United Nations
Educational, Scientific and
Cultural Organization

UNESCO INSTITUTE FOR STATISTICS

GLOBAL ALLIANCE TO MONITOR LEARNING

UNESCO Institute for Lifelong Learning

Monitoring adult numeracy for Indicator 4.6.1:

Numeracy Core Group Tentative Recommendations for GAML5

October 2018

UNESCO Institute for Lifelong Learning

Monitoring adult numeracy for Indicator 4.6.1:
Numeracy Core Group tentative recommendations for GAML5

Final draft: 29 September 2018

By the Numeracy Core Group:

Iddo Gal, University of Haifa, Israel, (Numeracy Expert Group Facilitator)
Diana Coben, King's College and the University of East Anglia, UK
Jean-Pierre Jeantheau, Agence Nationale de Lutte contre l'Illettrisme (ANLCI), France
Koji Miyamoto, World Bank, USA

Table of contents

Monitoring adult numeracy for Indicator 4.6.1:
Numeracy Core Group tentative recommendations for GAML5

Executive Summary

Overview

This paper presents recommendations and options for the assessment of the numeracy skills of adults, associated with SDG Indicator 4.6.1, which calls on countries to report the 'Percentage of population in a given age group achieving at least a fixed level of proficiency in functional (a) literacy and (b) numeracy skills, by sex'.

The paper is organized into six parts. Part 1 reviews some of the challenges associated with finding a fit between the general approach of the Programme for the International Assessment of Adult Competencies (PIACC) and the specific needs and constraints of assessing numeracy in all world countries Implications for the level of adoption of the Indicator 4.6.1 system are also discussed. Part 2 outlines the PIAAC definition and conceptual framework for numeracy and describes common domains and sub-domains. Part 3 discusses at what level (of numeracy) to 'fix' the proficiency level which will be reported for Indicator 4.6.1. Part 4 reflects on the extension of the PIAAC framework to include lower skills levels. Part 5 discusses a proposal for a UNESCO Institute for Statistics (UIS) reporting scale for numeracy. Finally, Part 6 broadly sketches two options for assessment methodologies and modalities.

The paper also includes four appendices with relevant details about the PIAAC frameworks of numerate behaviour and factors affecting task complexity, reporting levels in PIAAC and the Literacy Assessment and Monitoring Programme (LAMP), and examples regarding items and design issues in numeracy assessment.

The paper presents recommendations in four key areas:

Recommendation 1

Where to set the ('fixed') reporting level in numeracy. Based on statistical arguments and scholarly and policy considerations, it is recommended to set the 'fixed' reporting level in numeracy at the *minimal level* of cognitive skills which enable individuals to begin to successfully engage with and manage basic forms of written representations of mathematical information, in addition to having relevant 'mental maths' skills. This proposed minimal level is conceptually consistent with the lower end of the description of PIAAC Level 1, and in line with the proposed Global Alliance to Monitor Learning (GAML) 'Level C' which refers to Early Functional Literacy Level.

The proposal provides a tentative description of this minimal proficiency level:

> The respondent is able to carry out basic mathematical processes in common, concrete contexts where the mathematical content is explicit, with either little or no

text and minimal distractors. Tasks usually require simple one-step processes, and may involve understanding of representations of numerical entities (e.g., positions on a number line up to 100), performing basic arithmetic operations in reference to written or visual representations of quantities; understanding simple proportions (e.g., fractions or percentages such as 1/2 or 50%); locating, identifying and using elements of simple graphical or spatial representations; and understanding basic information about everyday measurement systems such as regarding time, length or weight.

Recommendation 2

Directions for extending the PIAAC framework. Based on a brief review relevant literature, it is recommended to extend the PIAAC framework downwards, if needed for reporting purposes, by *covering lower mathematical skills in the two numeracy content areas, 'Quantity and number' and 'Measures, dimension & space,'* which are most related to everyday life and can be implemented in test items that do not involve reading text. It is also possible to extend the PIAAC framework downward by *reducing literacy demands* of the assessment and of specific items, e.g., by using an oral administration mode and some text-free images or stimuli for selected items.

Recommendation 3

A tentative reporting scheme for numeracy

The paper sketches a possible scheme of 5 reporting levels (from A to E). Level C is the minimal ('fixed') level for reporting, i.e., the scheme assumes that the calculation of the proportion of persons who satisfy indicator 4.6.1 will be the percentage who achieved level C or above. The scheme is designed to allow for assessment and reporting of additional levels both above (Levels D and E) and below (Levels A and B) the 'fixed' minimal level, in line with the view of numeracy as a continuum, and in order to provide decision makers with more actionable information, and to satisfy local information needs.

(Note: the extensions to the PIAAC framework in Recommendation 2 may be needed to inform the creation of items focused on reporting level B, or on the lower end of level C.)

Recommendation 4

Two assessment options

To address the diverse needs of target countries, and to increase the chance for early adoption in light of the concerns noted in the paper, the proposed approach is to create a core item pool for numeracy, which can then be used for either of two implementation options: Option 1 follows the approach proposed by UIS in prior documents which is also the key modality used by PIAAC, i.e., computer-based, adaptive testing (which requires a core item pool with roughly 40 to 60 items).

Option 2 proposes a simpler design, employing a shorter 'base test' using 15–20 items selected from the core item pool. The test will be administered via short printed booklets

and some oral testing, using a simple adaptive mechanism controlled by the interviewer. The base assessment according to Option 2 would focus on assessing Levels B and C and hence require short testing time. Countries will be able to add an optional module to cover higher proficiency levels D and E.

Summary

Both assessment options can serve the reporting framework described in Recommendation 3 above, which revolve around the minimal ('fixed') proficiency level outlined in Recommendation 1. Both options can be launched in parallel in the field, after the core item pool is created and then the instruments piloted and calibrated in multiple countries and languages. Further planning work will also be required regarding the psychometric basis for proficiency estimates, and regarding the procedures for reporting indicator estimates that are produced via two different but related options.

Both options have advantages, but also costs and disadvantages which are discussed in the paper. Although Option 2 is inferior to Option 1 in terms of the precision of the proficiency estimates and the coverage of the numeracy construct, it requires fewer resources and lower technical capacity, and can enable faster implementation and reporting cycles. Thus, Option 2 may offer a sufficient base from which many countries that would not be able or willing to adopt Option 1 can still report for Indicator 4.6.1 in numeracy.

1. Introduction

1.1 **Background**

This paper focuses on issues and options for the assessment of the numeracy skills of adults, associated with Target 4.6[1] of the Sustainable Development Goals (SDGs). This target calls on countries to ensure that 'By 2030 ... all youth and a substantial proportion of adults, both men and women, achieve literacy and numeracy'. Indicator 4.6.1 to be reported regarding Target 4.6 is the 'Percentage of population in a given age group achieving at least a fixed level of proficiency in functional (a) literacy and (b) numeracy skills, by sex'. In 2017 UNESCO decided to measure both literacy and numeracy for indicator 4.6.1 via the approach used by PIAAC (i.e., OECD Survey of Adult Skills), with any needed extensions at the lower end. Further, UIS indicated that a revised or short version of LAMP could be developed for the actual assessment.

1.2 **The challenges in a nutshell**

The above decisions pose many challenges when thinking of a monitoring strategy for adult numeracy, for several reasons:

- PIAAC conceptualizes and measures numeracy (and other skills) on a continuum from low to high levels. This of course is essential, since numeracy should be viewed as lying on a continuum. However, Indicator 4.6.1 requires assessing a single 'fixed proficiency level' and this by itself has no operational meaning, since the Indicator could be 'fixed' at any level of the PIAAC (or other) proficiency continuum.
- The PIAAC approach is based on using computer-based adaptive testing for data collection, large national samples, and sophisticated Item Response Theory (IRT) psychometric models for national-level proficiency estimates. This approach has many advantages, and can provide policy-relevant, actionable information for participating countries. (LAMP has used a similar approach, on a smaller scale). However, of the 218 countries on the World Bank list,[2] fewer than 40 took PIAAC Cycle 1 (in 2013–2016), and so far fewer than 35 joined PIAAC Cycle 2 (to be reported in 2023). LAMP has been adopted by 4 countries.
- GAML expects the monitoring of Indicator 4.6.1 to be based on the PIAAC approach, but PIAAC itself is a 'moving target'. Much information has been published about the approach used in PIAAC Cycle 1 – but the full item pools are confidential. Only partial information has been released about planned changes in PIAAC Cycle 2, which is currently at development stage. Thus, it is difficult to fully fit the new planned monitoring system for 4.6.1 with the PIAAC approach.

1.3 The upshot of the above points is that further deliberations, value judgments and decisions are needed about what is the actual ('fixed') level of proficiency to be reported for 4.6.1. Also, additional options for assessment methodology have to be considered, to

[1] See UN website, Goal 4, target 4.6: https://sustainabledevelopment.un.org/sdg4

[2] According to the World Bank, at present (Sept 2018) there are 31 low-income countries, 53 lower-middle income countries, 56 upper-middle-income countries, and 78 high-income countries.

improve the chance for adoption of *any* monitoring process of 4.6.1 by the majority of the world countries, while retaining both validity (Messic, 1995) and policy relevance.

1.4 **Where we are**

The Numeracy Expert Group was formed in mid-2018 and has only just started its work, which is expected to take some time given the complexity of the many issues in its Terms of Reference. Hence, this paper presents preliminary ideas and points for discussion by GAML about a global framework for indicator 4.6.1 in numeracy. However it should be clear that the points presented here are tentative, incomplete, and subject to further modifications as the expert group continues its work and examines more issues.

1.5 **Organization of this paper**

The remainder of this paper is organized in five parts.

Part 2: outlines the PIAAC definition and conceptual framework for numeracy, describes common domains and sub-domains.
Part 3: discusses at what level (of numeracy) to 'fix' the proficiency level which will be reported for Indicator 4.6.1.
Part 4: reflects on the extension of the PIAAC framework to include lower skills levels.
Part 5: discusses a proposal for a UIS reporting scale for numeracy.
Part 6: discusses options for assessment methodologies and modalities.

2. The PIAAC numeracy framework: A brief review

2.1 **Definitions**

The conceptualization of numeracy used in PIAAC's first cycle (2013–2016; OECD, 2013*a*) involved a multi-faceted framework with three interlocking elements: a *definition* of the numeracy competency, a *model describing dimensions and sub-facets of 'numerate behaviour'*, and the *numeracy complexity scheme*. These are briefly discussed in this part. (Note: updates may be possible when more is known about PIAAC Cycle 2).

The PIAAC (Cycle 1) definitions of numeracy and numerate behaviour were as follows:

- *Numeracy* is the ability to access, use, interpret and communicate mathematical information and ideas, in order to engage in and manage the mathematical demands of a range of situations in adult life.
- *Numerate behaviour* involves managing a situation or solving a problem in a real context, by responding to mathematical content/information/ideas represented in multiple ways.

2.2 **Numerate behaviour**

Numerate behaviour was characterized by a model with four dimensions: *contexts, responses, content areas* and *representations*, and facets listed in *Annex A* of this paper. The facets specify the types of *contexts*, possible *responses* to numeracy tasks, four *content* areas (i.e., types of mathematical information and quantitative ideas) for which knowledge and skills are expected, and types of *representations* of quantitative or statistical information. *Annex A* also lists the proportion (percentage) of items used to cover each of the content areas, reflecting their relative importance in PIAAC Cycle 1.

2.3 **Complexity scheme**

This model was developed for Assessment of Literacy and Language (ALL) and adopted by PIAAC, for use as a tool for analysing cognitive task demands and estimating item difficulty even before a piloting phase. *Annex B* lists the five key factors that according to this model affect the complexity or relative difficulty of items, and describes in detail the *lowest* (easiest) level of complexity (see Gal et al., 2005, for full background). The simplest possible task in PIAAC would be given a score of 1 on all five factors in this model.

2.4 **PIAAC reporting levels**

In PIAAC, respondents were given estimates of their numeracy proficiency on a continuous scale ranging from 0 to 500. The scores were then clustered into five reporting levels, Levels 1 to Level 5 (highest). A sixth level, labelled 'Below Level 1', was added to capture individuals who were unable to cope consistently even with the simplest items included in the original assessment. *Annex C* lists the descriptions of the PIAAC Cycle 1 reporting levels. It also lists the three reporting levels (Levels 1 to 3) used in the LAMP assessment, which was fielded by UIS in some middle-income countries.

2.5 **The role of the PIAAC conceptual framework and reporting labels in assessment design, and what this means for GAML**

It is important to clarify that the two schemes described in *Annex A* and *Annex B*, i.e., the dimensions of numerate behaviour and the factors affecting item complexity, are not only conceptual tools but also frameworks that have guided scale development and enable us to evaluate the theoretical soundness of the item pool used for the actual assessment:

- In PIAAC, numeracy items intended to cover all combinations of the facets and all four content areas in the scheme of numerate behaviour in Annex A.
- In PIAAC, items were designed to span the full spectrum of numeracy proficiency needed in the PIAAC countries, i.e., items cover a range of cognitive skills, from simple (easy) to complex (difficult). However, given that PIAAC Cycle 1 was designed for high-income countries, most PIAAC numeracy items were designed to be at levels 2 to 4 (out of 5 main reporting levels); few items were designed to directly assess Level 1 skills. PIAAC Cycle 2 plans to cover lower levels of numeracy, but no specific information has been released or obtained in this regard up to this stage.

2.6 The description of the PIAAC reporting levels in Annex C, and in particular the description of Level 1, will become relevant in Part 3 below when we examine at what level to 'fix' the proficiency level to be reported for Indicator 4.6.1 in Numeracy. The model of complexity factors in Annex B (which informs item production) is where extensions that cover lower levels of proficiency can be implemented, i.e., simpler cognitive skills can be described (as discussed in Part 4 further below), and items can then be designed to capture those levels of complexity.

3. About a 'fixed' proficiency level for numeracy reporting

3.1 As noted earlier, Indicator 4.6.1 is defined as the 'Percentage of population in a given age group achieving at least a fixed level of proficiency in functional (a) literacy and (b) numeracy skills, by sex'. However, PIAAC and other skills surveys conceptualize and measure numeracy (and other skills) on a continuum from low to high levels. Hence, the Indicator could be 'fixed' at any level of the proficiency continuum. This means that the term 'fixed proficiency level' should be replaced with a more informative label.

Further, it should be clear that a decision on where (i.e., at what level) to actually set the 'fixed' proficiency level is a value judgment that carries many possible ramifications; hence it is not a simple decision that can be made based only on technical grounds. The set ('fixed') level for reporting Indicator 4.6.1 in numeracy will not only have policy-related impacts (e.g., on decisions about resources, adult education policies and programmes) but also political and economic implications. The way the Indicator is set will imply what proportion of the population in a country is described as 'lacking basic skills', and will affect how the country views itself, or is viewed by other stakeholders (or investors and donors) in terms of its human capital; it may cause stigmas to be attached to certain social groups, and so forth. It follows that a decision on the level that will be '... at least a fixed level of proficiency in functional numeracy skills' should consider multiple logics, and be informed by relevant evidence, scholarly arguments, needs of policy makers, etc.

3.2 **Statistical considerations**

We start by examining findings from PIAAC and LAMP, shown in Table 1, as these are the latest multinational surveys that provide comparative information about the distribution of numeracy proficiencies in at least some countries.

Table 1: Numeracy rates at different levels in PIAAC and LAMP

PIAAC results[1] (adults 16–65) *(Note: rounded percentages)*

	Below Level 1	Level 1	Level 2	Levels 3/4/5	Miss (rnd)	Total (rnd)
Chile	31	31	26	12	0.3	100%
Jakarta	26	34	29	11	0.0	100%
Turkey	20	30	33	15	2.0	100%
Spain	10	21	40	29	0.7	100%
Israel	11	20	30	36	2.4	100%
OECD avg	7	16	34	43	1.5	100%
Australia	6	14	32	46	1.9	100%
Germany	5	14	32	49	1.5	100%
Finland	3	10	29	58	0.0	100%

[1] PIAAC data taken from OECD (2016), *Skills Matter: Further results from the survey of adult skills*, Figure 2.12: Numeracy proficiency among adults

LAMP results[2] (adults 15+) *(Note: Rounded percentages)*

Country	% of all adults at Level 1 numeracy	% of all adults at Level 2 numeracy	% of all adults at Level 3 numeracy	% of *males* & *females* at Level 1 numeracy	% with *primary* education at Level 1 numeracy	% with *secondary* education at Level 1 numeracy
Jordan	25	45	30	M = 17 F = 34	55	23
Mongolia	17	45	38	M = 17 F = 17	65	17
Palestine	36	42	22	M = 26 F = 46	97	23
Paraguay	24	42	34	M = 21 F = 27	51	13

[2] LAMP data taken from *Table 6* of the relevant LAMP country reports (2009) at: http://www.uis.unesco.org/literacy/Pages/lamp-literacy-assessment.aspx

3.3 Table 1 presents the percentage of persons classified in the standard reporting levels of numeracy in eight selected PIAAC Cycle 1 countries, and in the four LAMP countries. Wording describing each of these reporting levels is shown in Annex C. The selected PIAAC countries in Table 1 include three countries above the OECD average, and five below it, of which three are middle-income (Chile, Jakarta [Indonesia], Turkey). All four LAMP countries were middle-income, but most PIAAC countries were high-income.

3.4 Table 1 shows that the four LAMP countries had, on average, between 17% and 36% of adults in Level 1. (Note that the breakdown by gender, as requested by Indicator 4.6.1, shows substantial differences between men and women, for some countries). The average LAMP percentages are quite similar to the combined percentages at 'Level 1' +

'Below Level 1' categories in the OECD average and the 5 high-income PIAAC countries (Australia, German, Finland, Israel, Spain) in Table 1. More important, the combined percentages of people in the PIAAC 'Level 1' + 'Below Level 1' categories rise sharply and reach 30% to 60% in Chile, Jakarta (Indonesia) and Turkey, all middle-income countries. We can assume that the percentages in PIAAC Level 1 + Below level 1 are likely to *rise* and be even higher in *low*-income countries with poorer educational systems and higher school drop-out rates.

3.5 What do these findings mean in the framework of the present paper? Table 1 may lead to multiple interpretations.[3] Our position is that if PIAAC is to serve as a basis for considering where to 'fix' a reporting level for 4.6.1 in Numeracy, based only on statistical considerations (others are listed below), the percentages in Table 1 imply that *the chosen 'fixed' proficiency level has to be no higher than the skills subsumed in PIAAC 'Level 1' and possibly on the lower end of Level 1.* Otherwise, the majority of entire countries will be classified as not having minimal numeracy skills, with all that this implies for policy makers and stakeholders.

3.6 **Scholarly and policy considerations**

In general, reporting frameworks of large-scale surveys of social or educational topics should be designed so as to respond to policy needs and provide useful information to stakeholders about the proportions and characteristics of key subgroups in the target populations. From this perspective, it can be argued that a decision on where (i.e., at what level) to set the 'fixed' proficiency level for reporting Indicator 4.6.1 in numeracy, should relate to the potential of persons to engage in or apply for at least entry-level jobs in a modern economy on the assumption that they can later benefit from training or adult education programmes related to numeracy. After all, modern workplaces (and social communities) expect employees (or citizens) to engage with at least some forms of text-based representations of mathematical or statistical information, as in, e.g., product and warning labels, task-related instructions, forms, manuals, e-mails, shipping costs on websites, etc. Likewise, to benefit from an adult education programme in numeracy, learners need to engage with study guides, assessment forms, and so forth.

3.7 ***Recommendation 1 - where to set the ('fixed') reporting level in numeracy***:

3.7.a Given the statistical arguments and scholarly and policy considerations listed above, it is recommended to set the 'fixed' reporting level in numeracy at the *minimal* level of cognitive skills which enable individuals to begin to successfully engage with and manage basic forms of written representations of mathematical information, in addition to having 'mental maths' skills. This minimal level is conceptually consistent with the lower end of the description of PIAAC Level 1 (which involved a score range of 175 to 225 on the PIAAC Cycle 1 scale). The rationale of choosing a minimal level of numeracy as described above is also consistent with the proposed GAML '*Level C*', which refers to *Early Functional Literacy Level,* and with the language of Indicator 4.6.1, which refers to '… achieving *at least* ...'

[3] Table 1 may suggest, for example, that the LAMP definition of "Level 1" is less demanding or differs from the PIAAC definition for that level, because PIAAC had 6 levels while LAMP had only 3 levels.

3.7.b The development of a detailed description of the proposed *minimal* proficiency level in numeracy requires further work by the numeracy group, including further study of level descriptors from different assessments, and examining the actual items that were on Level 1 in relevant studies such as PIAAC, LAMP, IVQ, LEO, etc. in order to analyse their cognitive demands. Such comparisons combining 'top-down' and 'bottom-up' analyses are essential, since, as Table 1 demonstrates, there are differences in what 'Level 1' covers in different studies.

For the time being, a tentative description of the proposed *minimal* proficiency level in numeracy is offered below, based on the original description of PIAAC Level 1 (see *Annex C*), with some additions and modest changes based on LAMP Level 1 and other sources:

Description (tentative) of the proposed *minimal* proficiency level in numeracy:

The respondent is able to carry out basic mathematical processes in common, concrete contexts where the mathematical content is explicit, with either little or no text and minimal distractors. Tasks usually require simple one-step processes, and may involve understanding of representations of numerical entities (e.g., positions on a number line up to 100), performing basic arithmetic operations in reference to written or visual representations of quantities; understanding simple proportions (e.g., fractions or percentages such as 1/2 or 50%); locating, identifying, and using elements of simple graphical or spatial representations; and understanding basic information about everyday measurement systems such as regarding time, length or weight.

4. Extending the PIAAC framework downwards

4.1 The rationale for examining ways to extend the PIAAC conceptual framework so as to cover lower-level skills is based on the need to provide more information about persons with low numeracy skills, who according to the information in Part 3 earlier, may constitute a sizeable portion of the population in some countries; such persons may perform below the minimal numeracy level proposed earlier as the reporting level for Indicator 4.6.1, but education planners may still desire to know what cognitive skills they do possess or what tasks they can cope with, hence the need to extend the PIAAC framework downwards.

The PIAAC complexity scheme already covers certain low-level skills in the lowest complexity category ('score 1') depicted in *Annex B*. It is possible to extend this framework downward further, however, in different ways, using two different logics discussed below, related to *covering lower mathematical skills* and *reducing literacy demands*.

4.2 **Extension in terms of covering lower mathematical skills**

The PIAAC complexity scheme relates to skills associated with four content areas:

- Quantity and number
- Pattern and relationship
- Measures, dimension & space
- Data & chance.

Examples for skills at the current *lowest* level (Score 1 in the complexity scheme) are the ability to recognize and handle one-step operations with numbers up to 1000, to recognize and work with benchmark fractions ($^1/_2$, $^1/_4$, $^3/_4$) or percentages (e.g., 50%) or key decimal fractions (0.5), and so forth – all in familiar contexts, with minimal text reading demands and virtually no distracting information. Many of these skills are incorporated into the proposed minimum proficiency level described above (see 3.7.b).

It is possible, however, to describe skills below those included in the current lowest level in the PIAAC complexity scheme, on the basis of resources regarding very low-level numeracy skills of adults, such as in (a) the Learning Progressions for Adult Numeracy in New Zealand[4], (b) the French certificate system CléA[5] which defines basic skills expected of adults in all vocational tracks in France and its territories, or (c) frameworks of adult numeracy in national surveys of adult skills such as in Kenya (KNALS, 2006) and Bangladesh (BBS, 2013).

Further, skills below those included in the current lowest level in the PIAAC complexity scheme are also covered in resources pertaining to elementary mathematical operations or rudimentary mathematical processes expected by most world countries to be mastered by school-age *children* in the first few grades of primary school. These are discussed in scholarly literatures on 'emergent numeracy' or 'elementary mathematics,' listed in many national curriculum statements, are often covered in 'early grade' mathematics tests for young children (e.g., Reubens, 2009), and are likely to be discussed by the GAML group for Indicator 4.1.1.

Based on such resources, we propose in Section 4.4 below some extensions to the PIAAC framework. However, first it is necessary to describe an additional rationale or approach, which pertains to literacy demands of test items with mathematical content.

4.3 **Reducing literacy demands**

The PIAAC approach demands that respondents read by themselves all questions (from a computer screen or booklets), and respond via a keyboard or in writing. This creates literacy-numeracy dependency, i.e., the numeracy score reflects to some extent literacy skills, not numeracy proficiency per se. As a result, some individuals may be misclassified and considered to have *no* numeracy skills, when in fact they may have *low* numeracy skills. They may possess mental strategies for handling some mathematical computations or other mathematical tasks 'in their head', i.e., have 'mental maths' skills that would be useful in functional contexts and may help them to learn mathematics more formally if they joined an educational programme. However these skills may be masked or remain hidden[6] when a test demands that test-takers read all questions and provide responses in writing. It follows that it is possible to improve the assessment of low-level mathematical skills by reducing the literacy (text-processing) demands of the test.

[4] https://ako.ac.nz/knowledge-centre/learning-progressions-for-adult-numeracy/lp-guide/
[5] https://www.certificat-clea.fr/le-dispositif-clea/le-referentiel
[6] See Gal (2016) for a discussion of how literacy-numeracy dependencies may have affected statistics about numeracy in the Kenya and Bangladesh adult surveys.

Although it may seem a technical issue related to assessment methodology, the issue of literacy demands is mentioned here because in fact it is about the conceptualization of adult numeracy in modern societies, and about the extent to which the assessment of numeracy of low-ability individuals should depend on literacy skills. If the text-processing demands are lowered, it may be possible to reduce literacy-numeracy dependency and improve the information provided by Indicator 4.6.1 compared to the PIAAC approach.

A potent example is the implementation of an *oral* testing approach in the Information and Daily Life (*Information et Vie Quotidienne;* IVQ) survey, conducted in France in 2002–2004 and again recently, to test the numeracy and literacy of 'low-level' adults. A unique feature of the IVQ was that questions were read aloud and responses recorded by the interviewer, with content for some items shown on separate cards (as drawings or photos), to minimize the impact of literacy skills on performance in numeracy tasks. A summary can be found in Gal (2016) and full details in Jeantheau (2005) and Murat (2008).

4.4 ***Recommendation 2 – Directions for extending the PIAAC framework***

4.4.a **Content areas.** Based on the considerations mentioned in 4.2 above, it is recommended to extend the PIAAC framework downwards in the two numeracy content areas, 'Quantity and number' and 'Measures, dimension & space', which are most related to everyday life and can be implemented in test items that do not involve text-reading. (Note: at this stage we are not proposing to extend downwards in the content area of Data and chance, since it overlaps with Document literacy, implying that text-reading demands can become a confounder when measuring understanding of some aspects of Data and chance).

Based on sources noted earlier, below are examples for relevant rudimentary numeracy skills which extend *below* the current 'Score 1' lists in the PIAAC Cycle 1 framework, and which could be adopted for extending the PIAAC framework downwards for Indicator 4.6.1:

In the area 'Quantity and Number':

- Say, hear and read numbers up to 100
- Compare numbers using idioms such as 'equal to', 'more than', 'less than'
- Show understanding of the number line by placing numbers on the number line or find the predecessor (number before) and successor (number after) of a number
- Subtract or add mentally with numbers up to 50 or 100

In the content area 'Measures, Dimension & Space':

- Recognize simple written landmarks related to time
- Read, compare and estimate lengths or weights

4.4.b **Literacy demands and oral administration**

Based on the considerations mentioned in 4.3, above, it is recommended that for the purposes of Indicator 4.6.1, some of the numeracy items will be designed and administered in a way that reduces literacy-numeracy dependencies, both by using stimuli without text (but with images or photos), by posing questions orally, and by capturing oral responses (e.g., to test the ability to do certain computations 'in the head' and other mathematical skills).

5. A proposed reporting framework for numeracy

5.1 **Background**

In general, reporting frameworks of large-scale surveys of social or educational topics should be designed so as to respond to policy needs and provide useful information to decision-makers and stakeholders about the proportions and characteristics of key subgroups in the target populations. The reporting levels should identify groups or social categories that can benefit from interventions or policy attention – and if the data collection is repeated, enable detection of trends over time. We should pay special attention to the lowest reporting level(s), as this is where key social or educational interventions could focus and where much policy-setting may occur. The practice in LAMP, PIAAC and most other adult surveys has been to lump together three different subgroups in the lowest reporting level in numeracy (i.e. in Level 1, and Below Level 1):

a. Adults who have *low literacy skills* and thus have difficulty with the numeracy test which uses written questions, ending up with a low numeracy score,
b. Adults who have *low numeracy skills* (they may also have low literacy skills, or higher literacy skills, a phenomenon highlighted by the IVQ survey in France)
c. Adults with *no numeracy skills*, i.e. truly 'innumerate' adults. With the exception of persons with severe intellectual or learning deficits, true innumeracy should be rare, based on results from studies of mathematical skills and practices of 'indigenous' or special populations who lack formal schooling and/or written scripts.

5.2 **How low is 'low'?**

There is a fundamental difference between the nature of literacy and numeracy at the lower end of the skills distribution. While people who lack reading skills can be considered 'illiterate', the literature suggests that people with little or no formal reading skills can still cope with selected mathematical tasks in everyday and work life, e.g. manage herds, conduct commercial transactions, plan construction, and perform selected tasks that may require counting, mental computing, estimation of time and distance, recognition of shapes, use of mathematical tools, or explaining with mathematical 'objects' using oral terms (Bishop, 1988; Lave, Murtagh and de la Rocha, 1984; Greeno, 2003; Straesser, 2015).

5.3 Functional mathematical skills of people with little or no reading skills may be masked or remain undocumented either because of the *conceptualization* of numeracy (i.e., if it is

defined as a skill that involves the ability to deal with *written* representations of numbers and quantities), due to *reporting* practices (i.e. combining literacy and numeracy in a single 'non-literate' category, as some countries do; see BBS, 2013), or because of *assessment* practices (i.e. using a test that requires respondents to *read* written numeracy questions or *answer in writing*, or via a laptop). Whatever the explanation, these undocumented numeracy skills are part of the construct of numeracy, and they do have both social and economic values, hence should be measured and reported.

Table 2: A tentative scheme of reporting levels for indicator 4.6.1 in numeracy

Note: Proposed as a basis for discussion. Needs further work by the numeracy core group

Level	Description/difficulty level	Comments
E	Skills related to PIAAC *lower* 'Level 3	'Adequate' level
D	Skills related to PIAAC Level 2 / LAMP Level 2	
C	**Skills related to PIAAC Level 1** ⇦ (discussed in Part 3)	**Minimal ('Fixed') level for reporting**
B	Knows few print-based or formal numeracy symbols and systems, though may be able to do very simply written maths problems. Can engage in some (possibly even advanced) mental calculations using indigenous number systems or measurement devices/techniques only.	Based on use of minimal or no text, i.e. in part on an *oral* assessment, in part on items with text-free stimuli which are read aloud
A	Cannot recognize the meaning of written digits or positions on a number line. Has no or few mental calculation skills beyond counting or adding of simple quantities.	

5.4 ***Recommendation 3*** **– A tentative reporting scheme for numeracy**

Based on the analysis above, for purposes of indicator 4.6.1 in numeracy, a possible scheme of 5 reporting levels is sketched in Table 2. The scheme is inspired in part by ideas first raised in Wagner, Sabatini and Gal (1999) who proposed a four-level system, in part by the conceptual frameworks for numeracy from PIAAC and LAMP, and is consistent with the scheme currently being considered by GAML which also involves five levels.

The scheme in Table 2 assumes that the calculation of the proportion of persons who satisfy indicator 4.6.1 will be *the percentage who achieved level C or above*. The scheme allows for reporting of additional levels both above and below the 'fixed' minimal level, in order to provide decision makers with more actionable information, and to satisfy local information needs.

The proposed scheme enables a separation between *minimal* formal skills in numeracy (Level C) and *very low or no* formal written numeracy (Level B and below). *Note: The extensions to the PIAAC framework discussed in Part 4 may be needed to inform the*

creation of items focused on reporting level B, or lower end of level C. See Gal (2018) for a discussion of how mental computation skills fit into reporting levels B and C.

Overall, the separation between levels A and B is based on (a) a theoretical necessity regarding the nature of numeracy and the ability of adults to activate mathematical reasoning – a key 'enabling process' for numeracy performance, even without any literacy skills, and (b) on the realization that persons to be designated in Level B cannot really engage in a numeracy programme without first acquiring working literacy.

6. Options for assessment methodologies and modalities

6.1 **Rate of adoption issues**

The proposal of more than one option for an assessment methodology for measuring numeracy for Indicator 4.6.1 is based on the realities in the field of national and international assessments regarding adoption pace and bottlenecks:

- *Pace*: We must assume that not all countries will want or be able to implement Option 1, either at all, or *in a timely fashion.* The data for studies of adult skills that use a household methodology support this assertion and show that adoption is not only a challenge, but a serious threat. Out of 215 world countries on the World Bank list, PIAAC was adopted by fewer than 40 countries in each cycle, LAMP by 4 countries. The STEP (World Bank) and MICS (UNICEF) assessment programmes run in waves, a few countries at a time, both because the coordinating agencies face slow adoption rates by countries, and partly due to bottlenecks.
- *Bottlenecks*: The UIS proposal implies that a single international agency (with the help of a technical consortium) has to design a multilingual computer-based assessment platform *and* a paper-based path for a huge number of scripts and cultural contexts, train personnel and conduct quality assurance regarding complex data collection operations, conduct centralized data reporting and analysis processes, and manage reporting in a timely manner. If most world countries join, this has to be done for a couple of dozen countries *each year*, a feat not tried before, which is likely to create various bottlenecks and capacity issues, since the target countries for Indicator 4.6.1 face huge linguistic and operational diversity

6.2 ***Recommendation 4 – two assessment options***

To address the diverse needs of target countries, and to increase the chance for early adoption in light of the concerns noted above, the general approach proposed is to create a core item pool for numeracy (by an international expert group), which can then be used for either of the two options: Option 1 follows the approach proposed by UIS in prior documents (computer-based, adaptive testing). Option 2 proposes a simpler design (short print booklets, a simple adaptive mechanism). These options are outlined below - see more details and technical explanations in Gal (2018).

6.3 **Core item pool for numeracy**

Both options described below require the creation of a *core item pool* with about 50–60 items that span the cognitive demands implied by the key reporting levels, based on the conceptual framework for numerate behaviour. The items could draw on existing LAMP items, PIAAC items (if shared by OECD), and items developed by the numeracy expert team or adapted from contributions by participating countries. Translation and adaptation guidelines should be developed, so national teams are clear on permissible flexibilities for adjusting content or stimuli of items so as to fit local cultures and needs, while retaining the underlying cognitive skills being measured. To ensure comparability, *the core item pool must be piloted* in multiple countries and languages, and (psychometric) comparability of items across countries has to be examined and validated, before the core item pool is released for use by all countries.

6.4 **Option 1 – the UIS approach** (computer-based, adaptive testing)**.**

Option 1 follows the key modality used by PIAAC, i.e., computer-based adaptive testing. This is also the approach proposed by UIS in prior documents. For example, UIS (2018) recommended to base the overall assessment for Indicator 4.6.1 on the LAMP methodology which already has an operational basis, using a computer-based (tablet or laptop) adaptive testing approach that employs a shortened or 'light' version of LAMP tests for literacy and numeracy, with needed extensions based on a revised conceptual framework. Data collection could be done via a dedicated survey, or by attaching the collection process as a module to an existing national survey. In any case, the assessment for Option 1 will have to rely on the development of the core item pool described above.

6.5 **Option 2 – paper-based test (with a simple adaptive process)**

This option is a variant on Option 1, and aims for a simpler assessment modality that takes advantage of the same core item pool, but not via a computer-based test. This option is proposed because it can be implemented by countries without the ongoing involvement of an international agency (beyond the creation of the core item pool, and provision of general guidelines) and has lower resource demands and other advantages discussed below.

As a background, the UIS options paper (2018), suggests that low- and middle-income countries can field on their own (i.e., without the assistance of an international agency) a large-scale household-based survey that affords 60–80 minutes of testing time per respondent. Within this timeframe and situation, Gal's (2016, 2018) reviews of operational aspects of national and international assessments argue that *it is possible to employ 15 to 20 relatively short numeracy items*, as Kenya and Bangladesh have demonstrated.

With Option 2, it is proposed that an expert team will construct a short ‘base test’ with 15-20 items selected from the core item pool, to offer some coverage of key dimensions specified in the conceptual framework of numerate behaviour. The base test would be short and focus on items suitable in terms of difficulty level to Reporting Levels C and B. Thus, a country which does not wish to use Option 1 can use this base test and then

report the results for indicator 4.6.1. However, Option 2 also calls for the development of an additional 'higher' module which can cover reporting Levels D and E, in order to offer countries a way to collect data about a broader range of proficiencies than the base test can provide.

Option 2 can use a simple adaptive mechanism. First, a short screener test will be used to detect how well the respondent can deal with very simple print-based numeracy tasks. The interviewer will then decide, based on pre-established rules, which of two numeracy testlets will be given: *Testlet 1* will include 15–18 items covering levels C, D and E (each with 5–6 items); it will be mostly a print-based test (in a booklet) but with a few oral items intended to cover selected mental maths skills. Those who receive a low score on the screener will be given *testlet 2* with lower-level items covering Level B-C, with some more oral items.

Additional development work is needed to establish the mode of score computation for Option 2, i.e., either via an IRT estimation, or a classical test approach (using a total score), Further empirical work (a field study) is needed to examine the correlation between scores computed via Option 1 and Option 2, to enable *calibration* of scores from both options.

6.6 **Examples for assessment design issues – see Annex D**

We have included in Annex D a few selected items taken from several national assessments and from ALL/PIAAC, and selected references to further literature. The items aim to illustrate some of the many issues related to assessment of low-end numeracy skills, e.g., what are rudimentary skills (at Level 1 in PIAAC or below), options for oral vs. written administration, how to reduce literacy-numeracy dependencies, and more.

The literature and item examples in Annex D also aim to alert readers that many *dilemmas*, *design conflicts,* and hence *value judgments* are involved in assessing numeracy of adults. The examples suggest that the construction of an effective item pool and a solid test requires work by experts that will take into account multiple factors which may affect the validity, reliability, cultural suitability and policy value of the assessment system developed for Indicator 4.6.1.

6.7 **Summary and comparison of the options**

Option 2 is inferior to Option 1 in terms of the precision of the proficiency estimates offered by a computer-based adaptive design and psychometrics that can be used. Yet, Option 2 also has advantages: it requires fewer resources and lower technical capacity, while basing the computation of a total score on performance on items drawn from the same core item pool, covering several content areas and levels of difficulty, albeit briefly. Thus, Option 2 may offer a sufficient base from which to report for Indicator 4.6.1 in numeracy, for the many countries which would not be willing to adopt Option 1, which is more costly and demands higher capacity.

We emphasize that both options can be launched in parallel. Hence, UIS may need to consider whether the statistics from the two options should be reported separately, in order to prevent misinterpretation of the results from each Option, which are based on items from the same core item pool, but on a different number of items and on different psychometric models.

Finally, we argue that offering to countries a choice between two options may be useful in terms of marketing the overall monitoring system for Indicator 4.6.1 and increasing long-term improvement in adoption and quality of the statistics collected, since countries will be aware from the very start of a possible or desirable 'upgrade path' and are not locked into a single option. This way, countries that cannot or will not adopt Option 1 may still be willing to adopt the simpler Option 2, but this can encourage later adoption of Option 1, once they get used to a new approach for measuring and reporting numeracy levels, and after their technical capacity is further developed.

Annex A

Numerate behaviour: Dimensions and facets in PIAAC Cycle 1

Numerate behaviour involves managing a situation or solving a problem …

1. in a real context:

- everyday life
- work
- societal
- further learning

2. by responding:

- identify, locate or access
- act upon, use: order, count, estimate, compute, measure, model
- interpret/evaluate/analyse
- communicate

3. to mathematical content/information/ideas:[7]

- quantity and number (30%)
- dimension and shape (25%)
- pattern, relationships and change (20%)
- data and chance (25%)

4. represented in multiple ways:

- objects and pictures
- numbers and mathematical symbols
- formulae
- diagrams and maps, graphs, tables
- texts
- technology-based displays

Numerate behaviour is founded on the activation of several enabling factors and processes:

- mathematical knowledge and conceptual understanding
- adaptive reasoning and mathematical problem-solving skills
- literacy skills
- beliefs and attitudes
- numeracy-related practices and experience
- context/world knowledge

[7] The percentages next to the four content areas reflect the number of *items* that should examine each area. The percentages imply that PIAAC deemed all content areas as important, with a slight preference for quantity and number, given the ubiquity and centrality of this area in adults' lives.

Annex B

PIAAC/ALL complexity scheme, with details of the lowest level

(Note: see Gal (2018) and Gal, van Groenestijn, Manly, Schmitt & Tout (2005) for the full scheme of complexity factors and all their levels)

Complexity factor	**What is the *lowest* score on this factor? (complexity score = 1)** Note: Based on *PIAAC Cycle* 1 complexity scheme (and on ALL)
1. Type of match/ Problem transparency *(Range: 1 to 3)*	In the question/stimulus, the information, activity or operation required: - is clearly apparent and explicit – and all required information is provided - is specified in little or no text, using familiar objects and/or photographs or other clear, simple visualizations - is about locating obvious information or relationships only - closed question – not open-ended
2. Plausibility of distracters *(Range: 1 to 3)*	No other mathematical information is present, apart from that requested – no distracters
3. Complexity of mathematical information/answer required *(Range: 1 to 5)*	**Context:** Very concrete, real-life activities, familiar to most in daily life. **Quantity:** Whole numbers: to 1,000 Fractions, decimals, percentages: Benchmark fractions ($^1/_2$, $^1/_4$, $^3/_4$); decimal fractions for a half only (0.5) and equivalent as a percentage (50%). **Pattern and relationship:** Very simple whole number relations and patterns. **Measures/dimension/space**: Standard monetary values; common everyday measures for length (whole units)**;** time (dates, hours, minutes)**;** Simple, common 2D shapes; simple localized maps or plans (no scales). **Data/chance:** Simple graphs, tables, charts with few parameters and whole number values**;** simple whole number data or statistical information in text.
4. Complexity of type of operation/skill *(Range: 1 to 5)*	**Communicate:** No explanation is needed for a response, or a single simple response required (orally, or in writing) **Compute:** A simple arithmetical operation (+, -, x, ÷) with whole numbers or money **Estimate:** *None at Level 1* **Use formula/model:** *None at Level 1* **Measure**: Knowing common, straight-forward measures**:** naming, counting, comparing, or sorting values or shapes

	Interpret: Locating/identifying data in texts, graphs and tables**:** orientating oneself to maps and directions such as right, left, etc.
5. Expected number of operations *(Range: 1 to 3)*	Only one operation, action or process.

Annex C

Reporting levels descriptions for PIAAC Cycle 1 and LAMP

Note: This table presents the definitions of 5 the 6 reporting levels in PIAAC. It does not show the definitions of the highest Level 5, which in the actual reporting was merged with level 4 since very few adults in any country achieved Level 5.

PIAAC (Cycle 1 definitions)

Below Level 1	Level 1	Level 2
Respondents classified at this level may have difficulty with many tasks at Level 1. They may be able to cope with very simple tasks set in concrete, familiar contexts where the mathematical content is explicit with little or no text or distractors, and that require only simple processes such as counting; sorting; performing basic arithmetic operations with whole numbers or money, or recognizing common spatial representations.	Tasks require the respondent to carry out basic mathematical processes in common, concrete contexts where the mathematical content is explicit with little text and minimal distractors. Tasks usually require simple one-step or two-step processes involving performing basic arithmetic operations; understanding simple percentages such as 50 per cent; or locating, identifying, and using elements of simple or common graphical or spatial representations.	Tasks require the respondent to identify and act upon mathematical information and ideas embedded in a range of common contexts where the mathematical content is fairly explicit or visual with relatively few distractors. Tasks tend to require the application of two or more steps or processes involving calculation with whole numbers and common decimals, percentages and fractions; simple measurement and spatial representation; estimation; interpretation of relatively simple data and statistics in texts, tables, and graphs.

Level 3
Tasks require the respondent to understand mathematical information which may be less explicit, embedded in contexts that are not always familiar and represented in more complex ways. Tasks require several steps and may involve the choice of problem-solving strategies and relevant processes. Tasks tend to require the application of number sense and spatial sense; recognizing and working with mathematical relationships, patterns, and proportions expressed in verbal or numerical form; interpretation and basic analysis of data and statistics in texts, tables and graphs.

Level 4
Tasks require the respondent to understand a broad range of mathematical information that may be complex, abstract or embedded in unfamiliar contexts. Tasks involve undertaking multiple steps and choosing relevant problem-solving strategies and processes. Tasks tend to require analysis and more complex reasoning about quantities and data; statistics and chance; spatial relationships; change, proportions, and formulas. Tasks in this level may also require comprehending arguments or communicating well-reasoned explanations for answers or choices.

LAMP definitions - see next page

LAMP (3 levels of proficiency)

Level 1	**Level 2**	**Level 3**
Can answer explicit questions requiring a one-step, simple operation; add 3 whole numbers with 2-3 digits or with decimals in a 'money' context; and subtract 2 whole or decimal numbers in a 'money' context when they are presented with material communicating information in a familiar context with easily accessible quantitative information due to its visual representations and minimal text. Questions contain no choices or distractors.	Can do everything in Level 1, and in addition can complete tasks involving some fractions and decimals; understand and use some simple fractions such as one-half (½) written with numbers or words; can demonstrate some understanding of the meaning of decimal numbers; and multiply a decimal number and a whole number when they are presented with material communicating information in a familiar context.	Can do everything in both Level 1 and Level 2, and in addition can perform multiple-step operations that require multiplication (maybe by repeated addition) and then division (maybe by repeated subtraction); subtract a per cent from an initial value; find a proportion by combining operations in a money context (sometimes with decimals); add 3 numbers (sometimes with decimals) after computing 2 of them through multiplying by 10 or 2; read time using clocks or in numeric form; interpret qualitative or quantitative data from tables or price tags with per cents, decimals and whole numbers; and can represent money and weight using appropriate measurement units when they are presented with complex tasks with several visual representations and asked explicit questions that may or may not have choices or distractors.

Annex D

Examples – items and design issues in numeracy assessment

This Annex includes three Examples with selected items taken from several national assessments and from PIAAC/ALL.

The examples aim to illustrate some of the many issues related to assessment of low-end numeracy skills, e.g., what are rudimentary mathematical skills (at Level 1 or Below Level 1 in PIAAC, such as regarding understanding of a number line and place value, single step arithmetic processes, mental strategies, and others), options for oral vs. written administration, literacy-numeracy dependencies in item construction, the use of simple images that can be locally adapted to increase cultural relevance, and more.

The examples overall also aim to alert readers to some of the *dilemmas and design conflicts* involved in assessing numeracy, which may ultimately affect the difficulty level of items, and impact the validity, reliability, cultural suitability, and policy value of the assessment system developed for Indicator 4.6.1.

Literature: There is of course extensive literature about conceptualization and assessment of adult numeracy, item design, and various factors that affect the suitability of items for comparative assessment purposes in the area of adult numeracy. For some of the many issues involved, see Brooks, Heath & Pollard (2005); Coben & Alkema (2017); Condelli, Safford-Ramus, Sherman, Coben, Gal, & Hector-Mason (2006); Evans, Yasukawa, Mallows & Creese (2017); Gal, van Groenestijn, Manly, Schmitt & Tout (2005); Gal & Tout (2014); Jeantheau (2005); Kirsch, Jungeblut, Jenkins & Kolstad (1993); Maddox (2015); Maddox, Zumbo, Tay-Lim, & Qu (2015); Tout & Gal (2015).

Example 1: 'Shirts' items from the IVQ (Oral administration)

This example is based on the IVQ survey, conducted in France in 2002–2004 and again recently, to test low-level numeracy and literacy of adults. The example shows three items that were implemented in an *oral* testing approach based on a single stimulus: The image was shown on a card, questions were read aloud, and responses recorded by the interviewer.

Q1. "Which shirt is the cheapest?"
Q2. "How much is it cheaper?"
Q3. "If you buy three blue shirts, what will the price be?"

Example 2: Bangladesh national adult literacy survey

Note: This example uses original texts taken from the national adult literacy survey which took place in 2011 in Bangladesh (BBS. 2013). This survey used 12 items to assess numeracy skills, all administered in a printed booklet which respondents had to read on their own.

Table 1 describes the content of the 12 items used in the numeracy module and shows their relative weights. Table 2 shows six sample items, as they were given to respondents. See Gal (2016) for an extended discussion of these items.

Table 1

Item description	Score derivation	Total Score
Reading out four numbers	$\frac{1}{2} \times 4$	2
One addition, one subtraction, one multiplication and one division	2×4	8
Counting of money	2	2
Arranging of five numbers from greater to smaller	2	2
Accounts after visit a market for buying some things	6	6
Knowledge about distribution of some fruits equally to some persons	5	5
		25

Table 2

[1] Read the following figures : 54, 99, 208, 368

[2] Add: 85
17
Answer :.....,

[3] Divide :
5) 25 (
Answer :.....,

[4] Arrange the following figures in descending order: 125, 280, 70, 300, 50
Answer : ....., ..., ..., ..., ...

[5] You went to the market with Tk.500 in your pocket. You bought rice for Tk. 200, meat for Tk.180, Potato for Tk. 20, and spices for Tk. 40. How much money will remain with you ?
Answer : Tk. .....

[6] How many lichis will be received by every persons, if 117 lichis are divided among 3 persons?
Answer : .....

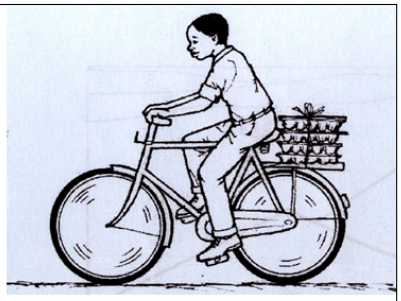
Example 3: Sample items from ALL/PIAAC and from the Kenya Adult Literacy Survey

Note: These items pose the same type of question to test-takers, regarding "A total number" of elements (eggs or bottles) in a three-dimensional everyday object (crates or packages stacked on top of each other). However, they differ in many subtle ways, and are shown in order to illustrate approaches for item design which can affect the extent to which items invite multiple response strategies, appear culturally relevant or "school-like", affect literacy-numeracy dependencies, and other factors that create design conflicts when developing a comparative assessment.

Note: Item 2 (Bottles) was the *simplest* numeracy item in ALL and PIAAC (i.e., a "Level 1" item) – it was administered as part of the screener test (before the adaptive test began). "Bottles" was originally used in the Adult Literacy and Lifeskills Survey (ALL), and later adapted for use in PIAAC by replacing the commercial product in the original stimulus below with a photo without a manufacturer name.

Item 1

(Source: Kenya's 2006 National Adult Literacy Survey)



Question: Mr. Wafula cycles every day to the market with 4 trays of eggs. Each tray contains 30 eggs. How many eggs does he carry to the market every day?

Answer: _____

Item 2

(Source: Adult Literacy and Lifeskills survey (ALL) released item, OECD and Statistics Canada, 2005)

Question: In total, how many bottles are there in the two full cases?

Answer: _____

References

BBS - Bangladesh Bureau of Statistics (2013). *Literacy Assessment Survey (LAS) 2011*. http://203.112.218.66/WebTestApplication/userfiles/Image/Latest%20Statistics%20Release/LAS_2011.pdf

Coben, D., & Alkema, A. (2017). The case for measuring adults' numeracy practices. *Journal of Research and Practice for Adult Literacy, Secondary, & Basic Education*, 6(1), 20–32.

Bishop, A. J. (1988). The interactions of mathematics education with culture. *Cultural Dynamics*, 1(2), 145-157.

Brooks, G., Heath, K., & Pollard, A. (2005). *Assessing adult literacy and numeracy: a review of assessment instruments.* London, UK: National Research and Development Centre for Adult Literacy and Numeracy. [online: www.nrdc.org.uk]

CléA (2003). Certificate framework. Online: https://www.certificat-clea.fr/le-dispositif-clea/le-referentiel/

Condelli, L., Safford-Ramus, K., Sherman, R., Coben, D., Gal, I., & Hector-Mason, A. (2006). *A review of the literature in adult numeracy: research and conceptual issues.* (Prepared by American Institutes for Research, for the Adult Numeracy Initiative of the US Department of Education's Office of Vocational and Adult Education). Online: http://eric.ed.gov/?id=ED495456

Evans, J., Yasukawa, K., Mallows, D., & Creese, B. (2017). Numeracy skills and the numerate environment: Affordances and demands. *Adults Learning Mathematics: An International Journal*, *12*(1), 17-26.

Gal, I. (2002). Adults' Statistical literacy: Meanings, components, responsibilities. *International Statistical Review*, 70(1), 1-25.

Gal, I. (2005). Towards 'probability literacy' for all citizens. In G. Jones (ed.), *Exploring probability in school: Challenges for teaching and learning* (pp. 43-71). London: Kluwer Academic Publishers.

Gal, I. & Tout, D. (2014). *Comparison of PIAAC and PISA Frameworks for Numeracy and Mathematical Literacy.* OECD Education Working Papers, No. 102, OECD Publishing. Online: http://dx.doi.org/10.1787/5jz3wl63cs6f-en

Gal, I. (2016). *Assessment of adult numeracy skills.* Paper commissioned for the *UNESCO Global Education Monitoring Report 2016: Education for people and planet*. Paris, UNESCO. Online: http://unesdoc.unesco.org/images/0024/002455/245573E.pdf

Gal, I. (2018). *Developing a monitoring scheme for adult numeracy as part of SDG Indicator 4.6.1: Issues and options.* A discussion paper prepared for the UNESCO Expert Meeting on Adult Literacy and Numeracy Assessment Frameworks, May 17-18, 2018, Paris. (Available via UIL).

Gal, I., van Groenestijn, M., Manly, M., Schmitt, M. J., & Tout, D. (2005). Adult numeracy and its assessment in the ALL survey: A conceptual framework and pilot results. In Murray, S. T., Clermont, Y., & Binkley, M. (Eds), *Measuring adult literacy and life skills: New frameworks for assessment* (pp. 137-191). Ottawa, Canada: Statistics Canada.

Greeno, J. G. (2003). Situative research relevant to standards for school mathematics. In J. Kilpatrick, W. G. Martin, & D. Schifter (Eds.). *A research companion to Principles and Standards for school mathematics* (pp. 304-332). Reston, VA: National Council of Teachers of Mathematics.

Jeantheau, J. P. (2005). Assessing low levels of literacy: IVQ Survey 2004-2005-Focus on the ANLCI module. *Literacy & Numeracy Studies*, *14*(2), 75.

KNALS - Kenya Bureau of Statistics (2007a). *National Adult Literacy Survey: Literacy assessment.* Retrieved Jan 15, 2016: http://statistics.knbs.or.ke/nada/index.php/catalog/58.

Kirsch, I. S., Jungeblut. A., Jenkins, L., & Kolstad, A. (1993). *Adult literacy in America*. Princeton, NJ: Educational Testing Service, National Center for Education Statistics.

Lave, J., Murtagh, M., & de la Rocha, O. (1984). In B. Rogoff & J. Lave (Eds.), *Everyday cognition: Its development in social contest*. Boston: Harvard University Press.

Lindberg, M., & Silvennoinen, H. (2017). Assessing the basic skills of the highly educated in 21 OECD countries: An international benchmark study of graduates' proficiency in literacy and numeracy using the PIAAC 2012 data. *Comparative Education,* DOI: 10.1080/03050068.2017.1403676

Maddox, B. (2015). The neglected situation: Assessment performance and interaction in context. *Assessment in Education: Principles, Policy & Practice*, *22*(4), 427-443.

Maddox, B., Zumbo, B. D., Tay-Lim, B., & Qu, D. (2015). An anthropologist among the psychometricians: Assessment events, ethnography, and differential item functioning in the Mongolian Gobi. *International Journal of Testing*, *15*(4), 291-309.

Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American psychologist*, *50*(9), 741.

Murat, F. (2008). Adult skill assessment: Emerging methods. *Education Formations*, 78, 81-92. http://media.education.gouv.fr/file/revue_78_Anglais/01/3/EDUC&FORM_EN_41013.pdf#page=80

OECD (2013a). *OECD skills outlook 2013: First results from the survey of adult skills*. Paris: OECD Publishing. doi: 10.1787/9789264204256-en.

OECD (2013b). *The Survey of Adult Skills: Reader's companion.* Paris: OECD Publishing.http://dx.doi.org/10.1787/9789264204027

Reubens, A. (2009). Early Grade Mathematics Assessment (EGMA): A conceptual framework based on mathematics skills development in children. USAID: Washington, DC. [Online]: https://pdf.usaid.gov/pdf_docs/Pnads439.pdf

Straesser, R. (2015). Numeracy at work: a discussion of terms and results from empirical studies. *ZDM*, *47*(4), 665-674.

Tout, D. and Gal, I. (2015). Perspectives on numeracy: Reflections from international assessments. *ZDM–International Journal of Mathematics Education, 47*(4), 691-706.

UNESCO (2004). *The plurality of literacy and its implications for policies and programmes*. Paris: Author. Online: http://unesdoc.unesco.org/images/0013/001362/136246e.pdf

UIS (2009). *The next generation of literacy statistics: Implementing the Literacy Assessment and Monitoring Programme (LAMP)*. Online: http://www.uis.unesco.org/Library/Documents/Tech1-eng.pdf

UIS (2017). *Implementation in diverse settings of the Literacy Assessment and Monitoring Programme (LAMP): Lessons for SDG 4*. Unpublished working paper.

UIS (2018). *Concept Paper on the Short Literacy and Numeracy Survey (SLNS)*. Unpublished working paper.

UIS (March 2018). *Reducing financial, technical and operational burden of monitoring progress towards SDG 4.6: Options Paper*. Unpublished working paper.

UNESCO (2017). Expert Meeting on SDG Indicator 4.6.1: Summary. Unpublished report.

Wagner, D. A., Sabatini, J., & Gal, I. (1999). *Assessing basic learning competencies among youth and young adults in developing countries: Analytic survey framework*

and implementation guidelines. Philadelphia: International Literacy Institute/UNESCO. (Online: http://literacyonline.org/products/ili/pdf/op9901.pdf]