Symposium Corresponding Author
Juliette Mendelovits
Australian Council for Educational Research
Juliette.Mendelovits@acer.edu.au

Symposium Chair
Jeaniene Spink
Australian Council for Educational Research
Jeaniene.Spink @acer.edu.au

Shaping Reading Pedagogy and Assessment for Sustainable Learning

1 SYMPOSIUM RATIONALE

Intergovernmental negotiations on the post-2015 Sustainable Development Goals will conclude in September 2015 and it looks likely that there will be an education goal underpinned by targets focused on improved and equitable learning outcomes. This will give new emphasis to reading development, as a foundation for successful learning and the futures it makes possible. To date, the recommendations of the 2000 US National Reading Panel (NRP) have played a large part in shaping ideas about reading pedagogy and assessment for children in developing as well as developed contexts. Although the NRP recommended treating comprehension as one of the five essentials of reading pedagogy, early reading interventions have been dominated by decoding. Further attention is warranted as to whether that emphasis is likely to be the most effective way to repair the main observed problem in developing contexts, namely comprehension deficits. To what extent is the framing of solutions for struggling readers in advanced economies applicable to contexts in which the gulf between home and school environment in culture, language and learning is much deeper and more pervasive?

This symposium assembles a number of studies that argue the importance of comprehension in reading pedagogy and assessment if improved learning outcomes from schooling are to be achieved. It addresses the subtheme questions concerning which assumptions and principles should underpin pedagogy and assessment for sustainable futures, and what the relationship should be between international assessments and local conditions and priorities.

2 SYMPOSIUM PAPERS

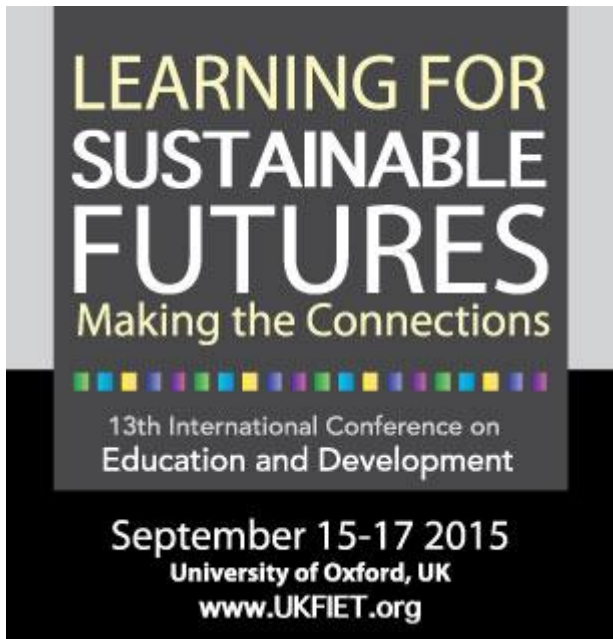### Paper 1 – The conditions of reading acquisition in contexts of low literacy

Prue Anderson	**Mary Fearnley-Sander**
Australian Council for Educational Research	Education Consultant
Prue.Anderson@acer.edu.au	mfearnleysander@mac.com

Paper 2 – Aligning reading assessment with national goals

Danielle Anzai
Australian Council for Educational Research
Danielle.Anzai@acer.edu.au

Maurice Walker
Australian Council for Educational Research
Maurice.Walker@acer.edu.au

Paper 3 – Monitoring reading globally for the post 2015 development goals

Juliette Mendelovits
Australian Council for Educational Research
Juliette.Mendelovits@acer.edu.au

Ross Turner
Australian Council for Educational Research
Ross.Turner@acer.edu.au

Ray Adams
Australian Council for Educational Research
Ray.Adams@acer.edu.au

Paper 1 - Prue Anderson
Australian Council for Educational Research
prue.anderson@acer.edu.au

Paper 1 - Mary Fearnley-Sander
Education Consultant
mfearnleysander@mac.com

Paper 1 - The conditions of reading acquisition in contexts of low literacy

1 ABSTRACT

A child's cultural, linguistic and cognitive distance from school language can make comprehension the critical issue in reading acquisition. Assessments of reading development tend to encounter ceiling effects in phonological skills within the first few years of schooling; but the gap in comprehension subsequently widens between advantaged and disadvantaged students in relation to school literacy.

The purpose of this paper is to show what kind of pedagogy, from early grades onwards, can result in 'beating the odds' for children in disadvantaged contexts. It draws on socio-linguistic theory, comprehension research and research into literacy development in a second language to show why the discontinuities between literate and social languages can affect the development of the cognitions necessary for word and text-level comprehension in the early grades; and the registers and content of academic genres from upper primary.

The development of comprehension bears on poverty reduction and access to sustainable futures because comprehension is associated with school success. The evidence is strong that particular home environments promote the kind of literacy that results in school success. However, research also shows that the success practices of such homes are replicable in schools in high poverty, low literacy contexts – not only in advanced economies but also in poor countries of Africa, Asia and the Pacific.

2 INTRODUCTION

The concern of development agencies about low levels of literacy acquisition has picked up pace over the last decade. Exemplifying the extent of the crisis in reading are the findings from EGRA Oral Reading Assessments in 20 African countries in the last five years that the 'majority of countries have more than 50% zero scores across all grades ... that is, more than half the tested

students could not read a single word' (Stern, 2015). The reasons these reading results cause such concern has been summarised by Abadzi (2008):

'In low-income countries children often learn to read in grades 4-6 (if ever) rather than grades 1-2
- Until they do, they cannot learn info from books

- They cannot solve written mathematics problems

- May drop out illiterate or graduate illiterate

- May get to grade 7 with reading speed of grade 3 and drop out of secondary. '

In response, reading interventions to date have been dominated by early grades decoding approaches to reading. However, further attention is warranted as to whether that emphasis is likely to be the most effective way to repair the main enduring problem in upper primary and secondary, namely comprehension deficits. This paper argues that in literacy-disadvantaged contexts, more, not less, attention has to be paid to the development of comprehension. It aims to show what kind of pedagogy and assessment from early years onwards can result in 'beating the odds' for children in disadvantaged contexts.

3 THE ARGUMENT OF THE PAPER

Data from international assessments and experimental studies have established the strong relationship between comprehension and home literacy. In particular PISA data shows 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 15' (OECD, 2012, p. 18). PISA found that, in almost all countries, students whose parents read books to them as they entered primary school are more likely to have higher reading scores at age 15. PISA data also show that practice of parents talking about books they read to their young children and using words in context is strongly associated with reading performance and even more with instilling a sense of enjoyment of reading in children.

These practices are not just a reflection of a family's socio-economic background and the resources available to the family (Hoff, 2006; Henning et al., 2010; Siraj-Blatchford 2010). Even among families with similar socio-economic backgrounds, PISA data show that better reading performance at the age of 15 is still strongly related to having had books read to them as young children[1].

Our view is that the development of reading comprehension skills requires extensive opportunities for children to understand the meaning of words typically found only in texts, and become familiar with the structures of text. We argue that this is best done through reading a wide variety of books to young children and interacting in talk that develops their understanding of abstract ideas. That is, children can learn how to interpret texts that an adult decodes for them well before they learn to decode for themselves.

The low literacy of many home environments in poor developing contexts means that school may be the only place where children have a chance to listen to and talk about written texts. 'Beating

---

[1] However, on average, socio-economically advantaged parents are 14 percentage points more likely to have engaged in the kinds of activities that are associated with positive outcomes for their children, such as reading books to their very young children

the odds', therefore, is about a pedagogy that can replicate in the schools and pre-schools of poor communities 'what school-oriented homes have'(Heath ,1982, p. 50).[2]

An inadequate background in comprehension at the start of school does not show up until later primary years. It is around upper primary that the effect of comprehension difficulties manifests itself in a slump in reading results. This is when learning begins to be text-based and literate language to differ most unrelentingly from spoken language. But the fact that comprehension deficit manifests as a critical issue in upper primary does not mean that it is not also critical from the time of emergent literacy.

The paper also raises the question of whether reading by Grade 4 or 5 in developing contexts constitutes a crisis. Expecting these children to be reading by the end of Grade 2 does not take account of the fact that by Grade 2 the majority of children in advanced economies have had 4-5 years of formal literacy preparation, (pre-school, kindergarten and primary foundation classes). They live in print-saturated environments, and have a far greater likelihood of being read to at home. For many students in developing contexts, Grade 2 is only the second year of any form of schooling. Many students start school unable to understand the language they are being taught in; and, with exposure limited to the classroom, take 2-3 years to understand it. Tardiness in reading acquisition has been viewed as a doom because of 'the Matthew effect' – the ever widening gap that is consequence of not reading well enough to read a lot (Stanovich, 1986, pp. 364-5). But is this concept relevant in developing contexts where the whole cohort follows the same trajectory?

4 COMPETING ACCOUNTS OF READING ACQUISITION

Both psychologists and socio-linguists have inferred from literacy disadvantage the need for a pedagogy for the poor, though with quite different orientations.

Psychologists privilege the mastery of the skills of the alphabetic principle: developing accurate grapheme-phoneme correspondences to enable rapid and accurate word identification in memory (Ehri, 2005, p. 137). Since the 2000 conclusions of the US National Reading Panel (NRP) on the evidence for effective early reading practice, priority has been given to decoding (Kuhn, 2010, p.242). Though vocabulary and comprehension were included in the NRP skills, in fact phonological awareness, the alphabetic principle and automaticity of word recognition have been dominant in the component skills approach to reading. The authors of the Dynamic Indicators of Basic Early Literacy Skills (DIBELS) program, which has been the most influential and widespread implementation of this approach in the US, cites those three skills as the important things to measure (Good et al., 2001, p. 7; cited in Kuhn, 2010, p. 241). The focus of instructional effort on meaning making and comprehension is viewed as the appropriate concern of later years, the job 'of reading to learn' after children have 'learnt to read'. Yet Snow, the lead NPR researcher, in fact advocated the *integration* of the alphabetic principle and activities supporting comprehension 'so that these two aspects of skilled trading should be going on *at the same time*' (Snow & Juel, 2005, p. 510).

Socio-linguists by contrast have emphasised the issue of a different kind of 'code' that has to be cracked to succeed in literacy: the elaborated language code of school literacy. The 'social gradient' of results such as those in PISA have shown that this code aligns better with some

---

[2] 'Beating the odds' summarises research connected with the UK's Effective Pre-school and Primary Education program (EPPE), showing disadvantaged families engaging in reading practices that result in school success (Siraj-Blatchford, 2010).

student backgrounds than with others – namely families whose language usages with children are closer to the language of school (Freebody, 2007; Bernstein, 1971; 1990; Gee, 2002; Halliday, 1994; Christie, 1999; Martin,1992).

Nevertheless the socio-linguistic explanation of literacy outcomes cannot replace the value of teaching children decoding strategies. As Purcell-Gates points out, 'however culturally specific literacy practices may be... the underlying symbol systems of written language are not tied to them' (2006, p. 127). There is also no contesting the correlation between automatic word recognition and comprehension (Stanovich, 1986; Ehri, 1991; Perfetti, 1999). It is in regard to the neglected domains of beginning reading – vocabulary and the constituent skills of comprehension – that the socio-linguistic account has force.

5 DEFINING COMPREHENSION

Accompanying the spread of early grades assessments in the measurement of the 'learning crisis' (EFA Global Monitoring Report, 2012, p. 130) is the advocacy of a matching pedagogy that focuses on the implementation of decoding skills.

Some decoding advocates write as if comprehension looks after itself if decoding skills are strong, early enough – by the end of Grade 2. They infer that a student able to read with sufficient speed and accuracy to be considered fluent, must also be able to understand the texts they can read aloud (World Bank Vanuatu EGRA Report, 2010).

But it been not been established that *automaticity* of decoding results in a fluency that is processing meaning at the same time. In fact Perfetti, a strong advocate of the role of phonological processing in reading ability, acknowledges 'when we look beyond the correlations working memory is not the critical factor in comprehension, or at least not the only one' (2005, p. 234).

The reduction of fluency to measures of speed and accuracy is also contested in the reading literature (Pressley, 2007). In phonetically transparent languages children can decode fast without understanding a word. If fluency is used as an indicator of reading comprehension it needs to exclude meaningless 'barking at print'. The definition of fluency of the US National Assessment of Education Progress includes 'expressiveness' which reveals whether the reader's prosody has captured the author's intention. Expressiveness – meaning-making – is integral to measures of fluency (Kuhn et al., 2010, p. 240).

Empirically, reading research also indicates that the relation between children's basic reading skills – that is, rapid word reading – and reading comprehension diminishes as children age and decoding skills become more automatised (Schwanenflugel et al., 2006; Paris, 2005, Droop & Verhoeven (2003).

Knowing the meaning of words entails knowing their nuanced meaning over a range of contexts (Perfetti et al., 2005, p. 241). It is this feature of vocabulary knowledge, as well as the greater number of words known, that differentiated children in Hart and Risley's famous study of the vocabulary differences of children from high income and welfare-supported families.

Another factor distinguishing skilled from poor comprehenders is what Perfetti calls the reader's standard of text coherence. This is the extent to which he or she reads for understanding, makes

inferences and monitors what is read for sense (2005, p. 233). Perfetti identifies sensitivity to text structure as a key variable in explaining differences in levels of comprehension (2005, p. 241-2). Developing sensitivity to text structure is what literacy educators of the socio-linguistic school have tried to do for disadvantaged students who have not acquired familiarity with text (Martin, 1992; 2014; Christie, 2005; Freebody, 1992).

6 ISSUES IN THE MEASUREMENT OF COMPREHENSION IN DEVELOPMENT

6.1 Contexts

The lack of clarity about the place of comprehension in early reading is matched by the confusing literacy data in development contexts. So far, most assessments do not tell us enough about what is being measured as reading comprehension to really understand what attainment levels are.

Table 1 illustrates this through a case study of assessment in Lao PDR. It shows the results of four different assessments: the 2013 Lao PDR government national Assessment of Learning Outcomes (ASLO) at Grade 3, supported by the World Bank; a PLAN-supported ASLO in two districts in Bokeo province; a 2011/2012 PASEC country assessment for Lao PDR at Grade 2 and at Grade 4; and a 2013 provincial Save the Children Literacy Boost Grade 2 baseline assessment. The two sub-national assessments targeted language minority populations, which altogether constitute around 45% of the population of Lao PDR. The national ASLO data is shown in the first column with disaggregated data for two provinces shown in the last column.

ASLO only provides an aggregate language score combining speaking, listening, reading and writing. PASEC combines oral and written comprehension, and writing at Grade 4.

Table 1: Differences of outcome by types of literacy assessment in Lao PDR

National ASLO (2013), Grade 3	Provincial ASLO (Bokeo) 2012, Grade 3	National PASEC 2011/2012, Grade 2 and Grade 4	Provincial Literacy Boost 2013, Grade 2 Save the Children	ASLO (2013) Grade 3 Attapeu Bolikhamxay
Aggregate result for Lao Literacy3: 83% functional or above: —59% functional —24% independent	Mean scores: (50 = minimum success level) Listening comprehension: 45 Consonant/vowels: 42 Vocabulary: 7 Reading comprehension: 24	Grade 2: 86% expected competencies

Grade 4: 41% expected competencies: —15% understand a text and write respecting syntax —26% analyse a long text with unknown vocabulary, * give their opinion in writing * mastered orthography and syntax. | Reading with comprehension (8 questions) 4% Attapeu province 2% Bolikhamxay province | Attapeu 93% functional or above: —28% independent —65% functional

Bolikhamxay 79% functional or above: —27% independent —52% functional |

---

3 ASLO reading comprehension is measured by 12 questions.

The most arresting aspect of these assessments is their extraordinary range, from two to four per cent of students able to read at Grade 2 in two minority provinces according to Literacy Boost, to national ASLO data for the same provinces claiming that 79 to 93 per cent of students were functionally literate or higher in Grade 3.

One explanation is that ASLO is an aggregate language score and the Literacy Boost focussed on reading. The low level of vocabulary knowledge in the Bokeo assessment (mean score of 7 – with 50 the minimum for success) suggested to the test analyst that students possibly did not even understand the test instructions. Students need enough language to understand the instructions, if the data collected is to be meaningful.

In the Literacy Boost assessment of letter identification the 'most difficult letters were combination letters that indicate tonal changes ... foundational learning to read words correctly' (Literacy Boost report, p.9). EGRA administration in Lao PDR in 2012 did not even assess the early grades; sampling from Grade 3 upward[4] possibly indicating that the trajectory of reading a phonemically and visually complex language – and a second language for nearly half of Lao children – does not conform to the Grade 2 norm for reading.

6.2   Validity issues with assessment data

The technical quality of some large-scale assessments of early reading is problematic. Tests need to be scrutinised to ensure correct or incorrect answers to questions do not reflect guess work, or general knowledge. An unpublished (2014) review for DFAT of the task quality in EGRA, PILNA, ASER, PIRLS and Pre-PIRLS, SACMEQ and Uwezo found that the tasks in EGRA, pre-PIRLS and PIRLS are generally of high quality while some of the other assessments contain some technical flaws.[5]

High quality reading assessments such as PISA and PIRLS provide a benchmark of quality that is not always reflected in national assessments of reading in developing countries.

As one of the sets of tasks in its toolkit, EGRA has five reading comprehension questions based on one text. This is broadly indicative, but a more comprehensive measure of students' emerging comprehension skills could begin with matching single words to pictures and interpreting some simple sentences before requiring students to interpret a 50 word passage.

The validity of the EGRA comprehension score is also confounded by the concurrent fluency measure. Students are given one minute to read aloud a 50 word passage. The passage is removed. Students are scored for fluency and asked questions about the passage up to the point where their reading aloud was stopped. Early readers' fragile capacity to decode and concurrently process the meaning of the text is likely to be overwhelmed by the demands of pronunciation and prosody as well as having to remember the passage. Tests such as PISA and PIRLS Literacy do not impose these kinds of read-aloud demands or remove texts after one minute and limit access to questions.

Improving the quality of assessments of early reading used in developing countries would help us to better understand the nature and extent of the problems and monitor the effectiveness of interventions.

---

[4] Results of the Lao EGRA test are not publicly available.
[5] The views expressed here are those of ACER as the authors, not DFAT

7 WHY IS THERE VARIATION IN STUDENTS' COMPREHENSION ACROSS DEVELOPED AND DEVELOPING COUNTRIES?

7.1 The iceberg theory

Remedial intervention for poor reading performance is appropriately focussed on strengthening the early grades. Educators who focus on teaching decoding skills and infer that comprehension success will follow achieving fluency can be missing the hidden iceberg. Figure 1 illustrates the hidden icebergs of literate practices below the line of classroom visibility of three different country contexts.
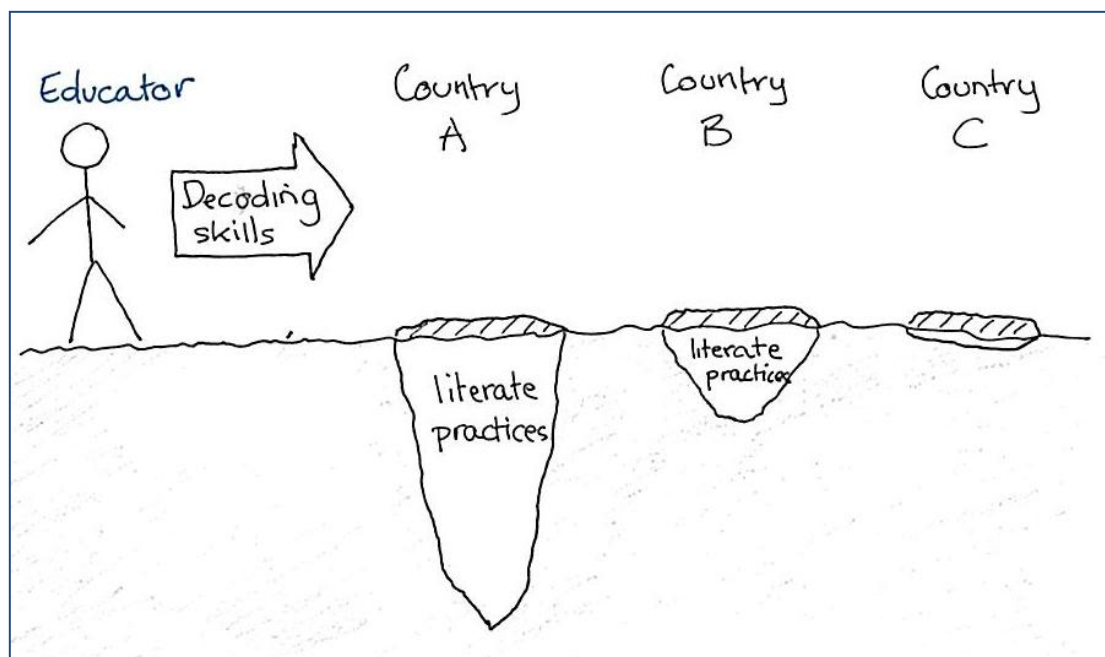

Figure 1: What the educator does not see

In this scenario, the eventual reading success of most students in Country A (say, an OECD country) is attributed to the effectiveness of the decoding program alone and fails to acknowledge the contribution of the literate practices students learned at home.

Consequently, students in Countries B (a lower middle income country) and C (a least developed country) are prescribed more and more decoding practice and the missing substratum of literate practices that these students need goes unnoticed. By the time these students eventually learn to decode fluently, years of valuable time have been wasted that could have been used to concurrently develop both decoding skills *and* literate practices.

Table 2 brings out the role that the 'iceberg' plays in comprehending the connected text of books. It compares the differences in cognitive demands for a young child processing an everyday oral conversation, with listening to an adult reading a story-book.

Table 2: Comparison of an oral conversation and listening to a story book

Oral conversation	Sharing an authentic story-book
Meaning typically contextualised by the immediate situation/speaker Explicit links usually made to listener's world knowledge	Meaning of text is de-contextualised: could relate to any situation including imaginary scenarios Listener has to link ideas in the text to their world knowledge
Responsive to the listener's needs and interests.	Monologue; no adjustment to reader
Short bursts of text are processed Listener can process bit-by-bit Can request clarification. Often much repetition.	Large slabs of text of text have to be processed Listener has to focus on main ideas and differentiate from details May be long and/or have complex sentences Listener has to maintain coherence Minimal repetition
Meaning supported by facial expressions, gesture and tone, volume, pace.	Meaning is mainly in words and may be supported by illustrations Reader's prosody supports meaning
Mainly explicit, literal, reinforced meaning	Meaning has some explicit elements but also may be layered or subtle
Simple, everyday vocabulary	A lot of less familiar vocabulary
Ephemeral: not easily re-examined or repeated verbatim to allow close examination of meaning	Permanent: can be closely examined to identify fine distinctions in meaning

In summary, interpreting the meaning of an oral conversation is less demanding, with more supports available for the listener, than interpreting the meaning of a story-book, which has much greater complexity and fewer supports. In respect of vocabulary, children's books typically contain fifty per cent more rare words than prime time television or even college students' conversations (Hayes and Ahrens, 1988).

Children have to develop their capacity to process the meaning in written texts. Parents who read to their children begin with very simple texts and focus on engaging the child. They gradually introduce more complex texts as the child is ready. Young children, who have been regularly read to, are capable of enjoying and interpreting quite sophisticated story books, like that described in Table 2, well before they go to school.

8 WHAT SUPPORTS COMPREHENSION

The following factors have all been identified as supporting reading comprehension.

8.1 Being read to and targeted conversations

Duursma, Augustyn and Zuckerman (2008) reference ample research which demonstrates that reading aloud to young children promotes the development of language and other emergent literacy skills, which in turn help students prepare for school. Adams (1990), Senechal, LeFevre (2002), Snow, Burns and Griffin (1998) and Storch and Whitehurst (2001) all found that shared book reading also promotes understanding of print concepts such as how to navigate texts, familiarity with the written language register and story structures.

The quality of the conversations between adults and young children is clearly linked to early reading success. Neuman (2004) found that children identified as early readers when entering school had educational contexts filled with lively, interactive conversation. The National Early Literacy Panel (2008) identified the importance of creating language-rich environments in early

childhood and of adults scaffolding this development through the integration of 'book language' in adult-child conversations. Conversations about quality children's literature in particular provide exposure to abstract language and abstract concepts not commonly experienced in typical adult-child conversations (Pollard-Durodola et al, 2011).

8.2 Contingent responsiveness

Sensitivity to the level of skill and the interests of the child determines the effectiveness of shared book conversations. Mol and Neuman (2013, p. 408) found that parents who were able to calibrate their interactions to children's interests and follow their lead through responsive interactions had the most significant effect on the development of their children's language skills. The effect negated the influence of socio-economic status on language production. Discussions that focused on describing and labelling illustrations during the reading of the story were found to be more effective in building vocabulary and print skills for children with initial low levels of vocabulary than discussions that focussed on the meaning of the story (Mol & Neuman, 2013, p. 408).[6]

8.3 Making inferences and contextualising vocabulary

The ability to infer information from an early age is particularly critical. Children who are able to make inferences at four to five years of age have improved listening comprehension skills at six years of age; and according to Vellutino listening comprehension predicts concurrent reading comprehension in 7 to 12 year-old children (Vellutino, Tunmer, Jaccard, & Chen, 2007).

The same study identified vocabulary as more important than grammar or short-term memory in supporting the ability of four to five year-old children to make inferences (Silva & Cain, 2015). The report of the National Institute of Child Health (2000), affirmed that teaching words in context is more effective than teaching isolated words. Repeated reading of storybooks is more effective in developing students' vocabulary than repeated testing of isolated vocabulary words with feedback (Damhuis, Segers & Verhoeven, 2015).

8.4 Explicit teaching of comprehension skills

Comprehension strategies need to be taught. These strategies are essentially the same from early childhood through to adulthood and include making predictions, relating the text to personal knowledge, checking the sense and coherence of ideas, retelling, enacting or visualising key ideas, adopting the vocabulary of the text to expand and elaborate ideas and critically reflecting on the text. Brown, Palinscar and Armbruser highlight the effectiveness of explicit comprehension instruction with children with adequate decoding skills, but low comprehension. A critical feature of this study is that these skills were modelled and embedded in the study of texts and not taught as isolated skills (2004).

---

[6] However children with higher initial levels of vocabulary profited more from discussions that focussed on the meaning of the story conducted after the story reading. Failing to increase the level of cognitive demand as children are ready may be counter-productive as 'peppering' children with low cognitive demanding questions about texts may distract from their learning and may not be supportive of language development.

8.5 Reading frequently

Finally, along with many other studies, the IEA study of Reading Literacy (IEA, 1991) demonstrated that reading improves comprehension: the amount of voluntary book reading that students did outside of school was positively related to their achievement levels.

8.6 A pedagogy for text comprehension

In Figure 2 we propose a model for teaching reading for comprehension. Figure 2 sets out four different skills involved in learning to read with understanding and their inter-relationships.
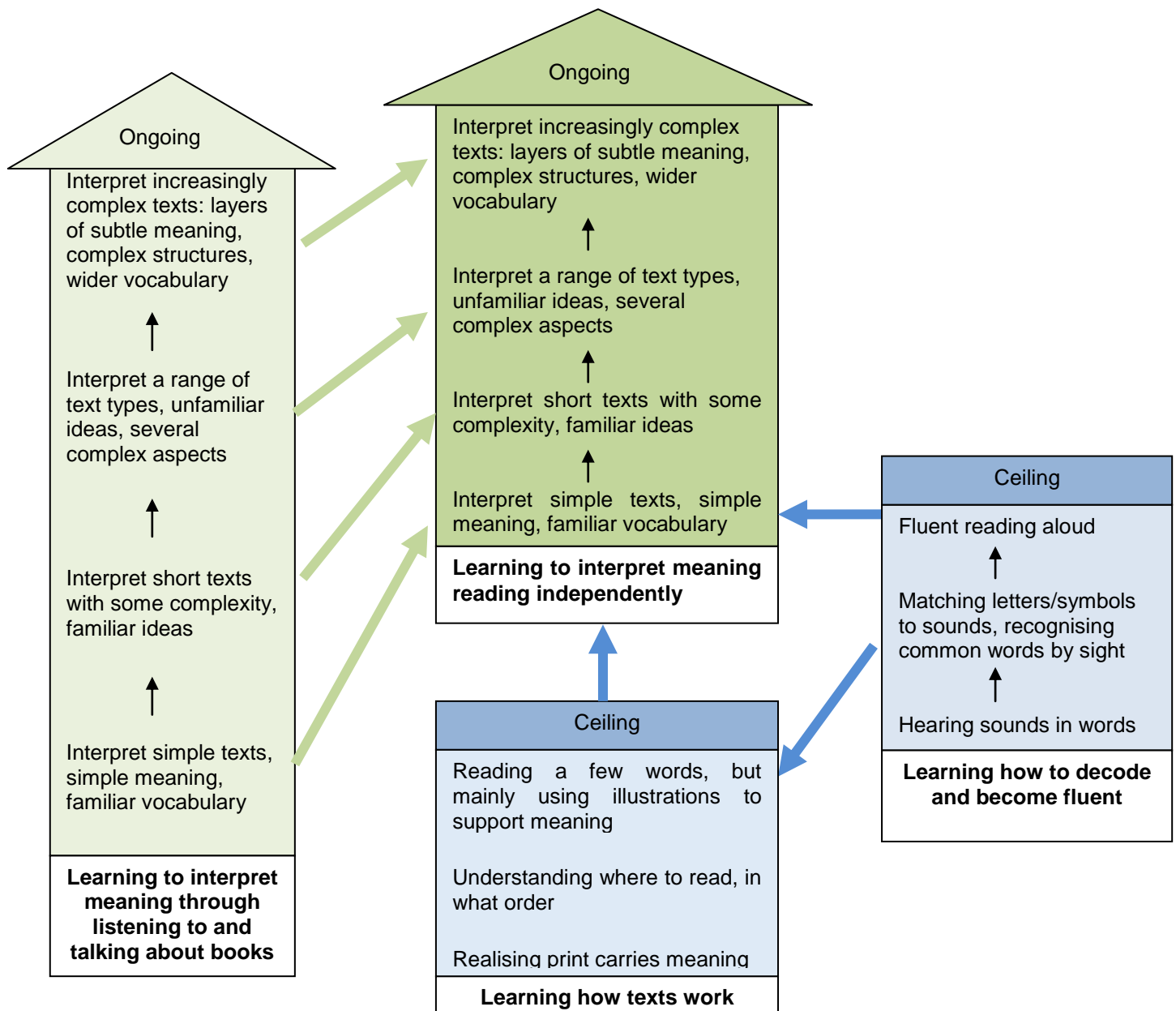
Figure 2: The inter-relationship of early reading skills

Beginning on the left, the model shows that students can start learning how to interpret meaning from texts through listening to and talking about books when this is tailored to their level of skill. The comprehension skills they develop will help them understand the texts they read

independently when they become fluent decoders. This is an ongoing skill as students can always benefit from listening to and discussing increasingly complex texts.

The lower centre pillar shows that students need to learn how texts work. Simple concepts about print are easily learned with explicit modelling and instruction and once understood, require no further instruction. Students who are learning these skills typically 'read' texts by interpreting the meaning of the illustrations alone or memorize the text as part of their emerging understanding of how texts work.

On the right, the model shows students need to learn to decode and develop fluency. They practice this by reading very simple books aloud. Once fluency is attained, further instruction is no longer required.

The upper centre pillar shows the growth of students' ability to read independently with understanding. This growth is supported by students having attained fluency and by having developed comprehension skills through listening to and talking about books. Development of this skill is ongoing as students read increasingly complex texts.

9 HOW COMPREHENSION IN LOW LITERACY ENVIRONMENTS COULD BE IMPROVED

The analysis of what supports comprehension shows why there are no icebergs beneath the tips of reading capability in poor Country B and C classrooms. The challenge of improving comprehension under these circumstances could not be greater. It throws some light on why efficient learning for the poor might focus on decoding as an achievable improvement (Abadzi, 2006). It may explain why compared with decoding interventions there are so very few initiatives for improving children's vocabulary and literate practices at the early primary level.

Those that do exist and that have been well evaluated, report effectiveness: the Shishuvachan study in Mumbai schools (He et al, 2009); Costa Rica kindergartners (Rollo et al. 2006); the Sa Aklat Sisikat program in the Philippines (Abeberese et al. 2013); South African reading programs (Sailors & Hoffman 2014; Pretorius & Mampuru, 2007).

There are systemic initiatives in developing contexts, but they are in pre-schools. Supporting agencies include the World Bank, UNICEF, Save the Children and other international NGOs as well as local NGOs. A feature of these approaches is their engagement with communities, parents and homes; and the establishment of book supplies and centres for borrowing, parent literacy programs and local book production. The pre-school curriculum frameworks put strong emphasis on adults reading with children, talking and listening, rich print environments as well as emergent alphabet knowledge. (UNICEF 2012; Save the Children, 2013; World Bank, 2015).

There is evidence of the positive impact of pre-school or school readiness on disadvantaged students' language capabilities in advanced economies. Programs such as the national Head Start in the US, specifically designed to narrow the achievement gap for disadvantaged children, have shown effect sizes of .09 for vocabulary gains sustained through to Grade 1 (Head Start Impact Study 2010). Longitudinal studies of cohorts from the 1970s show differences in earning capacities over time and other indicators of social success that are plausibly ascribed to Head Start (Gibbs et al., 2013). The research program Effective Provision of Pre-School Education (EPPE) (2004) reviewed all pre-school provision in the UK, and showed the beneficial effects of high quality

literacy experiences on children's literacy at the end of Years I and 2; and that this kind of experience was particularly advantageous to disadvantaged students.

Given this evidence, there are grounds for optimism that the effects of literate home environments on vocabulary and literate practices can also be reproduced in Country B and C environments. The large donor investments in these contexts are testimony to their confidence in the difference that early childhood care and education can make to children's literacy outcomes.

The problem for low income countries, however, is that very few of the really disadvantaged children can access pre-school programs, a sub-sector which in most cases is not resourced by governments. A compounding problem is that only high quality school readiness programs produce the desired effects, so inadequately resourced options are not likely to be effective.

Under these circumstances, it is all the more surprising that comprehension-supporting approaches have not been extended into primary schools. This is even more puzzling given the language contexts of many countries, where the conventional curricula can only start to be effective after children have developed some competence in the language of instruction.

Closer cooperation by donors in pre-school and school programs at a policy level would ensure continuity in a developmental curriculum that bridges pre-school and school. If institutionalising literate practices in government pre-school provision is feasible with often untrained teachers, the impact could be even greater in early primary with trained teachers (GMR, 2015, p. 68).

In summary the argument of this paper is that it is linguistically and culturally feasible to provide support for the early development of students' vocabulary and comprehension skills in the poor contexts of developing countries. The problem is one of raising awareness of the differences in these learning environments from those in advanced economies; and in identifying appropriate institutional approaches in contexts where many students need to acquire literacy and school language at school.

10 REFERENCES

Abadzi, H. (2008). Conceptual Framework for Early Reading Fluency: Insights and Monitoring Indicators from Cognitive Neuroscience. Presentation for a USAID Education Sector Professional Development Course on early grade reading, Washington, DC, 12 May, 2008. Retrieved from https://www.eddataglobal.org/documents/index.cfm?fuseaction=pubDetail&ID=140.

Abadzi, H. (2006). Efficient Learning for the Poor: insights from the frontiers of cognitive neuroscience. World Bank Washington, DC, USA.

Adams, M. J. (1990). *Beginning to Read*, MA: MIT Press.

Bernstein, B. (1971). *Class, Codes and Control,* vol.1. London: Routledge & Kegan Paul.

Bernstein, B. (1990). *The Structuring of Pedagogic Discourse: Class, Codes and Control,* vol.4. London: Routledge.

Brown, A.l., Palinscar, A.S., & Armbruster, B.B. Instructing comprehension fostering activities in interactive learning situations in Ruddell, R.B., Unrau, N.J. (Eds.) (2004). Theoretical Models and Processes of Reading Fifth Ed. International Reading Association Inc, USA.

Christie, F. (Ed.). (1999). Pedagogy and the Shaping of Consciousness London: Continuum Press.

Christie, F. (2005). *Language Education in the Primary Years Sydney*: University of New South Wales Press

Confemen 2013. Les résultats des évaluations PASEC en RDP Lao Retrieved 2 March 2015 http://www.confemen.org/4221/les-resultats-des-evaluations-pasec-au-cambodge-et-en-rdp-lao/

Damhuis, M. P., Segers E., & Verhoeven L., (2015). *Stimulating breadth and depth of vocabulary via repeated storybook readings or tests* in *School Effectiveness and School Improvement*, 26 (3), 383-396, Behavioural Science Institute, Radboud University, Nijmegen, The Netherlands.

DFAT. (2014). Evaluation of Reading Assessments: Desk-based research, Education Resource Facility, unpublished report commissioned by the Department of Foreign Affairs and Trade, Australia, May p7.

Droop, M. & Verhoeven, L. (2003). Language proficiency and reading ability in first- and second-language learners.*Reading Research Quarterly*, vol. 38, 1, 78–103.

Duursma E. Augustyn M., & Zuckerman B. (2008). *Reading aloud to children: the evidence*. Arch Dis Child, July; 93 (7) p554.

Ehri, L. (1991). Development of the ability to read words. In R. Barr, M.L. Kamil, P.B. Mosenthal & P.D. Pearson (Eds.), *Handbook of reading research*, vol. 2, 383–417. New York: Longman.

Ehri, L. (1999). Phases of development in learning to read words. In J. Oakhill & R. Beard, (Eds.) *Reading Development and the Teaching of Reading,* 79-108. Oxford Blackwell.

Freebody, P. (1992). A socio-cultural approach: Resourcing four roles as a literacy learner. In A. Watson & A. Badenhop (Eds*.), Prevention of reading failure*, 48-60. Sydney: Ashton-Scholastic.

Freebody, P. (2007). Literacy Education in School: Research perspectives from the past, for the future. Victoria: Australian Council for Educational Research.

Gee, J.P. (2002). A socio-cultural perspective on early literacy development. In Handbook of Early Literary Research. New York, NY: The Guilford Press.

Gibbs. C, Ludwig, J. & Miller, D. (2013). *Head Start Origins and Impacts in Legacies of the War on Poverty*. M.Bailey, & S. Danziger, (Eds.) Russell Sage Foundation.

Good, R.H., Kaminski, R.A., Simmons, D., & Kame'enui, E.J. (2001). Using Dynamic Indicators of Basic Early Literacy Skills (DIBELS) in an outcomes-driven model: Steps to reading outcomes. Oregon School Study Council, 44, 1, 6–24.

Hayes, D. P., & Ahrens, M.G. (1988). Vocabulary simplification for children: A special case of 'motherese'? *Journal of Child Language*, 15, 395-410.

Halliday. M.A.K. (1994). *Introduction to Functional Grammar*. London: Arnold.

Hart, B. & Risley, T. (2003). The Early Catastrophe: The 30 Million Word Gap by Age 3. *American Educator*. Spring: 4-9.

Heath, S.B. (1982). What no bedtime story means: narrative skills at home at school. *Language in Society*, 11, 49-76.

Henning, C; McIntosh, B., Arnott, W. & Dodd, B. (2010). Long-term outcome of oral language and phonological awareness intervention with socially disadvantaged preschoolers: the impact on language and literacy. *Journal of Research in Reading,* 33, 3, 231-246.

Hoff, E. (2006). How social contexts support and shape language development. *Developmental Review*, 26, 55-88.

IEA Study of Reading Literacy: Key Findings (1991). Retrieved from http://www.iea.nl/reading_literacy_study.html

Kuhn. M, Schwanenflugel, P. & Meisinger, E.B. (2010). Aligning Theory and Assessment of Reading Fluency: Automaticity, Prosody, and Definitions of Fluency. *Reading Research Quarterly*, 45, 2, April/May/June, 230–251.

Lao PDR Ministry of Education, (2012). Report of the National Assessment of Student Learning Outcomes (ASLO III) for Primary Grade 3.

Martin, J. (1992). *English Text: system and structure*. Amsterdam: John Benjamins Publishing Company.

Martin, J. (2014). Looking Out: Functional Linguistics and Genre. *Linguistics and the Human Sciences*, 9(3), 307-321.

Mol, S.E., & Neuman, S.B. (2014). Sharing information books with kindergarteners: The role of parents' extra textual talk and socioeconomic status in *Early Childhood Research Quarterly* 29, 399-410.

Morgan, P. & Meier, C.R. (2008) Dialogic Reading's Potential to Improve Children's Emergent Literacy Skills and Behaviour. *Preventing School Failure* 52: (4), 11-16.

National Early Literacy Panel (2008) Developing Early Literacy: Report of the National Early Literacy Panel, a scientific synthesis of early literacy development and implications for intervention. Washington DC: National Institute for Literacy.

National Institute of Child Health and Human Development (2000) Report of the National Reading Panel; Teaching children to read: An evidence-based assessment of the scientific literature

on reading and its implications for reading instruction: Reports of the subgroups (NIH Publication no. 00-4754) Washington. DC. U.S. Government Printing Office.

Neuman, S.B. (2004). The effect of print rich classroom environments on early literacy growth. *The Reading Teacher*, 58, 89-91.

NICHD Early Child Care Research Network (2005). Pathways to reading: The role of oral language in the transition to reading. *Developmental Psychology*, 41, 428-442.

OECD. (2012). *Let's read them a story: The parent factor in education*, PISA, OECD Publishing p.19. Retrieved from http://dx.doi.org/10.1787/9789264176232-en

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

Perfetti, C. (1999). Cognitive Research and the Misconceptions of Reading Education. In J. Oakhill & R. Beard, (Eds.) *Reading Development and the Teaching of Reading* 79-108. Oxford Blackwell, 42-59.

Perfetti, C., Landi, N. & Oakhill, (2005). The Acquisition of Reading Comprehension Skill. Snowling M & C. Hulme, (Eds.) *The Science of Reading: A Handbook*. Blackwell: UK. 227-248.

Pollard-Durodola, S. D. Gonzalez, J. E. Simmons, D. C. Kwok, O., Taylor, A. B., Davis, M. J. et al. (2011). The effects of an intensive shared book reading intervention for pre-school children at risk of vocabulary delay. *Exceptional Children*, 77, 161-183.

Pressley, M. & Fingeret, L. (2007). What we have learned since the National Reading Panel: Visions of the next version of Reading First. In M. Pressley, A.K. Billman, K.H. Perry, K.E. Reffitt, & J.M. Reynolds (Eds.) *Shaping literacy achievement: Research we have, research we need.* New York: Guilford.

Pretorius, E. & Mampuru, D. (2007). Playing football without a ball: language, reading and academic performance in a high-poverty school. *Journal of Research in Reading*, 30, 1, 38–58.

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

Sailors, M. Hoffman, J. Pearson, D. McClung, N., Shin, J, Mwale Phiri L., & Saka, T. (2014). Supporting Change in Literacy Instruction in Malawi. *Reading Research Quarterly*, vol. 49, issue 2, 209-231, April/May/June.

Save the Children. *Literacy Boost Lao PDR Baseline Report*. (2013). Retrieved from http://resourcecentre.savethechildren.se/library/literacy-boost-lao-pdr-baseline-report-2013.

Beyond School Walls. (2013) Retrieved from http://www.savethechildren.org/atf/cf/{9def2ebe-10ae-432c-9bd0-df91d2eba74a}/

Schwanenflugel, P.J., Meisinger, E., Wisenbaker, J.M., Kuhn, M.R., Strauss, G.P., & Morris, R.D. (2006). Becoming a fluent and automatic reader in the early elementary school years. *Reading Research Quarterly*, 41 (4), 496–522.

Silva, M., Cain, K. (2015). The relations between lower and higher level comprehension skills and their role in the 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.

Siraj-Blatchford I., & Sylva K. (2004). Researching pedagogy in English Preschools. *British Education Research Journal*; 30 (50), 713-730

Sylva,K. et E. Melhuish, Sammons, P. Siraj-Blatchford, I., & Taggart, B. (2004). The Effective Provision of Pre-School Education (EPPE) Project: Findings from Pre-school to end of Key Stage 1.

Senechal M., & LeFevre, J., (2002).Parental involvement in the development of children's reading skill: a five year longitudinal study. *Child Dev*; 73, 445-60.

Snow, C E., Burns S., & Griffin P., (1998).*Preventing reading difficulties in young children*. National Academy Press: Washington, DC.

Snow, C. & Juel, C., (2005). Teaching Children to read: what do we know about how to do it? In Snowling M & C. Hulme, (Eds.) *The Science of Reading: A Handbook*. Blackwell: UK, 501-521.

Stanovich, K.E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Psychology Today*. vol. 21, No. 4 Autumn, 360-407

Stern, J. (2015). The state of literacy in Africa.Presentation to the Comparative and International Education Society Annual Conference, Washington, DC March 10. Retrieved 29 July 2015. https://www.eddataglobal.org/documents/index.cfm?fuseaction=pubDetail&id=788.

Storch, S.A., & Whitehurst, G J. (2001).The role of family and home in the developmental course of literacy in children from low-income backgrounds. In: Britto PR. Brooks-Gunn J, eds. *New directions in child development: The role of family literacy environments in promoting young children's emerging literacy skills*. San Francisco, CA. Jossey-bass/Pfieffer, 53-71.

UNESCO. (2015). *EFA Global Monitoring Report: Education for All 2000-2015*. Achievements and Challenges. France: UNESCO Publishing.

UNICEF. (2012). *School Readiness: A Conceptual Framework*. Retrieved 16 May 2015. http://www.unicef.org/education/files/Chil2Child_ConceptualFramework_FINAL(1).pdf.

US Department of Health and Human Services. (2000). Teaching children to read: An Evidence-Based Assessment of the Scientific Research Literature on Reading and Its Implications for Reading Instruction. Report of the National Reading Panel.

US Department of Health and Human Service Office of Planning Research and Evaluation. (2010). *Head Start Impact Study*.

Vellutino, F. R. Tunmer, W. E. Jaccard, J. J., & Chen, R. (2007). Components of Reading Ability: Multivariate evidence for a convergent skill model of reading development. *Scientific Studies of Reading*, 11, pp3-32, doi:10.1207/s1532799xssr1101-2.

Whitehurst, G. J. Galco, F.L. Lonigan, C. J. Fischer, J.E. DeBaryshe, B. D. Valdez-Menchaca, M.C. et al. (1988). Accelerating Language development through picture book reading. *Developmental Psychology* 24, 552-559.

World Bank. (2012). Vanuatu Early Grade Reading Assessment; Tonga Early Grade Reading Assessment. Washington, DC. The World Bank.

World Bank. (2015). Early Childhood Development. *World Development Report 2015*, Washington, D.C. The World Bank

Wasik, B. A., Bond, M. A., & Hindman, A. (2006). The effects of a language and literacy intervention on Head Start children and teachers. *Journal of Educational Psychology*, 98, 63–74. doi:10.1037/0022-0663.98.1.63.
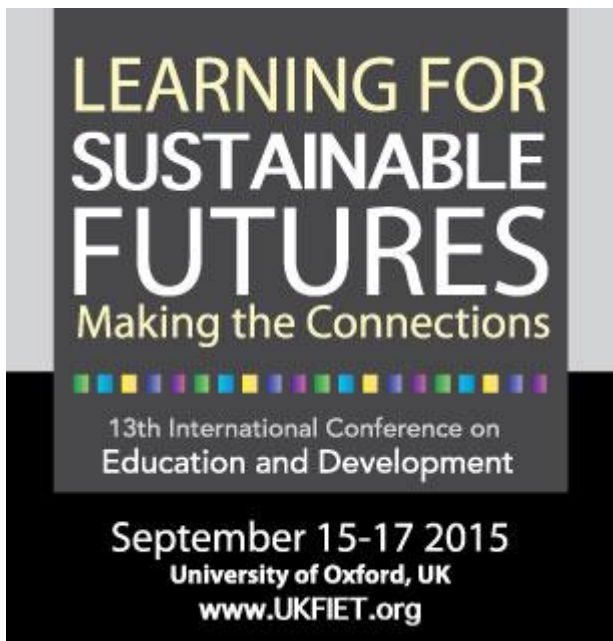
Paper 2 - Danielle Anzai

Australian Council for Educational Research
danielle.anzai@acer.edu.au

Paper 2 - Maurice Walker

Australian Council for Educational Research
maurice.walker@acer.edu.au

Paper 2 - Aligning reading assessment with national goals

1 ABSTRACT

Monitoring Trends in Educational Growth (MTEG), initiated in 2012, is an international assessment program especially appropriate to those countries where education systems are in rapid development. Key aims of the program are to provide policy-relevant information about learning and the factors related to it, focusing on the needs expressed by the country; and to track growth in learning over time. Developed in partnership with the Afghanistan Ministry of Education, MTEG was administered at Grade 6 in 2013, and will be administered at Grade 3 later in 2015. The assessment of reading at Grade 3 draws on the five components of reading recommended in the 2000 US NRP Report, but with a strong emphasis on listening and reading comprehension. The MTEG approach in the assessment of reading is distinguished from that of other reading assessment programs that operate in low income countries: the focus of these programs is largely, sometimes exclusively, on decoding. MTEG reading supports the Afghan national curriculum emphases, not only on the development of literacy skills, but also more broadly on promoting and strengthening children's ability to think and reason – abilities integral to reading comprehension. The aims articulated in the official curriculum are directed at bringing about social change, reducing poverty, and creating a more stable society. This paper will outline the MTEG reading assessment framework with particular reference to its application at Grade 3, present available results, and discuss issues arising from this approach to reading assessment in fragile situations.

2 INTRODUCTION TO MTEG

Monitoring Trends in Educational Growth (MTEG) is a service program run by the Australian Council for Educational Research. The program, which aims to establish or strengthen system-wide monitoring programs, is based on the following principles:

- A learning assessment should be targeted on identifying where students are in their progress along well-defined learning metrics. The sole purpose of the assessment should not be just to determine whether or not the student has met the curricular goals of the current year of instruction.

- Effective education develops student skills, knowledge and understandings with the intention that they can be applied both within and beyond the classroom, into work and other areas of life. Such an education is said to have a literacy orientation and MTEG aims to assess with a literacy orientation.

- Learning growth for all learners is the goal of all education systems and practitioners. Growth for all, regardless of their starting point, or their background, is the core of an equitable approach to education. MTEG aims to measure growth throughout schooling along well-defined metrics, from Grade 3 through to Grade 9.

- The research questions associated with measuring student learning must be policy relevant. The research questions are developed with the goal of informing current policy issues.

- Measures of student learning must be of high quality. Best et al. (2013) identified that high quality assessment programs were more likely to influence the policy process in developing countries; and, conversely, assessment programs perceived as having lower quality had little impact. Ensuring a high degree of technical quality gives weight to an assessment program by producing data that can be compared over time with reliability and a strong degree of validity.

2.1 The Afghanistan Context: national goals

In 2012, the Australian Council for Educational Research (ACER) and the Ministry of Education in Afghanistan established a partnership to develop a learning assessment system in Afghanistan, based on the MTEG service model.

An early step in that process was to determine how the national learning assessment would align with national educational goals. These goals were elaborated by the Ministry of Education as:

1. Acquiring and strengthening the learning skills as listening, speaking, reading and writing, using numbers and calligraphy in national and foreign languages.
2. Learning science, arts and modern technologies, and obtaining necessary individual and social skills.
3. Developing aptitudes for learning and self-evaluation in the learning process.
4. Promoting and strengthening the abilities of thinking, reasoning, study, research, diagnosis and innovation in academic, literary, cultural and technical contexts.
5. Gaining skills for solving and identification individual and social problems.

    (Ministry of Education, 1390 [2011], p. 30)

By measuring reading, writing and mathematical literacy, MTEG contributes to the monitoring of the attainment of these policy goals. Progress towards these goals is measured over time. The schooling system is investigated for the value it is adding at each stage of schooling from Grade 3 to Grade 9. The relative outcomes of sub-populations (such as girls and boys; and urban and rural dwellers) are benchmarked to be re-examined over time. Also important from a policy perspective is an analysis of the degree to which factors such as school resources and levels of teachers' professional development are associated with positive educational outcomes.

2.2 The Afghanistan context: capacity development

As noted, MTEG is a service rather than a product. That service is tailored to the specific needs of educational systems. Different systems require different levels of assistance with the design, implementation, analysis and reporting of a monitoring program. Within the Afghanistan Ministry of Education in 2012 there was virtually no capacity to develop and deliver high quality large scale assessment and only limited support available from external sources such as the University of Kabul. At the time, Afghanistan's assessment experience was largely limited to teacher developed examinations used to select (or exclude) students for the next stages of education.

ACER's MTEG provided a comprehensive service for Afghanistan that featured:
- design of an assessment framework;

- item writing for mathematical, writing and reading literacy;

- construction of contextual questionnaires;

- extensive trialling prior to the main study;

- item translations and linguistic quality assurance;

- booklet production, design and print (Grade 6 and Grade 9);

- computer tablets with built-in assessment application (Grade 3 only);

- sample design;

- scaling using an item response model;

- analysis using techniques appropriate for complex survey designs; and

- reporting that is appropriate for both technical and non-technical audiences.

A significant aspect of this service was the capacity development of Ministry staff in all of the above areas. Importantly though, MTEG was implemented as a partnership. It was the Afghanistan Ministry of Education who took the lead role in:

- establishing the Learning Assessment Team, dedicated to the implementation of MTEG, within the Ministry;

- providing feedback on content development for cultural and curriculum relevance;

- promoting the assessment program to local educational authorities and principals;

- administering the tests throughout the provinces;

- scoring the open-ended item responses; and

- liaising with policy makers, curriculum designers and teacher trainers.

In accordance with the design to measure progress over time, the first round of MTEG assessments was administered in 2013 for Grade 6 using pen and paper. The next round of testing is due in late 2015 for Grade 3 students using an offline app on computer tablets. Grade 9 testing is still to be scheduled. This paper focuses on the implementation of MTEG reading.

2.3 MTEG reading compared to other assessment projects in low-income countries

The MTEG approach to reading literacy is currently defined as 'understanding, using and responding to written texts, in order to meet personal, social, economic and civic needs.' The MTEG reading assessment is distinguished from that of other major assessment programs that operate in low income countries in two key ways:

1) **Assessment focus:** Many of the current testing programs, particularly those focused on early development in reading – for example, EGRA and ASER – are largely (sometimes exclusively) focussed on decoding. The ability to decode however does not automatically correlate with comprehension as outlined in the first of this symposium's papers, 'The Conditions of Reading Acquisition in Contexts of Low Literacy': '... early development of reading comprehension skills requires extensive opportunities to learn how to interpret written texts. Children need to understand the meaning of words typically found only in texts, and become familiar with structures of text that differentiate written language from the spoken language that children ordinarily hear.' and 'If fluency is used an indicator of reading comprehension it needs to exclude meaningless "barking at print"'. (Anderson & Fearnley-Sander, 2015).

   By not including comprehension in a reading assessment, the assessment structure runs the risk of not measuring the full extent of a student's overall reading ability. Assessing comprehension levels also has value for other subject areas. Newman (1977) argued that for a student to respond to a written mathematics question he or she had to be able to complete five stages:

   1. reading (decoding) the question or problem posed;
   2. comprehending what it meant;
   3. transforming it into mathematical equation;
   4. processing and solving that equation; and
   5. encoding (writing down) the answer.

   Newman and successive studies by Casey (1978), Clements (1980), Watson (1980) and Clarkson (1980) all found that about 50% of the errors made by students completing written mathematics problems occurred at the reading, comprehension and transforming stage, with errors at the comprehension stage contributing the largest proportion of the three.

   The MTEG reading literacy framework has been informed by the two pre-eminent international comparative reading assessment programs, the Programme for International Student Assessment (PISA) and the Progress in International Reading Literacy Study (PIRLS). PISA testing is targeted at 15 year olds and PIRLS is for students in Grade 4 (or equivalent). The MTEG framework targets Grade 3, 6 and 9 and '... is primarily described in terms of content (the text variables: text format and text type), context (the situation to which texts are relevant) and process (the cognitive processes readers use). As an adjunct, the inclusion of precursor skills contributes to elaborating the constituents of the domain at the early stages of reading development. The precursors are described in terms of constituent skills such as word recognition.' (MTEG Afghanistan Assessment Framework, Draft Version 4, 2015, p. 22)

The 'cognitive process' referred to in the MTEG framework involves the core skills that underpin reading comprehension. They are:
- locate information within the text
- interpret (make meaning from the text);
- reflect (the reader situates the text within the wider context of his or her experience) and
- recognise words (word knowledge).

(MTEG Afghanistan Assessment Framework, Draft Version 4, 2015, p. 23-26)

Like PISA and PIRLS, MTEG measures these skills across a range of fiction and non-fiction texts in continuous, non-continuous and mixed formats.

PIRLS has recently introduced PIRLS Literacy aimed at developing countries: 'Initiated in 2011, PIRLS Literacy (earlier known as prePIRLS) is based on the same view of reading comprehension as PIRLS but is designed to test basic reading skills that are prerequisites for success on PIRLS, for countries where most children are still developing fundamental reading skills. PIRLS Literacy can be administered at the fourth, fifth, or sixth grade' (PIRLS 2016). The PIRLS Literacy test however still requires students to decode the text independently in order to complete the comprehension test items.

Any reading assessment that relies on students' independent decoding in order to measure comprehension runs the risk of failing to validly measure reading in students with insufficient decoding skills. A non-reading or pre-reading student can be cognitively processing the comprehension skill set outlined in PISA and PIRLS (such as linking together ideas, interpreting a character's actions, and inferring the cause and effect of an event) *as they listen to a story*. Anderson and Fearnley-Sander argue that '... children learn how to interpret texts that an adult decodes for them well before they learn to decode for themselves' (2015). They cite the USAID 2015 report that showed that for the 20 African countries who conducted the EGRA Oral Reading Assessments the majority of countries had more than 50% zero scores across all grades: that is, more than half the tested students could not read a single word. The comprehension skills (as distinct from decoding skills) of these students are consequently not measured.

The MTEG reading assessment differentiates itself in one way from all the assessments referenced in this section through its ability to assess *both decoding and comprehension* in the early years, *including non- and pre-readers.*

2) **Assessment Delivery**: The standard way to administer early years reading assessments is through a one-to-one interview process, where the administrator reads aloud questions or instructions and records individual student responses. In low income countries, factors such as lack of staff and problems with transport infrastructure can make the delivery and data collection time-consuming and prohibitively expensive. Students can also feel anxious in an interview setting and administrators are likely to read out scripts with differing levels of expression and speed, which compromises the reliability of the data.

The MTEG Grade 3 assessment is delivered on lightweight tablet computers. This enables test administrators to carry (and manage) sets of tablet computers and consequently

simultaneously test small groups of students. The assessment is pre-loaded onto the tablet computers. Once loaded, the test can then be taken anywhere offline and multiple student response data can be stored on the app. Data stored on each tablet can be uploaded to a central ACER server when the test administrator can arrange internet access.

The administration load is minimal. The test administrator needs first only to model to the group how to use the tablet (via a practice program), and then to hand out headphones and monitor the students. All of the test instructions and questions are narrated to the students through audio files embedded within the test system.

3 MTEG'S GRADE 3 TABLET BASED ASSESSMENT

The MTEG Grade 3 tablet assessment model evolved from previous work by ACER in computer-based early years reading assessment. In 2012 ACER designed and implemented a tablet-based assessment with 539 early years students in the Northern Territory in Australia. The Online Assessment Year 1 (OLAY1) sample group was made up of many pre-reading students from urban through to extremely remote areas in the Northern Territory and about one-third of this cohort had never worked on a tablet or computer before. The trial was successful and laid the groundwork for the MTEG Grade 3 Assessment in Afghanistan and a smaller study in Lesotho. Although the main study of the Grade 3 Afghanistan is due to be carried out later in 2015, an extensive field trial was completed in 2014 with 1738 students (after a smaller scale pilot), which indicated that the both the design and content will work successfully with the full Afghan cohort.

3.1 How it works: design to match user experience

The MTEG Grade 3 Assessment program comes with a very simple practice program. This program situates the student in a digital context and guides them through what it means to be able to receive and manage the information they require and then respond accordingly. This not only includes demonstration of the program's two navigational features (see Figure 2), but also allows interactions that reflect any generic computer-based design: the ability to self-correct by 'undo-ing' an action, to re-read a text or re-listen to an audio (where provided) as often as needed, and to engage with some texts and tasks in a non-linear sequence. The program models the item formats students can expect to see in the assessment and allows them to interact with each one. The students initially watch the administrator go through the program and then try it themselves, encouraged to go through it at least twice.

Audio function: After tapping on the green 'Start' circle on the entry screen (see Figure 1), audio in either Dari or Pashto automatically begins and a male and female character explain the two navigation buttons (see Figure 2).
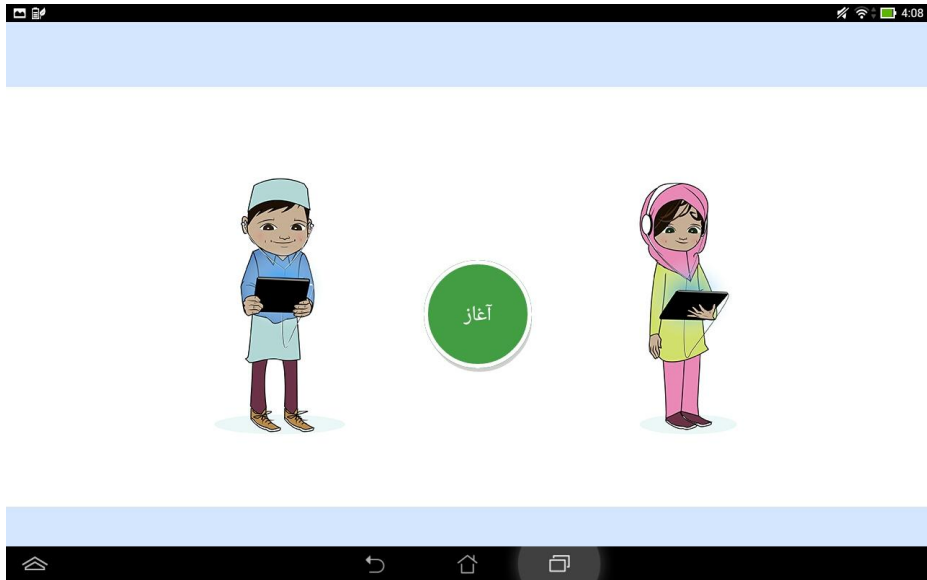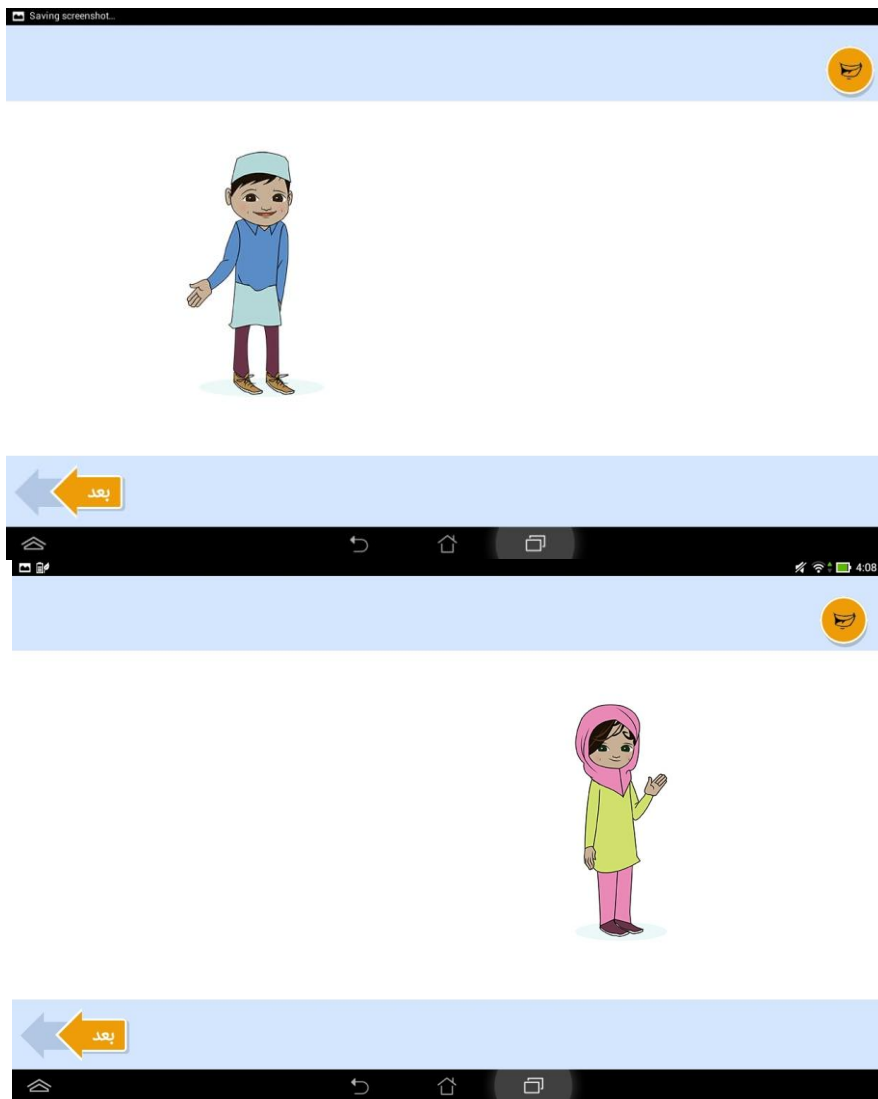
Figure 1: Entry screen to the Practice Program

Figure 2: Still images with automatic narration explaining how to use the audio and navigation icons

In most cases the test item instruction is accompanied with a 'mouth icon' that the student taps to hear the text read aloud.

The mouth icon audio provides two functions:

1. It tells the students what they are expected to do, eg. *Tap on the word that matches the picture.* or *Listen to each sound then tap on the one that is the letter 'a'*. In traditional assessments students need to be able to read in order to know what to do even though the skills being tested in the item may be of a lower cognitive or learning level.
2. It increases the variety of ways that decoding and comprehension can be tested. A few examples are:
   - matching sounds and words to written letters and text
   - building words that have been read out
   - listening comprehension texts with audio-supported response options
   - independent reading comprehension (no audio) with audio-supported response options

There are also texts and test items that test reading comprehension and decoding and provide no audio support except the single instruction: '*Read this story by yourself, then answer the questions about it.'*

Figure 3 shows an example of an item from the practice program that directs the student to the tap on the word 'mountain' in the sentence below. In the practice program the written word 'mountain' also appears in the instruction text. In the assessment proper, on the other hand, the word being tested would be read out in the audio instruction, but be replaced with an ellipsis in the instruction text, so students are not able to match up the words.

Figure 3: Item that directs the student to find a word in the sentence which highlights when tapped.

On-screen book: The program also introduces the on-screen book, which first appears without any test items attached. Students can tap on page tabs or 'swipe left' in the book to move to the next page.

Figure 4: On-screen book with tabs and audio button for each page

When students have finished reading the book they tap on the Next arrow and the test item slides in from the right-side of the screen, leaving the book still functional on the left-side of the screen (See Figure 5). This means that students have access to look through the 'pages' of the book to answer the questions and are not relying on recall or short-term memory to respond. They are using the book stimulus in the way it is used in real-life when searching for information. After listening to the options, students tap on the speech bubble to record their answer. Figure 6 shows a reading item that comes without audio support.

Figure 5: Text question appears next to the book when the student taps Next button
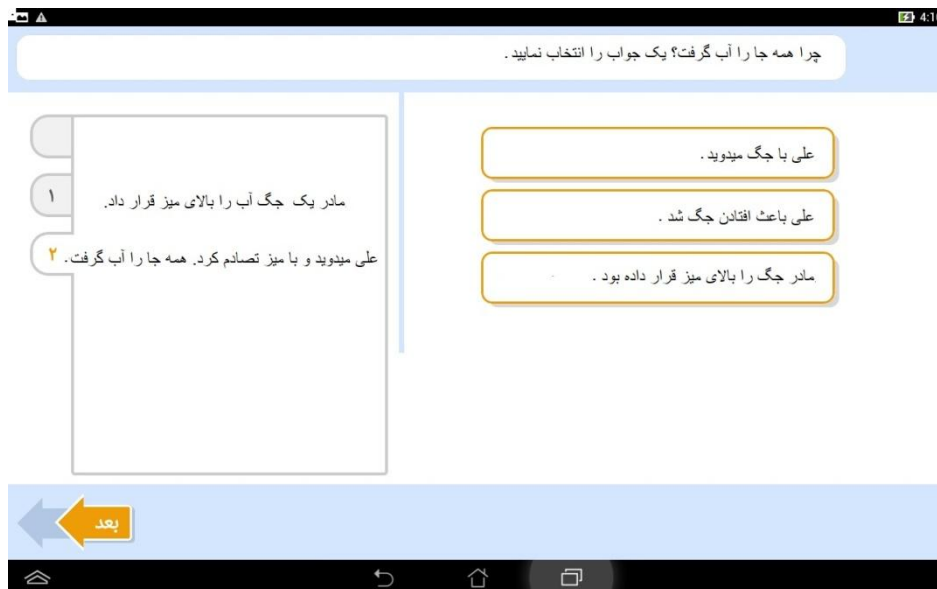
چرا همه جا را آب گرفت؟ یک جواب را انتخاب نمایید.

علی با جگ میدوید .

علی باعث افتادن جگ شد .

مادر جگ را بالای میز قرار داده بود .

۱ مادر یک جگ آب را بالای میز قرار داد.

۲ علی میدوید و با میز تصادم کرد. همه جا را آب گرفت.

بعد

Figure 6: Example of book and test items without audio support

3.2 How it works – design of content development

A significant aspect for MTEG is the alignment of research-based educational outcomes with each country's own culture and curriculum. There are two main reasons that it is important to strike this balance:

1. The look and feel of the content must be instantly familiar to the student, especially for the early years cohort. This settles the student into a context they can work with and helps improve the ability for cognitive processing, as shown in long-term research on schema theory (Bartlett, F., 1932; Anderson et al, 1977). In simple terms, schema theory suggests that new information is processed through the filter of the pre-existing schema of knowledge the individual has already created. 'Schema theory can help explain failure to comprehend: there may not be enough clues to activate existing schemata, the reader may not have appropriate schemata, and failure to comprehend may result when the reader makes an interpretation but not the one intended' (Clark, S., 1990). An example of this from the early development of the Grade 3 reading test was when a narrative set around children playing in a park with a parent supervising was proposed. Consultation with the Afghan team revealed that this context was unsuitable for the Afghan context. The team reported that parks are not set up in the same way as they are in many other countries, that boys and girls do not tend to play together, that boys and girls would play different types of games and it is unlikely that a parent would be present and supervising the children playing. An apparently simple question such as, 'What are the children doing at the park?' would therefore be likely to elicit confusion in Afghan students working through their own schema of behaviour in a park. The targeted reading skill can, however, still be measured by changing the setting of the question.

2. By establishing a familiar context in a reading test that matches the cultural background of the student, it helps alleviate student confusion when a less familiar concept then does appear in a test item. In Afghanistan, 'thinking' questions such as interpreting a character's actions are far less likely to be asked, but are an important part of

comprehension and appear in the MTEG assessment. 'To say that one has comprehended a text is to say that she has found a mental 'home' for the information in the text, or else that she has modified an existing mental home in order to accommodate that new information.' (Anderson, R.C., & and Pearson, D.P., 1984). Grade 3 field trial results showed that these types of 'thinking' questions were more difficult for students, with only 20–30% of the students responding correctly to items that involved inference in contrast 55–75% that involved locating directly-stated information.

The MTEG Grade 3 reading assessment draws on the five components of reading recommended in the 2000 US National Reading Panel Report, but with a strong emphasis on listening and reading comprehension.

The first three of the five components focus on decoding written language into speech.

1. **Phonemic Awareness:** In *Research-Based Methods of Reading Instruction* (2004) Vaughn and Linan-Thompson describe phonemic awareness as 'the ability to identify the phonemes (smallest identifiable units of sound) of spoken language, and how they can be separated (pulled apart or segmented), blended (put back together), and manipulated (added, deleted, and substituted).' When used together in a sequence, phonemes take on meaning or instigate a change in meaning (Cruttenden, A., 2008, p. 41). For example the three phonemes /d/ /o/ /g/ become meaningful as the word 'dog'; however if the phoneme /g/ is replaced with /t/, both the sound and meaning have been altered to make the new word 'dot'.

2. **Phonics**: This is a process of 'teaching reading that stresses the acquisition of letter-sound correspondences and their use to read and spell words' (Harris & Hodges, 1995).
In English, phonemes are directly connected to phonics instruction, 'For the purpose of reading instruction, a phoneme is a single sound that maps to one or more printed letters'. (Vaughn et al., 2004, p. 8). This is not the case for all languages, however, such as those that are character-based.

3. **Fluency:** The ability to decode at a high level. 'Fluent readers can read text with speed, accuracy and expression'. (National Reading Panel, 2000)

The final two components of reading address making meaning from what is read.

4. **Vocabulary:** 'The proportion of difficult words (vocabulary) in a text is the single most powerful predictor of text difficulty, and a reader's general vocabulary knowledge is the single best predictor of how well that reader can understand text' (Anderson and Freebody 1981). Building on from Anderson and Freebody's statement, other researchers have found that simply knowing the meaning of words alone is not enough, 'reading comprehension depends on a wealth of encyclopedic knowledge and not merely on definitional knowledge of the words in the text' (Nagy, 1988) and 'Vocabulary instruction should aim to teach students the relationships that exist among words so that they are learned as part of a network of ideas' (Moats, 2000, p. 192).

5. **Comprehension:** In its simplest terms, comprehension is the ability to draw meaning from spoken and written language (Vaughn et al., 2004, p. 136). The National Reading Report states that reading comprehension is a cognitive process that integrates complex skills and

cannot be understood without examining the critical role of vocabulary learning and instruction and its development. (Burns *et al*.)

What is apparent from all the researchers is that whilst fluency smoothes the way for comprehension to take place, it doesn't guarantee comprehension. And that the first three components are stepping stones to enable the ultimate goal of reading, which is making meaning from written text.

3.3   Assessment content

In 2014 93 individual reading items were trialled for the MTEG Grade 3 reading assessment. As detailed in Table 1 almost half of these items reflected elements of comprehension whilst the others were focussed on decoding, reading fluently and print conventions.

Table 1 Grade 3 MTEG Assessment field trial content 2014

	Reading Components	Number of test items
number of comprehension-based test items 48	Environmental print	4
	Independent reading comprehension (paragraphs, sentences and word level)	27
	Listening Comprehension of stories (paragraphs level)	17
	Vocabulary (sentences and word level)	9
number of decoding-based test items 45	Phonemic Awareness	9
	Phonics	27
	Print Conventions	9
	Fluency	read aloud a few short sentences

From this pool of 98 items, 58 have been selected to be in the main study. They will be delivered in a split-design where there are two separate tests that have a set of common items.

Independent Reader Test

20 reading and listening comprehension items at word, sentence and paragraph levels

20 Common reading items based mostly on Listening Comprehension

Emerging Reader test

18 decoding items including sound, letter recognition and environmental print
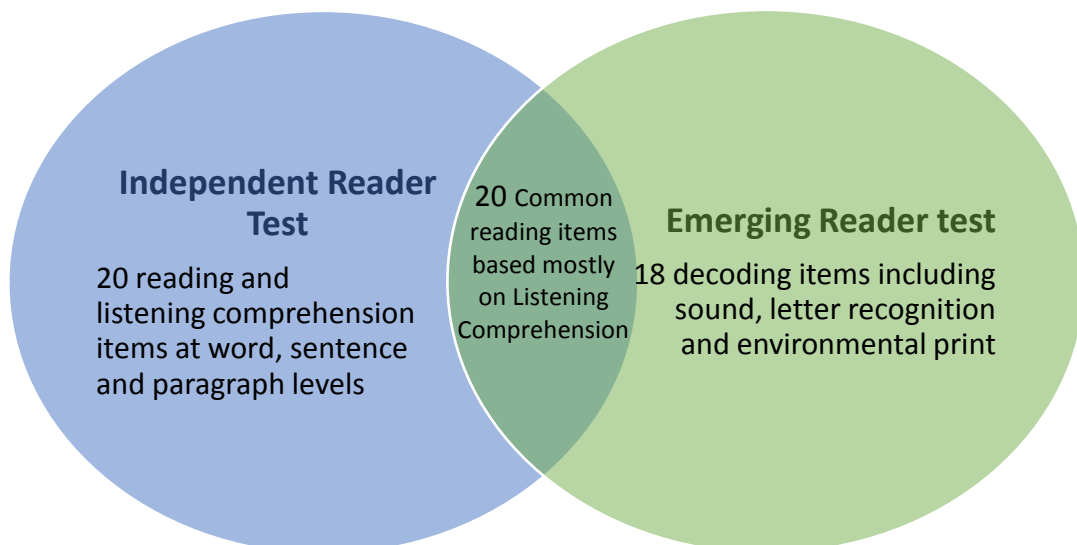
Figure 6: Split-design structure for MTEG Grade 3 Reading Assessment

Before the student sits the test, the administrator asks each student to read aloud 1 to 3 short sentences, making sure each one reads a different text from the other students in the group (they also do not need to read in front of other students). If a student does not start to read, he or she is asked, 'Are there any words you know?', and then encouraged to point to them and read them aloud. Students who correctly read any five words or more are assigned the Independent Reader Test on the app. Students who only read one or two words or cannot read any words at all are assigned the Emerging Reader Test, also on the app. The fluency check is to be stopped after one minute.

Having two tests available means a greater range of the cohort's abilities can be measured without the need for a single, much longer test. The overlapping set of 20 questions also means that students who are unable to read fluently are still able to be checked for both low-level decoding and the higher-level listening comprehension. These results can be directly compared to the fluent reader cohort through the commonly shared items.

3.4 MTEG Grade 3 field trial results

The field trial results showed that there was a good spread of item difficulties and the items showed good discrimination. The items were well targeted, with the average item difficulty and the average student ability of the sample being very similar. However, a significant proportion of items was too difficult for many students; these items will therefore not be included in the main study. The field trial results reveal that the majority of the items performed similarly for Pashto and Dari languages, and for male and female students. Items indicating gender or language bias have not been selected for the main study.

The field trial data also revealed that:

- both comprehension and decoding items were present across the full range of difficulty (easy, medium, hard);
- students found locating information an easier comprehension skill than interpreting meanings around characters' intentions and feelings;
- responses for a simple reading text that the Independent Readers read *without audio support* was generally found to be easier than a more difficult listening comprehension text that was fully audio-supported and required no decoding – demonstrating that the two skills can be teased apart for testing purposes;
- students' and administrators' feedback confirmed that the students found the Practice Program accessible and engaging and were consequently able to sit the test with a satisfactory level of confidence and independence; and
- 79% of the sample responded correctly to the easiest item, 16% responded correctly to the hardest item.

4 DISCUSSION OF ISSUES

The two most significant operational issues faced in delivering the program to Afghanistan were the relevance of content and the logistical aspects of delivering the tests (whether they were on paper or computer tablets). The partnership that was established early with the Ministry of Education staff meant that there was a pragmatic approach to feedback. Although Dari and Pashto are the official languages, there still exist many local dialects and vocabulary that differ by region.

Care was taken to not use terms that would cause confusion. Through the training provided, the Afghan Learning Assessment team became adept at interpreting the data and increasingly promoted the importance of including 'thinking-based' test items not traditionally included in their teaching or testing of reading. The placement of such items in the test modelled their value to both students and educators. As happened in the first administration of MTEG learning assessments in Grade 6, it is expected that the analysis of thinking-based outcomes will lead to the provision of practical pedagogical advice for teachers (see 'Suggestions for teaching' in Lumley et.al 2014)

The local administrators, trained and supported by the Kabul-based Learning Assessment Team, distributed the tests for the Grade 6 field trial and main study and the Grade 3 field trial. The administrators carried out their tasks with a great deal of efficiency both in the time-frame they had and collection of metadata. Training of the administrators was included in the MTEG program and they were able to report quickly on problems in the field. For example, during the Grade 3 field trial, unreliable internet connection in Kabul was interfering with the downloading of the tests onto the apps. As the connection dropped out, tests were coming through with missing images and text. The software developers were able to find another way to segment the data so that even when there was a disruption to the connection during the downloading of the tests, no data would be lost.

5 CONCLUSION

The MTEG reading assessment program has shown signs of success in the Afghanistan context via its implementation and reporting at the Grade 6 level and extensive trialling at the Grade 3 level. This has been accomplished by finding a number of innovative and culturally relevant ways to work towards the expectations of the national education goals. The close working relationship with the Learning Assessment Team at the Afghan Ministry of Education has been an important contributing factor to this achievement through their valuable feedback as well as promotion and implementation of the assessment nationwide.

The Grade 3 tablet-based approach demonstrates that it is possible to collect data on reading comprehension in a highly standardised and efficient way. Another significant feature of the Grade 3 assessment is that it has found a way to assess and collect data on all five components of reading assessment even with students who are unable to decode through its use of embedded audio. The added emphasis on listening and reading comprehension in the assessments also incorporate valuable thinking and reasoning skills that are a fundamental aspect to understanding text. The combined result of the assessment's content and design is the output of quality data that covers a large spread of abilities. With the analysis of outcomes, this is vital for policy-makers and practitioners to lead practical pedagogical advice for developing these aspects of learning to read.

6 REFERENCES

Anderson, R. C., and Freebody, P. (1981). *Vocabulary Knowledge. In Comprehension and Teaching: Research Reviews*, (Ed.) J. Guthrie, 77-117. Newark, Del.: International Reading Association.

Anderson, R.C., and Pearson, D.P., (1984) A Schema-Theoretic View Of Basic Processes In Reading Comprehension, Technical Report No. 306, University of Illinois at Urbana-Champaign.

Anderson, R. C., R. E. Reynolds, D. L. Schallert, and E. T. Goetz. (1977) Frameworks for Comprehending Discourse. *American Educational Research Journal* 14, no. 4: 367-81.

Bartlett, Frederic (Sir). (1932).*Remembering: A study in experimental and social psychology*. Cambridge University Press.

Best, M., Knight, P., Lietz, P. Lockwood, C., Nugroho, D., & Tobin, M. (2013). The impact of national and international assessment programmes on educational policy, particularly policies regarding resource allocation and teaching and learning practices in developing countries. Draft report 11 May 2012. Evidence for Policy and Practice Information and Co-ordinating Centre (EPPI-Centre), Social Science Research Unit, Institute of Education, University of London.

Burns, M. S. E., Griffin, P. E., Snow, C. E. E., & National Research Council (US). Committee on the Prevention of Reading Difficulties in Young Children. (1999). *Starting out right: a guide to promoting children's reading success,* M. Susan Burns, Peg Griffin & Catherine E. Snow, (Eds.); Committee on the Prevention of Reading Difficulties in Young Children, Commission on Behavioral and Social Sciences and Education, National Research Council.* Washington DC: National Academy Press.

Casey, D.P. (1978). Failing students: A strategy of error analysis. In P. Costello (Bd.), Aspects of motivation, Melbourne: Mathematical Association of Victoria, 295-306.

Clark, S. (1990). Schema Theory and Reading Comprehension. Distributed by ERIC Clearinghouse [microform]: Washington, D.C.

Clarkson, P.E. (1980) The Newman error analysis - Some extensions. In B.A. Foster. (Ed.), Research in mathematics education in Australia 1980, vol. 1, 11-22. Hobart: Mathematics Education Research Group of Australia.

Clements, M.A. (1980) Analysing children's errors on written mathematical task. Educational Studies in Mathematics, 11 (1), 1-21.

Cruttenden, A., (2008) Gimson's Pronunciation of English: Seventh edition NY: Hodder.

Harris, T. & Hodges, R. (Eds.) (1995). *The literacy dictionary*. Newark, DE: International Reading Association.

Lumley, T., Mendelovits, J., Turner, R., Stanyon, R. & Walker, M. (2015) *Class 6 proficiency in Afghanistan*. Camberwell: Australian Council for Educational Research.

Ministry of Education, (2011). Afghanistan Education Curriculum 1390.

Moats, L. (2000) *Speech to print: Language essentials for teachers*, Baltimore, MD: Brookes Publishing.

Nagy, W. E., (1988) *Teaching Vocabulary to Improve Reading Comprehension,* USA: Eric Clearinghouse on Reading and Communications Skills.

Newman, M.A. (1977a). An analysis of sixth-grade pupils' errors on written mathematical tasks. In M. A. Clements & J. Foyster (Eds.), Research in mathematics education in Australia, vol. 2, 269-87. Melbourne: Swinburne College Press.

PISA 2015 Draft Reading Literacy Framework, March 2013. Retrieved from http://www.oecd.org/pisa/pisaproducts/Draft%20PISA%202015%20Reading%20Framework%20.pdf

Vaughn, S., & Linan-Thompson, S. (2004) *Research-Based Methods of Reading Instruction*, Alexandria, VA: Association for Supervision and Curriculum Development ASCD.

Stern, J., (2015) The State of Literacy in Africa Comparative and International Education Society Annual Conference Washington, DC, 10 March 2015, retrieved from www.eddataglobal.org/documents/index.cfm/3%20State%20of%20Literacy_CIES%20Presentation_March%209.pdf?fuseaction=throwpub&ID=788

Watson, 1. (1980). Investigating errors of beginning mathematicians. Educational Studies in Mathematics, 11 (3), 319-29.
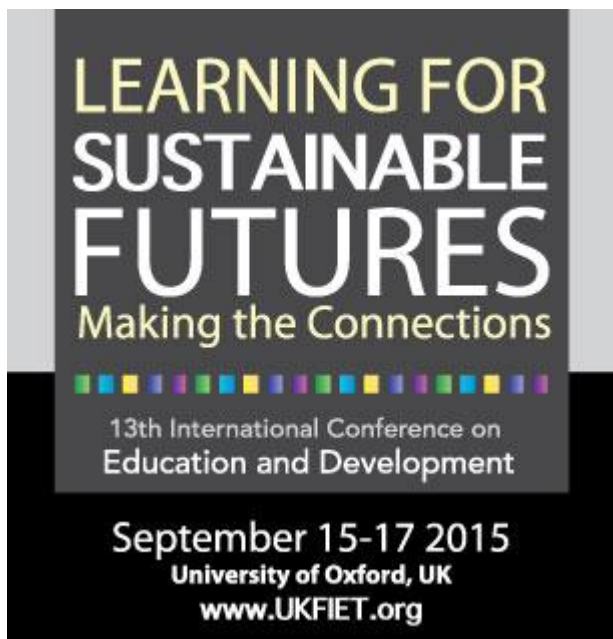
Paper 3 - Juliette Mendelovits

Australian Council for Educational Research
juliette.mendelovits@acer.edu.au

Paper 3 - Ross Turner

Australian Council for Educational Research
Ross.Turner@acer.edu.au

Paper 3 - Ray Adams

Australian Council for Educational Research
Ray.Adams@acer.edu.au

Paper 3 - Monitoring reading globally for the post 2015 development goals

1   ABSTRACT

The Global Education Monitoring Centre of the Australian Council for Educational Research (ACER-GEM) has initiated the development of a set of learning metrics, with the intention of developing a means of measuring and reporting on learning outcomes that will support the Post-2015 development goals. Initial activity was through what was referred to as the Learning Metrics Partnership (LMP), which was explored as a joint initiative of the UNESCO Institute for Statistics and ACER-GEM. The objective is to develop a set of nationally and internationally comparable learning metrics in mathematics and reading, and then facilitate and support their use for monitoring purposes, in partnership with interested countries. The empirically derived learning metrics will support national governments to effectively measure and monitor learning outcomes for policy purposes. This would not involve the development of a new test or testing program. Rather, it would support the use of existing assessments of various kinds (including citizen-led, regional and national assessments), and a pool of calibrated items that could be used to facilitate measurement and reporting of local learning outcomes against common metrics. A key output of this work would be a common learning metric for reading, spanning reading development from early primary school to early secondary school. In the conceptualisation of the reading metric, comprehension is represented as the central developmental continuum, with other components of early reading mapped differentially in relation to comprehension according to the linguistic specificities of different languages.

2   THE PURPOSE OF DEVELOPING A GLOBAL READING LEARNING METRIC

In September 2015 the UN General Assembly will set the development agenda for the next 15 years. There has been a paradigm shift in thinking about international education development, from the Millennium Development Goals of 2000–2015 (UN Secretary General, 2006), with their focus on *access* to education, to the Sustainable Development Goals (SDG) of 2015-2030, where

the emphasis is also on *quality* of education. The Incheon Declaration, formulated in preparation for the SDG at the World Education Forum in Incheon (May 2015), includes a commitment 'to quality education and to improving learning outcomes, which requires strengthening inputs, processes and evaluation of outcomes and mechanisms to measure progress' (UNESCO, 2015, Paragraph 9).

Despite the high level of participation in learning assessments at regional, national and sub-national levels, clearly defined descriptions of progression in learning, and intra- as well as inter-assessment comparability, remain limited. This presents particular challenges for the global development goals for the post-2015 discussions and reporting requirements.

In response to these challenges, the Centre for Global Education Monitoring of the Australian Council for Educational Research (ACER-GEM) has commenced initial work to develop a set of nationally and internationally comparable learning metrics. Learning metrics – sometimes called 'described scales' – encapsulate progression in a measured learning variable of interest, by combining quantitative information about the degree of progression observed at different locations along the metric with qualitative descriptions of what progression looks like at different parts of the metric.[7]

If the work of constructing comparable learning metrics is work is successful, and if there is interest in using these metrics globally, the aim then would be to facilitate and support their use for monitoring purposes, in partnership with interested countries. The development of common described scales will allow policy makers, education practitioners and education investors to not only quantify student proficiency, but also describe it in a meaningful way. A scale provides a means to assess the emerging competencies of younger children, and to explore cognitive growth and trends over time. There is a need for common described scales, spanning learning from early primary school to early secondary school, which is relevant and applicable to a range of developing country contexts. Such scales, starting with very early learning development, are currently unavailable.

While Paragraph 9 of the Incheon Declaration refers to several important areas of learning ('Quality education fosters creativity and knowledge, and ensures the acquisition of the foundational skills of literacy and numeracy as well as analytical, problem-solving and other high-level cognitive, interpersonal and social skills' (UNESCO, 2015)), the initial development and planning work has been based on the view that reading and mathematics, as universally recognised 'foundational skills', should be the focus of the first attempts to establish global learning metrics.

The two key features of the approach proposed are:
- Accommodating results from a range of different assessments of learning outcomes; and

- Yielding high quality data that are nationally relevant and internationally comparable.

The objective is to develop empirically supported learning metrics in mathematics and reading that will assist national governments in effectively measuring and monitoring learning outcomes for policy purposes. The proposal does not involve the development of a new test or testing program. Rather, it supports the use of existing assessments of various kinds, and a pool of

---

[7] For a more detailed description of learning metrics see:
http://www.acer.edu.au/files/Described_Proficiency_Scales_and_Learning_Metrics.pdf

calibrated items that could be used to facilitate measurement and reporting of learning outcomes against common metrics.

In early 2014, ACER-GEM proposed that a Learning Metrics Partnership be established with the purpose of developing a set of learning metrics for use by interested countries, regional assessment programs, and by the international development community, as a tool that could help to enhance existing practices for the monitoring of educational outcomes. In the first instance, development of metrics for reading and for mathematics was proposed.

As this idea took shape, a three-phase work program was envisaged. Phase 1 of the program has been to create draft reading and mathematics learning metrics. In Phase 2, the draft metrics will be validated at the country level. Data will be collected by administering combinations of items to children, which will enable the empirical determination of the relative difficulties of items across assessment programs. Phase 3 is the development and use of a set of tools and methodologies that permit the broad alignment of existing learning assessments and also the development and use of tools and methodologies to support the alignment of country-developed assessments with the common metrics.

The first phase was intended to develop the draft learning metrics.

In the original plan, development of the draft learning metrics for reading was to involve the use of test items from a wide range of existing assessment programs, with the development work to take place in the period from April/May to November 2014. So far the development work has used material found in the public domain, as well as relevant material from international and national assessment programs in which ACER has been directly involved, or for which specific approval has been obtained from item owners for this purpose. It is anticipated that in the second phase of the work a broader range of items will be used in order to further enrich the metrics as they are validated and tested empirically and further developed.

In this paper we describe the process followed thus far to develop the draft reading metric, including some of the theoretical and technical issues encountered.

3 THE PROCESS OF DRAFTING A GLOBAL READING LEARNING METRIC

3.1 Assembling multiple reading assessment instruments and identifying key and common features of items

The first drafting of the reading learning metric was accomplished in a number of stages.

Initially, a framework mapping the growth of early reading development was drafted on the basis of empirical work conducted in an Australian longitudinal study of literacy and numeracy that mapped literacy and numeracy of 1000 children for seven years, from school entry (aged about five) to the end of primary schooling (see Louden, et al., 2005; Meiers, et al., 2006).

Then, a range of reading assessment instruments was assembled for review, with the aim of further identifying features that could be generalised as indicators of increasing difficulty of tasks,

and concomitant progressive reading proficiency. Programs from which items were inspected were[8]:

- Annual Assessment of Education Report (ASER, ASER Centre, New Delhi, India)

- Early Grade Reading Assessment (EGRA, RTI International)

- Longitudinal Literacy and Numeracy Study (LLANS, ACER)

- Monitoring Trends in Educational Growth Afghanistan – Class 6 paper-and-pen assessment (MTEG Afghanistan, ACER)

- Northern Territory Online Assessment of Year 1 (OLAY, Government of Northern Territory of Australia)

- Program for International Reading Literacy Survey (PIRLS, IEA)

- Programme for International Student Assessment (PISA, OECD)

- Solomon Islands Standardised Test of Achievement (SISTA, Ministry of Education and Human Resources, Solomon Islands)

- Uwezo survey in Kenya, Zimbabwe and Uganda (Twaweza)

The assessment programs are diverse in their target populations, according to age, year level and stage of reading development; their methods of administration; and their constructs. Some of this diversity is worth noting here.

First, all of the assessments apart from PISA are targeted at children at early stages of reading development, or up to the end of primary school (Year 6). This is justified by the context in which the learning metrics project is being developed: to support the SDG.

Second, there is a very wide diversity of administration methods and response formats. ASER, EGRA, LLANS (reception and Year 1) and Uwezo are orally administered, one-on-one assessments. OLAY is a computer-based interactive assessment. The others are paper-and-pen group assessments.

Third, with regard to construct, all of the early stages of reading assessments (ASER, EGRA, LLANS Prep[9] to Year 2, OLAY and Uwezo) include tasks measuring children's proficiency in the constrained skills of reading. The term 'constrained skills of reading', coined by Paris (2005), is being used here to refer to those skills that underpin reading development and are constrained: that is, they are learned quickly and mastered entirely. Concepts of print, phonemic awareness and reading fluency are included in this category. 'Unconstrained skills of reading' are those that continue to develop and have no clear ceiling, in particular vocabulary and comprehension. Tasks addressing unconstrained skills constitute all or most of the paper-and-pen assessments (MTEG Afghanistan [Grade 6], PIRLS, PISA, and SISTA).

Further, while all of the programs style themselves as reading assessments, there is no single common component: some include listening comprehension, some fluency, some letter or word recognition – even reading comprehension is not a universal element.

Finally, even when there is an apparently common component, the way it is assessed may make comparison difficult. For example, EGRA includes reading comprehension (as do all of the paper-

---

[8] See Annex 2 for acknowledgement of the owners of these assessment tools.
[9] Australian term for 'Reception' or 'Kindergarten'

and-pen assessments), but a child answering the EGRA questions may not look at the text while responding, whereas the text remains available to children answering reading comprehension questions in the paper-and-pen assessments (and in some of the orally-administered assessments).

While acknowledging this diversity, researchers were nevertheless able to identify notable key features of the skills and knowledge addressed in the items, and to identify common features of the items where they existed.

3.2   Drawing on existing developmental scale descriptions

Described scales have become common as part of the reporting of assessment outcomes in a range of national and international assessment programs around the world. For example, they are used to report outcomes of the National Assessment of Educational Progress (NAEP) in the US; they are used to report outcomes of IEA studies such as TIMSS and PIRLS; and they are used by the OECD to report outcomes of the Programme for International Student Assessment (PISA). Each of the existing scales was constructed using empirical data from the tests' administrations that allow the calibration of items in order of difficulty. Using the locations of items on the constructed scale, generalised descriptions of the content and skills measured by the tasks in a given range along the scale are developed.

Extrapolating from these existing described scales, the researchers clustered descriptions of the several hundred reading items along a draft developmental continuum for reading. Item difficulty estimates from administrations of several of the assessments used in this study (LLANS, MTEG Afghanistan, OLAY, PISA, SISTA) were available, and these data were used to help arrange the descriptions of item content in a hierarchical ordering of difficulty. Descriptions of items from the other instruments, with comparable content and comparable estimated difficulty, were integrated into the same continuum of description.

This process yielded a very early first draft reading learning metric.

3.3   Implementation of the pairwise comparison study

While some of the work undertaken in initial drafting of the reading learning metric drew on empirical data from individual assessments, there was as yet no way to integrate the data from different assessments other than through subject-based expert judgement. That is, whilst the pre-existing metrics conveyed information about the relative difficulty of items *within* each study, there was no empirical evidence that addressed the issue of the relative difficulty of items *across* assessments.

To strengthen the validity of the draft, and to generate data that would permit comparison of the difficulty of the different item sets used, an additional methodology was employed: a 'pairwise comparison study'.

For that study, experienced markers (raters) of reading assessments were employed to rate the relative difficulty of a large number of pairs of items selected from the item pool. This methodology is based on a model proposed in Bradley and Terry (1952), and analysis of data generated by applying such a methodology has been implemented in ACER-ConQuest software (Adams, Wu, & Wilson, 2012).

A team of 12 expert markers was presented with pairs of items selected from those used in the first part of the process. The pairwise comparisons drew on 512 individual reading items. Markers were asked simply to judge which member of any given pair would be likely to be more difficult for students. The pairs were selected and assigned to the markers in a linked design such that relative difficulty across the entire set could be estimated.

In total, the team of 12 markers rated which item was more difficult in a little over 29,000 comparisons. The main purpose of this part of the study was to generate a preliminary estimate of the comparison across the different tests analysed. Analysis of data from this study also shed light on the suitability of having items calling on both constrained and unconstrained reading skills on the same scale.

The item analyses informed the development of proficiency descriptions that form the essence of the resulting draft learning metric for reading. Fourteen bands were defined along the metric, to cover the widest possible range of reading proficiency. The descriptions of item demand were used to build summary descriptions of proficiency at each of the defined bands. Illustrative items were identified for each level defined on the metric.

4 OUTCOMES

The initial draft reading metric referred to different skills, both constrained and unconstrained; a taxonomy of text types; and examples of tasks matching each band. It described growth in reading across 14 Bands. The full initial draft reading metric is shown in Annex 1. Two of the bands are reproduced in Figure 2 for illustration.

Band	Proficiency Descriptions by Band	Description of Text Types	Sample Items by Band
6	**Print Conventions** At this level, students are starting to recognise that dialogue is signalled by inverted commas.		
	Phonemic Awareness At this level, students delete, substitute and blend phonemes at the end of words.		Take the /d/ out of CARD. What word do you get?
	Reading Comprehension At this level, students are able to read one or two sentences/pages of **Text Types A and B**. While they have limited understanding of these texts they are able to retrieve very simple information with the support of a direct word match.	**Text Type A** Texts up to approx 150 words that offer high levels of support to the reader through: illustrations repetitive sentence structure many high frequency words and a limited range of additional vocabulary	Retrieve directly stated information by making a direct word match in an information text Meet My Mouse4
		Text Type B Short (up to approx approx 50 words) un-illustrated paragraphs that have: highly familiar contexts simple sentence structures many high frequency words	Retrieve directly stated information by making a direct word match in a simple short story King Chile
5	**Print Conventions** At this level students can identify basic punctuation eg full stops and question marks.		Show me a full stop.
	Phonemic Awareness They can delete phonemes from words and substitute phonemes at the beginning of words.		MEAT. Take the /m/ out of MEAT. What word is left? SIT. Swap the /s/ with a /p/. What is the new word?
	Phonics They start to read words that have simple letter-sound correspondence rules. They also expand their knowledge of high frequency words that do not conform to these basic rules.		Put the labels on the fruit and vegetables [accompanied by image].
	Reading Comprehension At this level, students are able to read short, repetitive texts that have a small range of vocabulary that is supported by illustrations.		
	Listening Comprehension In **Text Type 1**, students are able to recognise underlying ideas, such as inferring the main idea, theme or lesson of the text.	**Texts read to students** **Type 1** Simple texts containing familiar ideas and concepts	Retell a story by providing a short summary or focusing on the key aspect of the narrative Precious Eggs1 (Code 2) Identify a characters implicit intentions by linking illustrations to the plot Precious Eggs2 (Code 2)

Figure 2 Extract from the initial draft reading learning metric: Bands 5 and 6

The pairwise study was particularly useful in identifying the development of reading comprehension skills at the higher levels (Band 7 and above). When only paper and pen reading comprehension items were considered, the pairwise data showed strong correlations with existing data about the relative difficulty of the items (especially the LLANS data). The other components of reading (eg phonemic awareness and phonics) correlated less well, perhaps due to the great range of response types.

A further product of the initial descriptive work and the pairwise comparison study is a very rough mapping of the relative ranges of difficulty of the assessment programs included in the research. This rough mapping is shown in Figure 3. The boxes show the middle half of the distribution of item difficulty in each assessment.
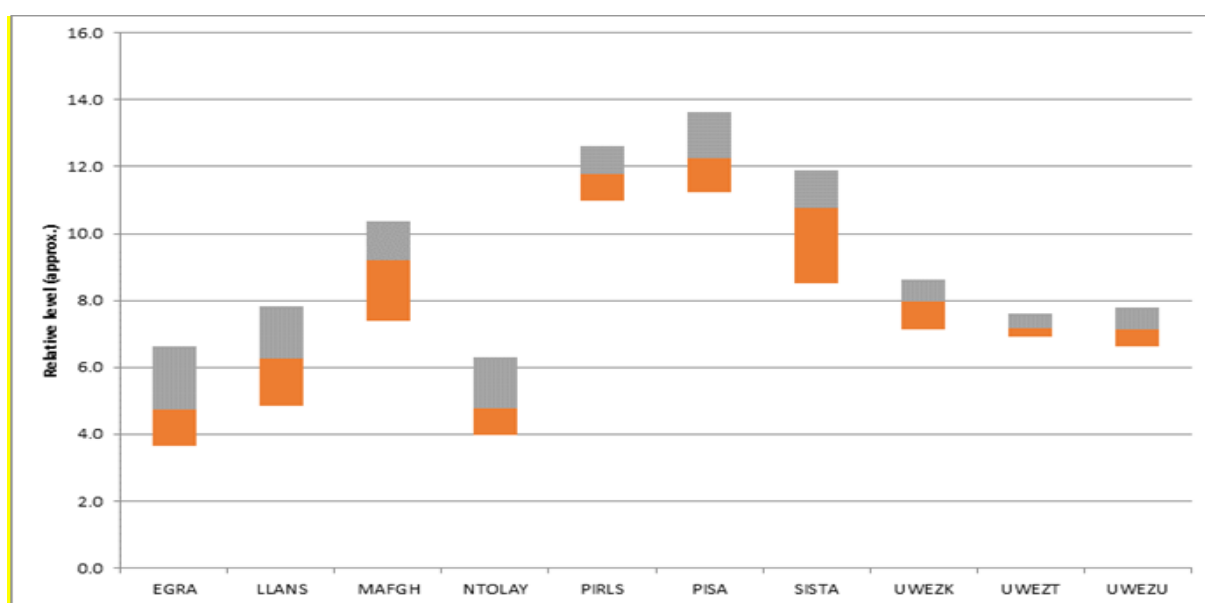
Figure 3 Approximate relative location of reading source test program items

Comment has been made that the relative difficulty of PISA and PIRLS shown here is implausible. This may be an artefact of the frame of reference: these two programs' tasks, compared with all the others, are considerably more demanding, so the difference between their levels of difficulty may have been obscured.

5 ISSUES AND CHALLENGES

A number of challenges in developing the draft learning metric for reading have been discovered and reflected upon as the project has unfolded. Feedback from stakeholders also raised further issues for consideration. Among the most important of these were:

- The complexity of the initial draft reading metrics, such that it was difficult to see a clear progression mapped out;

- The challenge of combining descriptions of the development of constrained and unconstrained reading skills in a single continuum, given the variations in their interaction in different languages and (possibly) in different pedagogical circumstances; and (connected with this); and

- The anglocentricity of the initial draft metric for reading.

Work undertaken by ACER researchers in the ensuing months has focused on addressing these three challenges.

5.1 Reducing the complexity of the draft reading metric

While the initial draft reading metric provided a wealth of information about reading development, with its references to different skills – both constrained and unconstrained; a taxonomy of text types; and examples of tasks matching each band, it was arguably too complex for easy communication.

An additional complication was the use of a coded taxonomy of texts for reference. In Figure 2, for example, 'Texts read to students Type 1', 'Text Type A' and 'Text Type B' are referenced in the second column and glossed in the third. Although this coding was intended to reduce repetition, it reduced the transparency and readability of the metric less transparent and less readable.

Concurrent work was being undertaken by a different research team at ACER on the construction of a reading metric spanning a similar range of ages and grades, but for an Australian context. This work on the ACER scale drew on the descriptions of skills and understandings, and of texts, developed by the learning metrics group. Reciprocally, the learning metrics group was able to incorporate from the ACER scale a more fluent approach to combining descriptions of skills and text types.

These enhancements are being incorporated into the draft reading metric. (See Table 2 below for a preliminary draft of this work.)

5.2 Accommodating multiple languages into the model of a reading learning metric

As described above, the first drafting of the reading learning metric was accomplished in a number of stages, using a range of reading assessments drawn from many sources. Although many of these assessments are administered in languages other than English, and most of these in multiple languages, the researchers' inspection of content was confined to English language versions, and their mapping of constrained skills instantiated in the items therefore focused on features of English language reading development. Similarly, the pairwise comparison study was conducted using English versions of the items only.

Anglocentricity, which underpins many studies of reading development that are then applied in an international context, has been remarked upon elsewhere (notably in Share, 2008). Looking at Figure 2, and in particular at the descriptions of 'Print Conventions', 'Phonemic Awareness' and 'Phonics' in the second column, and the sample items for these strands in the last column, it is readily apparent that the focus is on reading development in English. For example, the description of Print Conventions at Band 5 is, 'At this level students can identify basic punctuation eg full stops and question marks.' Full stops and question marks are not used in some languages. Again, the description of Phonemic Awareness at Band 6 is, 'At this level, students delete, substitute and blend phonemes at the end of words.' While 'blending phonemes' is an important marker of development in phonological awareness in English and many other languages, it is an irrelevant and probably incomprehensible skill in syllabic languages.

In developing the initial draft of the global reading metric the researchers were very much aware of the issue of linguistic differences, among languages or language groups, that may affect reading development. These differences manifest themselves most strongly at the early stages of reading, in the emergence and mastery of constrained skills.

The researchers outlined four possible approaches to dealing with the issue:

1. to include constrained and unconstrained skills on the same scale

2. to have constrained and unconstrained skills on separate scales

3. to have a single unconstrained skills scale and separate constrained skills scales for each writing system (syllabic, alphabetic, etc.)

4. to have a single unconstrained skills scale and provide descriptive statistics for constrained skills. These descriptions could be placed along a main scale for reading to indicate at which point a particular skill is expected to have been mastered. (ACER, 2015, p. 63)

At present, we are pursuing approach 2, with the hypothesis that, should empirical analysis support it, combined reading learning metrics (approach 3) may be used for different languages or language groups, though to represent a global metric, the constrained skills will need to be represented as separate subscales.

A focus of our work since the development of the initial draft of the reading metric has been on investigating the different features of languages, and consequent differences in the early progression of reading, in languages other than English.

This work has proceeded using a combination of strategies, including a literature search, inspection of assessment instruments, empirical assessment data where it is available, and consultation with speakers of other languages who are also assessment specialists. To date, investigation has begun for Arabic, Chinese, French, Hindi, Indonesian, Japanese and Spanish. The languages on which these early investigations have focused have been chosen for a variety of reasons including:

- The representation of major language groups, especially those language groups which are likely to be of key interest in countries that will use the learning metrics;

- Languages for which the researchers have access to assessment data;

- Languages that have unique characteristics that are likely to differentially influence children's progress in becoming proficient readers; and

- Languages spoken by individuals with assessment expertise to whom the researchers had access.

The next section of this paper gives an example of the way one of the languages (Japanese) is being analysed for the purposes of developing the reading learning metric, and how that analysis evolves into a description of a constrained skills subscale linked to the main text comprehension scale. While this is just work in progress, it is an example of preliminary research on the salient features (for our purposes) of a writing system in a language other than English.

6 ACCOMMODATING MULTIPLE LANGUAGES INTO THE MODEL OF A READING LEARNING METRIC: JAPANESE EXAMPLE

6.1 A brief description of Japanese language and early reading development

The Japanese use three writing scripts: *hiragana, katakana* and *kanji* (Chinese-based characters). Early readers of Japanese work with hiragana. As texts become more complex, the hiragana is augmented with katakana and *kanji*. *Hiragana* and *Katakana* each contain 46 characters with corresponding sounds but different characters. *Hiragana* is used for the native language and *Katakana* is used for foreign words introduced into the Japanese language. For example the word 'rice' (kome: こめ) would be written in Hiragana but the word 'bread' taken from the French word,

would be written in Katakana (pan: パン). The third writing system, *Kanji*, is based on the written form of traditional Chinese characters, but with different pronunciations and meanings.[10]

Because of the high level of regularity in Japanese (the fact that one character corresponds reliably to one specific sound), Japanese children tend to be more fluent with their decoding skills than children learning English as a first language. Some researchers (Mason, Anderson, Omura, Uchida, & Imai, 1989) have suggested that this has a direct bearing on the curriculum of the early years of education in Japan: 'The fact that hiragana are easy to learn to read may allow Japanese kindergarten teachers to focus on music, art, health, nature, communication, and social skills in kindergarten instead of on reading, and first grade teachers need only review letter sounds and can [focus more on] text recitation and text level comprehension' (Mason, et al., 1989, p. 4). Students consolidate their hiragana and katakana skills during their first year of elementary school (age 6), having begun learning them in kindergarten (age 5). By the end of the first grade, they will also be familiar with close to 80 written kanji letters.

There are fewer syllables in Japanese than in English – just over 100 in Japanese, compared to several thousand in English – which increases the number of verbal homophones that exist (although these are visually differentiated when written in *kanji*). Because of this, context arguably plays a more significant role in reading in Japanese than in English.

Syllabic awareness is an important prerequisite for the development of reading skills in Japanese, and according to Nagy and Anderson, 'measures of syllable awareness are highly correlated with reading ability for Japanese children... but not for American children' (1995, p. 4). This study, though, also points to a difference between early learners of Japanese and early learners of English: the role of syllabic awareness in literacy acquisition in a syllabic writing system 'cannot be completely analogous to that of phonemic awareness in an alphabetic writing system. Japanese beginning readers do not need the same kind of help with syllable awareness that American beginning readers need with phonemic awareness.' (Nagy & Anderson, 1995, p. 4). This supports the implication that the characteristics of a language and its writing system may significantly shape curricula and the rate of learning.

Figure 4 shows a sketch of the development of the constrained skills of reading in Japanese.

---

[10] There are thousands of characters in Kanji. Children will know some of them early (for example, numbers). By the end of the first year of school, they are expected to read and write eighty; by the end of primary school, around one thousand; and by the end of high school, around two thousand. Even educated adult readers do not know all the *Kanji*, and often use a dictionary. Learning Kanji appears to be a separate skill acquisition from the 'learning to read' in Japanese (that is, learning to read hiragana and katakana) that is parallel with learning to read in other languages.

6. Students can read and recite all 46 characters in both *hiragana* and *katakana*, and can therefore fluently decode all words regardless of whether they are familiar or unfamiliar due to singular pronunciation of each character/syllable. Students can delete, substitute and blend character/syllables. They are aware of differences in meaning depending on stressed syllables.

5. Students add or delete characters to or from a word. They identify differences between similar words. They spell simple words.

4. Students develop their knowledge of characters and use this to spell simple words. They understand that words or phrases may be separated by a space (in text for very young readers only). Students understand that words are grouped according to meaning.

3. Students identify some common characters. They identify sounds at the beginning and end of words, and orally blend syllables (characters) in short words.

2. Students are aware of the differences between a character and a word, and can differentiate between them. Students identify the first characters (syllables) in written words with distinctive sounds.

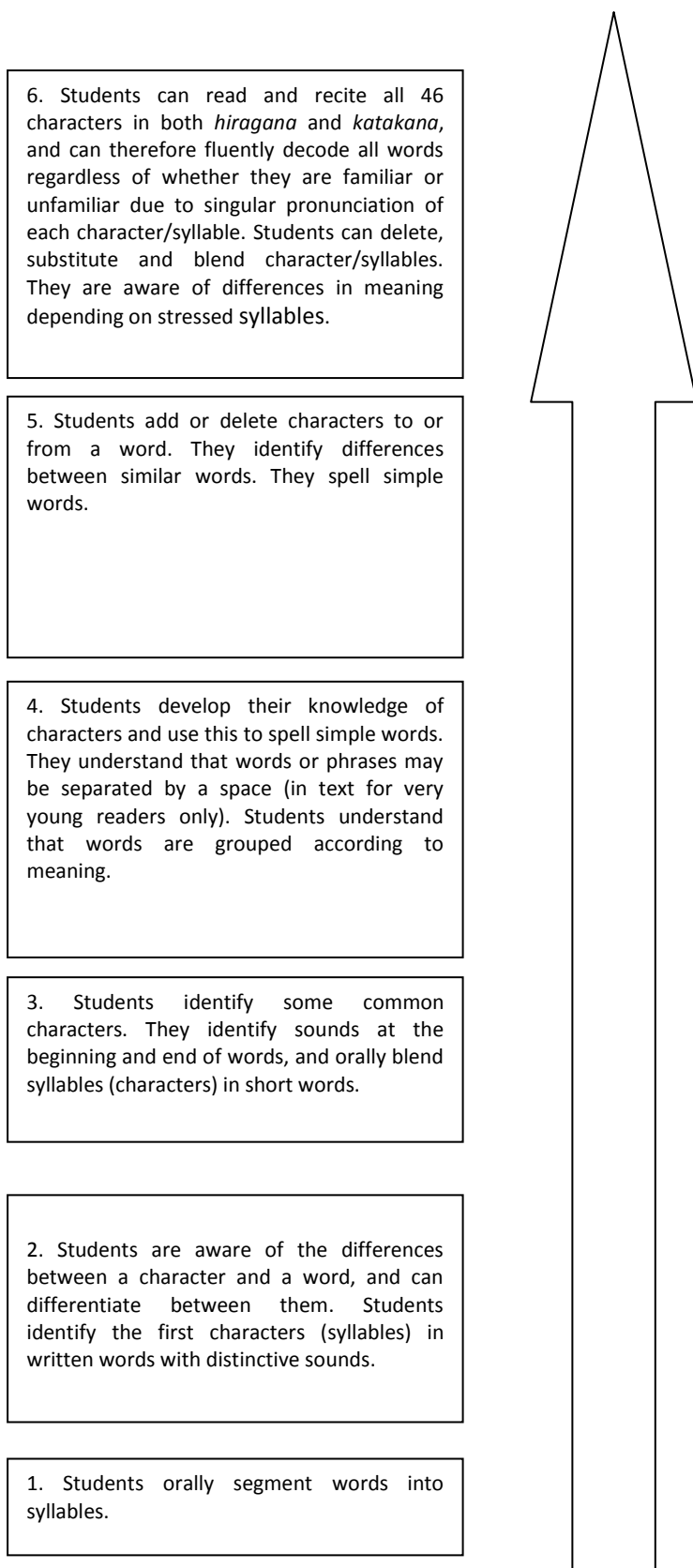1. Students orally segment words into syllables.

Figure 4 Development of constrained skills of reading in Japanese

7 REDRAFTING OF THE GLOBAL READING LEARNING METRIC: A SINGLE,UNIVERSAL TEXT COMPREHENSION SCALE AND LANGUAGE-DIFFERENTIATED CONSTRAINED SKILLS SUBSCALES

Text comprehension, an unconstrained skill that has no clear ceiling of skill development, may be assumed to be universal. While this assumption has not been tested at the earliest stages of development in text comprehension – those stages simultaneous with and even prior to the development of some constrained skills – the notion of universal progression in text comprehension is borne out by the international studies of reading, PIRLS and PISA. The extent to which this notion can be applied at the earliest stages of reading development will be tested empirically in Phase 2 of the learning metrics project. At this stage, we are working with that assumption.

Table 2 is an attempt to show how a global reading learning metric might look, with the main scale showing the development of text comprehension or meaning-making (unconstrained skills), and separate subscales for different languages showing the development of constrained skills as they are aligned with meaning making. The alignment between the *Text comprehension* scale and the *Constrained skills: English* subscale is based on the empirical work described in section 3. The alignment between the *Text comprehension* scale and the *Constrained skills: Japanese* subscale is more speculative, based on sourced literature and consultation with assessment specialists who are speakers of Japanese.

Table 2 shows only levels up to Band 9 of the 14 bands posited in the initial drafting of the reading learning metrics (see Annex 1). Band 9 is the highest band at which there are descriptions of constrained skills in English, while we hypothesise that Band 6 is the highest band for constrained skills development in Japanese. This makes clear the meaning of 'constrained' versus 'unconstrained' skills.
.

Table 2 Extract from combined draft learning metric: text comprehension scale with constrained skills subscales in two languages

Band	Text comprehension	Constrained skills: English	Constrained skills: Japanese
Band 9	Students interpret explicit information by making simple links within or across adjacent sentences, and make simple inferences, in a variety of common texts such as poems, lists and instructions, and longer narratives with familiar contexts but some linguistic complexity.	Students apply a range of strategies including syllabification and their knowledge of word families to a range of three- and four-syllable words.	
Band 8	Students understand underlying nuances and implicit meanings, such as the motivation behind a character's actions, in short simple texts. They begin to read a variety of common texts such as poems, lists and instructions, and longer narratives with familiar contexts but some linguistic complexity and can locate directly stated information in these texts. When listening to texts that present familiar ideas in a more complex or abstract way, students are able to draw inferences such as predicting outcomes and identifying character's feelings.		
Band 7	Students read and understand prominent ideas and some details in short, simple explicit texts with a small range of familiar words. They begin to retrieve information using direct or synonymous word matches and can make simple inferences. When listening to texts, students give reasons for events and explain unusual details. They support their personal opinions about a character's actions or the outcome of a story. They generalise about a character's behaviour and explain how simple problems are resolved.	Students can delete, substitute and blend phonemes in the middle of words. They can read a range of one- and two-syllable with phonetically regular patterns.	
Band 6	Students read and understand prominent ideas in very short, simple un-illustrated texts with a few highly familiar words. They mainly use direct word matches to retrieve information. When listening to texts, students sequence three or more events in a retelling. They make simple inferences about characters' feelings using clues in the text and the illustrations. They identify clear similarities between events and give reasons to support predictions or the likelihood of an event occurring when clues are prominent.	Students can delete, substitute and blend phonemes at the end of words.	Students can read and recite all 46 characters in both hiragana and katakana, and can therefore fluently decode all words regardless of whether they are familiar or unfamiliar due to singular pronunciation of each character/syllable. Students can delete, substitute and blend character/syllables. They are aware of differences in meaning depending on stressed syllables.

Band	Text comprehension	Constrained skills: English	Constrained skills: Japanese
Band 5	Students read and understand some simple, text-based explicit meaning in short, repetitive texts using a small range of vocabulary, with supportive illustrations. When listening to simple texts containing familiar ideas and concepts, students can infer the main idea, theme or lesson of the text.	Students fluently read some phonetically regular words and familiar, irregular (sight) words. They can delete, substitute and blend phonemes at the beginning of words.	Students add or delete characters to or from a word. They identify differences between similar words. They spell simple words.
Band 4	Students can identify a range of familiar, high frequency words in print.	Students name and give a sound for all the letters in the alphabet. They segment short words, including common blends, into phonemes.	
Band 3	Students interpret supportive illustrations to make meaning from texts. They understand the directionality of print. When listening to a simple text containing familiar ideas and concepts, students can recall information such as sequencing events and identifying characters' actions. They can make simple inferences such as predicting outcomes and identifying characters' feelings.	Students identify some common letters. They identify initial and end sounds in words and blend phonemes in short words. They understand the difference between upper case and lower case letters.	Students develop their knowledge of characters and use this to spell simple words. They understand that words or phrases may be separated by a space (in text for very young readers only). Students understand that words are grouped according to meaning.
Band 2	Students recognise their first name written down. They understand how to correctly hold a book and turn the pages. They recognise and interpret common environmental print such as Stop signs and toilet signs. When listening to a story, students are able to identify some basic information, eg an event in the story or the name of a character.		Students identify some common characters. They identify sounds at the beginning and end of words, and orally blend syllables (characters) in short words.

Band	Text comprehension	Constrained skills: English	Constrained skills: Japanese
Band 1	Students distinguish between writing and drawings or symbols. They start to understand that print carries meaning. For example, they realise their name can be written down and read. When listening to texts, students recall one or two pieces of basic information. They link illustrations to the text to explain a simple event. They locate the first event in a series, identify a prominent main idea, and make predictions or simple inferences that are supported by prominent clues in the illustrations. They give a reason to support a personal response about whether they like the text or not.	Students identify the first phonemes in words with distinctive sounds (sun, food), recognise prominent end-sounds (rhyme) and segment three-syllable words into syllables. They differentiate between a letter and a word.	Students are aware of the differences between a character and a word, and can differentiate between them. Students identify the first characters (syllables) in written words with distinctive sounds.
Pre Band 1	When listening to texts, students identify prominent information such as a key event. They may describe aspects of an illustration without connecting it to the main narrative.	Students segment two-syllable words into syllables by clapping.	Students orally segment words into syllables.

8 NEXT STEPS: EMPIRICAL RESEARCH: CONSTRUCTION OF SYNTHESISED TESTS TO INSPECT THE RELATIVE DIFFICULTIES OF ITEMS (KICE WORK)

As mentioned earlier, the second phase of the work of building common learning metrics is to empirically validate and further develop the draft metrics, with the involvement of a wider range of interested countries and assessment programs.

A first step has been taken with collaboration between the Korean Institute of Curriculum and Evaluation (KICE) and ACER, to include some items drawn from the set included in the pairwise comparison study in an equating study conducted with Year 6 students in Korea. In this study, reading items from the pool assembled by ACER have been combined with selected items from the national Year 6 monitoring program for Korea, and this combined test has been administered to a sample of Korean students. Analyses of these data are underway, and will further contribute to the learning metric thus far developed, thereby allowing further refinement and elaboration of the reading learning metric.

It is intended that similar studies will be implemented as the program develops.

9 REFERENCES

ACER. (2015). Phase I Outcomes: Draft learning metrics for mathematics and reading. Unpublished Meeting paper. The Learning Metrics Partnership Steering Committee.

Adams, R. J., Wu, M. L., & Wilson, M. R. (2012). ACER ConQuest 3.0.1. . Camberwell: Australian Council for Educational Research.

Bradley, R. A., & Terry, M. E. (1952). Rank analysis of incomplete block designs, I. the method of paired comparisons. *Biometrika*(39), 324-345.

Louden, W., Rohl, M., Barratt Pugh, C., Brown, C., Cairney, T., Elderfield, J., et al. (2005). *In Teachers' Hands: Effective Literacy Teaching Practices in the Early Years of Schooling*: Department of Education, Science and Training.

Mason, J., Anderson, R., Omura, A., Uchida, N., & Imai, M. (1989). *Learning to read in Japan* (No. 449). Urbana-Champaign: Center for the Study of Reading, University of Illinois at Urbana-Champaign.

Meiers, M., Khoo, S. T., Rowe, K., Stephanou, A., Anderson, P., & Nolan, K. (2006). *Growth in LIteracy and Numeracy in the First Three Years of School*. Camberwell: ACER.

Nagy, W., & Anderson, R. (1995). *Metalinguistic Awareness and Literacy Acquisition in Different Languages* (No. 618): Center for the Study of Reading, University of Illinois at Urbana-Champaign.

National Reading Panel. (2000). *Report of the National Reading Panel: Teaching children to read.* Washington, DC: US Government Printing Office.

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

Share, D. (2008). On the Anglocentricities of Current Reading Research and Practice: The Perils of Overreliance on an "Outlier" Orthography. *Psychological Bulletin, 134*(4), 584–615.

UN Secretary General. (2006). UN Millenium Project. Retrieved 4 August 2015, 2015, from http://www.unmillenniumproject.org

UNESCO. (2015, 19-22 May 2015). *Incheon Declaration.* Paper presented at the World Education Forum 2015, Incheon, Republic of Korea.

10 ANNEX 1 INITIAL DRAFT READING LEARNING METRIC

Band	Proficiency Descriptions by Band	Description of Text Types	Sample Items by Band
14	**Reading** **Comprehension** At this level, students understand the underlying nuances and implicit messages in **Text Type G**.	**Text Type G** Dense texts that have increasing levels of technicality and abstraction.	Interpret the main argument with dense language and unfamiliar vocabulary in a persuasive text; this involves locating a detail and relating it to the whole text
13	**Reading Comprehension** At this level, students understand the underlying nuances and implicit messages in **Text Type F**. They interpret information by linking multiple pieces of information, some of which may be deeply embedded and outside the main body of the text, and deal with strongly distracting competing information.	**Text Type F** A range of text types that have some structural and/or conceptual complexity.	Support an opinion by construing a character's motivation in a play The Play's the Thing Analyse the relationship between a text and a statement made about that text Mobile Phone Safety 1
12	**Reading Comprehension** Students are now reading **Text Type D** at a level that allows them to understand underlying nuances and implicit messages. They make inferences that involve linking information across a text where an understanding of the key ideas in the text is required, and deal with extensive and competing information that may be contrary to expectations. In **Text Type E** students are now able to integrate information across multiple text types to retrieve a detail.	**Text Type D** Lengthier texts (up to approx 800 words) that have less familiar content and vocabulary. **Text Type E** Texts that have a complex structure and some difficult vocabulary. This includes information texts with multiple components and technical procedural texts.	Identify a character trait and support with evidence from a long narrative text Fly, Eagle, Fly1 Recognise an assumption in an expository text Mobile Phone Safety2
11	**Reading Comprehension** Students can now read enough of **Text Type D** to retrieve a detail where the information is not prominent. They interpret explicit information by making connections between multiple ideas across a text and deal with competing information. In **Text Type E** they retrieve details and interpret explicit information where the links are simple, such as between a map and its key.	**Text Type D** Lengthier texts (up to approx 800 words) that have less familiar content and vocabulary. **Text Type E** Texts that have a complex structure and some difficult vocabulary. This includes information texts with multiple components and technical procedural texts.	Integrate information across paragraphs to explain a character's actions. Fly, Eagle, Fly2 Recognise the purpose of a table in an expository text Mobile Phone Safety3

Band	Proficiency Descriptions by Band	Description of Text Types	Sample Items by Band
10	**Reading Comprehension** Students are now reading **Text Type C** at a level that allows them to understand the underlying nuances and implicit messages. In **Text Type D,** where the information is prominent and there is little or no competing information, students are also able to retrieve directly stated information, usually relying on making direct or synonymous word matches. They can also make links across a short section of text to interpret explicit information.	**Text Type C** A variety of common texts including: poems, lists, instructions and narratives (up to approx 300 words) that have familiar contexts but some sentence and structural complexity. **Text Type D** Lengthier texts (up to approx 800 words) that have less familiar content and vocabulary.	Explain the trick that was played on a character in a picture story book Clever Bird1 (Code 2) Identify directly stated information by linking across sentences in the first paragraph of a long narrative text Fly, Eagle, Fly3
	Listening Comprehension In **Text Type 2**, students begin to show awareness of literary devices such as metaphor and repetition.	**Texts read to students** **Type 2** Texts that present familiar ideas and concepts in a more complex/abstract way	Infer the metaphoric significance of an illustration by linking to the story The Deep1 (Code 2)
9	**Phonics** At this level, students apply a range of strategies including syllabification and their knowledge of word families to a range of three and four syllable words.		Read unfamiliar words with phonetically regular segments [tactical]
	Reading Comprehension At this level students are able to read enough of **Text Type C** to allow them to interpret explicit information by making simple links within or across adjacent sentences in the face of some competing information. They can make simple inferences and are no longer dependent on direct word or synonymous word matches in the retrieval of information.	**Text Type C** A variety of common texts including: poems, lists, instructions and narratives (up to approx 300 words) that have familiar contexts but some sentence and structural complexity.	Recognise but can't explain how a trick has been played Clever Bird1 (Code 1) Infer the reason for an event from explicit clues in a short story Dog and Fox1
	Listening Comprehension In **Text Type 2**, students are able to recognise underlying ideas, such as inferring the main idea, theme or lesson of the text.	**Texts read to students** **Type 2** Texts that present familiar ideas and concepts in a more complex/abstract way	Interpret a character's feelings and provide a plausible, plot based reason for them The Deep2 (Code 2)

Band	Proficiency Descriptions by Band	Description of Text Types	Sample Items by Band
8	**Reading Comprehension** At this level, students understand the underlying nuances and implicit messages in **Text Types A and B**. They are able to recognise underlying ideas, such as the motivation behind a character's actions and can interpret explicit information by making links across several sentences. They are also able to read enough of **Text Type C** to allow them to retrieve information that is in a prominent position and is supported by a direct word match.	**Text Type A** Texts up to approx 150 words that offer high levels of support to the reader through: illustrations repetitive sentence structure many high frequency words and a limited range of additional vocabulary **Text Type B** Short (up to approx approx 50 words) un-illustrated paragraphs that have: highly familiar contexts simple sentence structures many high frequency words **Text Type C** A variety of common texts including: poems, lists, instructions and narratives (up to approx 300 words) that have familiar contexts but some sentence and structural complexity.	*These texts tend not to be long enough to have underlying idea* Retrieve directly stated information in a short story Dog and Fox2
	Listening Comprehension In **Text Type 2**, students are able to recall information, such as sequencing events and identifying characters' actions. They can make simple inferences, such as predicting outcomes and identifying characters' feelings.	**Texts read to students** **Type 2** Texts that present familiar ideas and concepts in a more complex/abstract way	Use illustrations infer a character's feelings without being able to explain the reason for them The Deep2 (Code 1) Give a literal interpretation of an illustration The Deep1 (Code 1)
7	**Print Conventions** At this level, students understand the purpose of more complex punctuation eg ellipsis and inverted commas.		What do these dots tell you? Meet My Mouse3
	Phonemic Awareness At this level, students delete, substitute and blend phonemes in the middle of words.		Listen to this word: BLUSH Swap the /l/ with /r/. What word do you get?
	Phonics At this level, students can read a range of one and two syllable words with phonetically regular patterns.		Read unfamiliar words with phonetically regular segments [ticket]

Band	Proficiency Descriptions by Band	Description of Text Types	Sample Items by Band
	Reading Comprehension At this level, students can read most of the words in **Text Types A and B.** Whilst their reading may contain errors that interfere with the exact meaning of the text, students understand the main details and surface features of the text, such as a literal message that is supported by prominent clues (eg the repetition of information). They begin to retrieve information using direct, or simple synonymous, word matches and can make simple inferences.	**Text Type A** Texts up to approx 150 words that offer high levels of support to the reader through: illustrations repetitive sentence structure many high frequency words and a limited range of additional vocabulary **Text Type B** Short (up to approx 50 words) un-illustrated paragraphs that have: highly familiar contexts simple sentence structures many high frequency words	Identify relevant section of an information book Meet My Mouse1 Link information in adjacent sentences to make a simple inference My name is Pat
6	**Print Conventions** At this level, students are starting to recognise that dialogue is signalled by inverted commas.		
	Phonemic Awareness At this level, students delete, substitute and blend phonemes at the end of words.		Take the /d/ out of CARD. What word do you get?
	Reading Comprehension At this level, students are able to read one or two sentences/pages of **Text Types A and B**. While they have limited understanding of these texts they are able to retrieve very simple information with the support of a direct word match.	**Text Type A** Texts up to approx 150 words that offer high levels of support to the reader through: illustrations repetitive sentence structure many high frequency words and a limited range of additional vocabulary **Text Type B** Short (up to approx approx 50 words) un-illustrated paragraphs that have: highly familiar contexts simple sentence structures many high frequency words	Retrieve directly stated information by making a direct word match in an information text Meet My Mouse4 Retrieve directly stated information by making a direct word match in a simple short story King Chile
5	**Print Conventions** At this level students can identify basic punctuation eg full stops and question marks.		Show me a full stop.

Band	Proficiency Descriptions by Band	Description of Text Types	Sample Items by Band
	Phonemic Awareness They can delete phonemes from words and substitute phonemes at the beginning of words.		MEAT. Take the /m/ out of MEAT. What word is left? SIT. Swap the /s/ with a /p/. What is the new word?
	Phonics They start to read words that have simple letter-sound correspondence rules. They also expand their knowledge of high frequency words that do not conform to these basic rules.		Put the labels on the fruit and vegetables [accompanied by image].
	Reading Comprehension At this level, students are able to read short, repetitive texts that have a small range of vocabulary that is supported by illustrations.		
	Listening Comprehension In **Text Type 1**, students are able to recognise underlying ideas, such as inferring the main idea, theme or lesson of the text.	**Texts read to students** **Type 1** Simple texts containing familiar ideas and concepts	Retell a story by providing a short summary or focusing on the key aspect of the narrative Precious Eggs1 (Code 2) Identify a characters implicit intentions by linking illustrations to the plot Precious Eggs2 (Code 2)
4	**Phonemic Awareness** They say most of the sounds in simple words and generate rhyme and manipulate the sounds at the beginning of words.		What are all the sounds in the word [mud]? Listen to this word 'cat'. Swap the /c/ with /r/. What word do you have now?
	Phonics They can name and give a sound for all letters of the alphabet. They can differentiate between capital and lowercase letters. They can also identify a range of familiar, high frequency words in print.		Select the letter that makes the sound /n/. What sound does this letter make? /t/ Read this word [are] [my]
3	**Print Conventions** At this level, students understand the directionality of print (where to start reading a book, where to read next) and understand there is a difference between capital and lowercase letters.		Show me where to start reading on this page. Point to a capital letter. This book is called [xx]. Show me where it says [xx].
	Phonemic Awareness They can identify the initial and end sounds in words.		What sound is at the end of [rat]?
	Phonics They start to identify some common letters and differentiate some familiar capital and lowercase letters.		Point to the letter [a]. This is the letter n. Show me a capital letter N.

Band	Proficiency Descriptions by Band	Description of Text Types	Sample Items by Band
	Reading Comprehension At this level, students interpret simple illustrations that do not rely on text for meaning such as identifying the main character in a story from the cover illustrations.		Identify the main character from the front cover of a picture story book.
	Listening Comprehension In **Text Type 1**, students are able to recall information such as sequencing events and identifying characters' actions. They can make simple inferences, such as predicting outcomes and identifying characters' feelings.	**Texts read to students** **Type 1** Simple texts containing familiar ideas and concepts	Retell some elements of a story without linking them to the main idea Precious Eggs1 (Code 1)
2	**Print Conventions** At this level, students recognise how to hold a book and turn the pages. They can identify a letter and a word. They can also recognise and interpret common environmental print.		What does this sign tell you? [image of signage with and without words eg image for toilet sign, road sign with the name of an appropriate city] Show me the front of the book. Draw a circle around a letter.
	Phonemic Awareness They understand that words are made up of letters and that letters make sounds and can recognise rhyme.		Do these words rhyme? [sandal / handle] Which words have the same first sound [bin lid box]? Which word begins with a different sound [back ten tin] What is the first sound in [up]
	Listening Comprehension At this level, students are able to recall some basic information eg an event in a story or the name of a character. They may rely heavily on illustrations.		Describe a character's actions as shown by the illustrations without linking them to the plot Precious Eggs2 (Code 1)
1	**Print Conventions** At this level, students can distinguish between writing and drawing.		Where is the writing on this box?
	Listening Comprehension At this level, students have a working, everyday speaking vocabulary. They recognise the names of familiar objects and can follow simple oral questions, requests and instructions, including positional language.		Everyday speaking vocabulary: Click on the hat/book pencil (6 illustrations to choose from) (drag and drop items) Put the pencil on the table Put the book in front of the boy

11 ANNEX 2 ACKNOWLEDGMENTS

Assessment items from a range of different assessment programs were used in the preparation of the draft learning metric for reading. The owners of those items are gratefully acknowledged for their willingness to permit the use of material for this purpose, including in some cases use of items to illustrate the metrics.

We thank the organisations responsible for the items sourced from the following assessment programs:

ACER, for use of items from the Longitudinal Literacy and Numeracy Survey (LLANS), and items from the MED Afghanistan assessment program

The ASER Centre, New Delhi, India, for use of items from the household-based ASER program

IEA, for the use of publicly released items from the Program for International Reading Literacy Survey (PIRLS)

The Government of the Northern Territory of Australia, for use of items from ACER's Online Assessment of Year 1 (OLAY)

OECD, for use of released items from the Programme for International Student Assessment (PISA)

RTI International, for use of items from the EGRA tools

Ministry of Education and Human Resources, Solomon Islands, for use of items from the Solomon Islands Standardised Test of Achievement (SISTA)

Uwezo at Twaweza, for use of publicly available material from the household-based Uwezo survey in Kenya, Zimbabwe and Uganda

The authors also gratefully acknowledge the work of their colleagues Prue Anderson, Lynne Darkin, Sandra Knowles and Trisha Reimers for their contributions to the development of the draft reading learning metric.