NOT FOR CIRCULATION

UNESCO
United Nations
Educational, Scientific and
Cultural Organization

UNESCO INSTITUTE FOR STATISTICS

SUSTAINABLE DEVELOPMENT GOALS

Compilation of Performance Level Descriptors across regional and international assessments

10-11 September 2018
UNESCO Headquarters
7, place de Fontenoy,
75352 PARIS 07 SP, France

Introduction

This paper aims at answering the following questions:

1. How do performance level descriptors from regional and international reading assessments compare to each other?
2. Which should be the MPLexpected in reading for Grades 2 & 3, end of Primary education and end of low Secondary education?
3. Which should be the MPLexpected in reading overall?

Based on the analysis of regional and international assessments of reading described on a previous paper, this document aims to facilitate the process of setting common expectations between different cross-national assessments to allow for international comparison.

In order to answer the first question, this paper shows the process by which the Performance level descriptors (PLDs) of all of the regional and international assessments on reading are analyzed, ordered according to difficulty and grouped into four performance categories. This is done for each educational level considered in the 4.1.1 indicator of SDG 4, which are Grades 2 & 3 (4.1.1a), the end of Primary education (4.1.1b) and at the end of Low Secondary education (4.1.1c).

However, consideration has to be given to the fact that this analysis is based on the PLDs only; therefore, the assessments' aims as well as the tasks used by them have not been considered. Even though some of the PLDs make explicit the type of text they use, most of them do not. Thus, the ordering process was done regarding the cognitive demand implied by the processes mentioned in the PLDs. This may lead to assumptions regarding difficulty that are not congruent if a task-based analysis is performed, altering consequently the order of the PLDs. For example, even though from a cognitive perspective making inferences is more difficult than retrieving explicit information, making an inference from a short narrative text is likely to be easier than retrieving explicit information from a long technical informative text. Therefore, an analysis of the PLDs together with task type information provided by the assessment frameworks may lead to a more precise ordering.

Furthermore, to answer questions 2 and 3, a Minimum proficiency level (MPL) was set for each educational level and for reading acquisition in general with accompanying policy descriptors.

Characteristics of the regional international assessments

Table 1 shows the cross-national assessments considered for this paper. Most of these tests are designed to evaluate formal learning, as is the case of reading. However, both ASER and UNICEF MICS 6 are broader questionnaires that aim at obtaining other development indicators at a personal, family, and environmental level, which include a section on reading that is the one considered in this analysis.

Table 1. Characteristics of the assessments

Name	Abbreviation	Grade/Age	Corresponding SDG 4 indicator	Minimum proficiency level	Observations
Annual Status of Education Report	ASER	6 to 14 year-olds	4.1.1.a;	Standard 2 (story)	Part of a household questionnaire in which the assessment is individual.
UNICEF Multiple Indicator Cluster Service	UNICEF MICS 6	5 to 17 year-olds	4.1.1.a;	Foundational Reading Skills	Part of a household questionnaire in which the assessment is individual.
UWEZO Annual Learning Assessment	UWEZO	6 to 16 year-olds	4.1.1.a;	Standard 2	Part of a household questionnaire in which the assessment is individual.
Early Grade Reading Assessment	EGRA	Grades 1 to 3.	4.1.1.a	Not specified	Individual assessment
Third Regional Comparative and Exploratory Study	TERCE	Grades 3 & 6	4.1.1.a; 4.1.1.b	Level 2	School-based assessment
Pacific Islands Literacy and Numeracy Assessment	PILNA	Grades 4 & 6	4.1.1.b	Level 4 (grade 4) and Level 5 (grade 6).	School-based assessment
Progress in International Reading Literacy Study	PIRLS	Grade 4	4.1.1.b	Low international Benchmark (second level)	School-based assessment
The Analysis Programme of the CONFEMEN	PASEC	Grades 2 & 6	4.1.1.a; 4.1.1.b	Level 3	School-based assessment. Partly individual assessment

Education Systems					
Southern and Eastern Africa Consortium for Monitoring Educational Quality	SACMEQ	Grade 6	4.1.1.b	Level 3	School-based assessment
Programme for International Student Assessment	PISA and PISA-D	15 year-olds	4.1.1.c	Level 2	School-based assessment

Characteristics of the regional international assessments

The initial step taken in this process to answer the first question was to develop a Proficiency Scale (PS) on reading. In this regard, all of the PLDs across the ten assessments analyzed were transformed into one-line descriptors by highlighting its main characteristics and those that differentiated them from the previous level.

After this, all of the descriptors were ordered according to their difficulty independently from the educational level they were designed for. This produced a 73 level PS that considers all of the PLDs provided by the ten assessments. It is important to note that the below level 1 descriptor from PASEC as well as the Level 0 descriptor from PILNA were not considered as there is no specific information regarding what the student can or cannot do in those levels.

An interesting finding that arises from the development of the PS is the incongruence between the expectations set by different regional and international assessments as well as the overlapping of PLDs designed for different educational levels.

Finally, in order to answer the third question, an overall MPL was set for reading in general. This was marked at the 50th level on the PS that corresponds to TERCE's Level 2 performance descriptor for Grade 3 which is summarized as: "Students understand the global sense of the text by distinguishing its central topic and making inferences regarding non evident information". If we analyze it from the Global Framework for Reading perspective, it assumes mastery of the decoding sub domain as well as explicitly includes the retrieve and interpret constructs from the reading comprehension sub domain. Even though the other constructs that correspond to the reading comprehension subdomain (reflect, metacognition and motivation and disposition) are desirable, these are not necessary for most of the reading tasks people are faced with in everyday life.

Figure 1 shows the PS and the overall MPL. Figure 2 shows the performance descriptors that are above the MPL.

Figure 1.
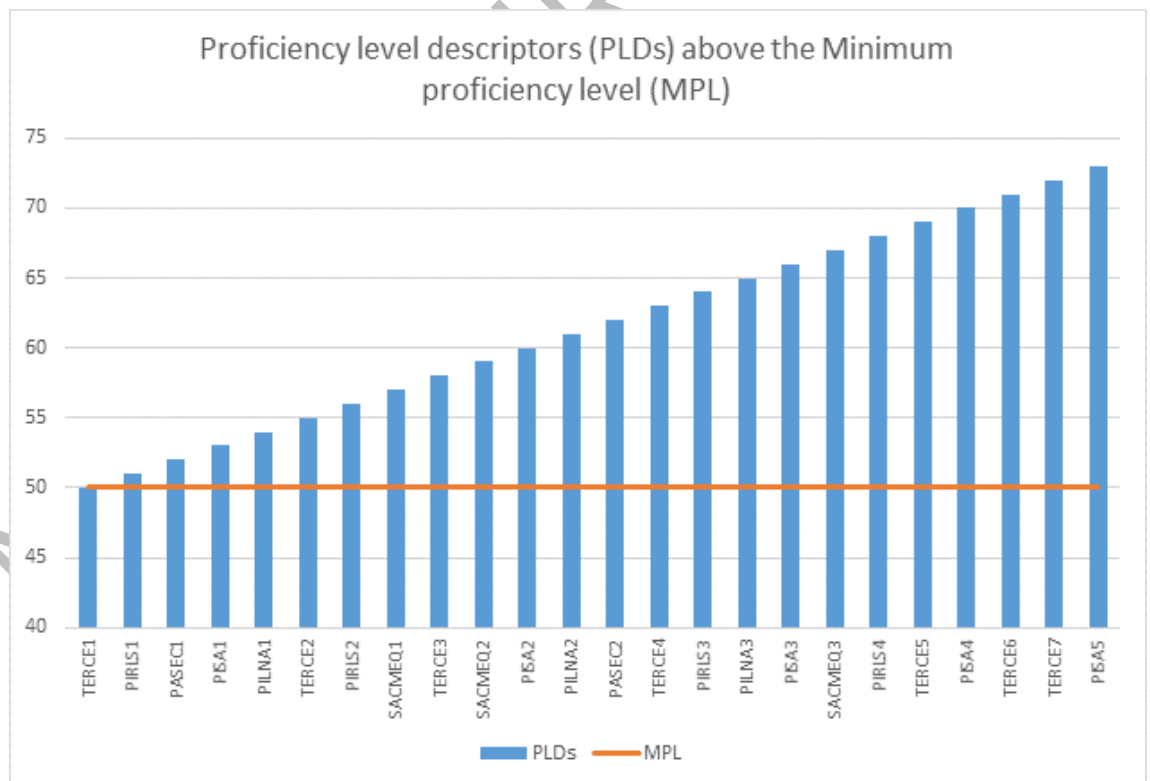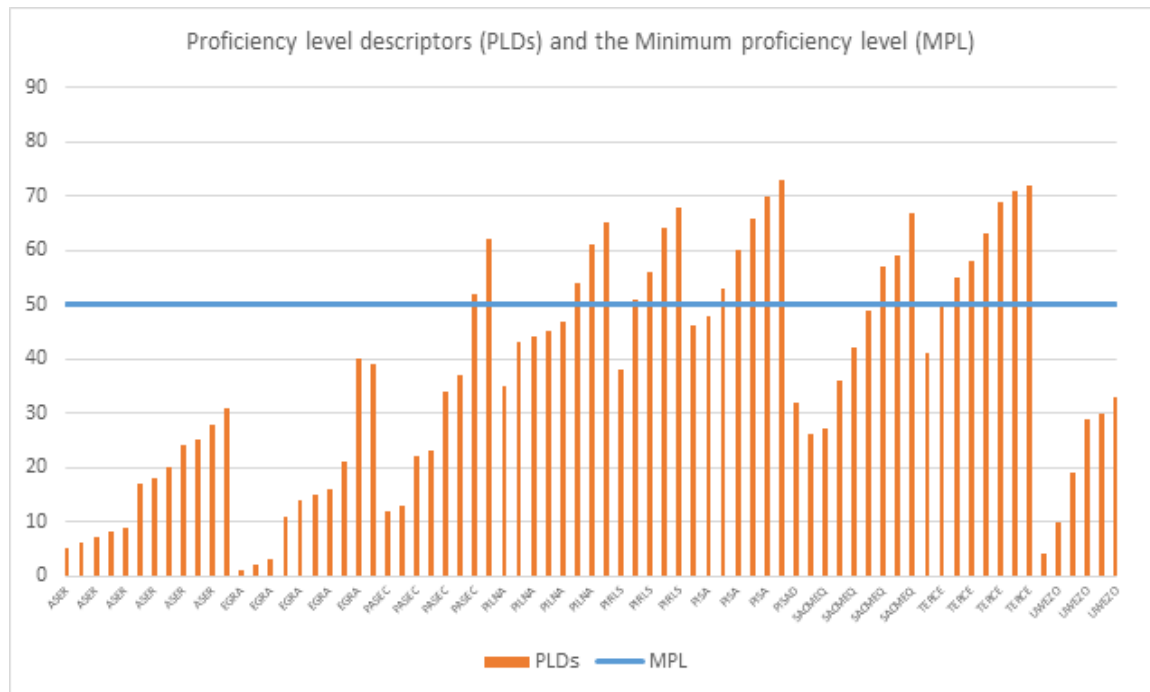
Characteristics of the regional international assessments

The 73 levels of the PS were divided into the three educational levels considering their levels of difficulty as well as the acquisition of skills these entailed. This constitutes the reference scales.

For all of the educational levels the descriptors included in the reference scale spanned from below basic level expected for that grade to advanced knowledge. Therefore, numerous performance descriptors overlap between educational levels.

Subsequently, the performance descriptors that compose each reference scale were divided into four categories according to difficulty. These categories are below basic, basic, proficient and advanced.

The below basic category is constructed based on descriptors that are expected to have already been achieved by the start of the educational level. The basic category, on the other hand, is composed by the performance descriptors that reflect the minimum skills to be acquired during that educational level. The highest descriptor of this category will constitute the MPL expected for that educational level. Moreover, the proficient category entails skills that, though being over the minimum expected, may be developed during the grade by an important percentage of students. Finally, the advanced category was developed in order to be able to consider those students that show very good reading skills.

The next three sections will answer the second question by describing how the PLDs from different assessments map into the reference scale developed for each educational level. A comparison between the MPLs set by each regional and international assessment and the MPL established according to the reference scale will be drawn.

Grade 2 & 3 (4.1.1.a.)

The reference scale for grades 2 & 3 is constituted by 20 PLDs that go from level 22 to 41 from the PS. Levels 22-25 belong to the below basic category, 26-32 to the basic category, 33-38 to the proficient category and 39-41 to the advanced category.

The MPL set for grades 2 & 3 is level 32 from the PS which corresponds to Level 1c from PISA for Development (PISA-D) which is summarized as "students understand the meaning of sentences and very short simple passages with familiar contexts". This is considered the minimum to be expected for this educational level because it implies having achieved mastery regarding precision in decoding, but not necessarily fluency in this sub domain. Moreover, it builds on students' linguistic knowledge by considering familiar contexts and assumes retrieving of simple explicit information.

Figure 2 shows how the different PLDs from the regional and international assessments map into the reference scale for this educational level. The assessment levels that are highlighted by black borders are the established as minimum proficiency by each assessment. While the performance level highlighted with red borders corresponds to the one explained in the previous paragraph.

Figure 2.
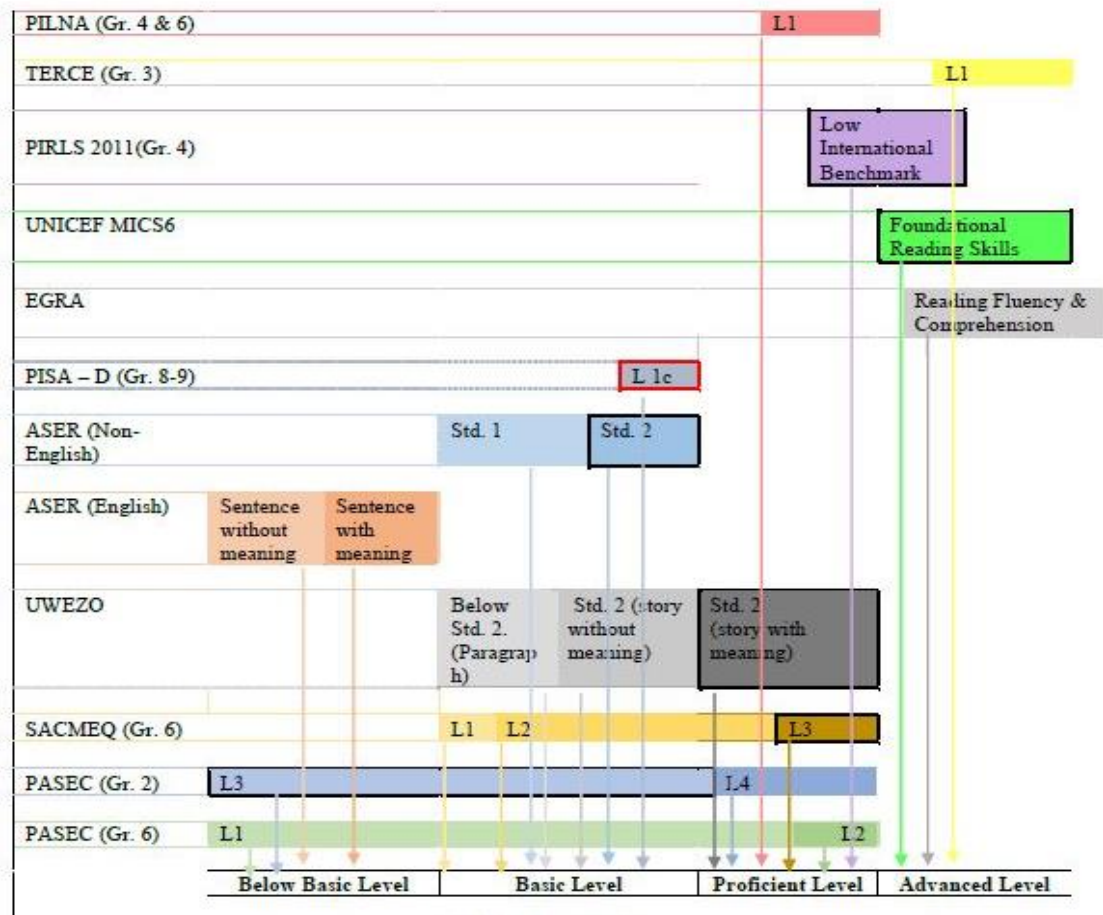
As can be concluded from the figure above, ASER's (Non English) and PASEC's (Grade 2) MPLs are easier than the MPL set in the reference scale. Moreover, TERCE's Level 1 is significantly more difficult than the MPL established for this educational level, being considered in the advanced category.

Finally, there is a surprising overlapping between assessments designed for different educational levels, being both SACMEQ's (grade 6) and PIRLS 2011's (Grade 4) MPL considered as proficient for grades 2 & 3 not far from this educational level's MPL. Moreover, considering ASER, UWEZO and UNICEF MICS 6 as the assessments with a broader application spectrum that cover up to the third educational level, it is interesting that their minimum MPLs correspond to the basic, proficient and advanced categories for grades 2 & 3 respectively.
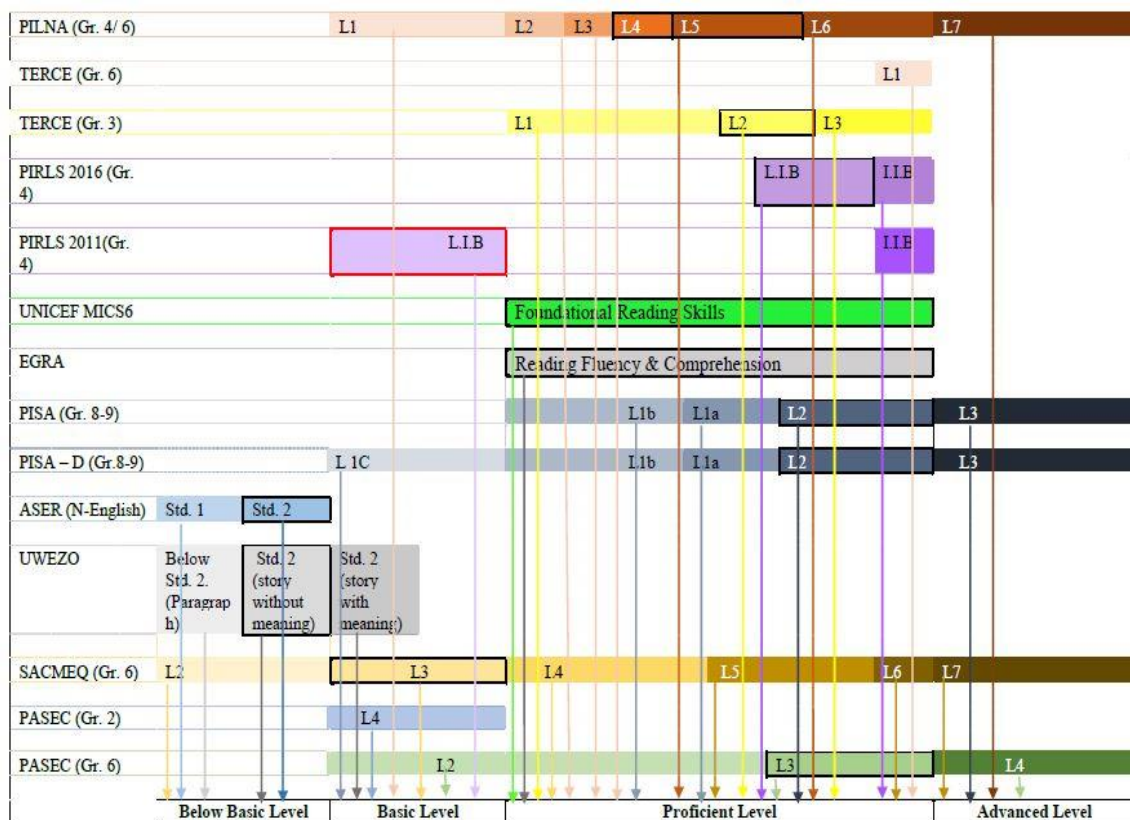
End of primary education (4.1.1.b.)

The reference scale for grades 4 & 6 is constituted by 36 PLDs that go from level 27 to 62 from the PS. Levels 27-31 belong to the below basic category, 32-38 to the basic category, 39-58 to the proficient category and 59-62 to the advanced category.

The MPL set for the End of Primary is level 38 from the PS that corresponds to Low International Benchmark from PIRLS 2011 which is summarized as "students identify and retrieve explicit information from informational and literary texts". This is considered to be the minimum to be expected for this educational level because it implies having achieved mastery regarding decoding as well as having developed at least the possibility of identifying different types of texts and retrieving explicit information from them.

Figure 3 shows how the different PLDs from the regional and international assessments map into the reference scale for this educational level. The assessment levels that are highlighted by black borders are the established as minimum proficiency by each assessment. While the performance level highlighted with red borders corresponds to the one explained in the previous paragraph.

Figure 3.

As can be concluded from the figure above, ASER's (Non English) and UWEZO's MPLs are easier than the MPL set in the reference scale for the end of Primary Education. Even though, the same happens with SACMEQ's, this is closer to the MPL set for this educational level. Moreover, PILNA's MPLs both for grades 4 & 6 are more difficult than the one that has been set, the same happens with PASEC's for grade 6. A very interesting difference is the one that exists between PIRLS 2011 and PIRLS 2016 Low International Benchmark, being the latter significantly more difficult than the former.

Finally, there is a surprising overlapping between assessments designed for different educational levels. The difference between the minimum levels of proficiency expected by

TERCE (grade 3), PASEC's (Grade 6) and PISA (Grades 8 & 9) is surprisingly small considering the grade variation.

End of lower secondary education (4.1.1.c.)

The MPL set for the End of Low Secondary is level 58 from the PS which corresponds to Level 1 from TERCE (Grade 6) which is summarized as "students make causal relations among information from a text and can identify the issuer of a text." This is considered to be the minimum to be expected for this educational level because it implies having achieved mastery regarding decoding as well as being able to retrieve explicit information, interpret the information given by relating it to previous knowledge and reflect upon information from the text as well as its author.

Figure 4 shows how the different PLDs from the regional and international assessments map into the reference scale for this educational level. The assessment levels that are highlighted by black borders are the established as minimum proficiency by each assessment. While the performance level highlighted with red borders corresponds to the one explained in the previous paragraph.

Figure 4.

As can be seen from the figure above, ASER, UWEZO and UNICEF MICS 6 do not appear, as the PLDs used by these assessments are significantly easier than what is expected for this educational level, even though the age range of application corresponds. Furthermore, it is important to note that PISA's MPL is also easier than the one established for this educational level.

Finally, there is evident overlapping between MPLs from different assessments, especially when considering all of the performance levels that correspond to the basic category, in which we can find the minimum proficiency levels expected by PIRLS for grade 4, PASEC for grade 6 and PISA for grade 9. Moreover, there is great overlapping in the advanced category between PISA's highest two levels and TERCE's (grade 6) highest two levels of performance, which is unexpected as there are two grades in between.

After analyzing the three educational levels separately, a summary of the MPLs set for each of them, as well as the overall one will be presented together with the policy descriptor for minimum proficiency.

Analyzing the minimum proficiency levels in the light of policy descriptors

In a previous paper, the process of developing policy descriptors was explained. From that process, a policy descriptor for achieving minimum proficiency in reading was created. That descriptor stated, "Students have developed the required competences for the described reading level. They have acquired the knowledge and skills necessary to decode written words, identify relevant information from written texts, understand their meaning and make inferences from their knowledge".
This section will look at the three MPLs in the light of the policy descriptor.
In the case of Grades 2 & 3 the MPL is "students understand the meaning of sentences and very short simple passages with familiar contexts". From this perspective, the required

competences to be developed in order to achieve this level are precision in decoding individual words and sentences as well as retrieving explicit information from very short passages.

For the end of Primary Education, the MPL set was "students identify and retrieve explicit information from informational and literary texts". In this regard, the competences necessary to achieve this level are precision and certain degree of fluency in decoding, as well as, the identification of different types of texts and retrieving explicit information from them.

Finally, for the end of Low Secondary Education the MPL established is "students make causal relations among information from a text and can identify the issuer of a text". Even though, not explicitly stated in the descriptor, this level implies having developed mastery in decoding regarding both precision and fluency, having achieved a literal comprehension of different types of texts, being able to interpret implicit information from different parts of the text as well as reflect upon the source of the text and its author.

Conclusions and recommendations

Overall, analyzing the MPLs for Reading at three educational cut-points allows for a better understanding on how the process of reading acquisition is expected to develop through formal schooling.

Even though, language and cultural differences may influence the rate of development, generating differences between countries at certain stages, it is believed that the MPL descriptors are specific enough to be measurable and at the same time broad enough to be adjustable to different languages and cultures.

In order to increase international comparability between assessments, agreement has to be reached related to the processes and skills being assessed and the level of development to be expected at each educational level.

In this sense, an option would be to create for each educational level a MPL, but at the same time, to separate that level into processes or skills, being able to assess student's achievements in those separately. In this model, different countries may achieve the MPL at a given educational stage in some processes and skills but not in all of them. This is similar to the model proposed by ACER's Learning Progression Explorer. Therefore, this would take into account country variability, while at the same time increasing comparison potential. Considering the constructs from the Global Framework for Reading in order to establish these processes and skills may prove to be useful.

Furthermore, another way of increasing comparability between regional and international assessments would be to make explicit in the PLDs some information regarding the tasks used. The main task characteristics that may affect PLD difficulty and therefore comparability would be:

a) Text type: continuous or discontinuous; narrative, descriptive, informational, etc.

b) Text length: overall text length as well as how sparse in the text is the information needed to perform the task.

c)   Text topic or meaning: is the topic of the text known to students, are they expected to have previous knowledge about it, would its meaning be clear to them.

d) Vocabulary: use of familiar or non-familiar words, use of technical vocabulary.

e) Different sources of information: does the task involve considering more than one source of information, for example: textual and paratextual information (images, tables, graphs, figures), more than one text.

A description of task characteristics, together with the processes being assessed, could aid in evaluating the overall cognitive demand and difficulty of any given performance level descriptor.