that it implemented in 15 of its 17 districts. This model succeeded *Guru BAIK* and responded to the lesson learned about teachers' need for a base of knowledge and skills in the domain before they could help students or recognise the nature of the problems they were facing.

This led to the Literacy 1 pilot to develop teachers' capability for teaching the elements of beginning reading in sequence. In seven modules the pilot introduces teachers to the component skills recognised in the 'science of reading': phonological awareness; construction and decoding of words from sounds; fluency and reading comprehension.[18] Vocabulary development was not targeted directly but the pilot promoted 'literate classrooms' — word walls and labelled pictures. The reading comprehension element in the course stressed the use of 'big books' for modelling and practising reading for meaning and for developing higher-order thinking skills, such as predicting and linking ideas across connected text.

INOVASI approached this technically-demanding agenda through modelling and having teachers practise practical strategies that aligned with the skills they were targeting. The emphasis was on a capabilities view of literacy that recognises the significance of being able to read well for children's overall development.[19]

The Literacy 1 course was delivered through the local institution of teachers' working groups. Local facilitators, drawn from the ranks of active supervisors, school heads and skilled teachers, delivered monthly sessions. In between the monthly teachers' working group sessions these facilitators mentored participants individually on delivering lessons based on what they had learned and facilitated reflection on the lessons.

Between *Guru BAIK* and Literacy 1 were a group of INOVASI pilots that started in the second half of 2017. Initially they were designed on the *Guru BAIK* model to meet specific problems: PELITA

[18] This term recognises the last 20 years consensus of reading researchers and practitioners as to what early skills need to be mastered for sufficient reading proficiency to succeed in school. See chapter 4 on the contextual literature review.
[19] INOVASI Literacy pilot modules

for specific issues in literacy; GEMBIRA on language transition; BERSAMA on community involvement and school attendance; and SETARA on inclusive education. After running for several months, they were transitioned into the Literacy 1 pilot. Literacy 1 pilots that began in this way lasted longer than the regular Literacy 1 pilots that went on for about a year. (These hybrid pilots were all in West Nusa Tenggara where INOVASI started eighteen months earlier than in North Kalimantan and Sumba.)

The Literacy 2 pilot, July–December 2019: This pilot reached 3,450 students. The aim of this pilot was to transform the introductory level of Literacy 1 into consolidated capacity for teaching reading. The course developed two pivotal skills across five modules: basing teaching on the diagnostics of student performance and developing strategies to support reading comprehension. Teachers learned to use a reading assessment process based on the Annual Status of Education Report (ASER) assessments to identify levels of proficiency in their class. They would then be able to group students according to the nature of the learning problem, plan lessons to target the problem area, identify readers to suit the children's level and give regular ongoing support through guided reading sessions and 'running records'.[20] Teaching reading comprehension focused on strategies to help students connect information and ideas across continuous text, and teachers asking the kinds of questions that help students retrieve information and develop their higher-order thinking skills.

Guru BAIK's legacy influenced this pilot development as it includes an additional focus on teacher reflection. Every group session required teachers to report back on their experience in implementing what they had learned in the previous session – about their own teaching and their students' learning. This systematic approach to teacher reflection met a criticism that the strong technical focus of Literacy 1 had come at the cost of the gains in teacher reflectiveness achieved in the Guru BAIK pilots.

Although Literacy 2 was designed to follow the same monthly sequencing in teachers' working groups as Literacy 1, this schedule had to be revised following the revival of MoEC support for the teachers' working groups. This was in the form of a module and grant-based continuing professional development program offered to districts from mid-2019**.**

The Literacy 1 and 2 pilots are known as 'short courses' since they offer teachers a sequenced series of sessions on early literacy teaching through the regular teachers' working group meetings. This name also conveys, particularly to district authorities, the importance of developing skills progressively and most effectively through the regular working group meetings. As an added advantage, teachers participating are eligible for career upgrades since the short courses are accredited by the education quality assurance agencies (LPMP).

While Literacy 2 intended to be a progression from Literacy 1, participation was limited since only 10 of the 15 Literacy 1 districts chose it as their priority for the second round of pilots.[21] This means that INOVASI has two levels of 'graduates' from its literacy pilots: those who completed the introductory Literacy 1 course and those who completed both Literacy 1 and 2 and therefore the full course to develop literacy teaching skills. Any assessment of the efficacy of the literacy pilots needs to analyse whether outcomes for teachers and students were affected by the continuum of teacher development over the Literacy 1 and 2 pilots.

[20] The Annual Status of Reading Assessment (ASER), developed by the Indian education non-governmental organisation, Pratham, is a model of reading assessment that enables non-specialists (community principals and teachers) to assess children's reading levels. The tool consists of four elements: letters, words, a short paragraph text and a longer 'story'. Children are marked at the highest level that they can do comfortably (https://www.asercentre.org/p/141.html).

[21] Batu city, Probolinggo, Sumenep, Pasuruan, Bulungan, Malinau, Bima, Dompu, East Sumba, West Sumba.

The multi-grade pilot

In addition to the Literacy 1 and Literacy 2 pilots, in 2019 INOVASI developed a multi-grade pilot in a sub-district of Probolinggo in East Java, at the request of the head of the education office. Since this covered the early grades, literacy and numeracy skills were the main targets of the teaching approaches in the multi-grade pilot and therefore it is one of the approaches to literacy trialled by INOVASI. It is a potentially important model because Probolinggo is the first district to see multi-grade teaching as a solution to the increasing financial pressure on districts to share the burden of funding teachers with the national government. The multi-grade approach is based on competency mapping in a subject area, as a basis for differentiating teaching and learning across the ability range of several grades. However until the core competencies for early literacy are clearly identified in the curriculum, it will be difficult for multi-grade approaches to follow the skills progression across the grades required for beginning reading.

Teacher pilots supported through grantee partnerships

Two main types of organisations are involved in these pilots: non-governmental organisations and university partners.

INOVASI had two key rationales for including the grantee pilots in its teacher development program. The umbrella rationale was to model to district authorities how to build partnerships to supplement and diversify sources of local technical support. A second reason was to introduce districts to the guidance available in specialist partnerships to address specific contextual difficulties in improving educational outcomes. In literacy these difficulties have mainly concerned language transition methodologies for teachers of local language speakers. A third rationale was to 'find out what works' – based on the strategy of diversifying the ideas base. Some non-governmental organisations might have good solutions that have yet to be 'proven'.

The grantee pilots supporting teaching have focused on different issues. University partners have implemented Literacy 1, in the dual interest of becoming service providers within provinces and acquiring valuable professional experience for their own pre-service teacher training. For example, the Sunan Ampel Islamic University (UINSA) in Surabaya is implementing the program in ten Islamic primary schools (*madrasah*s) in Pasuruan and looking at its potential for *madrasahs* in general. In North Kalimantan, the University of Borneo in Tarakan (UBT) and the State University of Makassar (UNM) are collaborating to train teachers in teaching literacy in both Bulungan and Malinau, based on an adaptation of Literacy 1. The two non-governmental organisations supporting language transition are the *Suluh Insan Lestari* (SIL) foundation in Southwest Sumba and the Sulinama Foundation in East Sumba. They both broadly derive their methodology from the Summer Institute of Linguistics. Within that methodology, SIL focuses on developing the orthography of the local language it is supporting, while Sulinama focuses on mother tongue comprehension and second language approaches to language transition. (Sulinama also developed the capacity of local facilitators to run the INOVASI-managed language transition pilot, GEMBIRA, in Bima, West Nusa Tenggara). Other teacher development pilots specialise in phonics for schools in poor communities (*Tunas Aksara Foundation*), literacy for children at risk of dropping out (*Dompet Duafa*) or whole-school approaches to literacy (Edukasi 101).

The book pilots

An important function for literacy was filled by non-governmental organisations and private partnerships that support book supply in schools as a critical element in improving reading. As with the teacher pilots, these initiatives demonstrate to districts a sustainable means of addressing the book deficits in classrooms and communities, particularly in remote districts. Several of these partnerships ran in the same schools as the Literacy 1 pilots, facilitating book-centred

approaches. This was particularly the case in North Kalimantan, where Litara, the One Person, One Book program (OPOB) and the Asia Foundation (digital books) also helped INOVASI support the district policy on implementing the literacy movement. In the four Sumba districts this function was filled by Rainbow Reading Gardens (*Taman Baca Pelangi*) that developed demonstration school libraries in each location.

Two of the book pilots combined book supply with teacher development. In Central Lombok, the Pen Circle Forum (*Forum Lingkar Pena*) developed books to support inclusive teaching, designing books with sign language and modelling inclusion though the diversity of characters included in the stories. In two districts of Sumba, the Indonesian Children's Literature Foundation (Yayasan Literasi Anak Indonesia – YLAI) provided a sample of balanced literacy teaching through the number of graded readers it has developed. It also modelled the shared and guided reading methodologies that enable teachers to put books at the centre of basic skills development and comprehension in reading.

4 Contextualised literature review

The affinity between Indonesia's vision of developing individual potential through reading and the broad definition of literacy as the means of realising human capabilities was discussed in chapter 2. We also profiled the level of achievement and areas of challenge in literacy acquisition for Indonesian students. The central issue is that large proportions of students in Indonesia do not adequately comprehend what they are reading. On either side of this finding are the other salient factors. One is that around fifty per cent of students in poor and remote areas struggling with the fundamental skill of relating sounds to letters so as to recognise words. On the other side is the limitedness of the higher level inferential skills. In other words, every step on the comprehension ladder to improved PISA results is implicated, beginning with the emergent skills of reading.

The following literature review examines the evidence for what works to tackle these different problems and the kind of curriculum and teaching practices for literacy the evidence implies.

The review also considers the evidence in relation to contexts of disadvantage. Research in literacy in the United States (US), United Kingdom (UK) and Australia has focused on disadvantage over the last three decades because of the apparently intractable association between high performance in literacy and high socioeconomic status (Freebody, 2007). These research findings on the key disadvantages in literacy learning and the most effective ways of addressing them can be applied universally.

Sources for this evidence include large-scale studies analysing literacy scores and teachers' effectiveness from international learning assessments, such as PISA; other OECD studies, meta-analyses and research syntheses of what supports the development of literacy; studies on specific relevant factors; and studies that have had a significant impact on countries' policies and practice.

These sources of evidence are largely from advanced economies and do not encompass the spectrum of disadvantage in lower-income economies. Much additional valuable information comes from the 'grey literature' of project-based learning assessments and analyses, especially in relation to the Indonesian context.

To target this analysis we need a profile of the disadvantages that may apply to students in Indonesia. This is strikingly provided in the USAID EGRA study summary of the children who bear the cumulative effects of disadvantage:

> '…poor, male, over-age/under-age, in a remote, public MoRA school, in MNP [Maluku, East Nusa Tenggara, West Nusa Tenggara and Papua islands], with no preschool and a home language that differs from the one used in school. The .. probabilities tell us that such students had about a 1 per cent chance of being in the top group and about a 91 per cent chance of being in the bottom group' (RTI and USAID Indonesia, 2014: 30).

These disadvantages point to the need for effective practices that mitigate the effect of poor homes, poor schools and poor regions, and particularly issues relating to remoteness – access to books, no pre-school exposure to language development and an unfamiliar language of instruction at school.

The evidence on effective programs for early grades literacy

INOVASI's intervention regions are not all like the profile cited above. As we saw at the outset of this chapter, there are two types of problem in literacy acquisition across INOVASI's regions. There are two types of problem in literacy acquisition across INOVASI's regions: decoding and

comprehension. The research evidence has important messages about the relationship of these issues in the early grades curriculum.

The curriculum dominates literacy research because of the high stakes for policy making and therefore for learners in terms of what choices are made. Over the last thirty years reading researchers have battled over what makes children successful readers and particularly on how disadvantaged children are affected by policy decisions.

One line of research suggests that early grades classes should emulate the advantages that middle class homes give their young children in literacy performance. This recommendation draws on research showing how parents' interactive story reading familiarises children with literate language, with the processes of pursuing meaning across a whole text and critically reading text in the light of their own knowledge of the world (Heath, 1982). These are the higher-order comprehension skills required for high performance on the PISA test. PISA studies also show that 'by far the strongest relationship is between reading to a child during his/her early years and better reading performance when the child is fifteen (OECD, 2012). This led many reading protagonists to the view that schools that disadvantaged children attend need to level this playing field by emphasising the story reading role in their early grade literacy curriculum.

Proponents of an alternative priority challenge this approach as underestimating the iceberg of cultural capital under the surface of these kind of interactions that impedes the success of children who can't access that capital. They cite the famous Hart and Risley study of the vocabulary gap between children from high income and welfare-supported families in accounting for the socioeconomic gradient in literacy outcomes (Hart and Risley, 2003).

The alternative approach to early reading emphasises phonics – mastering the relationships between sounds and letters. The value of phonics is that children can systematically work out unknown words and therefore read independently. Anyone who has watched children recite what is written on the board, without even looking at it, will recognise that teaching children how to work out what the words "say", will replace the practice of memorising them (Cincotta-Segi, 2011).

Speed of decoding leads to automaticity that releases the short-term memory to focus on the unit of meaning in what has been read and thus facilitates comprehension (Abadzi, 2006, 2008). The applicability of this approach to phoneme-grapheme relationships in most languages has led to interest in developing structured, replicable approaches to instruction in these foundational skills as the most effective way of getting all children reading. This clearly has value for early mastery of a language with a high sound and letter correspondence, like Bahasa Indonesia.

Since the 2000 conclusions of the United States National Reading Panel (NRP) on the evidence for effective early reading practice, more emphasis has been placed on phonics-related skills (phonemic awareness, phonics and fluency) than on vocabulary and comprehension, although the latter were also identified as crucial in the report (NRC, 1998; NRP, 2000). This dominance is evident in both the US (in the Reading First program mandated by the No Child Left behind Act, 2001) and in the UK (particularly in the 2013 national curriculum).

However, impact and curriculum evaluations from 2006 onwards in both the UK and US have yielded modest evidence for phonics-dominated curricula. The US Reading First impact study found comprehension scores had not been improved by dominantly phonics models of teaching reading (IES, 2008; Torgerson *et al.*, 2019:209). A systematic review of randomised control trial studies of phonics teaching since 2010 indicates statistically significantly overall positive gains. But in researchers opinion there is 'insufficient evidence to justify a "phonics only" teaching policy; indeed, since many studies have added phonics to whole language approaches, balanced instruction is indicated' (Torgerson *et al*., 2019). Researchers in the phonics tradition recognise

themselves that 'because early reading instruction emphasises word recognition rather than comprehension, the less skilled comprehenders' difficulties generally go unnoticed by their classroom teachers' (Snow and Juel, 2005:77). This may well apply in the Indonesian context where the very transparency of the orthography can mean children read at pace without understanding what they have read.

Moreover, as discussed in chapter 6, in the natural growth in beginning reading each year, phonological skills reach their ceiling within the first few years of schooling whereas comprehension and vocabulary are unlimited in growth and cause a steadily widening gap between advantaged and disadvantaged students in relation to school literacy (Paris, 2005).

Nevertheless, phonics approaches clearly assist beginning readers by enabling independent word recognition; and up to 50 per cent of children in Indonesia's eastern provinces need these skills.

'Balanced literacy' uses both explicit phonics instruction and reading for meaning approaches. Effective curricula for developing comprehension skills from grade one give the practice of reading books interactively a central place in reading instruction. Indonesia's policy around students attainment of the higher skills of comprehension skills requires this kind of approach: the same explicit and systematic programming as phonics in the classroom.

The kind of texts that balanced literacy requires includes quality graded readers, that authentically relate to children's experience of the world. Systematic instruction and assessment in all the beginning reading skills can be built around them: teachers can use interactions to diagnosis problems of decoding, word recognition, fluency, and word, sentence and text level comprehension through them. Rigorous evaluations report the effectiveness of initiatives trialling graded reader strategies for improving children's cognitive performance at the primary level (Banerjee *et al*., 2016; Abeberese *et al.*, 2011).

The balanced approach also requires texts that nourish cognitive growth. Narrative thinking is pervasive in children's early cognitive development and narrative texts are fundamental to developing comprehension at this stage. They develop the concepts of temporal and causal sequencing of events. Story schema familiarise students with narrative structures and facilitate comprehension across a whole text (Paris and Paris, 2003:40). Narratives expand ideas, vocabulary and world knowledge that are all critical for comprehension. As Snow summarises, '…*comprehension research has demonstrated clearly the importance of the reader's background knowledge for understanding* the main ideas in texts' (Snow, Burns and Griffin, 1998:62).[22]

Narrative texts – stories – encourage reading for pleasure. The findings in successive PISA reports on reading literacy are that 'On average in every country students at the high level of the index of positive attitudes towards reading had substantially higher reading achievement than those at medium or lower levels' (Mullis *et al.*, 2007:140).[23]

Appropriate curricular emphases for contexts of disadvantage

The language of instruction problem: The profile of disadvantage presented earlier in this review includes *a home language that differs from the one used in school* and the effects of this in Indonesia are seen in the low scores of students in the eastern provinces reported in the USAID's national EGRA study (2014). Any reform to reading literacy which does not address this and related problems in the contexts in which they occur, will be limited in its effect. Realising the

22 See also: Snow (2010); Siraj-Blatchford, (2010); Pressley and Fingeret (2007) Purcell-Gates, Jacobson and Degener (2004)
23 See also: OECD (2009).

inequality created by the choice of language of instruction – more precisely by neglecting mother-tongue based multilingual education – the UNESCO Institute of Statistics (UIS) has included this issue as one of its inequality measures in assessing performance on the Sustainable Development Goal for education (SDG 4) (UNESCO, 2016).

The research on language of instruction inequality emphasises both the dire effect of language incomprehension for learning and the near intractability of the problem (Cummins, 2001; Brock-Utne, 2010). This is due to the complex political, perceptual and implementation problems frustrating solutions. Politically, mother tongue instruction runs up against the hard issue of national identity to which a national language is integral — explicitly so in the history of Indonesia. Even where there is generous policy on mother tongue and multilingual learning, as in the Philippines, political opposition can obstruct its implementation. So also can public perception. This includes, for example, the erroneous belief that the earlier a child is introduced to the official language of instruction, the better they will master it. Community conflict can also erupt over which local language should be recognised in the curriculum provision. The scale of the implementation issues for poor countries is daunting: it can involve curricula and materials in multiple languages, matching teacher distribution to language demographics and multilingual teacher training. Nevertheless the UNESCO Institute of Statistics observes that governments are increasingly accepting that the imperative of improving education outcomes requires them to recognise the issue in their policy (Kosonen, 2017).

A pragmatic approach to the problem is to frame early grades mother tongue instruction as language transition support, as advocated in the 2017/18 Global Education Monitoring Report background paper on *Language of Instruction in Southeast Asia*. Influenced by the Summer Institute of Linguistics (SIL) expertise in the region, the model proposes transition to the official language of instruction in grade three. Children learn to read in their mother tongue first (grade one). The orthographies of these local languages are usually transparent and the decoding skills transfer readily to the second language. This transfer process begins in grade two and includes students developing oral language vocabulary in the target language of instruction. SIL has developed an expressive pedagogy (Total Physical Response) for teachers to communicate meaning to children. This depends on energised teacher talk in the classroom, including body language (Trudell and Young, 2016).

While such an early exit is not ideal (the consensus is around six years of mother tongue), the model is politically feasible. The main policy objective is to recognise that reading literacy in the official language of instruction should be deferred to grade two. Both curriculum progressions and national assessments need to be adjusted for contexts where language issues significantly hold back students' progress.[24]

Listening comprehension: Children in these contexts should not delay in developing their comprehension skills until grade two just because reading literacy is delayed. Listening develops comprehension before children can read. Reading and listening share many cognitive processes: syntactic and inferential processes, word knowledge and conceptual knowledge (Snow, 2010:64). Through listening children can learn how to interpret texts well before they learn to decode for themselves.

A balanced literacy curriculum accommodates these needs through 'read-alouds'. Listening to stories interactively – with questioning and conversation about them – is established in the research as 'especially important for children who would have had little storybook experience outside school' (Snow, 2010). Evidence shows that the quality of the conversations between

[24] Literacy assessors engaged in supporting the development of global benchmarks for Sustainable Development Goal 4.1.1 argue for rethinking a uniform grade benchmark for reading acquisition.

adults and young children around stories is the link with story reading and later reading success (Marulis and Neuman, 2013).

Vocabulary development: Comprehension depends on knowing the meaning of words; and research suggests that vocabulary is more important than grammar or short-term memory in helping five-year-olds to make inferences (Silva and Cain, 2015).

Knowing meaning includes knowing the *layers* of meaning that words acquire when they are encountered in a range of contexts (Perfetti, Landi and Oakhill, 2005). The report of National Reading Panel (2000), affirmed that teaching words in context is more effective than teaching isolated words and that reading storybooks is more effective in developing students' vocabulary than testing isolated vocabulary with feedback (Damhuis, Segers and Verhoeven, 2015).

The evidence for improving teacher practice

An effective curriculum for literacy implies the kind of teaching strategies for students' progress in literacy that teachers need to use, and cannot be implemented without teachers having the requisite knowledge and competence for using those strategies. Global research and research from developing countries, and Indonesia in particular, prioritise teachers' pedagogical knowledge of their subject in improving student outcomes in literacy and mathematics. It is the only finding of significant effect on student outcomes in the systematic review undertaken for INOVASI to guide its investment (Rarasati *et al*., 2016). [25]

In seeking the most effective investment in terms of students' learning outcomes, researchers combine evidence over a range of educational variables, including general classroom practice. In reviewing a decade of impact evaluations in South Asia (including Southeast Asia), Asim and colleagues found that investing in teachers, specifically in relation to teaching processes, consistently had the most effect (Asim *et al.*, 2015). Findings about which classroom practices are associated with student outcomes and to what extent, varies between the different systematic reviews, reflecting the broad local variations. While acknowledging the effect of different meta-analytical methodologies used in the different studies on effect sizes, Scheerens (2015:16) summarised the effect sizes of teaching variables from three meta-analyses of school effects (teacher practice counted as a school effect), including Hattie's review that synthesised 800 meta-analyses. The findings on the effects of teaching variables across the three meta-analyses are presented in table 3.

Table 3: Effect of teaching variables on students' outcomes: Scheerens' summary from three meta-analyses

Teaching level variables	Scheerens et al. (2007)	Hattie (2009)	Seidal and Shavelson (2007)
Time and opportunity to learn	.08	.34	.03
Classroom management	.10	.52	.00
Structured teaching	.09	.60	.02
Teaching and learning strategies	**.22**	**.70**	**.22**
Feedback and monitoring	.07	**.66**	.01

[25] See also: OECD (2019) and Glewwe et al.(2011)

Teaching and learning strategies— strategies that imply professional knowledge—have the largest effect, followed by feedback and monitoring, also large. Structured teaching and classroom management have medium effect sizes. In Indonesia, the USAID EGRA/Snapshot of School Management Effectiveness (SSME) supported the finding on feedback in its survey of teacher practices most strongly associated with student learning improvement (in this case (oral reading fluency-ORF). And it was emphased as a crucial skill of teaching in the OECD's report on the results of the 2018 Teaching and Learning International Survey (TALIS).

The USAID EGRA/SSME study looked more extensively at teaching processes (grade two level) than other studies considered here and found the largest effect associated with oral reading fluency was students' ability to state and defend their opinions (RTI International, 2014:57). This brings us back full circle to what the Teaching and Learning International Survey promotes in teaching practice: '…high leverage on student learning of cognitive activation'. This is exactly the kind of practice we discussed as necessary for developing comprehension skills.

However, these practices cannot be implemented effectively if they are not aligned with the prevailing cultures of teaching. And individual study findings on classroom practice have drawn attention to the persistence of memorization and recitation as teaching methodologies. The "deep embedding" of transmission models of education makes it very difficult for teachers to be consistently centred on the student (Vavrus and Bartlett, 2012). INOVASI's *Guru BAIK* pilot found that teachers' mindsets can be shifted at the level of *conscious* reflection; but the case studies (described in chapter 8) reveal how much of a challenge to change these teacher-centric habits can be.

The technical approach in student-centred teaching is to use formative assessment. If this is established as a core practice, other student-centred strategies naturally follow, such as differentiated planning for students according to their problems, and teaching and resourcing learning for readers at the right level. Competence in the subject pedagogy of literacy and effective classroom practice come together at this point.

5 Analytical approach

This chapter describes the analytical approach used to investigate the effectiveness of INOVASI's literacy interventions. We begin with a summary of the findings from the literature review developed for the study to provide a systematic basis for analysing effectiveness.

We then develop analytical constructs from the fit between these findings and what is suggested by the data from INOVASI's evaluation and monitoring process. This entails summarising the key features of the datasets used in the analysis and their respective contributions. The discussion moves on to explain the pilots targeted for the study and sets out the inquiry pathways to answer the key evaluation questions (KEQ). First, however, we address the question: *was there improvement in learning outcomes?* We then set out the more intricate pathway to address the question of what worked. This prepares the reader to follow the orientation in the remaining chapters of this study.

Developing the analytical focus for the study

Summary of findings from the literature review

The problems in Indonesian students' development of proficiency in literacy in Indonesia as presented in Chapter 2 fall into two categories. One category concerns problems in acquiring the component skills for learning to read, which disproportionately affect children in poor and remote regions in Eastern Indonesian. The other category of problem is comprehension: below minimal proficiency levels by Grade 4 for nearly half of the population, and very low performance on the higher order skills of comprehension.

The literacy research on effective approaches to literacy development in contexts of disadvantage is relevant to both these problem areas. As was suggested in the literature review, the long division reading researchers on effective approaches in these contexts, has given place to a general consensus that both mastering the code of written language and developing comprehension are vital from the start of school.

However, In disadvantaged contexts achieving a balanced approach faces challenges that better off contexts do not have to face. In the former, schools and systems need to be able to supply of books at the right level, enough of them to become the medium through which beginning reading is taught; and to enable children to read frequently and for pleasure. Th language issue has to be addressed in early grades. For many children the language of instruction is unknown at school entry. And the language issue is not just linguistic difference; the structures and words of school language are strange. Oral language development, listening comprehension and a focus on vocabulary development are all indicated by research as necessary instructional domains for developing the reading comprehension of disadvantaged children to the levels good readers have in advantaged contexts. These pre-requisites can only be satisfied by significant curriculum adjustments to give time for their sequenced development in such contexts. In second language situations this might mean a delaying reading instruction in the official language until the second year of school.

The research is also clear about the attributes that teachers must have to teach literacy effectively. Above all they need applied knowledge of subject pedagogy: of the phonological basis of reading, the cognitive levers of comprehension and the strategies for activating it. Of the more generic teaching skills, formative assessment is most critical to effective literacy teaching. This includes the capacity to use the assessment data to plan for differentiated teaching for the range of progress and problems any normal class of students presents. Tracking students' reading is the fastest route to them mastering decoding and fluent reading for meaning.

This emphasis on formative assessment has the potential to shift teachers' mindsets from teacher-centric to student centred. The transmission model of teaching is culturally embedded in many development contexts but integrating formative assessment into teaching routines disrupts rote-based instruction since student problems become the point of departure in teaching.

There are lessons, then, from the literature review for systems working to improve results like those in Chapter 2. These lessons are: the need for local adjustment of the curriculum to the contextual difficulties students face: centring early grades literacy teaching on direct instruction, from Grade 1, for component skills and comprehension; and an consequential investment in four pre-requisites for schools to be able to accomplish these changes: a major investment in appropriate reading materials; building a reading culture in schools; ongoing professional development of teachers in the subject pedagogies of early literacy and their application; and a system-led classroom and school focus on using data to monitor learning.

Orientation of the INOVASI data on literacy

The challenge of developing an analytical framework for this study is getting the best fit between the variables on which INOVASI has collected data; and relevant elements in the literature that are most associated with improved literacy performance in students and teachers. Such a fit would provide for a systematic analysis of literacy improvement aligned with variables known globally to be significant. It is also show whether INOVASI's context of operation yields similar or different findings.

To do this, it is necessary to describe the main sources of data in INOVASI for literacy.

There are three main quantitative databases this study can draw on to measure the achievements of the literacy pilots. In addition, the INOVASI education team reports and case studies of teaching practice provide qualitative data (chapter 8).

INOVASI's baseline education and learning survey – SIPPI

The baseline education and learning survey, referred to as SIPPI (*Survei Inovasi Pendidikan dan Pembelajaran Indonesia*) was INOVASI's baseline survey before it developed the pilots and before the program focused on early grades literacy and numeracy for learning improvement — one of the program's intended outcomes. As well as being committed to locally-responsive pilot variations, INOVASI needed to have measures of change at the program level. Evaluating activity at the pilot level only would make it impossible to aggregate the widely varying experimental results across the program, except at the learning outcome level. Thus it would be impossible to answer the question of 'what it was that worked'. The program's commitment to providing evidence also made it imperative to build baseline–endline comparisons that were statistically robust in terms of sampling. For all of these reasons, the SIPPI baseline–endline evaluation used generic variables that could apply to all the pilots.

The SIPPI database affords the literacy study five outcomes of interest in looking at student and teacher improvement: student literacy achievement in the component skills of reading and comprehension; teachers' literacy proficiency; classroom practice; teachers' mindset; and access to reading. This next section looks at how these variables fit with what is known about effective teaching and learning in literacy and how their attainment is measured.

Student literacy achievement in the component skills of reading and comprehension: The SIPPI has a student learning assessment at baseline and endline that draws on international assessment (PISA and PIRLS) constructs for reading comprehension (information retrieval, interpretation and evaluation/integration of meaning across text) and EGRA for the beginning

skills of reading. The survey therefore measures skills recognised as constituting proficiency in reading literacy.

Teachers' literacy proficiency: The SIPPI has teacher outcome proficiency measures based on the same tested comprehension constructs as those of the students. Teachers' own literacy abilities are not the same as pedagogical subject knowledge of literacy, the variable associated with a large effect on student outcomes. However it is reasonable to think while that personal competence in literacy does not entail being good at teaching it, not being very literate *is* likely to mean not being very good at teaching it.

Classroom practice: For teaching practices, the baseline drew on the literature on effective pedagogic and systemic practices in general rather than specifically in literacy. Thus we needed to identify the teaching practice variables most cogent for literacy to get a program-level picture of the effectiveness of INOVASI's literacy interventions. In this process we were guided by the fields of classroom practice identified in the literature review. These variables are the basis for an index of classroom practice we developed to assess teacher change on this dimension.

Box 3 sets out the SIPPI variables selected and their meaning for literacy teaching to make up the classroom practice index.

Box 3: Classroom practice index relevant to literacy

SIPPI variables	Significance in literacy teaching
1. Asking open questions	Developing comprehension skills including higher order reasoning
2. Giving feedback to students	Diagnostic approaches for differentiated planning in teaching reading
3. Using group/pair activities	Skills groupings of students for teaching at the right level
4. Using an appropriate teaching tool	Use of media to: • facilitate sound and letter matching • word recognition • comprehension and engagement through big books
5. Using the local language and Bahasa Indonesia alternately	Use of local language and vocabulary to aid comprehension and engagement
6. Displaying student work in the classroom	Student-centred approaches
7. Teacher circulating around all students in the classroom	Attention to individual learners

Teachers' mindset: The SIPPI collected data on teachers' mindsets to align with INOVASI's PDIA target and to avoid teachers and other stakeholders adopting the form but not the substance of the change. This has been labelled 'isomorphic mimicry' (Andrews, Pritchett and Woolcock, 2017). In SIPPI the perspective on changing teachers' mindsets uses Carol Dweck's trajectory of development from fixed to growth mindsets (Dweck, 2008). Indicators for a growth mindset include the following: being willing to embrace the problem as an opportunity to learn; having high expectations; valuing effort (as distinct from 'smartness'); and believing in the possibility of all minds learning. These SIPPI variables for mindset are consistent with the value the research literature puts on student-centred teaching, particularly in relation to teachers differentiating their

attention to students by identifying their particular problems. The index of SIPPI variables to measure mindset change in teachers, including those relevant to their views of their students is shown in box 4.

Box 4: Teacher mindset index

From the *Teacher self-administered questionnaire*

1. I can learn new things and I can change my intelligence
2. I don't have a certain intelligence level
3. I like to work where I can learn despite making plenty of mistakes in the process
4. I am very happy if given work that makes me think very hard

From the Classroom observation of teacher's behaviour

5. Praises students for their effort or performance
6. Encourages students to asking questions
7. Gives feedback to students

Access to reading: The student learning assessments in SIPPI drew on PISA and PIRLS for the literacy test and therefore include variables to measure influences from students' backgrounds. These are reading variables: availability of books in the home and in the classroom as well as frequency of reading and enjoyment of reading. These variables align with the emphasis in the literature review on access to reading material.

The constructs and variables for these five outcomes of interest were taken from different SIPPI instruments, including the student learning assessment tests and selected items from the classroom observation instrument and the teacher and student surveys.

The SIPPI baselines were administered to school communities on their entry into the program and this occurred at different times, reflecting INOVASI's graduated engagement with the different provinces. The survey was also carried out before each round of pilots. An endline evaluation, independently administered, was undertaken for the Literacy 1 (August 2019) and Literacy 2 (January 2020) pilots.

Spot-check data

In intent, the Spotcheck was for the collection of monitoring data on the pilots. In actual development by the INOVASI team, and in implementation, it has been used to make up for the generic nature of the SIPPI baselines by collecting data on the effectiveness of specific pilot characteristics.

The classroom observation instrument developed to collect spot-check data on the literacy pilots focused on teachers' use of subject pedagogy as well as their general practice. This data is the main source on the effectiveness of the pedagogical practice models that were trialled in the Literacy 1 and 2 pilots.

Spot-check data on the literacy pilots were collected twice: in May 2019 for Literacy 1 and in December 2019 for Literacy 2. The instrument items were designed specifically for the two different literacy pilots so no comparisons can be made over time with the spot-check data to establish whether teachers' literacy practice matured across the two pilots.

The "short course" pre and post test

For the Literacy 1 and 2 pilots the INOVASI education team developed pre and post tests and rubrics for scoring the tests, to establish what teachers understood about teaching literacy as a result of their participation in the short course and in the follow-up in the classroom. The INOVASI monitoring, evaluation, research and learning (MERL) team standardised the test scoring across the provinces.

The data from these tests reflect the extent of pedagogical subject knowledge teachers acquired through the pilots. In so far as knowledge of course content can be taken as subject knowledge, they also supply information, unavailable in SIPPI, about this aspect of teacher capability known to affect student learning outcomes the most.

Literacy 1 tested teachers' understanding twice – after units 1–3 and after units 4–7. The tests measured conceptual understanding of the key pedagogical strategies featured in the short course: phonological awareness; letter sounds; blending; fluency; and the concept of 'big books'. A single post test was administered for Literacy 2 that probed teachers' understanding of how to apply key pedagogies in supporting beginning reading and reading comprehension. The tests were administered by the local INOVASI education team.

Field monitoring

Field monitoring also contributed to tracking what was working in literacy. This took many forms and they all yielded qualitative data. While much of the knowledge at the field level was used for local adjustments to the pilots, this study extracted field monitoring data that has implications for the whole literacy intervention. Such data are mainly information from field work, looking specifically at literacy strategies through classroom observation or teacher interviews.

Teacher practices case study

Although the efficacy of the different intervention types for students or teachers' mastery of literacy are well established by global evidence, our study needed to check whether they also 'fitted' in our pilot contexts. We used three case studies of teachers' experiences in the classroom to explore how the teacher development intervention worked with the local culture of teaching. These qualitative case studies are only suggestive but nevertheless they show how new learning can be fitted into existing schema. This is useful to understand how to optimise take-up of effective practice and how to further develop teachers' capabilities in scrutinising the cultures and values that they operate in.

The inquiry pathways

Answering the overarching question of '*What works to bring about improvement in literacy outcomes in INOVASI's partner districts?'* requires two tiers of inquiry. Firstly we need to establish whether outcomes *did* improve – for both students and teachers. Secondly we need to investigate what among the different INOVASI interventions contributed to improvements and to what extent — in other words, we want to know what worked.

Inquiry 1: Was there improvement?

This first tier is explored by comparing baselines and endlines to see if there were gains in the endline outcomes that are not attributable to chance. This object of this inquiry is to find out the overall scale of improvement in outcomes for the targeted students and teachers. A second purpose is to identify provincial and district differences, starting point and end points, so as to better understand the significance of improvement.

The target pilots

The target pilots for this inquiry are those designed to develop teachers' knowledge and skills for teaching early grades literacy and that were directly managed by INOVASI. These are the Literacy 1 and Literacy 2 continuing professional development pilots.

As explained in chapter 3, these pilots test the efficacy of the teacher development *model* for literacy teaching. While government and some grantee pilots also focused on teachers' development, we examine only the outcomes of students and teachers who participated in Literacy 1 and 2 to see whether the pilots made a difference. These pilots are consistent in terms of content and management making them analysable as single interventions at program level. The contributions of grantee pilots to student and teacher outcomes are not discounted, however, as we compare them with results from Literacy 1 and 2 pilots in our second inquiry on what worked. Annex 1, section 1.2 sets out the pilots in each analytical category in this study**.**

As discussed in chapter 3, there are two sets of results from the INOVASI literacy pilots: those of participants in only Literacy 1; and those of participants in Literacy 1 who continued onto Literacy 2. Table 4 sets out how these two cohorts are treated in the analysis.

Table 4: Pilot participants, data and analyses for inquiry 1: Was there improvement?

<table>
<tr><th>Analytical focus and datasets</th><th>Target pilots</th></tr>
<tr><td>1. Program-wide baseline–endline comparison of students' literacy means on SIPPI student learning assessment:
Analysis for each grade level for each component:
1. Letter recognition
2. Blending
3. Word recognition
4. Basic literacy test (aggregation of scores on component skills 1–3)
5. Reading comprehension
6. Listening comprehension</td><td>INOVASI pilots only, including those with additional activity in pilot schools run by INOVASI or partners</td></tr>
<tr><td>2. Province-based baseline–endline comparison of students' literacy means on SIPPI student learning assessment
Analysis for each grade level for each component, as above</td><td rowspan="7">Step 1: Participants in Literacy 1 pilots

Step 2: Comparison of endline gains of participants of both Literacy 1 and Literacy 2 pilots</td></tr>
<tr><td>3. Program-wide baseline–endline comparison of teachers' literacy scores on SIPPI (MoEC assessment centre test)</td></tr>
<tr><td>4. Program-wide baseline–endline comparison of SIPPI teachers' mindset scores</td></tr>
<tr><td>5. Program-wide baseline–endline comparison of scores on composite SIPPI /spot check classroom practice index</td></tr>
<tr><td>6. Province-based baseline–endline comparison of teachers' literacy scores on SIPPI (MoEC assessment centre test)</td></tr>
<tr><td>7. Province-based baseline–endline comparison of SIPPI teachers' mindset scores</td></tr>
<tr><td>8. Province-based baseline–endline comparison of scores on composite SIPPI/ spot check classroom practice index</td></tr>
</table>

Inquiry 2: What worked?

What worked is a more complex question to answer. In this study we explore what worked in relation to the four key evaluation questions. The primary question (KEQ 1) is whether the teacher development pilots contributed to improving students' learning outcomes. This question is explored by examining the correlation of different constructs and variables with student outcomes on SIPPI.

Correlational studies also help to answer the other key evaluation questions: on the effect of book provision on student scores (KEQ2); on the efficacy of the language transition pilots (KEQ 3); and on the effectiveness of the pilots for improving higher-order thinking skills (KEQ 4).

We also look at relative effectiveness. While Literacy 1 and 2 were the dominant pilots in teacher development, there were variations on the theme in some implementations. We can compare outcomes of students who received these different treatments by comparing endline gains over baselines for the different interventions and with the overall performance of Literacy 1 and 2.

As described earlier, the study takes an additional approach to understanding what worked by focusing on the 'fit' of the teacher development intervention with the local culture of teaching, as exemplified in the case study explorations of three teachers' accounts of their teaching experience.

The evaluation questions and corresponding analyses and datasets for the second inquiry are set out in table 5.

Table 5: Analytical pathways for the question: What worked to improve students' and teachers' literacy outcomes?

Relevant evaluation questions	Analyses and datasets	Target pilots
KEQ 1: To what extent does training teachers to teach reading result in children's improved reading outcomes? KEQ 4: Is there any evidence that improved literacy resulting from the pilots will lead to better learning outcomes at higher levels/ across curriculum? Or better higher-order thinking skills (HOTS)?	**Analysis 1:** Correlational analysis of SIPPI student learning assessment scores with all teacher constructs: (i) teacher literacy scores (ii) classroom practice scores (iii) mindset scores Reading variables and student/school background variables	Step 1: Literacy 1 participants Step 2: Participants in Literacy 1 & 2

Relevant evaluation questions	Analyses and datasets	Target pilots
KEQ 3: To what extent does training teachers in mother tongue transition improve children's reading outcomes	**Analysis 2 Comparison of effectiveness of different training** Comparison of SIPPI SLA **Grade 2** student endline means of variant teacher development pilots: : INOVASI and key NGO partner pilots in teacher development including language transition	• Literacy 1 + Literacy 2 pilots • The language transition pilots: Sulinama, East Sumba and Bima, NTB • The Guru BAIK pilots: Southwest Sumba • Multi-grade: Probolinggo – Sukapura
	Analysis 3: Comparisons of SIPPI SLA Grade 2 student endline means from Literacy 1 plus book "add-on" activity; with endline means of students in Literacy 1 without book "add-on" activity.	Book provision pilots and book activity add-ons to Literacy 1 or 2: Indonesian Children's Literacy Foundation (YLAI) pilots West Sumba, Central Sumba Litara + OPOB (Malinau and Bulungan) Pen Circle Forum (Central Lombok) Rainbow Reading Gardens (East Sumba, Central Sumba)
KEQ 2: To what extent does providing appropriate books improve children's reading outcomes?	**Analysis 4:** Comparison of SIPPI baseline–endline student means on (i) classroom **reading corners** (ii) student **reading habit**	INOVASI and partnership literacy pilots: All schools with reading corner gains
	Analysis 5: Case study analysis **of teaching practice**	• Three video recorded literacy lessons and transcripts • Stimulated recall interviews on the lesson with the teacher Collaborative analysis of the lesson by INOVASI teacher mentors

The four chapters that follow work through these analyses and include more details about the samples, measures and procedures for evaluating the pilots.

6 Findings: Was there improvement in student outcomes?

This chapter presents the findings on whether students' literacy outcomes improved as a result of INOVASI's literacy interventions. There are two interests in this presentation of the data: what the data tell us about the nature of the problems in reading and what difference the INOVASI literacy pilots may have made.

The analysis is mainly of the aggregated performance across the program and at provincial level. Only at these two levels are the target populations of students large enough for evidence of effect that has statistical significance. Nevertheless, significant disparities within provinces can be brought to light.

The discussion is focused on findings from the Literacy 1 pilot, and to lesser extent from Literacy 2. Between them Literacy 1 and 2 reached far higher numbers of teachers than any of the grant-supported literacy pilots undertaken by partner organisations. As outlined in Chapter 3, grantee pilot inclusion in the analysis would introduce too great a variety of inputs to discern the clear effect of the pilot focus on score outcomes. In Chapter 9, however, in quest of the answer to "what works?" some grantee pilots are usefully compared with the approach of the Literacy 1 and 2 model.

As also explained in chapter 3, the Literacy 1 and 2 pilots were designed as a progression in teacher development for literacy teaching that provides a grounding in all the essential elements for a balanced approach. The value of Literacy 2 is as a successor to Literacy 1 but that progression did not always occur because some districts chose to implement Literacy 2 in schools that had not participated in Literacy 1. Only 78 of the 291 Literacy 1 schools sampled for baseline–endline comparisons followed through with Literacy 2. Literacy 2 has less analytical value as a stand-alone experiment than in the full implementation of the two-part model.

Therefore this inquiry into the effects of INOVASI's literacy pilots on student outcomes starts with the outcomes of Literacy 1, as the pilot that most participant students and teachers in INOVASI experienced. Following that, we look at the outcomes of those students and teachers who participated in both Literacy 1 and Literacy 2 and completed the endline survey.

Some students who participated in the Literacy 1 pilot were not available for the endline comparison. The move from one academic year to the next between the end of Literacy 1 and the beginning of Literacy 2 meant that there was a new Grade 1, (and their new baseline); and the loss of Literacy 1's Grade 3 students who graduated to Grade 4 at the beginning of July 2019. The number of sampled students who had experienced Literacy 1 and also Literacy 2 is 747—or 16% of the sampled students in Literacy 1. Ninety-five percent of their teachers also experienced both. This group has therefore has the advantage of constituting a panel, enabling a longitudinal study of effects on student outcomes over a longer period than a single pilot—between one year and eighteen months. The grouping of students and teachers in both Literacy 1 and 2 will be called the "Literacy 1+2 panel" in the analyses in this and following chapters.

The student literacy scores

To find out overall gains, we compared the student learning assessment scores of participant students in Literacy 1 with their endline scores.[26] This assessment tested grade one, two and three students on the component skills of early reading: (i) letter and sound recognition; (ii) sounding out words (blending); and (iii) word recognition. The assessment also included listening and reading comprehension. Item Response Theory (IRT) methodology was used to identify test

[26] The SIPPI version used for early grades student learning assessment was developed for INOVASI in 2017.

items at different levels, corresponding to expected grade level proficiencies. This meant the results could indicate proportions of students performing at, below or above their grade level.

The assessment drew on the EGRA test constructs for the component skills of early reading. For comprehension it drew on MoEC's student competence assessment (*Asesmen Kompetensi Siswa Indonesia* – AKSI) that is also based on PIRLS and PISA constructs (Puspendik, 2017). Reading comprehension was segmented into the three constructs of: (i) retrieving explicitly stated information; (ii) making straightforward inferences; and (iii) interpreting and integrating Ideas and information across text.[27]

The assessment was stratified so that only students who achieved two-thirds correct answers on all of the three basic component skills (that is, passed the program's benchmark basic literacy test) went on to do the comprehension tests. In the Literacy 1 and the Literacy 1+2 panel schools 4,784 and 747 children respectively were assessed on the basic literacy test. Of these 2,246 and 210 respectively were included in the comprehension test.[28] Thus we need to bear in mind that the results of the comprehension test represent the achievement of the top half of the students who participated in Literacy 1; and the top 39 per cent of the students who participated in Literacy 1 and 2.

Effect of the Literacy 1 intervention on student scores

Component skills of reading

Program level

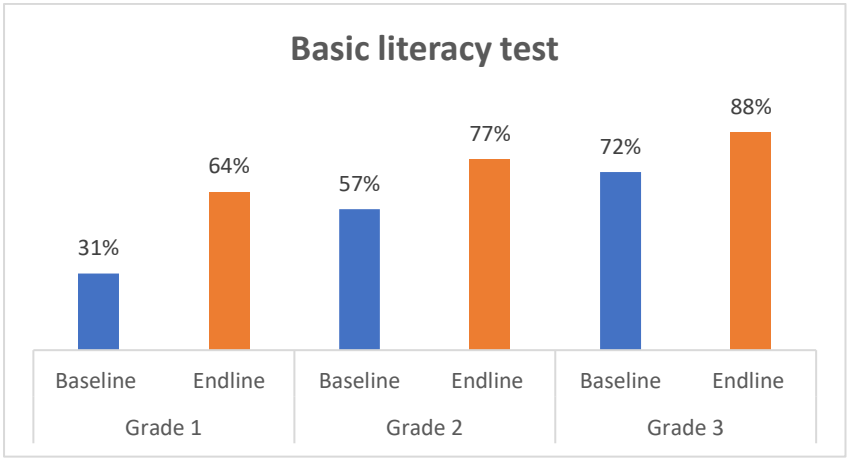
The number of students in Literacy 1 pilot intervention whose baseline and endline scores were so compared was in total 4,784. Of this number 1574 were in Grade 1, 1603 in Grade 2 and 1607 in Grade 3. In the following presentation of results those for the component skills of reading and for the comprehension tests are presented separately.

Figure 4 shows the results on the basic literacy test (the aggregation of the results on the component skills); and Figure 5, the results separated into each component. This helps to see where students' greatest challenges lie in these beginning skills of reading. All scores are rounded.

[27] The test was administered one-on-one. Instructions were delivered in the local language for children in these language contexts but the content was in Bahasa Indonesia.

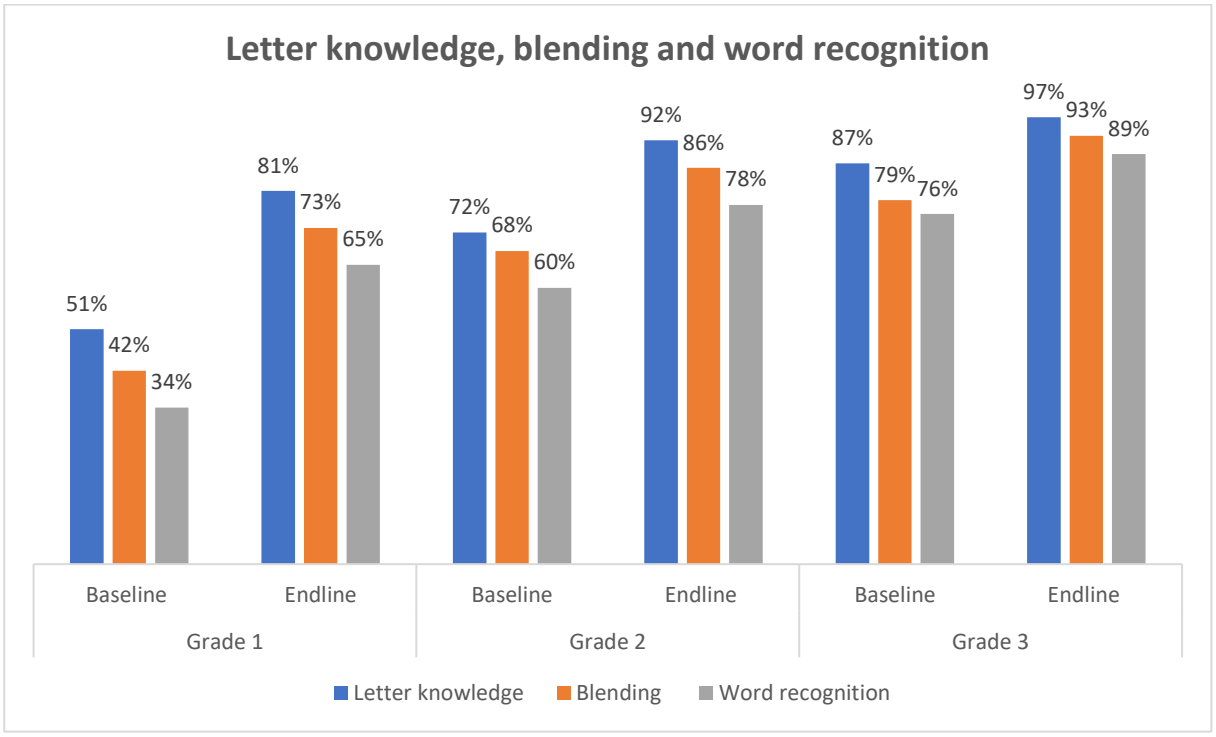
[28] Actual numbers who passed the basic literacy test were slightly higher in both Literacy 1 and Literacy 2: 2,521 and 292 respectively. However some students who passed were excluded because of data gaps in some of the student background variables needed for equating purposes.

Figure 4: Program-level baseline–endline comparison of Literacy 1 pilot student score means on the basic literacy test



At program level the grade two baseline is higher than in the USAID national EGRA study, which was 47 per cent for grade two students reading fluently. Figure 5 breaks the result down into the component skills.

Figure 5: Program-level baseline–endline comparison of Literacy 1 pilot student score means on the basic literacy test for each component (scores rounded)



The scores on word recognition are the lowest of the component skills, particularly at grade one. But the extent of natural growth between grades in all the skills, eliminates low scores in basic components for all but around 20 per cent of the tested population by the end of grade two and 10 per cent by grade three. This shows the attainability of these basic skills even without enhanced literacy teaching.

INOVASI did not have control schools for the Literacy 1 and 2 pilots. To find out what part of the difference between baseline and endline may have been caused by factors other than natural growth, we created a 'control' by treating the baseline of the next grade up as the untreated population. If the endline score of the previous grade (for example, grade one) is greater than the baseline score of the following grade (grade two) it indicates percentage gains not due to natural growth. The difference between the two is the size of the gain or deficit. [29] In this and following discussions these differences will be referred to as 'gains' to distinguish them from the improvements due to natural growth.

The gains are shown in Table 6.

Table 6: Program-level endline mean gains in the Literacy 1 pilot (after allowing for natural growth)

Gains in basic literacy test: Literacy 1 pilot		
	Grade 1	Grade 2
Basic literacy test	7	3
Letter knowledge	9	5
Blending	5	7
Word recognition	5	2

The gains are greater in Grade 1 than in Grade 2. The smallest increase across the components over the two years is in Word Recognition, where students perform lowest. The widened gap between results on blending and word recognition at the grade 2 endline may indicate problems associated with recognising word *meaning* as distinct from phonics-based word recognition.

Overall, these gains are not dramatic. Dramatic improvements would be unlikely in so short a time as a six month trial. The small scope that the competencies targeted in Curriculum 13 gives for component skills teaching, also reduced the opportunity for participant teachers to routinely teach these skills during the pilot; so the scores do not represent the effect of daily instruction.

Province profiles

The program-level analysis of data conceals extensive provincial variations in baselines and in gains from the Literacy 1 intervention. Baseline scores help make sense of provincial outcomes from the intervention. Table 7 shows the baseline range in the percentage of students passing the basic literacy test across provinces in the basic literacy test and the percentage gains at the end of the Literacy 1 pilot.

[29] No gain calculation for grade three is possible as INOVASI did not extend the tests to grade four teachers and students

Table 7: Baseline grade one and two percentage of student passes; and endline gains on the basic literacy test in the Literacy 1 pilot, by province

	% passes at Baseline and gains/deficits at endline	
	Grade 1	Grade 2
East Java baseline	**58**	**85**
East Java gain	**0**	**-2**
North Kalimantan baseline	**19**	60
North Kalimantan gain	13	9
West Nusa Tenggara baseline	33	62
West Nusa Tenggara gain	**6**	3
East Nusa Tenggara (Sumba) baseline	**3**	**21**
East Nusa Tenggara (Sumba) gain	**13**	11

Looking first at the baselines, the distance between provinces at the start of schooling is stark: for example, 58% of East Java students Grade 1passed the basic literacy test compared with 3% of students from from NTT (Sumba). That brings home the enormous difference in the circumstances of children in these two provinces and puts the spotlight on the distance Sumba children have to go to catch up with their peers elsewhere before schooling even begins.

A second finding that stands out is that the lower the baseline, the larger the gain – in most results. In the case of East Java, the impact in the basic literacy test is negative at grade two. It is also modest for West Nusa Tenggara. By contrast there are strong gains in North Kalimantan and Sumba where the baselines are much lower.

It is important to look at the individual skills at province level. In table 6 a pattern can be seen in the natural growth progression of students as they go through the grades. Columns 2 and 3 represent growth from of the previous year. Looking across from the Grade 1 baseline to the Grade 3 endline, on two of the component skills in most provinces, the large increases occur by the start of Grade 2, with a sharp drop in the percentage improvement by the end of Grade 3 — because the scores are already in the nineties.

Table 8: Grade progression in students' passes on the individual component skills, by province

		Grade 1 baseline (%)	% Growth by start of grade 2 (%)	% Growth by start of grade 3 (%)	Grade 3 endline (%)
		(1)	**(2)**	**(3)**	(4)
East Java	Blending	75	**16**	**5**	98
	Word recognition	61	**26**	**8**	97
North Kalimantan	Blending	30	**41**	**15**	98
	Word recognition	21	**43**	**16**	96
West Nusa Tenggara	Blending	43	**31**	**11**	92
	Word recognition	37	**30**	**13**	90
East Nusa Tenggara (Sumba)	Blending	8	**27**	**20**	86
	Word recognition	**3**	**21**	**25**	**77**

These data show the constrained nature of the component skills of beginning reading. Mastering sounds and letters and their combinations is finite learning, in quantity and complexity. Most students by mid primary school have acquired these skills, however slow the start, since the gaps and gains gradually narrow over the grades.

Sumba is an exception to the grade two pattern in important respects. There, children's acquisition of basic skills occurs as much in grade three as in grade two. These data suggest that the Sumba children are scholastically behind children in the other provinces. But It is important to note that in travelling from a baseline of 3 in grade one, they are within sight of the other provinces by the end of grade three.

Returning once again to the provincial picture, table 9 shows what the Literacy 1 pilot may have contributed to this skills progression. The table presents the endline outcomes on the Basic Literacy Test and the most fundamental of the skills—letter knowledge—to convey a comparison between the performance of the provinces on Literacy 1. The two lowest baselines provinces are coloured to highlight their higher gains.

Table 9: Endline gains in from Literacy 1 pilot on basic literacy and letter knowledge tests, by province

		Grade 1 gain	Grade 2 gain
East Java	Basic literacy test	0	-2
	Letter knowledge	-1	-1
North Kalimantan	Basic literacy test	13	9
	Letter knowledge	7	4
West Nusa Tenggara	Basic literacy test	6	3
	Letter knowledge	3	0
East Nusa Tenggara (Sumba)	Basic literacy test	13	11
	Letter knowledge	27	4

The larger effect was in grade one. The extremes in outcomes between East Java and Sumba again stand out, inviting reflection on the different effects of the intervention in provinces with high baselines of student capabilities in fundamentals compared with disadvantaged provinces. Considering how much of the gap Sumba closed with the other provinces by the end of grade three (table 7) these effects show how an intervention helps transform children's prospects of learning to read in such contexts.

There are important lessons from these different province profiles in baselines and gains for diversifying the inputs to take account of provincial differences. Negative results from Java Timur across both Grade 1 and 2 may indicate that in the target districts of this province there is already an effective pedagogy for beginning reading; perhaps even a little disrupted by a different approach to phonics introduced by INOVASI. In North Kalimantan the training in teaching beginning reading at both Grade 1 and Grade 2 level has added appreciably to children's progress. In a situation like Sumba's such training is crucial.

Comprehension skills

Program level

This section examines the provinces' performance in the three components of comprehension. Table 10 shows the baseline characteristics and the gains not attributable to natural growth for each of the comprehension components tested. (Note that 'reading comprehension' is the aggregation of the three component skills of comprehension, not an additional skill.)

Table 10: Program-level baseline and endline gains on student scores (rounded) on comprehension skills after the Literacy 1 pilot, allowing for natural growth

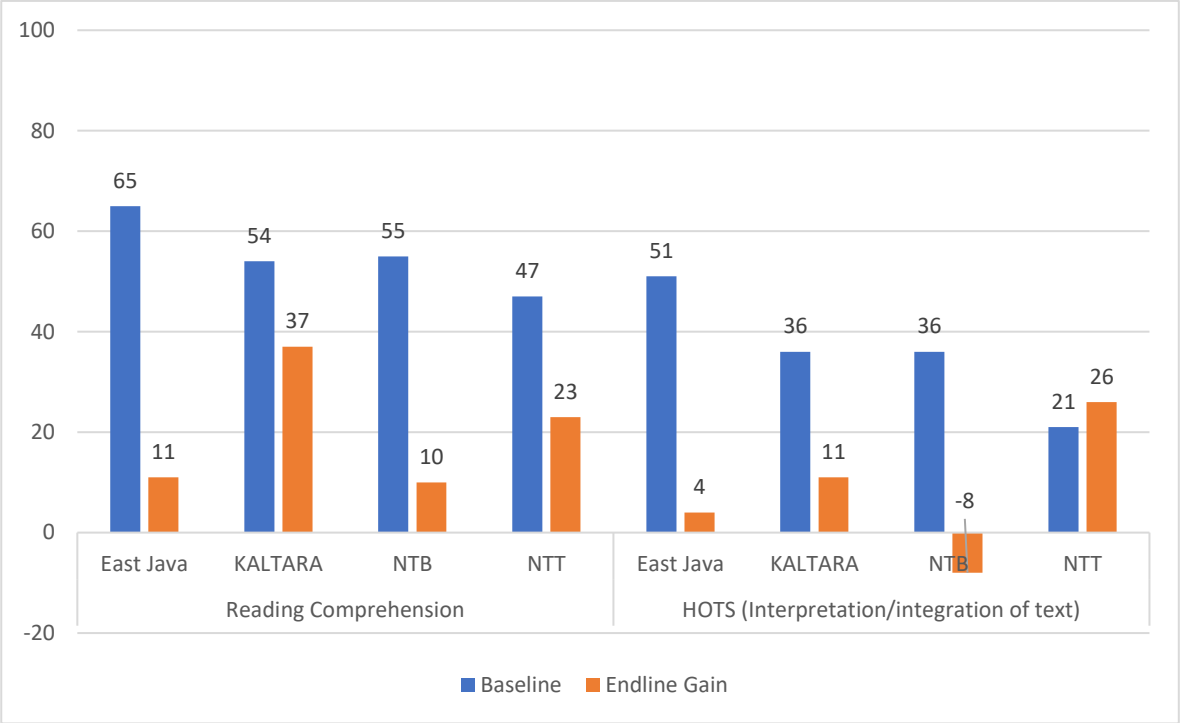
	Grade 1		Grade 2		Grade 3	
	Baseline	Endline gains	Baseline	Endline gains	Baseline	Endline
Focus on and retrieve explicitly-stated information	44	5	65	9	69	79
Make straightforward inferences	40	3	62	17	49	63
Interpret and integrate ideas and information	45	8	41	4	21	62
Reading comprehension	47	3	59	17	48	66
Listening comprehension	36	2	69	3	78	80

The endline grade two gains in comprehension are larger than the gains in the component skills of reading. While not uniform for all skills, these gains are strong for reading comprehension overall and high for straightforward inferencing – a critical skill for higher-order thinking. The higher gains for inferencing than for retrieving information may be because students are used to retrieving direct information in grade two Bahasa Indonesia classrooms. This takes the form of writing out answers to questions on the board or from the textbook passage and therefore it is not a new skill (INOVASI, 2017). Inferencing was the focus of a balanced literacy approach to comprehension teaching in Literacy 1 and this may be reflected in the students' results. On that basis, however, the gain for listening comprehension should also be high, as shared reading – listening to text – is integral to balanced literacy. But it is puzzlingly low, a reminder that the scores are not interpretable without more information. The low scores also on the highest of the higher-order thinking skills – Interpret and integrate ideas and information – may indicate that shared reading is not yet used to develop understanding across a whole text, a key function of shared reading activity in early grades classrooms.

Provincial profiles

Figure 6 shows the provincial range in reading comprehension and on the highest order thinking skill in comprehension: interpreting and integrating text. In this part of the analysis, we concentrate on results for grade two as the first year where these skills are appropriate for testing.

Figure 6: Profile of provincial mean baseline scores and gains in the student learning assessment test on reading comprehension and the highest higher-order thinking skills component



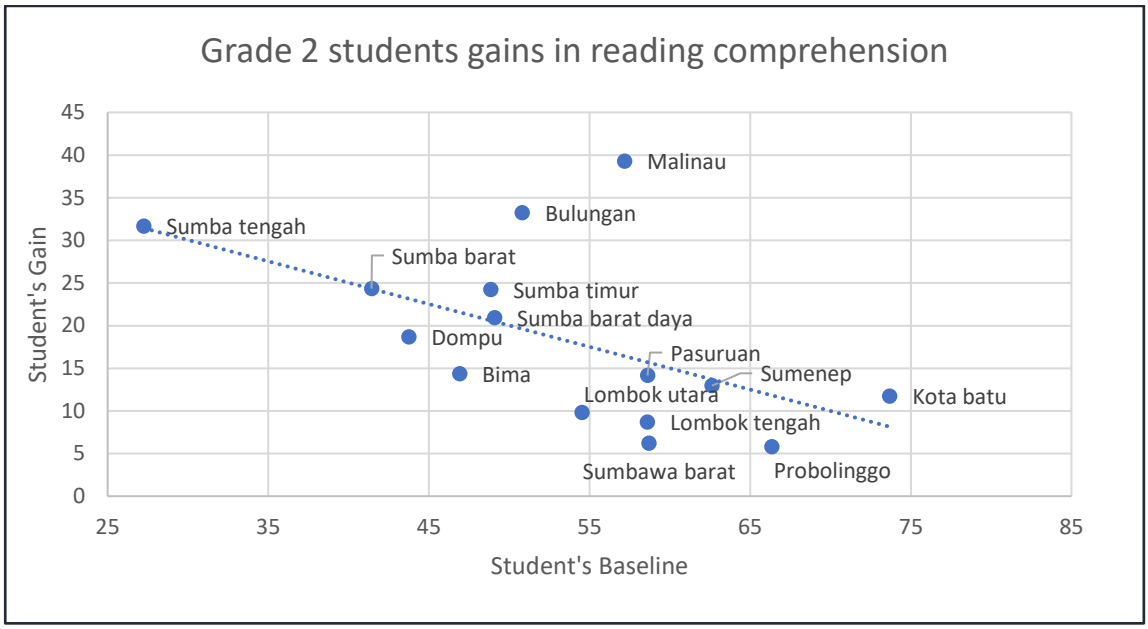
Notes: KALTARA = North Kalimantan; NTB= West Nusa Tenggara: NTT (Sumba) = East Nusa Tenggara; HOTS = higher-order thinking skills

Unlike with the component skills of reading, in reading comprehension all provinces gains at the endline. The pattern of bigger gains from low baselines recurs here too: North Kalimantan and Sumba have higher gains on reading comprehension than East Java and West Nusa Tenggara. The latter province even has a strong negative result on interpreting and integrating text. (This unexpected result cannot be interpreted without more information.) Sumba in East Nusa Tenggara has a gain of 26 points in higher-order skills, the highest of all the provinces. The mean scores for West Nusa Tenggara, North Kalimantan and Sumba on direct information retrieval were 74, 73 and 71 respectively, and Sumba scored highest of all provinces on listening comprehension with 93.

A finding that stands out is that on comprehension skills the gap between the provinces is less extreme than on the component skills of reading. This relates to the restricted eligibility of students for the comprehension tests: a high pass on all components of the basic literacy test. The proportions able to do that differed significantly in the provinces. Central Sumba, one district in East Nusa Tenggara had no students who passed, a reminder that disadvantaged children are blocked at the level of the component skills of reading. Without these children represented in the scores, the aggregate results for reading comprehension in the districts of Sumba are not so far from the district aggregates in the provinces of North Kalimantan and West Nusa Tenggara: 47, 55 and 54 respectively.

Figure 7 shows the pattern of relationships between low baseline and high gain for most districts on reading comprehension.

Figure 7: Distribution of district performance on comprehension against their student learning assessment



Effect of the Literacy 1+ 2 panel on student scores

Table 11 is about the value added to students' results of teachers' participation in the Literacy 2 pilot after completing Literacy 1. The table sets out the Literacy 1 baseline of students of such teachers to show the extent of the improvement at the endline of the first pilot; and then provides the percentage increase over the Literacy 1 endline in the Literacy 2 endline students' results of these teachers.

Table 11: The value-add of Literacy 2 to Literacy 1 endline gains at program level (Panel teachers)

	Literacy 1 grade 1 baseline	Literacy 1 grade 1 endline	Lit 2. grade 1 % endline increase over Lit 1 endline	Literacy 1 grade 2 baseline	Literacy 1 grade 2 endline	Literacy 2 grade 2 % endline increase over Lit 1 endline
Basic literacy test	29	68	12	57	74	13
Letter knowledge	54	88	6	71	98	0
Blending	38	78	8	65	87	4
Word knowledge	30	69	13	60	75	14

Compared with the extent of the improvement between Literacy 1 baseline and endline, the additional impact of teachers experience of Literacy 2 on student results is modest. The low results at the Literacy 2 endline for grade two are particularly disappointing as these students would have had the value of the Literacy 1 approach to teaching literacy in grade one whereas

the grade one for the Literacy 2 cohort is new. The low value added of Literacy 2 is concentrated in the two fundamental skills areas: letter knowledge and blending, while the significantly higher achievement of Literacy 2 over Literacy 1 is in word recognition. The different directions of these results may indicate the effects of the shift in Literacy 2 to focus on text comprehension and reading accuracy. However at least the gains from Literacy 1 are retained — and in most cases modestly added to, showing that what Literacy 1 developed continues to be applied six months later.

The lower achievement of Literacy 2 may also be explained by the limited budget and time that was available for this follow-on course. It had less than the full six months for delivery after the course adjustments required by budget cuts were finalised and in some places the course was only partially completed. Teachers' attendance at teachers' working group sessions however were as consistently high as in Literacy 1.

Literacy 2's value added for the comprehension results has not been included because some provincial results were unavailable.

Conclusion

The key evaluation question that this chapter set out to answer is whether there was improvement in student scores. The inquiry chose to answer the question by focusing on results from INOVASI's major experiment in literacy teaching development, the Literacy 1 and 2 pilots.

A robustness test, carried out to distinguish gain from natural growth, showed that there have been gains. They are appreciable rather than dramatic gains. Reasons for this may lie in the rapid acquisition of basic skills through *natural growth* in the early grades. This is a useful and unexpected finding obtained by tracking progress from grade one through to grade three. More than that this chapter does not speculate on why the gains are modest. However, important contextual information in evaluating the effectiveness of the pilots is that teachers had limited opportunity to implement the new methodologies in their classrooms as they had to continue to teach and evaluate students' performance on Curriculum 2013, limiting children's benefit from the skills teachers had acquired too.

For some provinces, notably East Nusa Tenggara and North Kalimantan, results on the different component skills (table 7) reveal larger effects than those at the aggregated basic literacy test level. This is a valuable finding diagnostically, showing for example, the low level of letter recognition in Sumba at grade one compared with the other provinces. The gains in that case also show the capacity of the program to overcome these key inhibitors to reading progress. Word recognition also lags behind the other foundational skills in most provinces but conspicuously in Sumba. This may point to language of instruction and vocabulary problems that are not fully recognised in early grades literacy methodologies and indicate that phonological approaches are not enough on their own to overcome these issues.

Overall, the findings also showed that the considerable variability in the extent of effect across the provinces is related to the different baseline levels of each. Importantly, the pattern of the lower the baseline the higher the gain was sustained through most findings, pointing to the particular value of interventions in disadvantaged locations.

The findings are positive on the effectiveness of the pilots on comprehension — more gains than in the component skills — and particularly on the higher-order thinking skill of inferencing. Perhaps that attests to the value of the balanced literacy approach adopted in the Literacy 1 and 2 pilots. East Java in particular gained more in comprehension skills than in component skills, suggesting that the emphasis needs to be on developing teaching methods in higher

order skills rather than in beginning reading skills. Teaching methods in the component skills already seem adequate in the province.

The extent to which a disadvantaged region such as Sumba closed the gap with other provinces on comprehension indicates the accessibility of higher-order thinking to children even while their basic reading skills are low. An important finding is Sumba leading the other provinces in the gains on listening comprehension, indicating the districts recognise the value of teaching methodologies that strengthen second language learners' understanding.

Nevertheless, the performance on comprehension, particularly in disadvantaged provinces, should not conceal the scale of the problem of reading poverty, represented by the small proportion of children in Literacy 1 and 2's catchment that proved eligible for the comprehension tests.

This chapter reported the outcomes for students without attempting to account for them. The following chapters suggest ways of making sense of these results, looking first at teachers' own improvement in chapter 7, further analysing this through the case studies in chapter 8 and then finally investigating whether those initiatives and the book pilots are associated with students' improvement in chapter 9.

7 Findings: Was there improvement in teacher outcomes?

> *'The teacher reads several sentences out aloud. The teacher writes parts of the sentences on the board for students to complete. For three quarters of the time, the students sit in silence copying the sentences from the board. Meanwhile the teacher sits or stands at the teacher's desk.'*
>
> *No student receives any help.'*
>
> *The Year 2 teachers observed seemed to assume that after year 1, teaching reading ceases.'*
>
> Source: INOVASI (2017)

Teacher practice is the heart of the change that INOVASI is seeking. Establishing the effectiveness of teacher practice is difficult to do in summative reporting. Evidence required to establish the representativeness of findings is based on quantitative data, derived from thin and inconclusive items in classroom observation schedules and other instruments. That way, it is not easy to see the breakthroughs that potentially transform practice. In presenting the data in this chapter therefore, an effort is made to evoke the achievement, where it occurs, that lies behind those thin descriptions. Contextual comment is intended to interpret and to moderate them in the light of field experience of pilot classrooms and stakeholder perceptions that INOVASI has accumulated. Hence also the opening to this chapter — a baseline picture of practice; and the inclusion of a case study of the practice of three teachers in the following chapter.

Effective teacher practice is a composite of particular capabilities, as the literature review reports. The most important among these capabilities that emerged are: teachers' understanding of the subject area itself, as well as pedagogical competence in it; classroom practice skills; and teachers' expectations of students' capability to learn. INOVASI's different pilot interventions were all concerned with developing these capabilities and the short courses in particular focused on subject and pedagogical knowledge and competence. The analytical framework in chapter 4 set out the constructs for these three capabilities that we look at in turn in this chapter.

This study uses all the INOVASI databases described in chapter 4. However, the spot-check data is particularly important because that database developed variables to look at characteristics of effective literacy teaching that are not included in the teacher or classroom observation instruments in the SIPPI database. Nevertheless discussion still draws on SIPPI variables in the constructs of classroom practice and teacher mindset, as baseline–endline accounts of teacher growth in practice are only available through the SIPPI data. This is supplemented by looking at growth in key practice variables in a small sample of schools where continuities between the two administrations of the spot check make this possible. [30]

The discussion continues to focus on the two short course literacy pilots as most relevant to teachers' literacy practice. We cover both short courses more evenly than in the previous chapter as each pilot covered different areas of teacher development in teacher practice and these differences were picked up in the spot-check data collection.

[30] Differences in schools participating in Literacy 1 and Literacy 2 contributed to the lack of continuity between the two implementations of the spot check, as did different sampling as between spot check 1 and 2.

Teachers' pedagogical understanding of literacy

As indicated in chapter 4, assessing how effective INOVASI's literacy pilots were in developing teachers' subject understanding was limited by the lack of baseline data in the spot-check mechanism. The only information we have on teachers' pedagogical understanding of literacy at the outset of the pilots is in their responses on the pre-test for Literacy 1**.** Although the test looks at conceptual understanding and not implementation, the responses still reflect teachers' *knowledge* of the appropriate pedagogies.

The open-ended items in the Literacy 1 pre-test were developed, scored and moderated by the INOVASI education and monitoring teams. In box 5 items from the two tests for Literacy 1 are arranged into the questions relating to the component skills of reading and reading comprehension (the two tests broadly represent the course division into these two areas).[31]

Box 5: Items testing subject pedagogical knowledge of teachers in the Literacy 1 pilot

	Literacy 1 units 1–3 test: Component skills of reading
1.	What does a classroom supporting literacy look like?
2.	What is meant by phonological awareness?
3.	How can phonological awareness be implemented in helping students learn to read?
4.	Why is it important for students to be able to match letters with sounds?
5.	Why is it important for students to know how to blend sounds into syllables?
	Literacy 1 units 4–7 test: Reading comprehension
6.	What is the effect on comprehension if a students is not a fluent reader?
7.	In your view can a student read if they are able to pronounce the words in a sentence?
8.	How are 'big books' used to help students learn to read?
9.	What is meant by 'predicting' and 'making connections' as strategies for supporting reading comprehension?

A total of 647 teachers (13 districts) participated in the units 1–3 test and 632 teachers (15 districts) in the units 4–7 test.

In the units 1–3 test, most teachers started from a low baseline and 90 per cent improved their understanding in the post-test. However only the East Java districts, West Sumbawa, Central Sumba and East Sumba scored above a 'pass' of 50 per cent and the others clustered around 30–45 per cent. In the analysis, questions were grouped into thematic categories. Teachers scored highest in the category of questions on reading for meaning and lowest on questions relating to decoding. In the test for units 4–7, 92 per cent of the participants improved their scores but the district performance was lower than in the first test, with only Batu, Bima and Central Sumba scoring 50 per cent or above. The theme areas that participants scored best on were in understanding reading comprehension and reading fluency.

[31] Pre and post tests were developed for Literacy 2 but are not used here. They probed teachers' application of what they learned through closed questions and the independent evidence of the spot check was considered more reliable than teachers' self-report in the post-test for Literacy 2.

The thematic areas that teachers scored best and least well in are important skills areas in early grades literacy and these results provide some indication of how teachers would be likely to cope with them in implementing learnings from the Literacy 1 course.

Teachers' literacy practice

The **Spotcheck** classroom observation instrument, used to monitor both the Literacy 1 and the Literacy 2 pilot, tells us about the extent of implementation of such knowledge, at a period towards the end of each of the pilot's duration. Data was collected by the facilitators who had conducted the training and mentored teachers. The observation was followed by an interview with the observed teacher, to obtain additional information and validate observations

Literacy 1

The 174 observations of Literacy 1 pilot teachers' in their classrooms covered all the Literacy 1 schools. Of these schools, 108 were 'plain' Literacy 1 pilots while the remaining 66 were either pilots that had transitioned into Literacy 1 pilots or pilots with additional modules.[32]

For Literacy 1 monitoring, the spot-check classroom observation reflected what teachers were doing in classrooms: how and to what extent the new learning was being integrated into practice. This was done through two questions. The first question was: *Which of the following activities were being conducted in the observed class*?[33] The results are shown in table 12.

Table 12:Strategies observed in classrooms during spot check 1 of Literacy 1 participants*

	Literacy teaching skills performed in observed classes	Implementing the activity %
1.	Learning the names of the letters	30
2.	Matching letters to their sounds	26
3.	Building words from syllables	30
4.	Practising reading aloud	72
5.	Practising writing	72
6.	Listening to a text and answering a comprehension question	45
7.	Storytelling/explaining personal information about themselves	29
8.	Reading a story aloud with the students	57

**Note: Other activities included: guessing words games, role-play games, using big books, making sentences exercises, questions and asking, counting words, matching pictures and words*

[32] These other pilots were: Literacy 1 and leadership; language transition (GEMBIRA); inclusive education (SETARA); community engagement (BERSAMA); and Literacy 1 with multi-grade learning
[33] Literacy 1 spot-check classroom observation instrument, question 32.

Decoding skills (items 1–3) were practised the least. This was the area where teachers performed lowest on the post test so this result may be because of difficulties they had implementing INOVASI's phonological approach.

Reading, writing and listening activities dominate. That may well reflect specifications in national teacher guides for Curriculum 2013 that the three macro skills of literacy (reading, listening, writing) should be covered in Bahasa Indonesia lessons. Speaking – the fourth literacy skill – is given less opportunity.

However, overall, the proportion of teachers implementing effective learning activities for literacy contrasts with the baseline snapshot at the beginning of this chapter. Most teachers (57 per cent) engage in shared reading of stories; and this is a promising result for improving comprehension and reading interest. The responses of teachers undertaking this practice also responded to the follow-up question in ways that show concern for improving comprehension: *If a story is read, does the teacher do any of the following activities with the children?*

Table 13: Frequencies of Literacy 1 sampled teachers integrating comprehension activities in shared reading

Asking students questions about the story		Asking students to predict what would happen		Asking students to re-tell the story		None of these activities	
#: 92	96 %	#: 52	54%	#: 45	47%	#: 2	2%

Literacy 2

For Literacy 2, 84 classrooms were observed. The spot check for Literacy 2 focused on the skills teachers were taught in the short course. These skills featured diagnostic teaching, formative assessment and reading comprehension, including higher-order thinking skills. The frequencies observed relating to those three skills areas are shown in table 14.

Table 14: Proportions of observed classrooms of Literacy 2 sample using reading strategies learnt in the pilot

	Literacy 2 short course emphases	Yes %	No %
1.	The teacher has implemented formative assessment	92	8
2.	The teacher conducts guided reading	56	28
3.	The teacher groups children according to ability levels for guided reading	51	6
4.	The teacher makes connections with the students' experience when reading to/with them	71	29
5.	The teacher models questioning that supports students' comprehension of the text: 'what, when, where, why and how' questions:		
6.	What, when, where questions:	54	
7.	How and why questions:	33	
8.	Only what questions	10	

This is an impressive snapshot of the skills developed in Literacy 2 being implemented in the classroom. The high percentage doing formative assessment is particularly notable. More than two thirds of teachers implement strategies to help students relate to the text (item 4) and most ask questions that help students retrieve information (item 6). There is less evidence of questions eliciting higher-order thinking skills (item 7). However, particularly positive is that most teachers practise guided reading with levelled groups (items 2 and 3). Behind this diagnostic activity lies the complex business of collecting data from hearing children read, assessing their reading levels; allocating them to groups depending on their reading problem, selecting appropriately graded readers for their level, as well as knowing what feedback to give on problems encountered during the guided reading session itself.

Taking the evidence of the post-test and particularly the implementation evidence of the classroom observations together, the Literacy 1 and 2 pilots seem to have given teachers a grasp of key elements in effective literacy practice. However, decoding and higher-thinking skills still seem to be mastered or taken up by a minority.

Teachers' general classroom practice

Featured in the literature review as relevant to effective teaching of literacy, are teachers' capabilities for student-centred teaching— the general practices of effective teaching. Key among student centred practices is differentiated teaching —teaching at the right level for the range of students in the class.

INOVASI has two sources of data relating to general classroom practice. The first is the Index of classroom practice, developed as explained in chapter 4 from classroom observation variables in the SIPPI baseline. In that chapter we also juxtaposed the variables from the SIPPI classroom practice index with teaching behaviour important to literacy to show their significance to this learning area.

The second source of data on classroom practice comes from variables of the spot-check database. These do not constitute an index; they are discrete variables with different scale values. However some of them overlap with important variables in the SIPPI classroom practice index and having the two sources of information means we can triangulate some results.

In the following section, we consider teachers' endline results on the SIPPI classroom practice index first – the performance of Literacy 1 teachers and then those who participated in Literacy 1 and 2.

Performance on the SIPPI classroom practice index

The Literacy 1 pilot: The following two tables set out, at program and provincial level respectively, the growth of teachers between baseline and endline on the classroom practice index from the SIPPI database. Results represent a sample size of 482 teachers.

Program level

Table 15: Literacy 1 pilot teachers' baseline–endline on SIPPI components of the classroom practice index: program level

Literacy 1 Pilot	Baseline %	Endline %	Gains %
Asking open questions	78	93	15
Giving feedback to students	62	60	-2
Group/pair activities	22	29	7
Using an appropriate teaching tool	37	61	24
Using local language and Bahasa Indonesia alternately	70	68	-2
Student work displayed in classroom	42	71	29
Spatially balanced teaching (teachers paying attention to all corners of the classroom)	44	75	31

Table 15 shows a range of outcomes on the different variables. The strongest gains at endline are in those that easily make a classroom more student-centred: *student work displayed in classroom* and *spatially balanced teaching* (29 and 31 percentage point gains respectively). *Giving feedback* — critical in INOVASI's literacy approach and more difficult to do — had a negative result. This result seems to be affected by the large proportion of participants from East Java where performance on this variable went backwards. It is not reflected in the results of other provinces, though their gains are small on this item.

> *Before we didn't really plan, just used the book. Now I have to study how to make the class work. I have to work out what I want the children to do.*
>
> *Teacher in Bulungan school, 2018*

The third highest scoring variable, however is one that indicates pedagogical understanding of the skill or concept being taught: *Using an appropriate teaching tool*. Arguably this is one of the most important developments that INOVASI has made to teacher practice. It is conceptually difficult to achieve, especially if there are no pre-existing resources for it. Lesson planning focusing on children understanding the concepts taught and identifying appropriate media to support them was included in pilot activity even before the short course pilots began. Facilitators and teachers' reflections recorded that this was the hardest skill to develop (INOVASI MERL team, 2018). The consensus view of principals and local heads of education is that effective use of media has made the most transformative shifts in the classroom. The idea most frequently used to describe the change is students being actively involved, as in this comment by a principal of a school in Sukapura:

> *'The old way was this: the teacher generally gave the class a task or got them to work a problem and then left them to it. With the new way, the teacher and the students are both much more active in the process of learning as well as more creative so that learning becomes more balanced between the teacher and the student — as well, the*

school head becomes interested in learning about this way of doing things' (Interview, school principal, Sukapura, 2019).[34]

The quality of the resources developed to help learning was also striking. Many participants from Literacy 1 have developed big books to read with children. Technical understanding is required for the choices of theme, word level, text length, font and supporting graphics in developing these crucial resources for beginning readers.

Provincial level

As with student outcomes in the previous chapter, the SIPPI findings show wide differences between the provinces on the classroom practice index.

Table 16 compares the baseline and endline performance of participants at the aggregate level of the SIPPI index, for the program and the provinces.

Table 16: Literacy 1 program and provincial means on the classroom practice index

Classroom practice index: Literacy 1 pilot		
	Baseline	Endline
Program (482 teachers)	45	68
East Java (153 teachers)	61	64
North Kalimantan (55 teachers)	27	54
West Nusa Tenggara (145 teachers)	40	72
East Nusa Tenggara (Sumba) (129 teachers)	41	76

The gain in the program mean at endline overall is high, despite the minimal gain for East Java with a large number of participants. The large gains are in all the other provinces. Sumba, rather than East Java that has the highest endline improvement, exceeding East Java's endline by 12 percentage points.

The other provinces also reflect large gains in using *an appropriate teaching tool.* In North Kalimantan, West Nusa Tenggara and Sumba the gains range from 27–41 percentage points in that sequence.

A variable with consistently low gains is: *Using local language and Bahasa Indonesia alternatively.* In East Java the score actually went backwards by 18 points. The gain is smallest in Sumba although two grantee language transition pilots took place there – but not in Literacy 1 schools. Nevertheless, the Literacy 1 short course lacked a focus on language of instruction issues and this may the reason for these results. This is a lesson learned for literacy pedagogy in regions where the largest stumbling block to literacy is the language of instruction.

Again in line with the pattern of student outcomes improvement in the previous chapter, districts with low baselines consistently improve the most and some of their improvements are sizeable, as shown in Table 17.

[34] Transcribed and translated from the original in Bahasa Indonesia, interview with school principal, Mr Agus Hartono, Ngampelsari primary school, Candi district, Sidoarjo, May 2019 field visit for fourth strategy testing.

Table 17: Lowest district baseline with highest improvement in Literacy 1 on the SIPPI classroom practice index

Districts with the lowest baseline and the biggest district improvement	Baseline	% improvement
East Java: Probolinggo (Multi-grade)	48	31
North Kalimantan: Bulungan	21	37
West Nusa Tenggara: Dompu	34	40
East Nusa Tenggara: West Sumba	37	46

A possible explanation could be that participants in provinces and districts that have had more exposure to established practice were less able or disposed to assimilate new ways of doing things than regions and districts that had no models in place for teaching literacy.

The Literacy 1+2 teacher panel: Table 18 shows the relative performance of Literacy 1 and Literacy 2 on the SIPPI classroom practice index at program and provincial level (136 teachers).

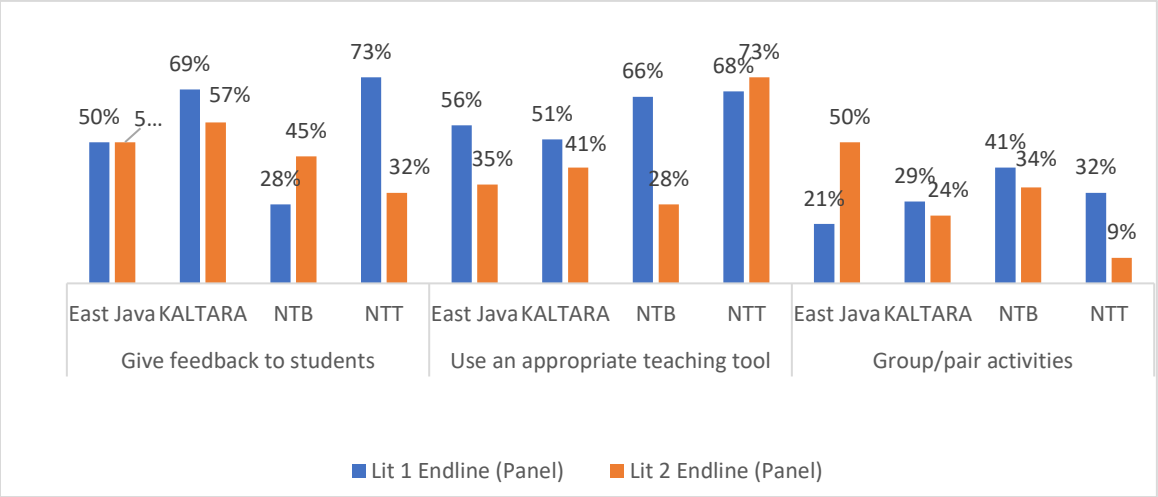
Table 18: Literacy 1 + 2 panel program and provincial profiles on the classroom practice index

	Baseline Literacy 1	Endline Literacy 1	Endline Literacy 2
Program level	42	64	66
East Java	68	62	80
North Kalimantan	27	54	72
West Nusa Tenggara	35	77	53
East Nusa Tenggara (Sumba)	44	75	59

At program level, comparing the endlines of Literacy 1 and Literacy 2 shows a minimal gain of only 2 percentage points, compared with the program performance in Literacy 1 on this Index.

Looking at the provincial pattern we selected three important variables from the index to see if there are differences between the Literacy 1 and 2 performance. The selected variables in figure 8 between them all require an understanding of literacy teaching and were the focus of Literacy 2.

Figure 8: Comparison of the performance on Literacy 1 and Literacy 2 (panel) on key classroom practice variables



Notes: KALTARA = North Kalimantan; NTB= West Nusa Tenggara: NTT = East Nusa Tenggara

Overall, the results for Literacy 2 are disappointing for variables that were central to it. West Nusa Tenggara is the only province to advance on giving feedback. The decline on *using an appropriate teaching tool* is consistent and across all provinces except for East Nusa Tenggara (Sumba).

Particularly disappointing in comparing the two endlines is the decline of West Nusa Tenggara and East Nusa Tenggara over their achievement in Literacy 1. The results for East Nusa Tenggara are erratic compared to its Literacy 1 achievement with a drastic decline in giving feedback and in group work down to 9 per cent.

This finding is at odds with the actual development in East Nusa Tenggara that included one of INOVASI's most successful classroom innovations.

This is the Pratham-like practice teachers developed once they learned how to recognise the problems students were having in decoding skills. The teachers who developed it describe the method in box 6. This is a high level of "giving feedback": formative assessment integrated into teaching. The excerpt also shows the whole school seeing the potential of getting data and using it to help the students. The education authorities in West Sumba and East Sumba have subsequently asked their schools to implement the practice.

Performance on the spot-check classroom practice variables

Table 19 presents findings across the two spot-check data collections on implementing practices to support effective literacy teaching in classrooms in the 34 schools that participated in both the Literacy 1 and the Literacy 2 pilots. (The general classroom practice variables were mostly the same in the two spot-check instruments as variations were only made for the literacy specific ones to reflect the different emphasis in Literacy 1 and 2.) The variables selected in Tables 17 and 18 look at the same practices as in the SIPPI classroom practice index: *giving feedback* and *use of group work*. We selected these because they are intrinsic to the effective literacy practice featured in Literacy 2 and yet findings on them in the SIPPI classroom practice index show such dramatic declines in several provinces.

Table 19: Spot check comparison of performance of schools participating in both pilots on informative feedback variable

'Giving informative feedback' in the spot check			
Spot check	**Frequent**	**Sometimes**	**None**
Literacy 1	4 or 12% schools	9 or 26 % schools	21 or 62% schools
Literacy 2	13 or 35% schools	6 or 18% schools	15 or 44% schools
Remarks	**Improvement: 23%**		

While the sample is small, one-third of teachers frequently giving informed feedback is an important development achieved at the end of Literacy 2, given that only a mark on a page was the longstanding practice up to this intervention. It is also an achievement that the proportion of teachers failing to give any feedback has declined by nearly half. The findings are at odds with the finding on feedback from the SIPPI classroom practice index – including its high baseline. A possible reason for the discrepancy in the two results is that the integration of formative assessment into the teaching activities through such practices as guided reading may not have been recognised by the enumerators for the SIPPI classroom practice index, whereas the facilitators collecting on the spot check had led the training and knew what should be included.

Table 20: Spot check comparison of performance in both pilots on the use of group work variable

'Use of group work variable' in the spot check			
Spot check	Yes, with differentiated tasks	Yes, with same tasks	No tasks at all
Literacy 1	12% schools	50% schools	38% schools
Literacy 2	41% schools	32% schools	27% schools
	Improvement: 29%		

On this variable as well, integral to effective reading practice, there are gains and decreases in the right direction. On this data collection, in 41 per cent of classes observed for Literacy 2, teaching to differentiated groups was taking place. The skill of managing a class with differentiated tasks is one of the highest skills a classroom teacher can have.

Box 6: Interview with early grades teachers SD Masehi 2

Waikabubak, West Sumba, 26 March, 2019

In my class 2, of 26 I have 12 students who had to stay back from class 3. Before INOVASI we had different ways of teaching children to read the letters of the alphabet. They learned it through to Z by rote. They could do that. But if we mixed up the letters they didn't recognise them.

After INOVASI came, we grouped children by the problem they had: a letter group, a syllable group, a word group. We split all the grades up to grade three into these groups. I had 14 children in my group who could not recognise the letters, Ibu Esy took another group. Ibu Nia took the word group.

We brought in the subject teachers to help out.

So that group of 12 children I had in my class who stayed back, after we did this, they all now have got their letters, they can work out syllables and now have moved up into the word group.

Improvement in teachers' own reading literacy scores

Teachers' own reading literacy was not directly targeted by any INOVASI intervention although it is relevant to students' outcomes. Nevertheless, we can see if there was an improvement in teachers'

own proficiency by comparing their reading literacy score in the SIPPI teacher test and at the endline of the panel teachers participation in Literacy 2. By the end of the Literacy 2, learning strategies for improving students' reading comprehension may have influenced teachers' own performance on comprehension.

Table 21 shows this comparison.

Table 21:Panel teachers' endline increases in literacy proficiency

Teachers' reading literacy score	Baseline Lit 1	Endline Lit 1	Endline Lit 2
Panel teachers' reading Literacy 1 & 2 (140 teachers)	51	55	59

There is an increase in the panel teachers' reading literacy at the end of Literacy 2 — and it is not insignificant. It is similar to the increase in literacy proficiency found at the end of the Sumbawa *Guru BAIK* experiment where there was an increase of 10 per cent. In neither cases had literacy been explicitly addressed but possibly influenced by engagement with literacy learning.

Teacher mindset improvement

As discussed in the analytical framework in chapter 5, INOVASI developed a teacher mindset index to measure change in teachers' outlook over the program. This index initially used Carol Dweck's growth–fixed mindset indicators to measure teachers' mindset change (Dweck, 2008). The result was a teacher-focused instrument, so the INOVASI team added variables from the SIPPI (*Give feedback to students; Encourage students* to ask questions) and the spot check (*Students receive praise for their effort or performance)* to help orient the inquiry to teachers' behaviour towards ***students'*** capacity for growth. Table 22 presents the results on mindset change for Literacy 1 and for the beneficiaries of the Literacy 1 + 2 panel.

Table 22: Program results on items in the mindset construct for Literacy 1 and Literacy 1+2 panel teachers

	Lit 1 Baseline	Lit 1 Endline	Lit 1+2 panel endline
The Dweck construct of mindset (teacher self-administered questionnaire)			
I can learn new things and I can change my intelligence	64.50	66.26	68.61
I don't have a certain intelligence level	65.30	65.95	66.17
I like to work where I can learn despite making plenty of mistakes in the process	80.22	79.00	78.57
I am very happy if given work that makes me think very hard	73.40	72.84	74.25
Classroom observation (teaching application)			
Praise students for their effort or performance	74.20	84.63	81.20
Encourage students to asking question	34.54	41.05	30.83
Give feedback to students	62.47	59.58	48.87
Teacher mindset index	46.38	54.24	52.62

On the Dweck component of the Index there is minimal gain – two and four percentage points respectively at the end of Literacy 1 and Literacy 2 for the panel teachers. At provincial level the same stability is consistently shown across the three time points in the Dweck component.

The decline in the education component of mindset between Literacy 1 and the panel participants contradict the findings from the spot check on formative assessment and differentiated working groups. (We saw this discrepancy between the two surveys on these two variables in the earlier discussion of the classroom practice index.)

Developing teacher reflectiveness had been the basis of the approach to mindset change developed in the *Guru BAIK* pilots. As explained in chapter 3, INOVASI restored an emphasis on teacher reflectiveness in the Literacy 2 pilot because lack of attention to it in Literacy 1 was viewed as a loss. Therefore, the increase in mindset change on the Dweck variables at the end of Literacy 2 reflects some success. However, the kind of reflectiveness that teachers need to develop, to support growth in their students' mindsets, is a more important inquiry and will be featured in the following chapter.

Conclusion

What conclusions can we draw from this chapter on the key evaluation question: did INOVASI improve teacher practice?

There is no unified answer to that question and four main reasons suggest themselves for this.

One reason is that some critical teacher practices for improving literacy improved while other equally critical practices did not show clear progress. Those that evidently improved are practices supporting reading comprehension, including a start on higher-order thinking skills. There is also strong evidence of teaching to the problem in both general teaching and literacy-specific practices. Many teachers have the capabilities to target lessons and media for a concept and to organise differentiated lessons for different reading levels. Teaching approaches associated with decoding did not make much progress. This is exemplified in both the Literary 1 post-test and in the lower take-up of these practices compared with other literacy teaching activities reported by the Literacy 1 spot check (Table 17). It also corresponds with the findings on student outcomes in chapter 6 showing that gains in the component skills of reading on the student learning assessment, while achieved, were more modest than the gains in reading comprehension.

We also saw in chapter 6 how many children in some regions do not have basic skills in reading and so improved practice in teaching the component skills is an issue that needs probing. INOVASI's sequenced approach to decoding is based on established global practice. However the approach may have competed in some regions, such as East Java, with different, established methodologies — syllabic rather than phoneme based — that evidently work well with Bahasa Indonesia, given the rate of students' acquisition in that province.

Provincial differences in response to the two literacy pilots is another reason there is no single answer to the question of improvement. The range of practice take-up between provinces broadly followed a similar pattern to the range in student gains in chapter 6: provinces (and districts) with low baselines seemed to benefit more from the pilots. But the pattern does vary with variables –seemingly erratically sometimes, as shown in Table 17 – so more scrutiny is needed of the areas of literacy pedagogy that different provinces responded well to, or did not.

This apparent erratic nature of the data is a third reason why a unitary answer cannot be given to the evaluation question. An underlying theme of this chapter is: *what are appropriate ways of*

reporting at scale on so wholistic a process as teacher practice? The field verdicts are so much stronger than the scores. This is not likely to be because of partisan reporting but because it is impossible to see the depth of change in a summary variable statement on a questionnaire.

This may explain the widely different picture given of the success of the pilots in the SIPPI data collection and in the spot check. An interesting question is whether observation is more reliable when it is disinterested – as in external enumerators looking at the classroom for the SIPPI – or when it is interested – as in the training facilitators in the spot check looking at the classroom for evidence of take-up. At the very least, variables such as 'give feedback' could mean two different things to the different sets of evaluators. This becomes an issue when a whole pilot program seems to have come out with little value added, as Literacy 2 does on the SIPPI panel collection.

The fourth reason why there is an incomplete answer to whether teaching practice improved is that the literacy pilots were not adequately represented in the SIPPI database. This database is the only source of baseline–endline comparisons that could definitively answer the question of success. The literacy pilots had to make do with baseline measures that did not reflect their core work because those pilots and the core contributions had not been designed at the time the SIPPI baseline was set. A full answer to the question of whether INOVASI's approach to *literacy* teaching and learning improved practice relies on other ways of knowing than the data we have.

8 Teacher practice case studies

These literacy practice case studies are developed from reading lessons delivered by three different teachers and the INOVASI education team's subsequent analysis of the lesson, as well as the teachers' own reflections.

Purpose

The purpose of the activity was to better understand teachers' beliefs about teaching reading and the constructions they put on the approach to teaching reading in INOVASI's Literacy 1 and 2 pilots. Looking at actual teaching behaviour in relation to INOVASI's effectiveness clarifies, supplements and provides plausible ways of interpreting the behaviour reported via quantitative items in the previous chapters. As part of understanding the relevance of INOVASI's intervention in the broad cultural sense, the study seeks to know how the new learning interacts with the local culture of teaching and teaching literacy in particular.

The study concept

This study draws broadly on the design of the World Bank Trends in International Mathematics and Science Study (TIMSS) video study of mathematics teaching (World Bank, 2015). This entailed making a video-recording of lessons for analysis of the practice by the teacher and the researcher. The video record enables intensive analysis of small interactions, including body language, that in the following analysis afforded glimpses of larger, unspoken drivers of teacher behaviour. Research into this 'ethnographic' use of videos cautions against interpreting actions or demeanours, that might be due to the intrusion of the video camera, as typical behaviour or cultures (Jewitt, 2012). Anyone analysing the video therefore needs to be informed about the process and background to avoid any misinterpretations. This proved true during the analysis of these three lessons and highlighted the worth of including team members who worked on the pilots with these teachers.

The case study also built on the approach that Louden and Wallace (1995) used in developing teaching case studies for the Australian National Project on the Quality of Teaching and Learning. They layered the interpretation of the lesson and the teacher's reflection by additional analysis from peers and subject experts. In this case study additional analyses came from the education program development field mentors in the teachers' districts and then analysing and synthesising the learnings from the whole ensemble of materials in an education team workshop.

Methodology

Selection of teachers

Teachers were selected based on a stratified random sample. The criteria were full participation in Literacy 1 and 2, and for one teacher of the three to be randomly selected from the sub-group of teachers who scored highly on the Ministry's teacher test. This was to see if teachers' own literacy proficiency made a difference to their teaching of literacy. A short list was developed on these criteria and submitted to the relevant INOVASI education advisors to ensure that all those chosen were reasonable performers in the classroom.

This selection process yielded one teacher from a district in an urbanised district that had low results on the student tests and two other teachers from a remote and disadvantaged district. Though unintended, having two teachers from the same district context helped to avoid

ascribing cultural influences on teaching when they were simply displaying individual traits. As it happens, two of the selected teachers had high scores on the literacy test.

The process

The key elements in preparing for the lesson study started with the local mentor from INOVASI's education team obtaining consent for the study from the teachers, school heads and the parents. Then the mentor briefed the teachers on the strategy to exemplify in the lesson and the lessons were professionally video-recorded. Some days later, the mentor interviewed the teachers, after a joint viewing of the video-recording to stimulate recall. The purpose of the interviews was to have the *teachers* identify moments in the lessons that they thought significant to see what these judgments indicated about the participants' teaching and learning values. The mentor's role in this process was to ask neutral questions to elicit these thoughts, offering no opInion or comment. This way the teachers' selection would be uninfluenced.

The mentors, through their knowledge of the context, were able to enhance understanding of the lesson and they also wrote an analysis of the lesson they had observed after reviewing the video. These analyses, together with the original video and the teachers' reflections, were used for the workshop analysis of the videoed samples of practice. The full INOVASI education team participated in the discussion.

The focus strategy

The reading strategy chosen for the lesson focus was guided reading with a group of four students. This strategy draws on many skills in the teaching of reading and well exemplifies INOVASI's approach: integrating formative assessment and targeting reading instruction to the right level. The strategy consists of listening to students read in turn and provides a wholistic context for strengthening reading: supporting decoding, word knowledge, fluency and comprehension at the level of the sentence and the whole text. Immediate feedback is an essential part of the process and is often used to make a 'running record' of the students' mistakes and correction strategies to guide later strengthening. The case study teachers were asked to complete a running record for each child in the lesson session. The books the children read were graded readers, chosen by the teachers as appropriate to the level of the guided reading group.

The analytical framework

The analytical framework for the lesson was adapted from the constructs used throughout this literacy study to analyse effective teacher practice. Key to the case study is the extent of pedagogical subject understanding the teachers manifested in the lessons. This has two aspects: the depth of understanding behind the implementation of strategies taught; and the extent to which the INOVASI model fits with strong, existing drivers of classroom practice.

Mindset is also central. This construct also has two aspects in relation to the cases: teachers' beliefs about reading and teaching reading; and the tenor of teachers' interactions with their students. General Classroom skills, the third construct used in the Literacy study, is less appropriate for a study of a teacher with four students and not part of the analytical framework for the case studies.

Box 7 sets out the constructs organised in a sequence that best fits the analysis and discussion.

Box 7: Summary of the analytical dimensions for the case study lesson analysis

Analytical constructs for the case studies
1. Mindset: teachers' beliefs about reading and teaching reading 2. Mindset: Teacher–student relationships 3. Pedagogy for literacy: technical understanding and its fit with context

In the following discussion we present a description of the lesson in each case first. This is a synthesised description, based on a log of each lesson. The description is intended to provide evidence of what teaching looks like after an INOVASI pilot, as well as to enable the reader to understand the focus of the analysis that follows.

After the description of the lesson is a synopsis of the teacher's reflection on it, followed by a summary of the discussion and conclusions reached in the workshop organised according to the constructs in box 7.

The lessons

The three lessons are thirty minutes long, a regular time allocated for a Bahasa Indonesian lesson in Indonesia. They were conducted with four grade three children seated with the teacher around a table. We intended to have children at the end of grade two but to fit with the end of teachers' participation in Literacy 1 and 2 (late November 2019), we had to use grade three students, at that stage nearing the end of their first semester.

Lesson A

About the lesson

The book selected for the guided reading was *Di mana Telurmu?* (*Where are your eggs?*), a narrative exploration supporting children's science knowledge about animals that lay eggs, with a surprise encounter at the end of the story of an animal that has live births. It was a graded reader at level D-2 appropriate for children at the start of grade 3, with 5–9 sentences per page.

The lesson had six distinct teaching segments: (1) the teacher introduced the task and gave instructions on punctuation; (2) the children read in turn while the teacher jotted notes and gave feedback at the end of each reading; (3) the teacher summarised the narrative after the first child read, asking comprehension questions about the story content; (4) the children reading in turn resumed; (5) the students read the story aloud in turns for the second time, while the teacher completed running records on their reading; and (6) the teacher wrapped up the session, reinforcing what they had learned and encouraging them to do more reading. Each child read for around four minutes, finishing a page of text, twice.

The guiding of the reading

In the Introduction the teacher distributed the small readers and quickly introduced the book by having children read the title. She moved on to explain the punctuation they would encounter in reading the story, evidently revising what children already knew. She had cards she made herself illustrating each punctuation mark. Her explanation was confined to indicating the reading behaviour (intonation and pausing) symbolised by the punctuation mark and demonstrating this behaviour for different marks.

She then returned to introducing the story: looking at the cover, asking the children what they saw on it — the bee (main character) and the bird and eggs in the nest (topic)—and had them read out the name of the writer and illustrator, explaining the word, 'illustrator'.

The teacher used this moment to build knowledge from experience children had in relation to the story. She asked the children if they knew what bees made and what the bird was doing in the nest; she then used their answers – 'sleeping' – to develop the concept of a nest, making an analogy between a nest and a home.

Before the reading of each child commenced, she asked the children to study the picture, asking them questions about the characters on the page and the setting — where the action was taking place. Each child read a page. This teacher also used the page (as was intended by the story lay-out) as a frame for the comprehension of the event on that page.

While the child read, the teacher jotted down things needing correction for feedback that she gave straight after the reading. All but one of the children read by sounding out each word, mono-tonally drumming it out, syllable by syllable, often running over full stops. In the feedback the teacher modelled the correct prosody, word or letter, and asked the child to repeat it after her. When the child did that, she said, *Pinter!* (clever), and moved on. She did not comment on students' self-corrections as they read, which were frequent.

Over the course of all the reading the teacher addressed three kinds of problems. The most frequent was punctuation – question intonation missing, comma pauses not observed. A punctuation mistake that stood out was in a sentence read this way: *Hutan tempat tinggal Lili adalah rumah bagi. banyak binatang Lili tidak pernah merasa kesepian tinggal di hutan itu (*The forest where Lili lived was a home for. Many animals Lili never felt alone in that forest). But the teacher did not correct sentence run-ons past full stops.

The second correction related to issues of blending and misreading letters. There was not much opportunity for instruction at the letter and decoding level because children made few mistakes here. However, one child showed repeated problems in sounding out long words and with particular letters. The teacher modelled the multi-syllabled word but did not pick up on his pattern of the missing 'n' sound in words he read.

The third correction related to a child relying on memory to answer a question about what she had read, rather than looking for the word in the text. This was the only occasion when the teacher referred a child to the text to check the meaning by reading the words.

The teacher asked questions before or after the child read the page and also in the closing section of the lesson. The questions served a variety of purposes. One purpose was to monitor children's understanding of the story by asking a question to see if they could make straightforward inferences to understand the text. One example was: *What is the bee's name?* when the bee was introduced on a page for the first time by her name, rather than as Bee.

Another kind of question supported reasoning skills. *What would happen if the bird did not have a nest to put her eggs in*? Questions also enlarged children's understanding of their own experience by connecting it with the story. *Have you ever seen a nest? Where do you find a nest*? *What happened when you found a nest?*

These questions led to a spontaneous 'teaching moment' that the teacher could not have anticipated, revealing her ability to use the reading lesson for the "widening horizons" function of literacy in its introduction of children to issues of empathy, perspective taking and preparation for social life. After discussion on why birds put eggs in nests the teacher continued:

Teacher: *What do you do when you find a nest?*

Child: *You take it. You take the nest.*

Teacher : *Oh! The poor bird, it wouldn't have a home. If the nest is taken, it wouldn't have a home any more. Where could it put its eggs? Poor thing. So, let the nest be, in the tree. So the bird can have eggs. So they can hatch. He, he. Clever!*[35]

Other questions were to build knowledge or concepts. *What kind of tree is that*? *Is it a river or is it a pool?* – looking at the page where the fish and the frog had laid their eggs – and quickly shaping a long river and a round pool with her hands. The main knowledge-building questions came at the end of the story when suddenly the pattern of animals who lay eggs was broken as Bee came across a dog and its puppies. That led to the teacher summarising the basic classification that had been encountered between animals that lay eggs and those that have live offspring.

The lesson concluded with the teacher 'high-fiving' the children, seeking confirmation that they had enjoyed reading and exhorting them to take books and read them as that way they would gain more knowledge.

Teacher's reflection: Lesson A

This teacher's reflection lent itself most to addressing the question of what literacy and reading meant to her. These beliefs are inferred from her answers to the questions about what struck her the most about the lesson, what she thought about the students ability to learn and the insight she gained from the lesson.

In summary she replied that what stood out most for her about the lesson was that: 'Even though we went through the reading again, the students made the same mistakes. This is because they did not pay enough attention to the writing but relied on remembering what the teacher had stressed.'[36]

This answer is important in understanding her beliefs about teaching reading because it shows the value for her of getting the meaning from the written word.

Two other beliefs about literacy and teaching reading emerged in her responses. The first one was her belief that reading is the gateway to knowledge. The insight she gained from the experience was that 'reading is the way to increase children's knowledge and that's why they have to do a lot of it.' She believes the best way of getting that knowledge is by reading a lot.

The third belief she indicated she held about literacy and reading was 'the enthusiasm with which the children read, making them make quick progress'. For her the most rewarding moment of the lesson was discovering that one of the students had read on to the end of the book during the lesson and so knew the surprise on the last page. She was also struck by the excitement with which students answered her questions: the engagement that the contents of a book can produce.

The key value she held of reading the text with understanding was reinforced by her answer to the question of what she thought of her students' capacity to learn to read. She thought their

[35] In Bahasa Indonesia: *Kasihan Burungnya, tidak punya rumah. Kalau sarangnya diambil burungnya tidak punya rumah. Kalau meletakkan telurnya dimana? Kasihan. Jadi biarkan saja sarangnya tetap di atas pohon. Biar bertelur. Biar menetas. He eh. Pinter.*

[36] In Bahasa Indonesia: *Tetapi menginat apa ditetapkan guru.*

critical progress was that ‘they are understanding what they read’. This she amplified in her evidence for that view:

‘At the start the children answered the questions based on the pictures. But after they had read the text they answered based on the actual words in the text.’

Lesson B

About the lesson

The book selected for the guided reading was *Tersesat* (Lost), a graded reader at level B-2 with one or two simple sentences per page for most of the book. The story was of a child, Sita, lost in the forest, and her mounting distress, until her mother woke her from her dream. Although the text was simple and repetitious (where is …my mother?, where is my father?), the pictures in the book presented conceptual challenges by trying to convey the feelings of the little girl through images of home inset into the forest in which she was lost.

This lesson involved five segments: (1) the teacher introduced the activity, including some teaching on punctuation, and read out the rules for the session; (2) the teacher read the story to the class; (3) the children being asked to predict the story from the cover; (3) the children reading in turn, accompanied by interventions from the teacher; and (4) the childing doing a task of placing the right punctuation marks for sentences from the story. The completion of the Running Records took place during the reading of each child.

Guiding the reading

The teacher introduced the lesson by explaining that it was a guided reading lesson to help them learn to read fluently. The lesson began with going through class rules: don’t interrupt each other while we learn; keep the noise down; don’t going in and out of the classroom. Books on the table. When I give you the books, do not pick them up.

Next the teacher explained two of the points of punctuation they would encounter in the story, exemplifying on a board sentences that were a question and a statement, with the appropriate punctuation mark.

He then read the story to the class, telling them to listen well. The readers were not distributed. He read from a small book and the students were not able to see the pictures or the text; the intent was for them to listen.

His carefully articulated reading of the story was followed by an activity to predict the story from the cover, a picture with all the animals encountered in the forest and the girl in the middle, looking downcast. His question: *What do you see on the page?* elicited a list of items, without the class ‘seeing’ the message of the picture itself. When the teacher asked them to guess what the story was about, one child said immediately: ‘*Tersesat*’, the title of the story, seeming to indicate that the book was already known.

The guided reading proper started 12 minutes into the lesson. The reading turns had the same pattern. Each one lasted between 2 and 5 minutes and during that time each child read on average two sentences, while the teacher completed the running record.

After each child had read, the teacher directed the student’s attention to the facing picture and asked questions in quick succession about what each student had read (*What was Sita doing? What did Sita ask? Who did she ask?*) and repeated expressively what the child had just read.

His main concern appeared to be that students follow the developing story: what it feels like to be lost.

On one occasion the teacher responded to the text a child had read when a word had been wrongly guessed at. The child read *kecil* (small) instead of *kelinci* (rabbit). (Another child had previously confused these two words in the story). The teacher interpreted the problem as a decoding mistake. He decoded *kelinci* with the class, pronouncing first the letter name and then the letter sound, for each letter. Decoding *kelinci* this way took one minute and thirty seconds.

The teacher depended on questions mostly directed at the pictures for building comprehension of the story and mostly *What's happening*? questions or anticipating what was to follow. Some of these showed he had difficulty in articulating questions that students could understand. An example of an interaction was: *Sita lagi apa di situ? Sita lagi apa di situ? Sita lagi apa di situ? Ada apa Sita di situ? Sita ada buat apa di situ*? (all variants of the question: *What's Sita doing there?*). This difficulty was compounded by the challenge of pictures that required inferential understanding. Questions such as *How many people do you see in the picture*? *Who are those two people?* did not help children get close to the meaning of the insert of an image of Sita's parents in the picture of the forest.

An incident of unclear significance occurred when toward the end of the story he asked a question to relate getting lost to the children's own experience:

Teacher: *Have anyone of you ever got lost on the road?*
Students: *Yes*
Teacher: *Who got lost? Experienced the same thing as Sita? Yes? You have?*
Who has been lost? Kornelis, have you ever been lost?
What does lost mean? Lost. Lost. What does it mean?
What does lost mean? Who can tell me? Got lost. What does it mean?
It means taking the wrong way.

It is not clear what produced students' reticence on the meaning of lost when initially they all acknowledged having had the experience. Was it concern at having to answer or did they in fact not know the meaning of the key word in the story?

The final segment of the lesson returned to the punctuation target of the lesson, with materials the teacher had prepared to see if the children could allocate the right punctuation mark to a selection of sentences from the text. Then each child read the sentence with the appropriate intonation.

The lesson concluded with the teacher asking the children collectively if they had any difficulties with any of the words and they said no, so he brought the lesson to an end.

The teacher's reflection

This teacher's reflection mostly revealed aspects of mindset about his teaching and his students' learning.

His main reflection was that he had been far too dominant in the lesson. He described his constant telling of the story as taking away the children's opportunity to read – to read the text independently and carefully. In particular he spent too much time on the pictures in his attempt to 'translate' the meaning of the story. In reading the story at the beginning he thought he had done enough to give them the flow of the story. By asking all those questions he stole the time from their reading. These are perceptive observations, showing his grasp, even if retrospective, of an efficient flow of teaching that facilitates learning.

Several parts of his reflection related to students' learning. His first realisation from the experience was that even in a guided reading group there are children of different ability levels and he had not sufficiently catered for those who couldn't do what the lesson aimed at, which was to read whole sentences fluently. He felt he had left them behind and a lesson he took away was the importance of planning for such students learning in future lesson plans. This may refer to one child who had greater difficulty than most in answering the questions.

For him also the most interesting event in the lesson occurred with the punctuation task and this was seeing students helping others who didn't understand so well, to complete the task – and this in spite of the teacher's own instructions that they were not to interfere with others' learning. He felt proud that in this situation there was a student who was not prepared to leave the others behind. This may have been an oblique way of compensating for his reprimand during the lesson of the same child for helping the other child during the reading, as breaking the *not interfering with others' learning* rule. If so, it gives an insight into what opening up teachers' awareness of problems in their teaching involves, especially where the teaching culture gives them an unassailable authority over their students.

Lesson C

About the lesson

This lesson used the book, *Saat Saya Sakit (When I'm sick)*, a graded reader also at level B-2 with on average two sentences per page.

The lesson was clearly structured and activity was varied over eight segments. It was similar to Lesson B in its initial focus on the teacher reading the story but differed from both the others in continuing the teacher's modelling of the reading through most of the children's practice. The only occasion on which children read by themselves was the final revision – a segment with total duration of 4 minutes – during which the teacher recorded their performance in the running records. Another feature unique to this lesson was the postponement of the questioning for comprehension skills development until the final, longest, segment of the lesson where it continued for eight minutes.

Guiding the reading

The purpose of the lesson was clearly explained in terms that the children would most likely understand: to practise their reading and to find out how much they understand what they read.

As with the other lessons, the guidance started with revising punctuation – only the full stop and the comma were used in the story. But there was no further reference to or monitoring of punctuation in the lesson. The single-clause brevity of the sentences did not give much opportunity for intonation and appropriate pausing to be observed.

This teacher was the only one who also preceded the reading activity by explaining vocabulary items that occurred in the story: *istirahat, menemani, sakit (take a break, be together with*, be sick*)*. However for two of these this was done with dictionary definitions (for the first two: *take a break; be together with*). This was a missed opportunity for opening up the meaning of *menemani,* which has the word for 'friend' in Bahasa Indonesia as its stem. He got children to help define '*sakit'* (sick) by asking them what it felt like when they were sick and aggregated the symptoms they volunteered (headache, tooth ache, feeling feverish) by defining sickness as suffering from something wrong with the body.

Before beginning the reading the teacher also had the children predict what the story would be about from the cover, with a simple question (invited by the explicit message of a child in bed on the cover picture): *What's happening to the child in the picture?* that the children were able to answer. He asked them to suggest an appropriate title for the story and they replied with the exact title, indicating that they already knew the book.

The reading sessions themselves took this form: the teacher read through the story while the children listened, as with Lesson B, without text or pictures. Then teacher and students read through the story together. After that, individual turn taking occurred but the teacher read the first sentence of each page and the child the second (the second sentence on the page introducing the new word or concept). Finally the children read in turn a couple of sentences by themselves.

Most of the children pronounced words successfully. Interactionally, the teacher always ordered the child to repeat the sentence – *Ulang! (Stop!)* – even if there had only been a stumble that was self- corrected.

Though there was little opportunity for word correction, on three occasions in the lesson the teacher embarked on decoding instruction. The first was spelling out the title, by inviting attention at the word level because of the same initial letter in the three title words. Accompanying the children decode, the teacher used sounding out techniques but the sounds did not blend easily into the words.

The second case was with *menemani* (be together with) where each child was required to say the word, by splitting it up into syllables, albeit a little blurred. However the teacher did not persevere with a child who could not get past the first syllable. The third was when a child read *kamar* (room) instead of *rumah* (house). It was an intelligible mistake that meant the child was focused on meaning because all the pictures in the story were of the bedroom. The teacher treated it as a decoding mistake and embarked on an elaborate and unnecessary decoding of *rumah*. He mixed spelling and sounding techniques to decode the word that only emerged at the end of it when he resorted to syllabising it: *ru-mah*.

The teacher focused on comprehension at the level of the whole text and his questioning was clear, structured and served a variety of purposes. In the long sequence at the end of the lesson, he made a life-skills theme out of '*sakit*', trying to build up a picture from children's experiences of what made them sick and behaviour to avoid and to adopt. However the content limitations of this very early years reader and the children' inarticulateness, limited the extent to which he could extend children's conceptual exploration of illness. Again using children's knowledge he used inferential questions about setting and the character's feelings. He concluded the lesson by using the five "wh" questions – *who, what, why, when* and *where* – to structure children's recall of what the story had been about. The children answered all of them except the "who" question. He prompted them. '*Saya'* (me) they answered. Both children and teacher seemed aware that this who needed some explanation because the book was not about them! But the teacher accepted their laugh of puzzlement and did not embark on explaining how and by whom stories can be written.

The teacher's reflection

This teacher's reflection contained three main ideas. On the overall insight he gained from the lesson, he had been mainly impressed by the way guided reading gives a good idea of the extent that a child can read, because 'ordinarily we don't know from the class who can and who can't read'.

The decoding instruction emerged as a key experience. '*Membunyikan* (sounding out) the letters to make a word: it's a challenge for me and for the children at grade three. They don't accept it. They are used to *spelling* out a word.'

On students' learning he thought, fairly, that the students were quite competent readers: 'They can read and they can understand what they read.' This view does not explain why he spent so much time scaffolding the two sentences they read. He added that they just need to be motivated. It is not clear what aspect of the lesson gave rise to that reflection but it may have been the increased animation that the students displayed during the relaxed questioning at the end of the lesson.

Discussion

This account of the discussion is a synthesis of the main themes that emerged over the workshop in the interpretation of the three lessons, organised to fit the three analytical constructs.

Mindset: teachers' beliefs about reading and teaching reading

The difference between Lesson A and the two other lessons was marked for all participants in the discussion. The INOVASI education team approached this difference from a technical perspective since their role was to offer technical support to teachers through the pilots. These specifically technical matters are discussed under the third section of this analysis. However the *key* difference that the team noted is relevant to the issue of how the teachers understood what it is to read and what approach they took to teaching it.

They contrasted Lesson A with the others as a case of the latter two having misunderstood the technique of guided reading by having first read the story to the students – and as a listening exercise, unaccompanied by tracking the written words. By contrast, the Lesson A teacher went straight into reading turns after the lesson introduction. The team pointed out that guided reading is to be able see how the children are progressing and where they are having problems in their reading. This requires children to approach the text without prior familiarisation with it. The observations triggered the idea that in Lesson B and Lesson C the teachers were setting the children up for memorising the text. This seemed particularly true of Lesson C where children read by themselves only after three teacher-led readings of the text. Also the book seemed already known to the children in both Lesson B and C whereas it was not known in Lesson A. Also with only one or two sentences per page, and often repeated sentence stems, memorising was possible.

The discussion then went into other indications of the teachers' attitude to the reading of text. Particularly in Lesson B they observed that the teacher relied on his own re-tellings to scaffold the children's understanding of the story and on the *pictures* to question students on what they had read. His comprehension strategies were aural and picture based, as for a pre-literate class. He seemed to prioritise comprehension of the story over comprehension of written text.

At that point the team recognised a fundamental similarity in all the teachers' response to written text, including in Lesson A, although less marked there. This similarity was their not checking at the end of the turn, whether the students had understood the sentences they had read. Even the lesson A teacher was critiqued by team members for getting a child to repeat the correct reading of the sentence after she had modelled it, rather than getting the child to *read* it again from the text. None of the teachers asked a student to re-tell what they had read. Across the

three lessons all but one of the teacher interventions in the reading was either related to expression or word errors treated as issues of decoding.

Nevertheless, that one intervention was telling. It was the incident reported in the description of Lesson A, where the teacher told the student who had answered an information retrieval question wrongly, to look at the word that was written on the page. That teacher's reflection had also focused on her (optimistic) pleasure at finding that 'at the start the children answered the questions based on the pictures. But after they had read the text they answered based on the actual words in the text.'

The team also recognised that practices scaffolding students' making meaning of the words, such as discussing pictures, questioning and summarising, was effective teaching and that the Lesson A teacher had done this as had the other teachers.

Some of the team members argued that what the teachers in Lessons B and C were doing in reading first, was effectively adapting the guided reading procedure *to context* – not misunderstanding the guided reading strategy. A context in which children's reading, even by grade three, was not strong enough to be able to retrieve the sentence meaning from the written words, required additional scaffolding from the teacher.

Which of the two interpretations of Lesson B and C was plausible became a central debate in the workshop. The debate turned on the question of what these teachers meant by reading. In this debate the team referred to the two goals of teaching reading: *membaca lancar* (reading fluency) and *membaca pemahaman* (reading comprehension) – as if they were distinct. That all teachers were engaged in supporting reading fluency in the reading turns makes sense of their concern with intonation and with treating all mistaken words as decoding problems – even when they were issues of what words fitted and didn't fit, in terms of sense.

If this is an accurate interpretation of teachers' construction of reading in Lesson B and C, it remains unclear what their strategies are for teaching comprehension at the sentence level, as distinct from teaching comprehension at the level of the whole text; or whether they recognise that teaching reading at the level of the word and sentence requires more than decoding strategies. The inclusion of punctuation in the lesson was a comprehension strategy, a way of identifying the unit of sense in a sentence by copying the pattern in speech. But only the Lesson A teacher used intonation instruction in this way. Key also to sentence-level comprehension is understanding words and concepts but in Lesson B and C there was no meaningful exploration of vocabulary in the context of the sentence, or in Lesson B, explanation of concepts that made sense in the larger context of the story.

Lesson A was the exception. Her excitement in seeing the children progress towards mastery of reading came from the thought of building their knowledge this way. Her lesson was an attempt to exemplify how reading can build knowledge.

Another unknown from studying Lesson B and C was whether the children *could* decipher text. There was certainly a clear struggler in each group but over both, there were nearly no mistakes in the decoding. And most in each group read the sentence fluently without breaking down words into syllables. The teacher of Lesson C was confident that his children could read – with understanding. That raises another important question: if these children had reading fluency, why was the level of reader and the teaching strategy so limiting of their opportunity to learn? If they were not competent decoders why did the teacher pre-empt this becoming apparent, when the point of guided reading is to enable diagnosis of the problems?

These reflections raise important issues for understanding both the strengths and the blockages to understanding what it takes to teach children to read in these contexts. They raise the question of whether — and why — comprehension is being well achieved as *listening* comprehension in some places— much better than others, where reading comprehension is higher.

Most importantly the question arises of whether there is a mistaken presumption about what it takes to read on the part of teacher participants and that INOVASI has not focused on in its approach so far. These teachers seemed to share the assumption that retrieving meaning in reading continuous text follows from being able to decode. Yet understanding the semantics of sentences requires developing a distinct set of skills, particularly in non-literate cultures, and in classroom settings where *oral* sentences aren't even formed by children.

The neglect of the text by two of the teachers raises another serious question related to mindset: their attitude to students' progress and attainment — the aspect of mindset associated with teachers' expectation of students. Why are fluent grade three readers in some parts of Indonesia reading two sentences of a grade one reader when in other places they have rich text and extended opportunity to practise?

Mindset: teacher–student interactions

One of the main differences between Lesson A and the others is the tenor of the interactions between teacher and students. All the teachers were authoritative but there seemed to be a difference in the source of the authority. In Lesson A it seemed to derive from the intensity of the teacher's investment in students' learning from the lesson. The students were rivetted. In the case of the other two teachers it seemed more to derive from their entitlement as a teacher.

We need to proceed cautiously in interpreting behaviour as norm-based when it may well derive from the external stresses of being videoed. Nevertheless, differences in teacher authoritativeness had a marked effect on students opportunities to learn in these lessons. In addition, if it is *practices* that produce these effects, these cannot be wholly attributed to the intrusion of the camera.

One of the bases for the interaction of Lesson A that the workshop team remarked was a kind of automatic respect shown to the students. This was manifested in the way the teacher did not interrupt their readings with corrections but waited until the turn was finished to go methodically through them, having kept track by her jottings. When she had done this and modelled the right way and had the student repeat her model, she would conclude by saying: '*Pinter!*' (Clever!). Often this practice of uncalled-for praise, is criticised as the wrong mindset (effort should be praised, not cleverness) but this teacher seemed consistently to use it to build up students after focusing on their mistake.

These practices seem to be connected with the extent to which students were such active participants in the lesson. An explanation of the difference between their eager responsiveness and the monosyllabic answers of students in the other classes is that they were not afraid of making mistakes ('*takut salah*'), as one mentor put it. They had not been commanded to stop or repeat, or interrogated for their comprehension in their interactions with this teacher.

The INOVASI team who worked in these contexts put a perspective on the interpretation of these practices. This was to say in some contexts children are acculturated to such peremptory styles in school as well as in the wider local culture – and rarely articulate more than one-word answers to adults. That is a larger problem of mindset than can be resolved by supporting teachers to reflect on whether they hold a growth mindset view of student learning.

Children's expressiveness is critical to reading literacy and higher-order thinking skills. In another important respect, none of the three teachers afforded children the opportunity to talk about their experiences or to re-tell the story in their own words or to respond to it with their own thoughts. The teachers took all the initiatives and the teacher's language dominated the literacy classroom. No teacher included children's self-corrections in their running records, although these are good evidence of decoding comprehension skills and a route to independent reading. All teachers valued students' comprehension but by and large they were not enabling children to comprehend independently. None of them recognised how teacher-dependent the children were. This is the opposite of student centredness, the underpinning orientation of effective teaching of reading literacy.

Pedagogy for literacy: the pilots' fit with context

The previous section has already outlined INOVASI's fit with context that arises from issues related to the teaching culture and teachers' mindsets.

The lesson descriptions showed a range of difficulties that the case study teachers had with specific practices that they learned through the Literacy 1 and 2 pilots. As practitioners, the education team observed a number of problems in the way these teachers implemented what they had learned. Such aspects are about understanding specific technical issues before the model can be said to fit, for example: recognising that guided reading and running records have two different purposes and teachers should not try to do them at once; avoiding too many instructional foci during a guided reading session; not reading the story to the children first; choosing a reader that challenges children as well as well as fits their level.

But there are a several larger technical issues about fit with context that imply INOVASI needs to review some aspects of its literacy model.

The first among these is INOVASI's approach to aspects of the beginning skills of reading: word construction and decoding. The 'science of reading' approach discussed in the literature review and adopted by INOVASI, stresses phonemic awareness and sound and letter matching in decoding. As demonstrated in two of these case studies, teachers find this approach difficult to master. This is especially the case for those teachers who come from a background where words are approached by spelling out the names of the letters. Sounding out as well as naming the letters ends up being combined in their 'word attack' — which, as we saw, led to nearly one and half minutes to decode *kelinci* and losing the word in the process. According to the INOVASI education team, doubts about the value of sound and letter matching are often articulated in INOVASI training. However, widespread in Indonesia and in INOVASI's provinces is an established method for word construction and decoding that is syllable based — the *suku kata* (syllables) approach. Children seem to make rapid progress with this. In Bahasa Indonesia it is easy to learn consonants paired with a vowel. Many Indonesian words in children's lexicon have only two syllables, so being able to sound out syllables quickly, leads to quick word acquisition. INOVASI needs to research the relative suitabilities of the syllabic and the phonological approaches for beginning reading in Indonesia.

Three other related technical issues with wider implications emerged from these case studies. One is that teachers' difficulty in framing questions has been underestimated in teacher development so far. This relates not only questions to elicit higher-order thinking but also questions to help students retrieve the meaning and the information from what they have just read. The value of a systematic approach to this was shown in Lesson C where the teacher efficiently led students through a summary of the story using the structured "wh" questions. Much more than this is required for questions eliciting inferential understanding, as can be seen

in the Lesson B teacher's struggles with an inferential text. INOVASI's balanced literacy approach that places reading authentic texts at the centre of learning to read also requires teachers to have questioning skills if the approach is to help develop children's comprehension skills.

Related to this is the issue already touched on in connection with teachers' understanding of what it is to read. This is the missing middle between being able to decode and being able to comprehend at the level of sentences and short continuous text. Teacher development for literacy needs to include understanding the distinctiveness of the wording and the structures of literate language. Practices which support this understanding – children's retelling in their own words, children's writing sentences from the story – are also ways of putting students at the centre of learning.

Finally these issues all have implications for teachers' own literacy proficiency. Do some teachers have problems in formulating questions to retrieve information or surface an implicit meaning; in recognising the connectives between words in a sentence or in continuous text, due to their own levels of literacy proficiency? In the following chapter on what works, a strong correlation between students' achievement in reading comprehension and higher-order thinking skills emerges with teachers' own literacy performance. As noted, unless teachers themselves are sufficiently competent in literacy, teaching practices that pre-suppose such competency are not likely to be meaningfully sustained. Should INOVASI be thinking of supporting *teachers'* literacy development to realise the full benefit of the model?

INOVASI's approach to improving literacy teaching has had two macro goals. One was teachers' competence in pedagogies that put students at the centre of learning through diagnostic assessment and teaching at the right level. The other was enabling teachers to help students progress beyond beginning skills to comprehension. This was through the balanced literacy orientation to widening horizons and raising levels of thinking, envisaged in Indonesia's national goals for literacy achievement.

These case studies have shown a continuum of teacher capabilities on these goals. The teacher of Lesson A, with her integrated command of the INOVASI model, shows that the project is feasible. Deeper understanding of the underlying meaning of reading and the pedagogies that support it is the remaining challenge for all teacher participants in the pilots.

9 Findings: What worked for improving teaching and learning?

This chapter is presented in two parts. Part one covers the interventions associated with teacher development to improve literacy. Part 2 covers the INOVASI interventions to support children's access to reading.

Part one: The effectiveness of teacher development pilots for improving teaching and learning

Part one responds to the key evaluation questions 1, 3 and 4. These questions take up different aspects of INOVASI's interventions in teacher development for literacy teaching. The first question – *To what extent does training teachers to teach reading result in children's improved reading outcomes?* – is addressed by looking for associations between students' endline results on the student learning assessment and the teachers' literacy proficiency, pedagogical skills for teaching literacy, classroom practice skills, and the mindset they have acquired through the pilots.

The Literacy 1 pilot is again the default pilot in exploring effect, as it engaged teachers in a full (though introductory) understanding of teaching literacy and it reached the largest number of teachers in the program. The first step in this inquiry is to examine the impact of Literacy I on students' results.

We then compare the effectiveness of variants on teacher development on students' literacy scores, as outlined in the analytical pathways section in chapter 5. There are three comparative inquiries. The first compares students' gains when their teachers have completed both the Literacy 1 and 2 pilots with their gains when their teachers have completed only Literacy 1.

The next comparison is between the various pilots that adapted teacher development to contextual priorities, including in comparison to Literacy 1. This inquiry includes the issue of language transition and covers key evaluation question 3: *To what extent does training teachers in mother tongue transition improve children's reading outcomes?* These comparisons involve not just INOVASI-supported pilots but also grantee pilots designed to support these specific contextual needs.

For the pilots in each of these inquiries, see annex 1: literacy pilots by analytical category.

Key evaluation question 4 is also addressed in most of these comparisons: *Is there any evidence that improved literacy resulting from the pilots will lead to better learning outcomes at higher levels/ across curriculum? Or better higher-order thinking skills?*

The study mainly draws on SIPPI data from the student learning assessment, the teachers' reading literacy test, classroom observation instrument and the student questionnaire (on reading). The SIPPI database makes correlational testing on learning and other outcomes possible.

Evidence of the effect of teacher variables on student learning outcomes in the Literacy 1 pilot

Associations were explored between students' literacy outcomes and the three domains of teachers' capabilities, constructed out of the SIPPI variables: teachers' own reading literacy scores, classroom practice and teachers' mindset. While the SIPPI data is not specifically on literacy teaching practice, particular classroom practice skills in the index are relevant to effective literacy teaching.

The term ‘association’ means that a change in the independent variable, for example, an increase in a particular teacher practice, will increase the probability of an increase in the dependent variable, the student scores. Regression also enables all other observed factors that determine the students’ scores (socioeconomic status, gender, teacher and school level characteristics) to be controlled for in identifying the strength of association with the variables of interest. The large variation in INOVASI school contexts made random effect regression the appropriate analytical technique as it minimises bias in the data produced by between-and-within school variance (Clark *et al*., 2010).

As well as exploring the predictive power of the variables on change in the student scores, the size of the effect was also calculated, using partial eta squared to measure effect size (Richardson, 2011). A rule of thumb for the strength of both the regression coefficient and the effect size used is set out in box 8.

Box 8: Rule of thumb estimates for the significance of effect sizes

	Small	Medium	Large
r (correlation coeff.)	0.1 - 0.3	0.3 - 0.5	0.5 or larger
Partial eta squared	0.01	0.06	0.14

Source: Draper (2002)

The following analysis focuses on grade two. As explained elsewhere, grade two is the first year where the school makes a significant difference to learning outcomes across all the provinces and grade two or three is the level that early grade reading literacy is usually assessed internationally— including for the Sustainable Development Goal 4.1.1 on education. Selecting this grade provides comparability with other assessments.

The results of the regression analysis are summarised in the following two tables. Table 23 sets out correlations and effect sizes on the component skills of reading and Table 24 sets out correlations and effect sizes on the skills of comprehension. Student, teacher and classroom characteristics have been included as control variables and also presented in the table to show the relative predictive power and effect size of the teacher attribute variables. Figures without brackets are random effect regression coefficients and bracketed values are the effect sizes.

Table 23 presents all the variables in the SIPPI database that correlate positively or negatively with students’ learning outcomes in the component skills of reading, with low to high levels of probability.

Table 23: Literacy 1 pilot: correlations with grade two endline student outcomes on the component skills of reading in the SIPPI student learning assessment.

	(1)	(2)	(3)
VARIABLES	Letter recognition	Syllable recognition	Word recognition
Teacher reading literacy score	0.004	0.007	0.008
	(0.002)	(0.002)	(0.003)
Classroom practice index	0.061*	0.055*	0.080**
	(0.015)	(0.013)	(0.014)
Teacher's mindset index	0.008	-0.004	0.021
	(0.000)	(0.000)	(0.000)
Raven score	0.170***	0.194***	0.214***
	(0.023)	(0.035)	(0.042)
Socioeconomic status index	0.200***	0.255***	0.311***
	(0.033)	(0.058)	(0.089)
Gender of student (female=1)	0.235***	0.240***	0.289***
	(0.016)	(0.019)	(0.028)
Reading corner with non-textbooks	0.365***	0.294***	0.220***
	(0.024)	(0.018)	(0.009)
Civil servant status teacher	-0.059	-0.004	-0.064
	(0.001)	(0.000)	(0.002)
Certified	0.031	0.020	0.088
	(0.000)	(0.000)	(0.002)
Textbook availability in class	0.217***	0.212***	0.112
	(0.011)	(0.014)	(0.008)
Availability of books at home	0.175***	0.124***	0.133***
	(0.011)	(0.007)	(0.009)
Observations	1,903	1,903	1,903
R-squared	0.233	0.279	0.333

Note: Low to high levels of probability: *** = p<0.01; ** = p<0.05; * = p<0.1

Overall, the effect of the teacher variables on student scores for the component skills of reading are not evidenced in the regression analysis, apart from a weak correlation, with a small effect size, of the classroom practice index with student outcomes. Though the effect size is small, this is evidence that this aspect of the teacher professional development strategy to improve student outcomes has worked.

The control variables of socioeconomic status, gender, availability of books in the home (likely to be implicated in socioeconomic status) and student ability (Raven score) are correlated with student learning assessment scores at the highest level of probability, although the effects are also small, apart from a medium effect size of socioeconomic status on word recognition. The

variable with the strongest correlation is reading corners with books that are likely to engage children, though curiously, the lowest effect size is on word recognition. Textbook availability is also correlated but not quite so strongly and not at all with word recognition.

Table 24 presents all variables on the SIPPI database correlated with grade two student outcomes on the comprehension elements of the SIPPI student learning assessment.

Table 24: Literacy 1 pilot: correlations with grade two endline student outcomes on comprehension skills from the student learning assessment test

	(4)	(5)	(6)	(7)	(8)
VARIABLES	Reading comprehension	Listening comprehension	Direct recall	Inferencing	Interpreting text
Teacher literacy score	**0.152*** **	0.044	**0.115** **	**0.128*** **	**0.233*** **
	(0.029)	(0.001)	(0.018)	(0.016)	(0.059)
Classroom practice index	-0.006	**0.101** **	0.048	0.042	-0.033
	(0.000)	(0.011)	(0.005)	(0.002)	(0.000)
Teacher's mindset index	0.000	0.049	0.048	-0.006	-0.012
	(0.000)	(0.003)	(0.005)	(0.000)	(0.000)
Raven score	**0.192*** **	**0.079** **	**0.138*** **	**0.132*** **	**0.204*** **
	(0.038)	(0.006)	(0.022)	(0.016)	(0.040)
Socioeconomic status index	**0.146*** **	0.034	**0.073** *	**0.174*** **	**0.236*** **
	(0.016)	(0.001)	(0.004)	(0.020)	(0.038)
Gender of student (female=1)	0.081	0.097*	**0.120** **	0.080	**0.179*** **
	(0.002)	(0.003)	(0.005)	(0.002)	(0.010)
Reading corner with non-textbooks	**0.248** **	**0.235*** **	**0.286*** **	**0.188** *	0.119
	(0.017)	(0.014)	(0.024)	(0.009)	(0.003)
Civil servant status	**-0.211** **	-0.129	**-0.176** **	-0.134	**-0.279*** **
	(0.014)	(0.004)	(0.009)	(0.005)	(0.021)
Certified	**0.278*** **	**0.174** **	**0.218** **	**0.231** **	**0.280*** **
	(0.020)	(0.007)	(0.012)	(0.012)	(0.016)
Textbook availability in class	0.089	0.104	0.104	0.035	0.087
	(0.003)	(0.003)	(0.004)	(0.000)	(0.003)
Availability of books at home	**0.119** *	**0.066**	**0.159*** **	0.036	-0.018
	(0.004)	(0.001)	(0.008)	(0.000)	(0.000)

Observations	1,098	1,098	1,098	1,098	1,098
R-squared	0.189	0.084	0.173	0.125	0.223

Note: Low to high levels of probablity: *** = p<0.01; ** = p<0.05; * = p<0.1

The picture is different with comprehension skills. Teachers' own reading literacy is highly significant for all aspects of comprehension except listening, with the strongest effect (though still not large) on interpreting text, the highest order skill. This is logical as to assist students in developing inference and interpretational skills, teachers would need to be able to do it themselves. Teacher's literacy score and this higher-order thinking skill have one of the highest correlations of all variables with student learning outcomes.

The absence of SIPPI variables on literacy teaching is felt in this part of the analysis. There is no opportunity to look at effect that could be directly ascribed to INOVASI's main literacy pilot. Teachers' classroom practice is too general to impact on comprehension and only affects listening comprehension. Mindset has no significant association — only a slight negative result!

The variable reading corner with non-textbooks shows a highly significant correlation with reading comprehension, as it did for the component skills of reading. Student background variables have a high association but, except for socioeconomic status and students' ability on the higher-order thinking skills, with mainly small effects. Teacher background variables have not been part of this study but correlations here with student learning outcomes are striking: negative for civil servant status but positive for teacher qualification or certification.

Though the gains attributable to INOVASI are mainly small, so are all significantly correlated effects and it is an achievement that the literacy related effects are stronger than student background effects, that usually eclipse educational effects.

The regression results for the Literacy 1 + 2 panel for the component skills of reading show no appreciable difference from those of Literacy 1 in all variables. However for the comprehension skills, teachers' literacy scores show less impact than for Literacy 1, weakly influencing inferencing and interpreting text. Curiously, the strong correlation of non-textbook reading corners with student scores, has disappeared.

Comparison of teacher development contextual adaptations

In answering the question of what worked? it is possible to find out "what worked best?" by comparing the variations on approaches to teacher development in INOVASI that responded to different contextual priorities.

There are three main variants. First is the *Guru BAIK* pilot in Southwest Sumba, East Nusa Tenggara (16 schools). The second is an INOVASI-supported teacher development in multi-grade methodologies in Probolinggo, East Java (29 schools). The third variant is the language transition pilots. There were three of these. One, the pilot in Bima, West Nusa Tenggara (56 schools) was an INOVASI-supported pilot that added language transition methodologies to the delivery of Literacy 1. The second was a grantee pilot in East Sumba (14 schools), designed and run by Sulinama, an Indonesian foundation specialising in the use of local language in education. A third language transition pilot was run in the Kodi sub-district of Southwest Sumba in East Nusa Tenggara by the *Suluh Insan Lestari* foundation.

Through coverage of the language transition pilots this part of the discussion also helps to answer key evaluation question 3: *To what extent does training teachers in language transition from mother tongue to the language of instruction improve children's reading outcomes?*

The findings from this part of the study are only suggestive of possible differences in effectiveness. The populations of variant pilots are too small for generalisability. In addition, variations in implementation and in the specific district context of implementation – differently difficult mother tongues, different class sizes, different rates of student absenteeism, different levels of teacher qualification – will have contributed to results in each case.

Table 25 compares the variants of teacher development pilots on outcomes on the basic literacy test.

Table 25: Comparison of gains to teacher development from different pilot approaches

Basic literacy test	Grade1	Grade 2
	Gains	Gains
Literacy 1	7	5
Guru BAIK	19	2
Language transition	11	9
Multi-grade	6	3

Two of these pilot types, *Guru BAIK* and language transition show much higher gains than Literacy 1, although in the case of *Guru BAIK* this is confined to grade one. The multi-grade variation is the only one that did not exceed the Literacy 1 result. A monitoring report on the pilot suggested that the effectiveness of the multi-grade implementation for improving component skills of reading was constrained by using the basic competencies framework of Curriculum 2013 as the basis for differentiating children's learning. These competencies provide little scope for developing the skills of sound and letter matching and decoding (INOVASI, 2019:27).

The gains of the *Guru BAIK* and language transition pilots are even more dramatic at the comprehension level, as table 26 indicates. However these results are minus the SIL foundation experiment in language transition in Southwest Sumba. This is because none of the students in its schools passed the baseline basic literacy test and so a baseline for comprehension could not be generated. The student learning assessment basic skills tests were all in Bahasa Indonesia. This is a lesson in itself about the excluding the effect of language on the scores of children who do not know the language of the test.

In the presentation in table 26 the results for the two language transition pilots are separated out to show the different effects of a language transition addition to a Literacy 1 pilot (the Bima pilot) and the Sulinama pilot that focused throughout on language transition.

Table 26: Comparison of variants of teacher development on endline gains in comprehension

	Reading comprehension	Listening comprehension
	Endline gains	**Endline gains**
All Literacy 1 pilots (180 sample schools)	6	12
Guru BAIK pilots (13 sample schools)	**18**	**39**
Literacy 1/language transition Bima pilot (7 sample schools)	3	**36**
Full language transition pilot (Sulinama — 6 sample schools)	**15**	**32**
Multi-grade pilot (7 sample schools)	5	6

Even taking into account the small and possibly distorting effects of comparing such small pilots against the large Literacy 1, the gains — sustained for both grade one and two — in *Guru BAIK* and Sulinama as so high as to amount to an important finding from INOVASI as to what works.

The special effects of *Guru BAIK* are corroborated in another study undertaken for INOVASI of its value added over 'plain' Literacy 1 (Purba and Sukoco, 2019:8). This is a comparison of the performance of students of Literacy 1 schools in Southwest Sumba that were preceded by a *Guru BAIK* pilot and of those that were not. The sample was 401 students and 38 teachers. Besides showing the considerably larger gains of students whose teachers had experienced this approach, the analysis also shows the differential value of the *Guru BAIK* approach for different kinds of learning disadvantage.

The authors attribute the greater effect of the *Guru BAIK* pilot to its student-centredness, systematised by a reformed classroom action research model. The focus of the model is on understanding the nature of the student problem through use of formative assessment data, and trialling and evaluating the solutions. A key part of the methodology addressed teachers' mindset to come to see 'that a student's struggle to learn is an opportunity for growth instead of incapability for learning' (Purba and Sukoco, 2019:3).

Box 7: Performance of a Literacy 1 pilot alone compared to Literacy 1 and *Guru BAIK* pilot combined in Southwest Sumba

% of students who passed basic literacy test (letter, syllable, and word recognition):	INOVASI – Literacy 1 pilot	INOVASI – Literacy 1 and Guru BAIK pilot
	% Increase	% Increase
All students	76%	113%
Gender		
Male	81%	135%
Female	72%	97%
Student with special needs		
Student with special needs	41%	193%
Student without special needs	84%	100%
Socio-economic status index		
Top	88%	49%*
Middle	53%	58%
Bottom	88%	121%
Student's mother tongue		
Indonesian	41%	59%
Local Language	111%	131%

Source: ICEAP study (2019)

Student-centredness could also be said to sum up the distinctiveness of the Sulinama language transition approach, with exceptional results like those of *Guru BAIK*. This student-centredness is produced by a different methodology, that focuses on teachers' expressiveness in communicating with the students.

The rapid language development of students could well be attributable to how expressively teachers explain in the classroom. These teachers enabled students to grasp the meaning of words in the new language not just by their vivacious bilingual media but also through the physical energy of their teaching to convey the meaning of words and concepts: for example jumping to explain 'jump', kneeling at the level of the children's desks to talk with children as they went around the groups. Contrasted with the verbal passivity of many teachers these were extraordinary transformations of teaching that struck many who visited these classrooms (INOVASI 2018, 2019).

Bilingual word wall at Sulinama school, East Sumba

As the literature emphasises this language experience focus in teaching vocabulary is associated not just with quicker second language mastery but higher levels of reading comprehension in later years.

Part two: The effectiveness of book provision for reading outcomes

Part two of this chapter addresses key evaluation question 2: *To what extent does provision of appropriate books improve children's reading outcomes?*

Through advocacy and policy support interventions and by supporting non-governmental organisation partnerships, INOVASI has raised the profile of book supply as a crucial element in improving reading outcomes. This section looks first at the achievement of three interventions at the school level to change the resourcing of reading.

The effect of two book provision variables on students' learning assessment scores has already been shown in the regression analyses in tables 23 and 24. This section explores evidence of improvement on two other reading outcomes of interest. One is the increase in books in classrooms. The other is the increase in the interest in reading, a strong predictor of reading proficiency and its effect on students' learning assessment scores.

Improvement in books in classrooms

Reading corners with and without non-textbooks

INOVASI is able to measure the increase in both book corners with textbooks and also book corner provision that include non-textbooks – story books and readers – in Literacy 1 pilot schools. Table 27 shows the findings on these two measures.

Table 27: Percentage increases in reading corners with non-textbooks and reading corners with textbooks only in Literacy 1 pilot schools

Comparative increase in reading corner provision			
	Baseline	Endline	**Endline gains**
Classroom reading corner (with non-textbooks)	20%	55%	35%
Classroom reading corner (with only textbooks)		68%	48%

The increase of book provision overall is considerable. The classrooms with non-textbook reading corners has increased from 20 per cent to more than half the schools in the pilot. Classrooms with textbook only reading corners have increased from the same low base to over two thirds of the schools.

The low baseline for textbooks located accessibly in the classroom particularly gives an idea of how little literacy classes were conducted with reference to reading matter in most of INOVASI's target schools before the program intervention. The lower gain in reading corners with non-textbooks also shows how difficult it is for schools to acquire suitable reading matter that will stimulate students' interest in reading. It also shows there is still some way to go in persuading school heads to allocate the schools' operational funds (BOS) for purchasing books.

Inexplicably, the Literacy 2 endline showed this achievement going backwards from the Literacy 2 baseline: 9 per cent for corners with non-textbooks and 2 per cent for textbook corners. An explanation to the evaluation team from teachers in schools where this occurred is that books

were removed from classrooms for fear of loss of the asset if they were left over time. That problem usefully draws attention to the unexpected complexity in such contexts of a simple seeming solution to book supply for teaching reading.

Effect on interest in reading

Did the Literacy 1 increase in books in the classroom have any effect on students' interest in reading? The SIPPI database includes a variable on students' reading interest: *proportion of students who say they love to read.* Table 28 shows endline gains over the baseline for the Literacy 1 pilot on this variable.

Table 28: Literacy 1: endline gains on students' reading interest

Student reading interest			
	Baseline	Endline	Increase
Grade 1	80	92	12
Grade 2	86	92	6
Grade 3	89	92	3

Source: SIPPI student questionnaire

Counterintuitively, while all start from a high baseline, the gains occur more at grade one than in later grades – when children can read. Could that be connected with the kind of reading material available for grade three as compared with grade one? An alternative suggestion is that Grade 1 interest in reading, starting from a lower baseline than the other grades, has caught up by endline-- possibly Grade 1 strugglers overcoming their difficulty with books through increased familiarity with them.

To find out whether students reading interest was activated more by non-text book material than by text books, students' reading interest gains in schools with non-textbook reading corners were compared with interest gains in schools with only textbook material in their reading corners. However no pattern emerged in the analyses and the differences between the two were negligible.

The value of the book pilots

In some districts, grantee pilots tried various types of book provision in Literacy 1 schools (and elsewhere): supplying school libraries or classroom reading corners; providing levelled readers; and big books to support literacy instruction. Under INOVASI's direction these grantee pilots went mainly to remote locations where schools and families have difficulty accessing books. These book pilots and locations are identified in table 29.

Table 29: Pilots supporting the teaching of literacy through books in Literacy 1 pilot schools

Location	Schools affected	Pilot	Focus
Bulungan	7	Litara	Library/community book centres
	7	One person, one book program (OPOB)	
Malinau	13	Litara	
	13	OPOB	
East Sumba	2	Rainbow Reading Gardens	Libraries
Central Sumba	2		
West Sumba	2		
	10	Indonesian Children's Literacy Foundation (YLAI)	Literacy Instruction with levelled readers and big books
Southwest Sumba	10		
	2	Rainbow Reading Gardens	
Central Lombok	19	Pen Circle Forum	Inclusive levelled readers

To see the effect of this focus on students' literacy scores, table 30 compares the endline gains of the Literacy 1 pilot schools with and without these additions.

Table 30: Comparison of endline gains of Literacy 1 pilot schools with and without book pilots

Literacy measures	Literacy 1 baselines	Literacy 1 endline gains	Literacy 1 + book pilot baseline	Literacy 1 + book pilot endline gains
Reading comprehension	61	**6**	50	**9**
Listening comprehension	68	**11**	67	**17**

With similar baselines, the value added by the book pilots to teacher pilots at endline for both grade one and two, compared to the plain Literacy 1, shows student outcomes in reading comprehension increasing by one third. The larger gain in listening comprehension is particularly pleasing because it is evidence of these books being read to children and used to develop comprehension skills. Given this result, It would be valuable to conduct a study of the specific impact of the Indonesian Children's Literacy Foundation (YLAI), because of its emphasis not only on providing big books for shared reading but its exemplification of their use in a balanced literacy approach. This non-governmental organisation provided an enriched version of the Literacy 1 methodology.

Evidence for effect on student scores

Table 31 shows the results of regression analysis to explore the association of reading interest and book access variables with students' learning assessment scores in the basic literacy test and the reading and listening comprehension tests.[37]

[37] The rule of thumb is: 0.01=small, 0.06=medium, 0.14=large (Cohen 1988). http://www.psy.gla.ac.uk/~steve/best/effect.html (Cohen, 1998)

Table 31: Effect of interest in reading and book availability on performance in student learning assessment tests

VARIABLES	Basic test	Listening comprehension test	Reading comprehension test
Love to read (student interest)	**0.370*** **	0.041	0.011
	(0.022)	(0.000)	(0.000)
Time to read at home	0.019	0.012	-0.005
	(0.001)	(0.000)	(0.000)
Parents said that the students love to read	**0.157*** **	-0.021	-0.039
	(0.005)	(0.000)	(0.000)
Availability of books at home	**0.150*** **	0.054	**0.129*** **
	(0.007)	(0.001)	(0.004)
Reading corner with non-textbooks	**0.281*** **	0.216***	**0.348*** **
	(0.018)	(0.011)	(0.033)
Textbook availability in class	0.115** **	0.072	0.137*
	(0.006)	(0.001)	(0.006)
Student's raven score	0.012***	0.007***	0.007***
	(0.060)	(0.015)	(0.018)
Student's socioeconomic status index	0.010***	-0.001	0.008***
	(0.059)	(0.000)	(0.026)
Gender of student (female=1)	**0.216*** **	0.032	**0.159*** **
	(0.015)	(0.038)	(0.007)
Constant	-1.776***	-0.615***	-1.100***
	(0.052)	(0.124)	(0.128)
Observations	5,622	2,847	2,847
R-squared	0.331	0.036	0.143

Note: Figures without brackets are random effect regression coefficients while those inside the brackets are partial eta-squared (effect size). The rule of thumb is: 0.01=small, 0.06=medium, 0.14=large

Students' interest in reading has a strong correlation with the student learning assessment scores on the basic literacy test, higher than any other variable on this basic literacy test, including the control variables of socioeconomic status, student ability and availability of books

For interest in reading, the actual survey items are: *Apakah kamu suka membaca?* (Do you like to read?) (SIPPI student survey BB.9); *Apakah anak Ibu/Bapak suka atau senang membaca?* (Do you like or enjoy reading?) (SIPPI parent survey D.2a) Book availability items are:
1. Availability of books at home: *Selain buku pelajaran, berapa banyak buku/majalah di rumah yang sesuai untuk anak usia 5-12 tahun?*(Besides textbooks, how many books / magazines at home are suitable for children aged 5-12 years?) (SIPPI Parent's questionnaire)2. Textbook availability in class: *Berapa banyak siswa yang menggunakan buku pelajaran saat pembelajaran berlangsung?*(How many students use textbooks during learning?)(SIPPI classroom observation)

in the home. This includes the control variable of gender which is only one of the control variables that is correlated, in line with well researched differences in reading interest based on gender.

This is an important finding for what works. But equally important is to find out why the association cuts out for comprehension where it *should* have the effect.

Reading corners with non-textbooks is the variable with the next highest correlation with students' scores on the basic literacy test and the association extends through to listening and reading comprehension at the highest level of probability. The correlation between this variable and reading comprehension emerges as the strongest correlation in the whole analysis.

The consistent association of non-textbook reading corners with students' outcomes across different regressions makes it an important finding for what works in improving literacy outcomes. Nevertheless it is not possible to interpret the results without more information. More trials and analyses are needed to establish whether it is because it engages children in reading and/or engages teachers in reading to them or using comprehension-enhancing reading material in their teaching. If these possibilities turn out to be the case, then the policy message is clear: the factor that will clearly affect the development of higher-order thinking skills lies within the routine responsibility of government for educational provision. They need to ensure the supply of textbooks and reading material so that children can learn to read and think by reading.

Alternatively, there is a possible underlying factor in the variable influencing its success, for example, the possibility that non-textbook reading corners occur in better-resourced schools with better-off students. The policy message on resourcing higher-order thinking skills would in that case need to be heavily modulated in favour of equity: prioritising the needs for reading resources of poor and remote schools.

Conclusion

The answers to the three key evaluation questions to be addressed in this chapter are all interconnected: *What worked for reading outcomes?*; *What worked for higher-order thinking skills outcomes?;* and *What worked for literacy for second language learners?* All of the approaches in the INOVASI pilots worked in terms of the evidence in this chapter. And they all worked because across them, they all delivered on some fundamental elements of student centredness. These were: classroom practices organised for teaching at the right level; problem-driven support to students by analysing assessment data; a foundational understanding of how to meet the specific literacy needs of children whose home language is different from the language of instruction; and an idea of mindset development that positively casts the work of teaching as the challenge of student-centred learning.

The limitations of the data collection meant that correlational evidence for the prerequisite importance of literacy pedagogies could not be provided but was attested in the previous chapters, including in the consistency of gains to students outcomes from this focus of Literacy 1. Added to the importance of pedagogy is the realisation that books made a considerable difference – for developing love of reading, widening the horizons in shared reading experience and developing higher-level comprehension skills.

An important implication in the convergence of all these strategies on making a difference to teachers' capacity for student-centred teaching is that combining established knowledge of how children learn to be literate with a grounded understanding of context is productive. This kind of understanding of context has been achieved through iteration and reflection on the different kind

of pilot successes; but from a substantive platform of knowledge of what learning to be literate requires. Combining all these different contributions is a potential approach for future piloting – not through module add-ons or add-on techniques for teacher reflectiveness, but through deepening and integrating the different approaches evidenced here to work into a literacy curriculum for early grades in Indonesia that fits the implementation context.

10 Implications

INOVASI set out on its journey to discover what works for improving literacy outcomes, particularly through teaching improvement, with no suppositions and equipped only with a Geiger counter for detecting local problems and solutions. At the end of the first part of this journey the program has developed a fair map of what and where the key problems are in student learning, teaching and teacher support. A model is emerging of what works to enable teachers to teach literacy, along with the implications for systems behind that model. Equally valuable is the sense of what is needed to make solutions work, a 'take' on how to achieve a contextual fit that includes the mindset context. This is a practitioners' sense, a way of working, not reducible to a program of policy or system development but that nevertheless might result in *stakeholders* developing innovative policy and systems that fit their contexts.

This final chapter considers the implications of INOVASI's level of achievement by the end of the first phase of the program from the perspective of what needs strengthening in practice. We review the most significant of the teaching and learning results and conclude with what those findings imply about planning for the future.

Significant findings on early grades learning in literacy

There are four significant findings on student outcomes. The least *unexpected* finding is the difference at baseline on beginning reading skills between the provinces, although it is worth looking again at the extent the differences can run to, for example: 58 per cent of students in East Java passed the basic literacy test that measures beginning reading skills while just 3 per cent of students in Sumba passed this same test.

That leads to the second and most surprising finding: that there is not so much difference between the provinces at the comprehension level. Not only is the baseline gap smaller than for the beginning skills of reading but the gains from the program are higher in disadvantaged provinces compared to the more advantaged provinces. Sumba, the most disadvantaged region, attained the highest gain in one of the higher-order thinking skills. This is not peculiar to Sumba – in every province the district that had the lowest baseline had the highest gains.

These two findings have several implications. The first, arising from the relatively high performance of the disadvantaged provinces on comprehension, is to avoid a deterministic assumption about poor and disadvantaged contexts meaning lower performers. Low performance does not mean children cannot develop to the full potential of literacy like children elsewhere. Rather, the finding highlghts the disproportionate numbers of students in these regions who are excluded at the threshold of comprehension by not having the most basic skills of letter and word knowledge. Initial and decisive exclusion from the benefits of literacy happens in grade one and grade two for most children in a region like Sumba.

The study found how natural growth, without interventions, eventually resolves beginning problems for most children: only 10 per cent still struggle to master these skills by grade three for most provinces. This mastery however takes an extra year in a place like Sumba. That extra year comes at a cost if curriculum expectations do not make allowances for it in such regions – students lose out by being unable to read to learn at a point — around grade 3 —when this becomes the school agenda.

That points to the need for solutions to equalise the rate that children acquire these skills. The extent of the gains over baseline in grade one and two as a result of pilot adaptations to contextual circumstances in disadvantaged districts (language transition, book-based pilots, problem focussed pilots) suggest that a more precise and intense focus on grade one and two learning is needed in such contexts.

It is likely also that what worked emphatically in those adaptations are methodologies that early graders everywhere would benefit from, in that they enhance regular practice and provision. An example of this is the expressiveness of the instruction in language transition pilots where teachers encouraged students to talk and explore words. Word knowledge, which is critical to comprehension, lagged behind other beginning reading skills in all districts, perhaps because generally vocabulary teaching — if it occurred—was not focussed on students understanding the meaning of the concept in the word.

A further significant finding arises from the smaller gap between the provinces on students' comprehension baselines, but this time to put more advantaged districts into the spotlight. Why are provinces like East Java and West Nusa Tenggara not able to perform better on reading comprehension and higher-order thinking skills considering their strong base in beginning reading skills? Was there a lack of fit with local professional practice in INOVASI's methodology or or does this tell us something about a more universal problem in teaching approaches to reading comprehension, particularly at the higher level of comprehension skills?

Significant findings in teacher practice

This study provides evidence that most pilot teachers have a working grasp of some key elements of effective literacy practice. This is mainly evidenced by the frequency in the spot-check data of teachers' use of core literacy strategies that integrate a range of skills.

Most teachers are starting to include the balanced literacy strategies that support comprehension in their teaching and slightly less than 50 per cent were using shared reading to build comprehension skills through questioning routines. Around 50 per cent are able to use diagnostic techniques to track students' progress in reading to the extent of organising classes into levelled groups for targeted skills and reading practice.

One important general capability of particular value in literacy teaching is developing and using appropriate media. In literacy, teachers produced big books for shared reading, a process that demands and heightens their understanding of grading text to students' levels. Their school heads celebrate this creativity and the interactivity it produces in classrooms. On this use of target media the SIPPI classroom practice index found a large endline gain of 24 percentage points – much larger still in the disadvantaged regions of North Kalimantan and Sumba.

These are all demanding skills. They are also the skills of problem-based teaching, essential for tracking the different levels of progress children normally are at in a reading classroom. They are consistent with the whole problem-based approach to improving teaching that was the distinctive insight of INOVASI's development theory for approaching what works.

In contrast to this level of success in comprehension teaching, there is evidence in many of the data sources, of teachers' difficulties with aspects of the pilot approach to beginning reading skills; specifically with the phonemic approach to decoding written language. This difficulty may be an instance of well-credentialed methodologies *not* working in context. In regions in Indonesia where there are professional traditions of teaching reading, such as East Java, teachers use a successful local methodology, the syllabic (*suku kata*) approach to word building. More investigation is needed into whether this is a more appropriate approach for

Bahasa Indonesia than the phonemic technique of sound and letter matching. A finding from INOVASI's experience is also, however, that being local is not a sufficient criterion for adoption – an approach also has to be successful. The program found one problem that hampers children's mastery of beginning skills in Sumba is a decoding system that doesn't work (spelling out a word rather than sounding it out).

Problem-based teaching is also the entry point into student-centred teaching in the way it upturns the relationship between teaching and learning. The pilot participants made much progress in the visibility of learning, through displays of students' work and visual monitoring of the whole class. However the lesson analyses in the case studies show a different challenge to achieving student-centredness that is not reachable through teachers' technical development.

This is teacher-centredness as an unexamined mindset. The case studies showed an assumption operating that students learn through repeating what the teacher tells them. The long tradition of having children learn by memorisation is hard to shift, especially when it is accompanied by a teaching culture where the teacher has the prerogative to talk and shape all interactions. Even in a teacher who was respectful of the children's efforts, this dominant idea interfered with the children's comprehension of the *written* text, both at the level of the sentence and in the overall meaning.

The study found that this is the teacher mindset to reckon with for bringing about transformative teacher change.

While the findings established that teachers can put many key practices of literacy teaching into effect, we do not yet have evidence on the *quality* of this practice. This cannot be established by quantitative instruments or even in classroom observation. Frequencies on a survey instrument question such as *asks open questions,* for example, or *gives informed feedback,* do not disclose whether the open question made sense or was relevant, or whether the informed feedback was appropriate to the problem. (We saw in the case studies instances of both of these problems with such strategies.)

One way of knowing about quality is the effect on student outcomes. Unfortunately we do not know whether the literacy pedagogies learnt through the pilots were effective enough for that impact, as literacy specific variables were not included in the baseline measures for INOVASI. We do know however that the regression analyses showed only the slightest association between teacher practice and student literacy outcomes on the SIPPI index for the beginning skills of reading and for comprehension only listening comprehension scores were affected.

We also know that the teacher attribute variable that was significantly associated with student comprehension outcomes was teachers' own reading literacy: the higher the teacher score, the higher the students' score, particularly in higher-order thinking skills. Although INOVASI did not include improvement of teachers' literacy in its piloting, this finding may throw light on whether effective teaching for comprehension and higher-order thinking skills is achievable without it. In chapter 2 the mean score for teachers at program level on reading comprehension on Grade 4 international test questions was 54 per cent and no province, except East Java, scored 50 per cent on higher-order thinking skills.

To be able to develop reading comprehension, teachers need themselves to be comprehenders of explicit and implicit meaning in text. However, the case studies suggested that teachers think that reading is about reading fluently, with fast and accurate pronunciation. Hence, to help students understand they had recourse to the facing pictures and scaffolding questions but only rarely directions to look at the words in the text. There may be limited awareness among teachers that the structures and wording of literate language are different from speech patterns

and especially those that children hear. If so, teachers' capacity to develop students' higher-order comprehension will be curtailed by this limitation.

The pilots were mainly six months long and were delivered by local facilitators – teachers and supervisors — rather than outside experts. Teachers' learning took place in between their other work, in most cases in monthly facilitator visits. Opportunities to practise were also constrained by the curriculum. These circumstances explain the modest gains from the literacy pilots and make the extent of uptake a considerable achievement.

However teachers may not yet have a deep understanding of the complex processes involved in acquiring reading literacy. Teachers are at the beginning of change at the end of this first phase of INOVASI. For these beginnings to be able to affect students' results, teachers' understanding needs to be deepened and supported.

The emerging model of teacher development for literacy and its systemic implications

The following discussion makes suggestions about 'what next' from INOVASI's pilot achievements. The purpose of the pilots was to influence district, provincial and national government take-up of what has worked. In that take-up the key issues are what should they take up and how can they ensure that the piloted strategies work at scale to improve teaching and learning outcomes.

Chapter 9 included a comparison of variants on the teacher development pilots. In addition to the skills platform that participants acquired through Literacy 1 and 2, there were three variants that had added value. These were the language development pilots, the book pilots supporting Literacy 1 pilot schools and the *Guru BAIK* pilot in Southwest Sumba, building an increased focus on problem-based teaching onto the pedagogical skills learnt in Literacy 1.

The aspects covered by these pilots are all indispensable conditions of effective literacy teaching and learning. A suggestion for take up of what worked in INOVASI is for a model integrating these different facets of support for literacy learning – language transition being included in areas where children are unfamiliar with the language of instruction. In view of the evidence that teachers' own literacy proficiency influences student outcomes, consideration could also be given to teacher development that develops teachers own literacy skills.

This emerging pilot model has systemic implications: for the professional development mechanism of the teachers' working groups used to deliver the literacy training; for school support for learning improvement; for book provision to support the balanced literacy approach that underpinned the model; and for the balanced literacy approach itself as a curriculum model. Attending to these will ensure that piloted strategies work at scale to improve teaching and learning outcomes.

INOVASI has worked alongside the pilots on all these systemic dimensions. Much of the policy, regulatory and funding initiatives that districts have taken to support literacy or quality of learning improvement was brought about through INOVASI's advocacy, technical support and provider partnerships. Chapter 2 detailed the breakthrough changes in these implicated systems that have taken place. In different districts these have made teachers' working groups more functional and accessible; made district, village and school funding more available to purchase non-textbooks for schools; put tracking of learning progress at the centre of school activity; and resulted in a growing number of districts setting up a formal monitoring system for early grade learning.

Also described in chapter 2 is the progress of national curriculum and assessment reform around learning outcomes improvement, including in early grades literacy and numeracy. This also has been in step with and influenced by INOVASI's decentralised innovations. If there is alignment between a new curriculum framework and the successful teaching of literacy that has emerged from the pilots there is potential for systemic integration of these approaches with the work of teachers, school heads, district and national education authorities.

This concluding chapter opened by saying that one of the program's achievements was a fair sense of what is needed to make solutions work: a "take" on how to work for contextual fit. This applies even more to systemic support than it did to supporting improved teaching. The program recognised early, for example, that the quality of the model would be only as good as its professional development delivery system; that principals' funding support for materials for teachers to make media lagged behind their praise for its transformative effect. Much effort went into trying to orchestrate district interest in acting on these and other systemic problems. However what evolved in the process —and through a political commitment to the principles of stakeholder ownership —are solutions that stakeholders have been worked out themselves, for fit with the opportunities and constraints of their different situations. There are many of them, in each of the systems implicated in teacher support. More importantly, in many districts there are now systematic processes for deciding what to prioritise and how to go about change. Understanding context means avoiding specification what districts systems should look like in bringing piloted strategies work at scale to improve teaching and learning outcomes.

What does need precision, however, is recognition that at the end of INOVASI, teacher practice has not yet transformed; but that we now know that, and how, it variously can. Transformed teacher practice is for districts to achieve. The point of this resumé of the significant findings of this study, is to show the importance of a careful reading of context to get the right fit. With that we might avoid a difference between provinces in terms of learning performance; and have instead a difference in ways of achieving learning.

References

Abadzi H. 2006. Efficient learning for the poor: insights from the frontiers of cognitive neuroscience. Washington DC: World Bank.

Abadzi H. 2008. 'Conceptual framework for early reading fluency: insights and monitoring indicators from cognitive neuroscience'.. Powerpoint presentation for a USAID education sector professional development course on early grade reading. Washington DC, 12 May 2008. https://www.eddataglobal.org/documents/index.cfm?fuseaction=pubDetail&ID=140. Accessed 29/07/2015

Abeberese AB, TJ Kumler, LL Linden. 2011. Improving reading skills by encouraging children to read in school: a randomized evaluation of the Sa Aklat Sisikat reading program in the Philippines. Cambridge MA: National Bureau Of Economic Research.

Andrews M. L Pritchett and M Woolcock. 2017. Building state capability: evidence, analysis, action. Oxford: Oxford University Press.

Asim S, RS Chase, A Dar and AD Schmillen. 2015. Improving education outcomes in South Asia: findings from a decade of impact evaluations, 2015. Policy research working paper 7362. Washington DC: World Bank Group. http://documents.worldbank.org/curated/en/254411468188375892/Improving-education-outcomes-in-South-Asia-findings-from-a-decade-of-impact-evaluations

Assessment Centre (Puspendik). 2017. [Indonesian student competence assessment] (Asesmen Kompetensi Siswa Indonesia – AKSI). Jakarta: MoEC.

Australian Council for Education Research (ACER), UNESCO Institute for Statistics. 2017. Principles of good practice in learning assessment. Melbourne: ACER.

Banerjee A, R Banerji, J Berry, E Duflo, H Kannan, S Mukherji et al. 2016. Mainstreaming an effective intervention: evidence from randomized evaluations of 'teaching at the right level' in India. NBER working paper No 22746. Cambridge MA: National Bureau of Economic Research. https://www.nber.org/papers/w22746.

Brock-Utne B. 2010. 'Research and policy on the language of instruction issue in Africa'. International Journal of Educational Development 30(6): 636–645. https://doi.org/10.1016/j.ijedudev.2010.03.004

Cincotta-Segi A. 2011. 'Talking in, talking around and talking about the L2: three literacy teaching responses to L2 medium of instruction in the Lao PDR': Compare: A Journal of Comparative and International Education 41(2): 195–209.

Clark P, C Crawford,F Steele and A Vignoles. June 2010. The choice between fixed and random effects models: some considerations for educational research. Working paper 10/240. Bristol: Centre For Market And Public Organisation, http://www.bristol.ac.uk/media-library/sites/cmpo/migrated/documents/wp240.pdf

Cohen J. (1988). Statistical power analysis for the behavioral sciences. New York: Routledge Academic.

Cummins J. 2001. 'Bilingual children's mother tongue: Why is it important for education? Rights and responsibilities of educators of bilingual-bicultural children'. In LD Soto (ed). Making a difference in the lives of bilingual-bicultural learners, Bern: Peter Lang.

Damhuis CMP, E Segers and L Verhoeven. 2015. 'Stimulating breadth and depth of vocabulary via repeated storybook readings or tests, school effectiveness and school improvement'. An International Journal of Research, Policy and Practice 26(3):382–96.

Draper SW. 14 May 2002, Effect size. http://www.psy.gla.ac.uk/~steve/XXXX.html (visited 2020 April 17)

Dweck C. 2008. Mindset: the new psychology of success. New York: Random House.

East Nusa Tenggara team. 2019. Report on the visit of East Sumba vice-regent to schools. Jakarta: INOVASI.

Freebody P. 2007. Literacy education in school, 2007: Research perspectives from the past for the future. Melbourne: ACER.

Glewwe PW, EA Hanushek, SD Humpage, R Ravina. 2011. School resources and educational outcomes in developing countries: a review of the literature from 1990 to 2010. National Bureau of Economic Research working paper. Cambridge MA: NBER. 17554 http://www.nber.org/papers/w17554 .

Hart B and T Risley. 2003. 'The early catastrophe: the 30 million word gap by age 3'. American Educator Spring: 4–9.

Heath SB. 1982. 'What no bedtime story means: narrative skills at home and school'. Language in Society 11(1): 49–76.

INOVASI MERL team. 2018. Teacher reflections from 2018 pre-pilots. Jakarta: INOVASI.

INOVASI, 2017. Indonesian National Assessment Program West Nusa Tenggara 2016, NTB: What NTB students know and how the government, school, teachers and parents support them. Jakarta: INOVASI.

INOVASI. September 2017. Rapid participatory situation analysis of support for learning: North Kalimantan province. Jakarta: INOVASI.

INOVASI. December 2017. Guru BAIK process evaluation report. Jakarta: INOVASI.

INOVASI. August 2018. Third strategy testing. Jakarta: INOVASI.

INOVASI. November 2018. Emerging evidence and policy recommendations. Jakarta: INOVASI.

INOVASI, May 2019. Batu City Perwali: the Mayoral regulation on Batu as a literacy city. Jakarta: INOVASI.

INOVASI. July 2019. Fourth strategy testing. Jakrta: INOVASI.

INOVASI, December 2019. APBD analysis report on the function of education. Jakarta: INOVASI.

International Conference on Education and Psychology (ICEAP). 2019. ICEAP proceedings book. International Conference on Educational Assessment and Policy. Jakarta, Indonesia, 24–25 September 2019.

Jewitt C. 2012. An introduction to using video for research. NCRM working paper 3/12. London: National Centre for Research Methods.

Kosonen K. 2017. Language of instruction in Southeast Asia. Background paper prepared for the Global Monitoring Report. Paris:UNESCO.

Leer J, M de Fretes and G Rinjani. 2016. Literacy Boost Belajar Indonesia. Jakarta: Save the Children.

Louden and Wallace J. 1995. 'Case methods and teacher education', Paper presented at the Annual Conference of the Australian Association for Research in Education, Hobart, 26-30 November, 1995.

Marulis l and S Neuman. 2013. 'How vocabulary interventions affect young children at risk: a meta-analytic review'. Journal of Research on Educational Effectiveness,6(3):223-262.

Ministry of Education and Culture (MoEC). 2015. Strategic plan for education, (Renstra), 2015–2019. Jakarta: MoEC.

Ministry of Education and Culture (MoEC). 2016. [Guide to the school literacy movement in primary schools] (Panduan gerakan literasi sekolah di sekolah dasar). Jakarta: MoEC.

Ministry of Education and Culture (MoEC). 2019. [National education reconstruction: a reference frame] 6 April 2019. Jakarta: MoEC.

Ministry of National Development Planning (Bappenas). 2015. [National medium-term development plan (RPJMN) 2015–2019]. Jakarta: Bappenas.

Ministry of National Development Planning (Bappenas). 2019. [National medium-term development plan (RPJMN II) 2020–2024]. Jakarta: Bappenas.

Ministry of National Planning and Development (Bappenas). 2019. [Technical plan] (Rancangan Teknokratik) (RPJMN 2020-24). Jakarta: Bappenas.

Mullis IVS, MO Martin, AM Kennedy and P Foy (2007) PIRLS 2006 international report, Chestnut Hill: TIMMS and PIRLS International Study Center

Mullis IVS, MO Martin, AM Kennedy, KL Trong, and M Sainsbury. 2011. Progress in International Reading Literacy Study PIRLS 2011 Assessment framework. Chestnut Hill: TIMSS & PIRLS International Study Center.
https://timssandpirls.bc.edu/pirls2011/downloads/PIRLS2011_Framework.pdf

National Research Council. 1998. Preventing reading difficulties in young children. Washington, DC: The National Academies Press. https://doi.org/10.17226/6023.

Organisation for Economic Co-operation and Development (OECD). 2010. PISA 2009 Results: Executive summary. Paris: OECD Publications.

OECD. 2012. Let's read them a story: The parent factor in education. Paris: PISA OECD Publishing. http://dx.doi.org/10.1787/9789264176232-en

OECD. 2017. PISA 2015 assessment and analytical framework: science, reading, mathematics, financial literacy and collaborative problem solving. Paris: PISA OECD Publishing. https://doi.org/10.1787/9789264281820-en.

OECD. 2017. 'PISA 2015 reading framework'. In PISA 2015 assessment and analytical framework: science, reading, mathematics, financial literacy and collaborative problem solving. Paris: PISA OECD Publishing. https://doi.org/10.1787/9789264281820-en Paris: OECD.

OECD. 2019. Programme for international student assessment (PISA) results from PISA 2018 Country note: Indonesia. Paris: PISA OECD Publishing https://www.oecd.org/pisa/publications/PISA2018_CN_IDN.pdf

OECD. 2019. TALIS 2018 results, volume 1: Teachers and school leaders as lifelong learners. Paris: Teaching and Learning International Survey (TALIS), OECD Publishing.

Paris S and A Paris. 2003. 'Assessing narrative comprehension in young children', Reading Research Quarterly 38(1): 36–76 January/February/March

Paris, S. G. (2005) Reinterpreting the development of reading skills. Reading Research Quarterly, 40(2), 184–202.

Perfetti CA, N Landi and J Oakhill. 2005. 'The acquisition of reading comprehension skill'. In MJ Snowling and C Hulme (eds). Blackwell handbooks of developmental psychology. The science of reading: A handbook (p. 227–247). Hoboken: Blackwell Publishing. https://doi.org/10.1002/9780470757642.ch13

Pressley M and L Fingeret. 2007. 'What we have learned since the National Reading Panel: Visions of the next version of Reading First'. In M Pressley, AK Billman, KH Perry, KE Reffitt and JM Reynolds (eds.) Shaping literacy achievement: Research we have, research we need. New York: Guilford Press.

Purba R and G Sukoco, G. 2019. 'Rethinking classroom action research: a worthy effort for classroom learning', Paper presented to the International Conference on Education and Psychology, June 2019.

Purcell-Gates V, E Jacobson and S Degener. 2004. 'The seemingly incommensurability of the social and cognitive'. In V Purcell-Gates, E Jacobson and S Degener. Print literacy development: uniting cognitive and social practice theories. Masachusetts: Harvard University Press.

Rarasati N, D Nugroho, D Suryadarma, S Kurniawati, M Akrom, RP Artha et al. 2016. Interventions, practices and contextual factors linked to Indonesian students' literacy and numeracy outcomes: a systematic review. Jakarta: INOVASI.

Regional Economic Development Institute (REDI) and INOVASI. 2019. Final report: Guru BAIK building teachers' capacity in West Nusa Tenggara, Indonesia. Jakarta: INOVASI.

Richardson JT. 2011. 'Eta squared and partial eta squared as measures of effect size in educational research', Educational research Review 6(II): 135–47.

RTI International and United States Agency for International Development (USAID). 2014. National Early Grade Reading Assessment (EGRA) and Snapshot of School Management Effectiveness (SSME) survey: report of findings. Jakarta: RTI International and USAID Indonesia.

RTI International and USAID. 2017. PRIORITAS final project report, volume 1 main report. Research Triangle Park: RTI International and USAID Indonesia.

Save the Children. 2015. Literacy Boost Belajar Indonesia Baseline report . https://resourcecentre.savethechildren.net/sites/default/files/documents/indonesia_belajar_baselinereport_20151030.pdf

Scheerens J. 2015. 'Theories on educational effectiveness and ineffectiveness', School Effectiveness and School Improvement 26(1):10–31.

Silva M and K Cain, 2015 'The relations between lower and higher level comprehension skills and their role in prediction of early reading comprehension', Journal of Educational Psychology 107(2):321–331.

Siraj-Blatchford I. 2010. 'Learning in the home and at school: how working class children succeed against the odds'. British Educational Research Journal 36 (3): 463–82.

Snow CE and C Juel. 2005. 'Teaching children to read: what do we know about how to do it?'. In MJ Snowling and C Hulme (eds.). Blackwell handbooks of developmental psychology. The science of reading: A handbook. pages 501–520. Hoboken: Blackwell Publishing.

Snow CE. 2010. Reading comprehension: reading for learning. Cambridge MA: Harvard Graduate School of Education. cal.org/create/conferences/2012/pdfs/reading-comprehension.pdf;

Torgerson, C, G Brooks, L Gascoine and S Higgins. 2019. 'Phonics: reading policy and the evidence of effectiveness from a systematic "tertiary" review'. Research Papers in Education 34(2):209.

Trudell B and C Young (eds). 2016. Good answers to tough questions in mother tongue-based multilingual education. Dallas: SIL International. http://www.sil.org/literacy-education/good-answers-tough-questions-mother-tongue-based-multilingual-education

UNESCO Institute of Statistics. 2006. Education for all: literacy for life; EFA global monitoring report, 2006. Paris: UNESCO.

UNESCO. 2016. Education for people and planet: creating sustainable futures for all, Global education monitoring report, 2016. Paris: UNESCO.

US Department of Education Institute of Education Science (IES). 2008. Reading First impact study. Washington DC: IES.

Vavrus F and L Bartlett. 2012. 'Comparative pedagogies and epistemological diversity: social and materials contexts of teaching in Tanzania'. Comparative Education Review 56(4):634–58.

World Bank, 2015. Indonesia: A Video Study of Teaching Practices in TIMSS Eighth Grade Mathematics Classrooms. Understanding What Teaching Practices are Used, Why They are Used and How They Relate to Student Learning. Main Report.

Legislation

Ministerial regulation No 23 of 2015 on developing character

Mayoral regulation on Batu as a literacy city No 93 of 2018

Minister of Home Affairs Circular Letter 420/9240 SJ of 1918, concerning the implementation of school literacy education in the regions

No Child Left behind Act of 2001 (United States)

Ministerial regulation no 11 of 2019 on priorities for the use of village funds in 2020

Annex 1: Literacy pilots

1.1 Districts with literacy pilots

PROVINCE	DISTRICT
East Java	Batu city
	Pasuruan
	Probolinggo – Paiton
	Probolinggo – Sukapura
	Sumenep
North Kalimantan	Bulungan
	Malinau
West Nusa Tenggara	Bima
	Dompu
	Central Lombok
	Lombok Utara
	Sumbawa Barat
East Nusa Tenggara	West Sumba
	Southwest Sumba
	Central Sumba
	East Sumba

1.2 Pilots by analytical category

A. LITERACY 1 PILOTS (INOVASI-managed)		
Province	**District**	**Additional interventions in districts**
East Java	Batu city	Leadership
	Pasuruan	
	Probolinggo – Paiton	
	Sumenep	
North Kalimantan	Bulungan	Books (Litara and OPOB)
	Malinau	Books (LITARA and OPOB)
West Nusa Tenggara		
	Bima	Gembira (language transition)
	Dompu	& Bersama (community engagement)

	Central Lombok	& Setara (inclusion)
	North Lombok	
	West Sumbawa	
	West Sumba	Leadership & Books (INOVASI,Rainbow Reading Gardens)
	Southwest Sumba	Guru BAIK, Books (INOVASI, Rainbow Reading Gardens)
	Central Sumba	& Books (INOVASI, Rainbow Reading Gardens)
	East Sumba	& Books (INOVASI, Rainbow Reading Gardens)

B. LITERACY 1 AND 2 PILOTS (INOVASI-managed)	
Province	**District**
East Java	Probolinggo – Sukapura
	Pasuruan
	Sumenep
North Kalimantan	Bulungan
	Malinau
West Nusa Tenggara	
	Bima
	Dompu
East Nusa Tenggara	West Sumba
	East Sumba

C. ONLY LITERACY 2 PILOT (INOVASI-managed)	
Province	**District**
East Java	Probolinggo – Sukapura

D. ADDITIONAL TEACHING FOCUS PILOTS (INOVASI-managed)		
Province	**District**	**Focus**
East Java	Probolinggo – Sukapura	Multi-grade (phase 1), Literacy (phase 2)
West Nusa Tenggara	Bima	Gembira (Literacy & Mother Tongue)
East Nusa Tenggara	Southwest Sumba	Literacy, Guru BAIK, Books (INOVASI,Rainbow Reading Gardens)

E. GRANTEE PILOTS —LANGUAGE TRANSITION		
Province	District	Focus
East Nusa Tenggara	East Sumba	Language transition
	Southwest Sumba	Language transition

F. GRANTEE PILOTS- BOOKS			
Province	District	Grantee	Focus
North Kalimantan	Bulungan	Litara	Literacy and books
		One Person One Book (OPOB)	Literacy and books
	Malinau	Litara	Literacy and books
		OPOB	Literacy and books
West Nusa Tenggara	Central Lombok	Pen Circle Forum	Literacy inclusion and books
East Nusa Tenggara	West Sumba	Indonesian Children's Literacy Foundation (YLAI)	Literacy and books
		Rainbow Reading Gardens	Books
	Southwest Sumba	YLAI	Literacy and books
		Rainbow Reading Gardens	Books
	Central Sumba	Rainbow Reading Gardens	Books
	East Sumba	Rainbow Reading Gardens	Books
		Rainbow Reading Gardens	Books

1.3 All Grantee pilots

PROVINCE	DISTRICT	PARTNER	KIND
East Java	Batu city	UINSA	Literacy
	Pasuruan	UNESA	Literacy
North Kalimantan	Bulungan	Universitas Borneo Tarakan (UBT)	Literacy
		Litara	Literacy and books
		OPOB	Literacy and books
	Malinau	Universitas Borneo Tarakan (UBT)	Literacy
		Litara	Literacy and books
		OPOB	Literacy and books

West Nusa Tenggara	Central Lombok	Forum Lingkar Pena	Literacy inclusion and Books
	Sumbawa Barat	Edukasi 101	Literacy and Numeracy
		Edukasi 101	Literacy and Numeracy
East Nusa Tenggara	West Sumba	YLAI – West Sumba	Literacy and books
		Rainbow Reading Gardens	
	Southwest Sumba	YLAI – West Sumba	Literacy and books
		Rainbow Reading Gardens	
		SIL	Language transition
	Central Sumba	Rainbow Reading Gardens	
	East Sumba	SULINAMA	Language transition
		Rainbow Reading Gardens	
		Friends of the Islands (Sahabat Pulau Indonesia – SPI)	Literacy and numeracy
		Dompet Duafa	Literacy and Leadership
		Tunas Aksara Foundation (YTA)	Literacy

INOVASI
Innovation for Indonesia's School Children
Australia Indonesia Partnership

Ratu Plaza Office Tower 19th Floor,
Jl. Jend. Sudirman Kav 9,
Jakarta Pusat, 10270
Indonesia
Tel : (+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 Australian Government
Palladium
MAKE IT POSSIBLE