UNESCO INSTITUTE FOR STATISTICS

GLOBAL ALLIANCE TO MONITOR LEARNING

SUSTAINABLE DEVELOPMENT GOALS

Final Report of the Results of the Consensus Building Meeting on Proficiency Levels

This report was prepared by Anthony Nitko, Professor Emeritus, Department of Psychology in Education, University of Pittsburgh.

Table of Contents

Executive Summary

Preamble

The UNESCO Institute for Statistics' (UIS) goal as a custodian agency for reporting against the Sustainable Development Goals in Education (SDG4) is to develop standards, methodology and guidelines to enable countries in the production of data for the reporting of indicators. Indicator 4.1.1 requires member countries to report on the "*proportion of children and young people in Grade 2 or 3 (4.1.1a), at the end of primary education (4.1.1b), and at the end of lower secondary education (4.1.1c) to achieve at least a minimum proficiency level in reading and mathematics".*

This will include the establishment of the reporting mechanism that will enable national governments to effectively report the indicator in a comparable manner; to support the global education community and national governments to measure and monitor students' learning outcomes in mathematics and reading against SDG indicator 4.1.1 over time, and to utilize the data for making informed policy decisions. It is a further goal to support the use of existing national assessments and cross-national assessments to facilitate measurement and reporting for learning outcomes.

In order to report on the three education levels (*in Grade 2 or 3 (4.1.1a), at the end of primary education (4.1.1b), and at the end of lower secondary education (4.1.1c)* in two subject areas (Reading and Mathematics) as specified in indicator 4.1.1, there is a need to define performance or skills needed to achieve proficiency.

Date and venue of the meeting

The meeting took place on 10-11 September 2018 in Paris, France and was kindly hosted by the UNESCO headquarter,7, place de Fontenoy, 75352 PARIS 07 SP, France

Objectives of the meeting

The purposes of the meeting were for the participants to assist UIS to:

- Seek consensus for the reading and mathematics proficiency scales.

- Seek consensus for the alignment of the educational levels (Grades 2-3, Grades 4-6, Grades 8-9) to the respective reading and mathematics proficiency scales.
- Seek consensus for the proficiency levels of each educational level in reading and mathematics, respectively.
- Seek consensus for the placement of the "minimum proficiency level" for each educational level in reading and mathematics, respectively.

Meeting Date, Venue and Agenda

The meeting was held at UNESCO Headquarters, Paris, 10-11 September 2018.

The meeting agenda is shown in Appendix A. The list of participants and partners attending is shown in Appendix B.

Meeting Procedures and Processes Leading to the Consensus Recommendations

1. Prior to the meeting the UIS consultants in mathematics and reading conducted an analysis of the proficiency level descriptors (PLDs) of cross-national, regional, and community-led tests. The tests they analyzed are shown in Appendix C. The analysis resulted in:

- An ordered list (proficiency scale) for mathematics and for reading of essentially all of the proficiencies that were represented in the tests shown in Appendix C.
- An alignment of each of the tests' proficiency reporting levels in relation to the ordered proficiency scale.
- Ordered lists of proficiencies that are\ aligned to each of the three levels of education (lower primary, upper primary, and lower secondary).
- A suggested description of minimum proficiency at each of the three levels of education.

2. During the meeting the participant partners reviewed and discussed each of the above results, offering suggestions for improvement as appropriate.

3. After review, discussions, and suggestions the participants reached consensus on the following:

- The reading and mathematics proficiency scales were developed in a logical and systematic manner and are reasonable ways to relate the many cross-national, regional, and

community-led tests that are now in use to a common ordered list of proficiencies against which users of these tests may come to understand the proficiencies that each assesses.

- The alignment of the different tests reporting levels to the proficiency scale allows countries to use those alignments to help them report attainment of SDG 4.1.1 using one or more of the existing tests shown in Appendix C.
- Minimum proficiency level at each of the three education levels can be demonstrated by attaining one or more of the results shown in Appendix D.
- The minimum proficiency level content descriptions that were drafted at the meeting are appropriate but should be (a) rewritten for clarity and appropriate level of content detail and (b) reviewed by the meeting participants before presenting them at the GAML meeting.
- Follow-up implementation of the consensus meet are shown in Appendix E and were presented at the GAML% meeting in Hamburg, October 2018.

Concluding Consensus

At the end of the meeting a formal questionnaire was administered to the participants. The results are summarized in Appendix G. The appendix shows that the participants reached a consensus that the proficiency scale, the alignment of the assessment program levels, and the minimum proficiency levels defined were satisfactory and that after a final review by the partner participants, they should be brought to the GAML meeting.

Final Report of the Results of the Consensus Building Meeting on Proficiency Levels[1]

I. Introduction

The UNESCO Institute for Statistics' (UIS) goal as a custodian agency for reporting against the Sustainable Development Goals in Education (SDG4) is to develop standards, methodology and guidelines to enable countries in the production of data for the reporting of indicators. Indicator 4.1.1 requires member countries to report on the "*proportion of children and young people in Grade 2 or 3 (4.1.1a), at the end of primary education (4.1.1b), and at the end of lower secondary education (4.1.1c) to achieve at least a minimum proficiency level in reading and mathematics".*

This will include the establishment of the reporting mechanism that will enable national governments to effectively report the indicator in a comparable manner; to support the global education community and national governments to measure and monitor students' learning outcomes in mathematics and reading against SDG indicator 4.1.1 over time, and to utilize the data for making informed policy decisions. It is a further goal to support the use of existing national assessments and cross-national assessments to facilitate measurement and reporting for learning outcomes.

In order to report on the three education levels (*in Grade 2 or 3 (4.1.1a), at the end of primary education (4.1.1b), and at the end of lower secondary education (4.1.1c)* in two subject areas (Reading and Mathematics) as specified in indicator 4.1.1, there is a need to define performance or skills needed to achieve proficiency.

The September 2018 consensus meeting at UNESCO Headquarters in Paris was convened to present the work that UIS had completed and to seek consensus from UIS partners on setting minimum proficiency levels (MPLs).

Date and Venue of the Meeting

The meeting took place **on** 10-11 September 2018 in Paris, France and was kindly hosted by the UNESCO headquarter, 7, place de Fontenoy, 75352 PARIS 07 SP, France.

Objectives of the Meeting

[1] 10-11 September 2018, UNESCO Headquarters, 7, place de Fontenoy, , 75352 PARIS 07 SP, France

The purposes of the meeting were for the participants to assist UIS to:

- Seek consensus for the reading and mathematics proficiency scales.
- Seek consensus for the alignment of the educational levels (Grades 2-3, Grades 4-6, Grades 8-9) to the respective reading and mathematics proficiency scales.
- Seek consensus for the minimum proficiency levels of each educational level in reading and mathematics, respectively.
- Seek consensus for the placement of the "minimum proficiency level" for each educational level in reading and mathematics, respectively.

Meeting Agenda

The meeting agenda is shown in Appendix A. The list of participants and partners attending is shown in Appendix B.

II. Brief Overview of Meeting Procedures and Processes Leading to the Recommendations

1. Prior to the meeting UIS conducted an analysis of the proficiency level descriptors (PLDs) of cross-national, regional, and community-led tests in mathematics and reading. The tests analyzed are shown in Appendix C. The analysis resulted in:

- An ordered proficiency scale for mathematics and for reading of essentially all of the proficiency level descriptors (PLDs) that were represented in the tests shown in Appendix C.
- An alignment of each of the tests' proficiency reporting levels in relation to the ordered proficiency scale.
- Ordered lists of proficiencies that are aligned to each of the three levels of education (lower primary, upper primary, and lower secondary).
- A suggested description of minimum proficiency at each of the three levels of education.

2. During the meeting the participant partners reviewed and discussed the methodology and each of the above results, offering suggestions for improvement as appropriate.

3. After review, discussions, and suggestions the participants reached consensus on the following:

- The proposed methodology was deemed adequate and pragmatic.
- The reading and mathematics proficiency scales were developed in a logical and systematic manner and are reasonable ways to relate the many cross-national, regional, and community-led tests that are now in use to a common ordered list of proficiencies against which users of these tests may come to understand the proficiencies that each assesses.

- The alignment of the different tests reporting levels to the proficiency scale allows countries to use those alignments to help them report attainment of SDG 4.1.1 using one or more of the existing tests shown in Appendix C and as demonstrated by the diagrams in Appendix F.
- Minimum proficiency level at each of the three education levels can be demonstrated by reaching the level taken as minimum in either of the alternatives shown in Appendix D.
- The minimum proficiency level content descriptions that were drafted at the meeting are appropriate but should be (a) rewritten for clarity and appropriate level of content detail and (b) reviewed by the meeting participants before presenting them at the GAML5 meeting.

III. Mathematics PLD Analysis Prior to the Meeting[2]

The following is a description of the work done prior to the meeting and which what presented to the partner participants during the meeting. This material has been extracted from the UIS (2018a, September) document.

Definition of Performance Level Descriptors

Each assessment in Appendix C has a number of performance level descriptors (PLDs) associated with it. These PLDs delineate one or more mathematical skills and/or processes that are associated with test takers who achieve that performance level. The number of PLDs varies by assessment, as does the format in which the PLD's are written. Examples of mathematical skills include counting, adding fractions, solving equations; examples of mathematical processes include employing basic formulas, interpreting problem situations, and communicating reasoning.

Analysis, Comparison, and Ordering

The primary, if not sole, criterion for analysing PLDs is the *cognitive demand* required by the mathematical skills and/or processes contained in each PLD. This is complicated by the fact that most, if not all, PLDs contain multiple skills and processes. Thus, comparing PLDs becomes a matter of determining and comparing the *overall* cognitive demand of each PLD. This requires a high level of careful analysis, and is as much art as science. Successively comparing PLDs against each other eventually resulted in a list of

[2] The mathematics consultant preparing this work was Michael Bell.

the PLDs within each measurement point (i.e., educational level), arranged from lowest to highest overall cognitive demand. As an additional point of information, each PLD was given a one-sentence summary, which may facilitate easier comparison for future work.

Proficiency Scale

Once the list of PLDs for each measurement point was completed, it was then necessary, and possible, to create the overall Proficiency Scale for mathematics. This was begun by placing *all* the PLD's from *all* three measurement points into a single list, from the lowest of grades 2-3 to the highest of grades 8-9. However, it could not be assumed that the highest-level PLD of one measurement point had a lower cognitive demand than the lowest-level PLD of the next-highest measurement point. The next step was, therefore, to compare the high-level PLDs of grades 2-3 against the low-level PLDs of grades 4-6, utilising the same process of comparing the overall cognitive demand of the PLDs, and re-arranging PLD's as appropriate. This was then repeated with the PLDs at the border of grades 4-6 and grades 8-9. This resulted in a list of *all* PLDs across all three measurement points. (See Appendix F.)

Ordering within Measurement Points

The final step after creating the Proficiency Scale was to identify which PLDs contained grade-level appropriate (GLA) skills and processes for each measurement point. For this step, cognitive demand was *not* a criterion, as each measurement point contains a range of skills from low to high cognitive demand. The Proficiency Scale includes a number of PLDs that did not contain GLA skills or processes even at the lowest measurement point. It also included many PLDs that were GLA at more than one measurement point.

Once the Proficiency Scale was complete, it was then possible to set the performance levels at each measurement point, using the list of GLA PLDs. Each measurement point used the same four performance levels—Below Basic; Basic; Proficient; and Advanced. As with the first step in the process, determining where to set each performance level required a good deal of careful analysis, especially since the skills and processes taught at each grade can vary, in some cases widely, from nation to nation. Finally, at each measurement point, the lowest PLD in the proficient performance level was marked as the dividing line between proficient less than proficient test takers. (See Appendix F.)

Policy Level Descriptors

Previously, policy level descriptors in the area of mathematics were developed to characterize (in general terms) the difference in ability between mathematically proficient test takers and those who

achieve at a level below proficiency. These policy level descriptors reflect the dividing line between proficient and non-proficient test takers, even though they do not delineate between the two sub-categories at each level: Below Basic vs Basic, and Proficient vs Advanced. The policy level descriptors are an exceedingly useful and important tool that can be used to validate that the content described at each measurement point is an accurate reflection of the mathematical skills and processes for which students around the world should be expected to demonstrate a certain degree of mastery.

IV. Reading PLD Analysis Prior to the Meeting[3]

The following is a description of the work done prior to the meeting and which what presented to the partner participants during the meeting. This material has been extracted also from the UIS (2018a, September) document.

Development of a Proficiency Scale

The first step taken in this process was to develop a Proficiency Scale on reading. In this regard, all of the performance level descriptors 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 Proficiency Scale that considers all of the performance level descriptors 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 Proficiency Scale is the incongruence between the expectations set by different regional and international assessments as well as the overlapping of performance level descriptors designed for different educational levels.

Finally, an overall **minimum proficiency level** was set for reading in general. This was marked at the 50th level on the Proficiency Scale 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

[3] The reading consultants preparing this work were: Ariel Cuadro Cawen, Carola Ruiz Hornblas, and Anna Laura Palombo Segredo

topic and making inferences regarding non-evident information". If we analyze it from the reading global framework's 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.

Development of a Reference Scale and Minimum Proficiency Level for each Educational Level

The 73 levels of the Proficiency Scale 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 **minimum proficiency level** 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 describe how the performance level descriptors from different assessments map into the reference scale developed for each educational level. A comparison between the minimum proficiency level set by each regional and international assessment and the minimum proficiency level established according to the reference scale will be drawn.

Grades 2 & 3 (4.1.1a.)

The reference scale for grades 2 & 3 is constituted by 20 performance level descriptors that go from level 22 to 41 from the Proficiency Scale. 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 minimum proficiency level set for grades 2 & 3 is level 32 from the proficiency scale 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.

V. Participants' Discussion of the Proficiency Scales

After the content experts presented the proficiency scales and the methodology for developing them, there ensued a lively discussion among the partner participants and UIS. Among the topics discussed were clarification and further explanation about the process and methods used by mathematics and by reading; whether there should be better assessment at the lower educational levels, especially to distinguish between grade 2 and Grade 3 rather that putting them together; whether there should be different minimum proficiency levels for countries who are far behind many more developed countries in implementing their educational improvement programs; how different cultures and language groups acquire reading skills; whether it would be better to have only two proficiency levels for each measurement point (i.e., proficient and not proficient); and whether proficiency levels setting would result in a large majority of student being classifies as non-proficient for some less developed counties.

After the discussion, the consensus of the participants was:

- The proficiency scales were developed using a logical and systematic approach.
- The proficiency scales show a way forward so that countries could relate their results from cross-national, regional, or community-based assessments to the same mathematics and reading frameworks and scales.
- There should be an attempt made to review and validate the proficiency scales by either a complete independent replication or an independent review of the process and procedures the current scales were developed by.
- It was suggested that an attempt be made when measuring proficiency to consider response format (e.g., oral verse written), language difference in acquiring reading skills, and cultural difference in how the students express their skills.

- The consensus was to have only two proficiency levels per each of the three educational levels: proficient and not proficient.

Issues for which consensus was not reached and would need further discussion during the second meeting day:

- How high or how low should the minimum proficiency levels be set. The proportion of a country's student population meeting the minimum proficiency level can be estimated from the cross-national and regional test results since the proficiency scales have been aligned with these tests.
- There may be a need to estimate the impact on countries where the minimum proficiency level is too high or too low. Setting the minimum proficiency level too high might discourage teachers.
- There may be a need to write policy level descriptors that correspond to the minimum proficiency and proficiency levels to make them understandable for educational policy makers.
- There remained a question concerning whether the minimum proficiency level should (a) reflect what should be learned after X years of school or (b) reflect what needs to be learned to be successful at the next level of education.
- The wording of the description of the performance levels just below and just above the minimum proficiency level need so be clearly written so they are clearly differentiated.

It was noted that several participants had not had the opportunity to reads the complete documents about PLD analyses and the proficiency scale development prior to the first day's meeting. Thus, the participants were urged to study these documents before the second day's meeting was convened.

VI. Development of the Consensus Minimum Proficiency levels

The second day's meeting was devoted to answering any remaining questions about the mathematics and reading proficiency scales and to the development of minimum proficiency levels for each of the three educational levels.

The participants were presented with suggestions from the content experts concerning which cross-national, reginal, and community-based proficiency levels constituted minimum proficiency at each of the three educational levels. The participating partners discussed these and

great length and made several adjustments and changes. One of the changes was to allow a country to report at either Grade 2 or Grade 3 for the lower primary educational level because the nature of schooling and the curricula varies greatly among countries.

VII. Minimum Proficiency Levels Defined

At the conclusion of the discussions minimum proficiency levels were set for math and for reading at each of the three educational levels. These MPLs and descriptors are shown in Appendix D. (See Appendix E also.)

VIII. Follow up for Clarifying Minimum Proficiency Levels

Before concluding the meeting, the partners suggested several follow up activities:

- Review the verbal performance descriptors for each MLP at each educational level to present them in a general content language so they are understandable by educational policy makers and so they would not be too technical. This may entail having a second statement for each descriptor to the "unpacks" the general descriptor statement for those interested in the specific skill encompassed.
- Accompany each MLP with example test items that illustrate the types of skills each MLP implies that students should be able to answer. This will further clarify the meaning of thee MPLs.
- Given the short time remaining before the GAML5 meeting in October, it was recommended that the proficiency scale and MPLs be reviewed by outside experts and then UIS have a discussion with them about any suggestions they may have. There would be no time for a comprehensive independent evaluation.

IX. Concluding Consensus

At the end of the meeting a formal questionnaire was administered to the participants. The results are summarized in Appendix E. The table shows that the participants reached a consensus that the proficiency scale, the alignment of the assessment program levels, and the minimum proficiency levels defined were satisfactory and after a final review by the partner participants, they should be brought to the GAML meeting.

X. Next Steps

The work will encompass two additional steps. First, one is the unpacking of the general PLD into examples of tasks that could help to operationalize the concept. The second steps are to add examples of items in current assessment that embed the proficiency that is required in the PLD with the potentiality to add some generic items as examples as well. The results thus far from the follow-up of the consensus meeting are shown in Appendix E. This appendix was included in the GAML5 meeting in Hamburg and submitted for consensus there.

References

UIS (2018a, September). Methodology for Mapping Performance Level Descriptors (Mathematics). (Mapping Paper_Drfat 1)

UIS (2018b, September) Compilation of Performance Level Descriptors across regional and international assessments). (Reading).

UIS (2018c, September). Methodology for Analyzing/Ordering Performance Level Descriptors (Mathematics)(Methodology Paper_Draft 1.docx)

UIS (2018d, September). Methodology for Ordering Performance Level Descriptors (Mathematics). (MATH_ordered_PLD_methodology_paper.pdf)

UIS (2018e) Minimum proficiency levels (MPLs): outcomes of the consensus building meeting Background papers (GAML5/REF/4.1.1-29).(4.1.1_29_Consesnus building meeting package.pdf)

Appendix A. Agenda for the Paris Meeting September 2018

Day 1: Monday, 10 September 2018	
13:00 – 13:30	Registration
13:30 – 14:00	**1.** Opening session a. Welcome b. Introduction of participants c. Objectives of the meeting d. Work plan for the meeting
14:00-15:30	**2.** Status of the work a. Overview of what has been done thus far, Dr. Anthony Nitko b. Review of reading proficiency scales, Ms. Carla Ruiz c. Review of mathematics proficiency scales, Mr. Michael Bell DISCUSSION and CONSENSUS
15:30-16:00	Coffee Break
16:00-17:00	**3.** Educational levels alignments to the proficiency scales a. Review of reading alignment, Ms. Carla Ruiz b. Review of mathematics alignment, Mr. Michael Bell DISCUSSION and CONSESUS

Continued on the next page.

Day 2: Tuesday,11 September 2018	
09:15 – 10:45	• Proficiency Level Descriptors, Suggested performance levels and descriptors for each educational level • Introduction, Dr. Anthony Nitko • Mathematics, Mr. Michael Bell • Reading, Ms. Carla Ruiz DISCUSSION and CONSESUS **Moderator**: Dr. Anthony Nitko, Mr. Michael Bell, Ms. Carla Ruiz
10:45 – 11:00	Coffee Break
11:00 – 13:30	**4.** Minimum proficiency at each educational level a. Minimum proficiency levels for mathematics, Mr. Michael Bell b. Minimum proficiency levels for reading, Ms. Carla Ruiz DISCUSSION Moderator: Dr. Anthony Nitko **5.** Summary of consensus DISCUSSION Moderator: Dr. Anthony Nitko
	6. Concluding discussion a. Value of these inputs b. Next steps c. Concluding remarks **Chair**: Silvia Montoya, UIS

Appendix B. List of Participants at the Paris Consensus Meeting September 2018

First Name	Last Name	Organization
Maurice	Walker	ACER
Ketan	Verma	ASER Centre
Baela	Jamil	ASER/ITA
Juliane	Hencke	IEA
Oliver	Neuschmidt	IEA
Michael	Ward	OECD (Pfd)
Miyako	Ikeda	OECD (Pfd)
Hilaire	Hounkpodote	PASEC/CONFEMEN
Labass	Lamine	PASEC/CONFEMEN
Ethel Agnes	Pascua-Valenzuela	SEAMEO
Silvia	Montoya	UIS
Friedrich	Huebler	UIS
Ariel	Cuadro Cawen	UIS consultant
Carola	Ruiz Hornblas	UIS consultant
Anna Laura	Palombo Segredo	UIS consultant
Michael	Bell	UIS consultant
Anthony	Nitko	UIS consultant
Atilio	Pizarro	UNESCO Santiago (OREALC)
Camilla	Woeldike	UNICEF (SEA-PLM Secretariat)
Manuel	Cardoso	UNICEF
Marguerite	Clarke	World Bank
Caine	Rolleston	Young Lives

Appendix C. Assessment Programs whose PLDs were Analyzed

	Assessment Name	Type of Assessment	Level of Assessment
ASER	Annual Status of Education Report	National Citizen-Led	Grades 2-3
EGRA	Early Grade Reading Assessment	Cross-national	Grades 2-3
PASEC	The Analysis Program of the CONFEMEN Education Systems	Regional	Grades 2-3
TERCE	Third regional Comparative and Exploratory Study	Regional	Grades 2-3
UNICEF MICS6	UNICEF Multiple Indicator Cluster Service	Household Survey	Grades 2-3
Uwezo	Capacity Annual Learning Assessment	National Citizen-Led	Grades 2-3
PASEC	The Analysis Program of the CONFEMEN Education Systems	Regional	Grades 4-6
PILNA	Pacific Islands Literacy ad Numeracy Assessment	Regional	Grades 4-6
PIRLS	Progress in International Reading Literacy Study	Regional	Grades 4-6
SACMEQ	Southern and Eastern African Consortium for Monitoring Educational Quality	Regional	Grades 4-6
PILNA	Pacific Islands Literacy ad Numeracy Assessment	Regional	Grades 4-6
PIRLS	Progress in International Reading Literacy Study	Cross-national	Grades 4-6
TERCE	Third regional Comparative and Exploratory Study	Regional	Grades 4-6
TIMSS	Trends in International Mathematics and Science Study	Cross-national	Grades 4-6
PISA, PISA -D	Progress in International Reading Literacy Study	Cross-national	Grades 8-9
TIMSS	Trends in International Mathematics and Science Study	Cross-national	Grades 8-9

Appendix D. Minimum Proficiency in Reading and in Mathematics in relation to Results on Existing Cross-national, Regional, and Citizen-led Tests that emerged from the September 2018 Consensus Meeting.

READING		
Educational Level	Descriptor	Assessment PLDs that align with the descriptor
Grades 8 & 9	Students establish connections between main ideas on different text types and the author´s intentions. They reflect and draw conclusions based on the text.	• PISA 2015 – Level 2 • PILNA 2015 – Level 6 • TERCE 2014 (Gr. 3) – Level 3 • PIRLS 11/16 – Intermediate • SACMEQ 2007 – Level 6 • TERCE 2014 (Gr. 6) – Level 1
Grades 4 & 6	Students interpret and give some explanations about the main and secondary ideas in different types of texts. They establish connections between main ideas on a text and their personal experiences as well as general knowledge.	• SACMEQ 2007 – Level 3 • PASEC 2014 (Gr. 6) – Level 2 • PIRLS 2011 – Low • SERCE 2006 (Gr. 6) – Level 2
Grade 3	Students read aloud written words accurately and fluently. They understand the overall meaning of sentences and short texts. Students identify the texts' topic.	• PISA-D – Level 1c • Uwezo – Std. 2 (Story with meaning) • PASEC 2014 (Gr. 2) – Level 4 • TERCE (Gr. 3) – Level 1 • UNICEF MICS 6 – Foundational Reading Skills • EGRA – Level 9 • ASER – Std. 2 (story)
Grade 2	Students read and comprehend most written words, particularly familiar ones, and extract explicit information from sentences.	• PASEC (Gr. 2) – Level 3

Continued on to next page

MATHEMATICS		
Education Level	Descriptor	Assessment PLDs that align with the descriptor
Grades 8 & 9	Students can solve computation and application problems involving whole numbers, and interpret the results. Students can match tables to bar graphs and pictographs. They can make use of formulae and algebraic representations.	• PISA 2015 -- Level 2 • TIMSS 2015 -- Low International
Grades 4 & 6	Students can perform the four basic operations with whole numbers, fractions, and decimals, including in application problems. They can order whole numbers and decimals. Students can interpret common measurement units, tell time on an analogue clock, and estimate the weight and length of real-world objects. Students can read, interpret, and construct tables and graphs. Students can identify position, direction, and coordinates on maps and graphs. They can visualize three-dimensional shapes from two-dimensional drawings. Students can identify rules for, and continue, number patterns. They can solve simple application problems involving proportional reasoning.	• SACMEQ 2007 -- Level 3 • SACMEQ 2007 -- Level 4 • PASEC 2014 -- Level 1 • PILNA _2015_ Level 6 • TERCE 2014 __ Level 1 • TIMSS 2015 __ intermediate international benchmark
Grade 2/3	Students can read, write, and compare whole numbers up to 100. They can add and subtract numbers up to fifty and solve application problems with numbers up to twenty. Students can recognize simple shapes and their elements. They possess foundational knowledge of spatial orientation, and can appraise the relative size of real-world objects.	• TERCE 2014 -- Level 2 • PASEC 2014 -- Level 1 • PASEC 2014 -- Level 2

Appendix E. Minimum Proficiency in Reading and in Mathematics in relation to Results on Existing Cross-national, Regional, and Citizen-led Tests that Implemented the Recommendations from the Consensus Meeting

Minimum Proficiency Levels for Mathematics

Educational Level	Descriptor	Assessment PLD's that align with the descriptor	MPL's in the Assess-ments
Grades 2-3	Students demonstrate skills in number sense and computa-tion, shape recognition and spatial orientation.	PASEC 2014 – Level 1 PASEC 2014 – Level 2 TERCE 2014 – Level 2	Level 2 Level 2
Grades 4-6	Students demonstrate skills in number sense and computa-tion, basic measurement, read-ing, interpreting, and construct-ing graphs, spatial orientation, and number patterns.	PASEC 2014 – Level 1 SACMEQ 2007 – Level 3 SACMEQ 2007 – Level 4 PILNA 2015 – Level 6 TERCE 2014 – Level 1 TIMSS 2015 – Intermediate Interna-tional	Level 2 Level 3 Level 5 Level 2 Intermediate International
Grades 8 & 9	Students demonstrate skills in computation, application prob-lems, matching tables and graphs, and making use of alge-braic representations.	PISA 2015 – Level 2 TIMSS 2015 – Low International	Level 2 Intermediate International

Unpacking of the general descriptors

- Number sense: skills such as reading, writing, comparing, and ordering numbers.
- Computation: math problems presented without context, in arithmetic form, such as 38 + 67 or 23×92.
- Spatial orientation: position and direction on a diagram, map, or graph, often described by words such as "above", "below", "left", "right", "inside", "outside", etc.
- Application problems: also known as "word problems" or "story problems", these are problems that are presented in context, without explicitly telling students *which* mathematical opera-tion(s) to use.
- Algebraic representations: examples include expressions, equations, and inequalities, all of which contain one or more variables.

CONTINUE ON THE NEXT PAGE

Minimum Proficiency Levels for Reading

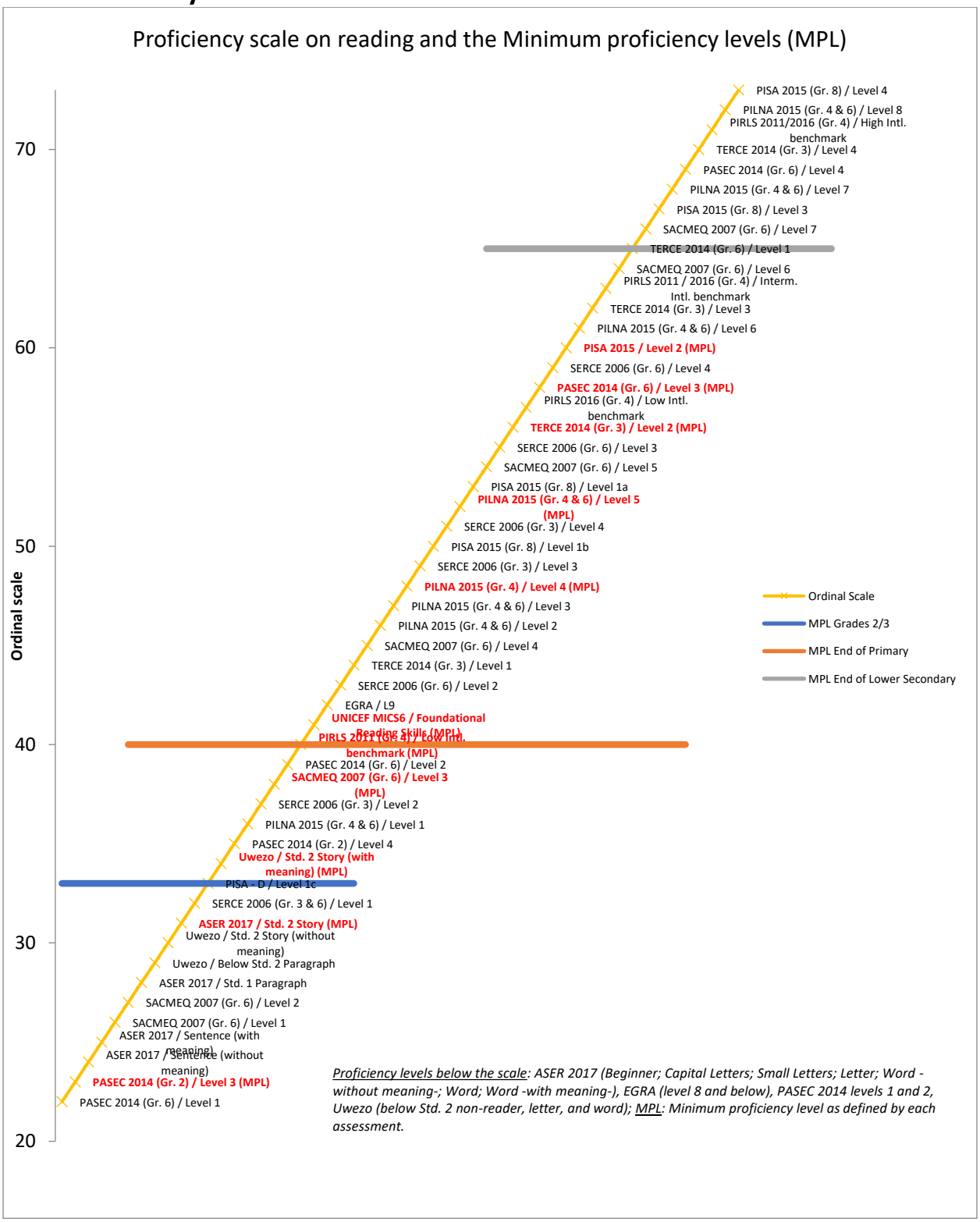
Educational Level	Descriptor	Assessment PLDs that align with the descriptor	MPL in the assessment, if available
Grade 2	They read and comprehend most of written words, particularly familiar ones, and extract explicit information from sentences.	• PASEC (Gr. 2) – Level 3*	• Level 3
Grade 3	Students read aloud written words accurately and fluently. They understand the overall meaning of sentences and short texts. Students identify the texts' topic.	• PISA-D – Level 1c	• Level 2
		• Uwezo – Std. 2 (Story with meaning)*	• Std. 2 (Story with meaning
		• PASEC 2014 (Gr 2)–Level 4*	• Level 3
		• TERCE (Gr. 3) – Level 1*	• Level 2
		• UNICEF MICS 6 – Foundational Reading Skills*	• Foundational Reading Skills
		• EGRA – Level 9*	• Not specified
		• ASER – Std. 2 (story)*	• Std. 2 (story)
Grades 4 & 6	Students interpret and give some explanations about the main and secondary ideas in different types of texts. They establish connections between main ideas on a text and their personal experiences as well as general knowledge.	• SACMEQ 2007 – Level 3	• Level 3
		• PASEC 2014 (Gr. 6) – Level 2	• Level 3
		• PIRLS 2011 – Low	• Low
		• SERCE 2006 (Gr. 6) – Level 2	• Level 1 (appears that way from Technical reports)
Grades 8 & 9	Students establish connections between main ideas on different text types and the author´s intentions. They reflect and draw conclusions based on the text.	• PISA 2015 – Level 2	• Level 2
		• PILNA 2015 – Level 6	• Level 4 (grade 4) and Level 5 (grade 5)
		• TERCE 2014 (Gr. 3) – Level 3	• Level 2
		• PIRLS 11/16 - Intermediate	• Low
		• SACMEQ 2007 – Level 6 • TERCE 2014 (Gr. 6) – Level 1	• Level 3 • Level 2

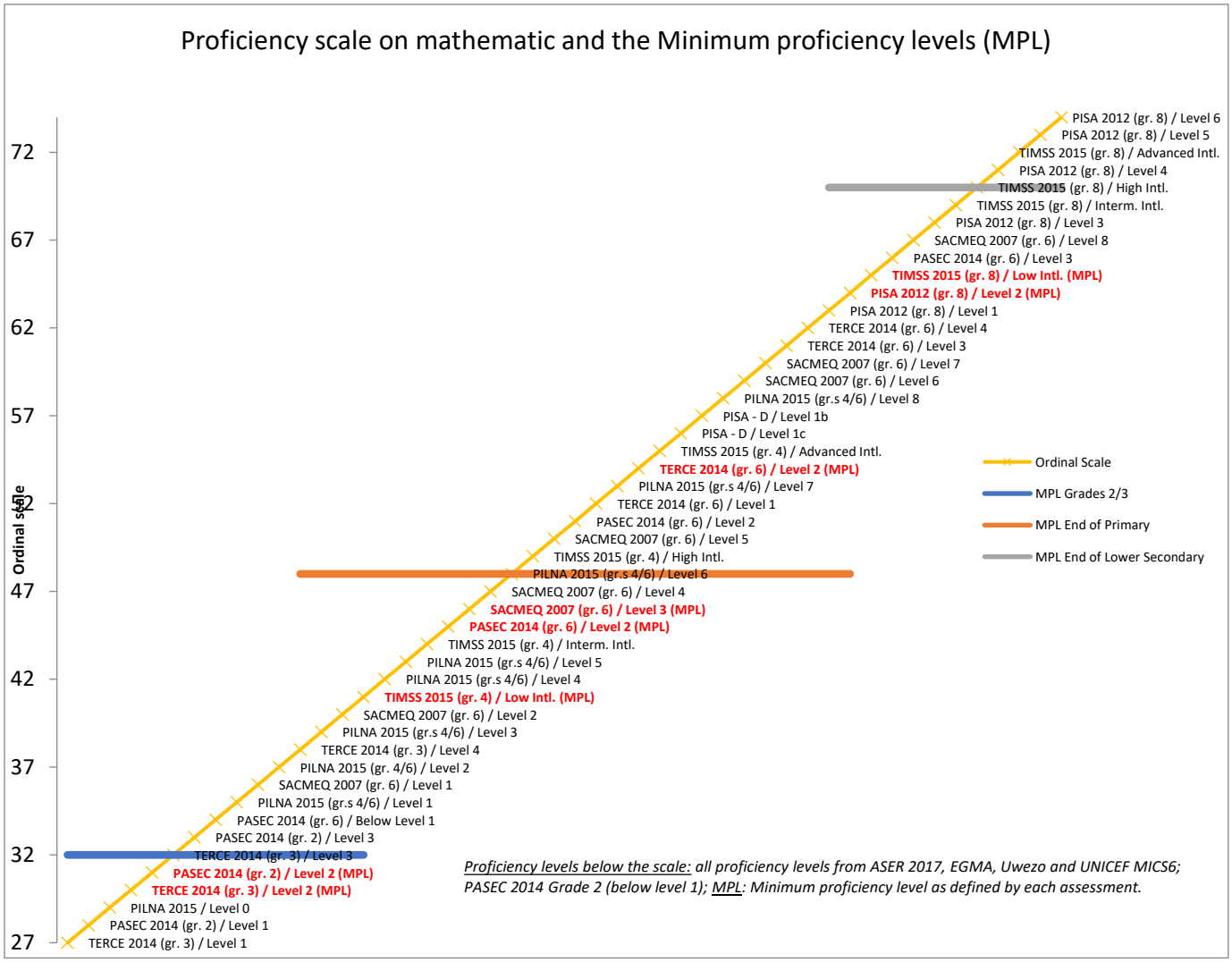
**Subject to alignment to the MPL.*

Unpacking of the general descriptors:

- Familiar words: words that are part of the students' vocabulary and that have been read before more than once.
- Explicit information: information that is presented in the text.
- Accuracy/Precision (in decoding): Correct recognition of the phonological form of a word based on its orthographic form.
- Fluency (in decoding): Presupposes accuracy and speed in word recognition. It can also include qualities such as volume (reading at a volume that is adequate to the instructions given or the audience), pace (adjusting the pace to the instructions, to improve precision or comprehension), expressiveness and tone (adjusting it to the audience' characteristics, to the content and the characters).
- Short texts: texts that are between 60-80 words in length.
- Overall meaning of a text or sentence: refers to the most relevant information of the text.
- Topic of a text: an identified theme or subject.
- Interpret: Extract and recognize implicit and explicit information from a written sentence or text to relate it with other information or apply it to new situations or problem solving.
- Text types: narrative, descriptive, expository, procedural, verbal interaction, that report a central paragraph and complementary information and reference texts.
- General knowledge: previous knowledge that the student has in reference to everyday life and world affairs.
- Author´s intentions: may include the author´s choices (literary resources, title, words, etc.); the author´s feelings or motivations when/for writing, the author´s aim when writing, the author´s intentions when sharing a text in social media or publishing online.
- Reflect: Critically analyze and give an opinion about the information presented in a written sentence or text and the consequences the information may have.
- Draw conclusions: Generate conclusions from a text; generate conclusions about a topic considering different sources of information; generate conclusions about a character´s motivations or intentions.

Appendix F. Proficiency Scales for Reading and Mathematics with Minimum Proficiency Level marked for each Measurement Point




Proficiency scale on mathematic and the Minimum proficiency levels (MPL)
Ordinal scale
72
67
62
57
52
47
42
37
32
27
PISA 2012 (gr. 8) / Level 6
PISA 2012 (gr. 8) / Level 5
TIMSS 2015 (gr. 8) / Advanced Intl.
PISA 2012 (gr. 8) / Level 4
TIMSS 2015 (gr. 8) / High Intl.
TIMSS 2015 (gr. 8) / Interm. Intl.
PISA 2012 (gr. 8) / Level 3
SACMEQ 2007 (gr. 6) / Level 8
PASEC 2014 (gr. 6) / Level 3
TIMSS 2015 (gr. 8) / Low Intl. (MPL)
PISA 2012 (gr. 8) / Level 2 (MPL)
PISA 2012 (gr. 8) / Level 1
TERCE 2014 (gr. 6) / Level 4
TERCE 2014 (gr. 6) / Level 3
SACMEQ 2007 (gr. 6) / Level 7
SACMEQ 2007 (gr. 6) / Level 6
PILNA 2015 (gr.s 4/6) / Level 8
PISA - D / Level 1b
PISA - D / Level 1c
TIMSS 2015 (gr. 4) / Advanced Intl.
TERCE 2014 (gr. 6) / Level 2 (MPL)
PILNA 2015 (gr.s 4/6) / Level 7
TERCE 2014 (gr. 6) / Level 1
PASEC 2014 (gr. 6) / Level 2
SACMEQ 2007 (gr. 6) / Level 5
TIMSS 2015 (gr. 4) / High Intl.
PILNA 2015 (gr.s 4/6) / Level 6
SACMEQ 2007 (gr. 6) / Level 4
SACMEQ 2007 (gr. 6) / Level 3 (MPL)
PASEC 2014 (gr. 6) / Level 2 (MPL)
TIMSS 2015 (gr. 4) / Interm. Intl.
PILNA 2015 (gr.s 4/6) / Level 5
PILNA 2015 (gr.s 4/6) / Level 4
TIMSS 2015 (gr. 4) / Low Intl. (MPL)
SACMEQ 2007 (gr. 6) / Level 2
PILNA 2015 (gr.s 4/6) / Level 3
TERCE 2014 (gr. 3) / Level 4
PILNA 2015 (gr. 4/6) / Level 2
SACMEQ 2007 (gr. 6) / Level 1
PILNA 2015 (gr.s 4/6) / Level 1
PASEC 2014 (gr. 6) / Below Level 1
PASEC 2014 (gr. 2) / Level 3
TERCE 2014 (gr. 3) / Level 3
PASEC 2014 (gr. 2) / Level 2 (MPL)
TERCE 2014 (gr. 3) / Level 2 (MPL)
PILNA 2015 / Level 0
PASEC 2014 (gr. 2) / Level 1
TERCE 2014 (gr. 3) / Level 1
Ordinal Scale
MPL Grades 2/3
MPL End of Primary
MPL End of Lower Secondary
Proficiency levels below the scale: all proficiency levels from ASER 2017, EGMA, Uwezo and UNICEF MICS6; PASEC 2014 Grade 2 (below level 1); MPL: Minimum proficiency level as defined by each assessment.

Appendix G. Rating Results of Questionnaire and Feedback from Participants at the Consensus Meeting, Paris, 10-11 September 2018

		Total Number	Omit	Strongly disagree	Disagree	Agree	Strongly agree
Do you agree that in general, after discussions and your suggestions, that the processes and the outputs presented in the past two days are a useful way for UIS to identify skills and abilities needed to achieve proficiency in order to report on the three education levels?	(1) In Mathematics?	13	8%	8%	0	62%	23%
	(2) In Reading?	13	8%	8%	0	62%	31%
Do you agree that in general, after discussions and your suggestions, that the processes and the outputs presented in the past two days are a useful way for UIS to support the use of existing national assessments and cross-national assessments for measuring and reporting student learning outcomes in order to report on the three education levels?	(3) In Mathematics?	13	0	8%	0	77%	15%
	(4) In Reading?	13	0	8%	0	77%	15%
Do you agree that the *minimum proficiency levels* for the three educational levels appropriate for UIS to use in helping countries to report progress on SDG 4.1.1?	(5) In Mathematics?	13	8%	8%	0	69%	15%
	(6) In Reading?	13	8%	8%	0	69%	15%