Organisation de Coopération et de Développement Économiques
Organisation for Economic Co-operation and Development

26-May-2010

English - Or. English

DIRECTORATE FOR EMPLOYMENT, LABOUR AND SOCIAL AFFAIRS
DIRECTORATE FOR EDUCATION
BOARD OF PARTICIPATING COUNTRIES FOR THE PROGRAMME FOR THE INTERNATIONAL ASSESSMENT OF ADULTS COMPETENCIES

PIAAC ASSESSMENT FRAMEWORKS

4th meeting of PIAAC BPC
2-3 April 2009
Paris

The framework documents represent deliverables from the consortium and provide some refinement and extension to the draft versions that were distributed for review in October 2008. These documents serve to communicate to interested stakeholders what PIAAC measures and as guidance for the development and interpretation of the assessment instruments. The consortium, Education Testing Service (ETS), will be happy to clarify issues and implement corrections raised by the BPC (Contact: Claudia Tamassia: Tamassia@ETS.ORG).

Contact: Andreas Schleicher; Email: Andreas.Schleicher@OECD.org; Tel: + 33 (1) 45 24 93 66
Contact: William Thorn; Email: William.Thorn@OECD.org; Tel: + 33 (1) 45 24 78 04

JT03284135

TABLE OF CONTENTS

PIAAC ASSESSMENT FRAMEWORKS

Introduction

1. The approach to the measurement of adult literacy, numeracy and problem solving skills used in PIAAC is based on the development of frameworks which define the domain of competence to be assessed, what ought to be measured and identify characteristics that can be used in the construction and interpretation of assessment tasks.

2. Given the role of the assessment frameworks in defining what will be measured by PIAAC and in structuring the interpretation the results of the assessment, it is important that participating countries share a common understanding of the frameworks and their implications for the measurement and reporting of proficiency. To this end Members of the BPC are invited to:

- **DISCUSS** the frameworks for the domains of literacy, reading components, numeracy and problem solving in technology rich environments,
- **ADOPT** the definitions proposed for these domains, and
- **AGREE** to the proposed distributions of assessment tasks.

3. Draft versions of the PIAAC frameworks covering literacy, reading components, numeracy and problem solving in technology rich environments can be found in the annexes to this document. To aid the deliberations of members, this paper provides a summary of the definitions proposed for the domains, the dimensions or aspects of the domains and the distribution of tasks across the various dimensions or facets.

Definitions

4. The definitions proposed for the three domains of assessment are set out below.

Literacy: 'Literacy is understanding, evaluating, using and engaging with written texts to participate in society, to achieve one's goals, and to develop one's knowledge and potential.'

Numeracy: '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.'

Problem solving in technology rich environments: 'Problem solving in technology rich environments involves using digital technology, communication tools and networks to acquire and evaluate information, communicate with others and perform practical tasks. The first PIAAC problem solving survey will focus on the abilities to solve problems for personal, work and civic purposes by setting up appropriate goals and plans, accessing and making use of information through computers and computer networks.'

Task Dimensions

Literacy

5. Literacy is assessed in relation to:

- The *types* of *text* adults must read – this includes different text *formats* (continuous texts, non-continuous texts, digital texts, combined and multiple texts)*,* the range of *rhetorical stances* found in texts (e.g. description, narration, exposition).

- The *contexts* in which texts are read – those of work, personal life, community and citizenship, and further education.

- The *cognitive operations and representations* involved in adult reading – i.e. to access and identify information in the text, to integrate and interpret (relate parts of text to each other) and to evaluate and reflect on texts. Elements of texts that make these processes more or less difficult include semantic complexity, the transparency of information in the text, and the prominence of the information.

6. The reading components aspect of literacy will be assessed in relation to:

- word meaning (print vocabulary);

- sentence processing; and

- basic passage comprehension.

Numeracy

7. Numeracy is assessed in relation to four facets of numerate behaviour:

- the *contexts* in which adults face numeracy demands – those of everyday life, work, societal, further learning
- the types of *responses* required to manage numeracy problems – identify, locate or access, act upon, use (order, count, estimate, compute, measure, model), interpret, evaluate/analyse, communicate
- the *mathematical content* which characterise the mathematical demands faced by adults – quantity & number, dimension & shape, pattern, relationships, change, data & chance
- *representations* of mathematical information - objects & pictures, numbers & mathematical symbols, formulae, diagrams & maps, graphs, tables, text, technology-based displays

Problem solving

8. Problem solving in technology rich environments is assessed in relation to:

- The *cognitive dimensions* of problem solving – goal setting and progress monitoring, planning and self-organisation, accessing and evaluating information, making use of information.
- The *technological dimensions* of information problems – i.e. hardware devices, software applications, commands and functions, representations (e.g. text sound, video)
- *Task dimensions* – the intrinsic complexity of a problem, the explicitness of the problem statement.

Proposed distribution of tasks across dimensions

9. The proposed distribution of tasks across the main dimensions or facets of literacy, numeracy and problem solving is summarized below.

Literacy

Dimensions	***Distribution (%)***
Text types	
Print-based texts	
As used in previous studies	60-70
New to PIAAC: texts which combine prose and document elements, texts with images, and diagrams.	10-15
Digital texts	20-25
Contexts	
Work/occupation	15
Personal	40
Community and citizenship	30
Education and training	15
Cognitive aspects	
Access and identify information in the text	30-40
Integrate and interpret (relate parts of text to each other)	40-50
Evaluate and reflect	15

Numeracy

Dimensions	***Distribution (%)***
Responses	
Identify, locate or access	10
Act upon, use: order, count, estimate, compute, measure, model	50
Interpret, evaluate/analyse, communicate	40
Mathematical content	
Quantity & Number	30
Dimension & Space	25
Patterns, Relationships, Change	20
Data & Chance	25

Problem solving

Dimension	***Distribution (%)***
Contexts	
Personal	40
Work	30
Civic life	30

Distribution of field test tasks as a function of environment and cognitive dimensions

	Web	Spreadsheet	Email	Multiple
	9	4	6	6
Goal setting and monitoring progress	2	1	1	1
Planning	2	2	2	4
Acquiring and evaluating information	3	0	0	0
Making use of information	2	1	3	1

ANNEX 1: PIAAC LITERACY- CONCEPTUAL FRAMEWORK

Introduction

10. With national economies facing growing unemployment as the new century ends its first decade, the issues of human capital development rise in importance. In a series of studies in the 1990s, the OECD, Statistics Canada, and Educational Testing Service demonstrated the importance of literacy skills for the effective functioning of labor markets and for the economic success and social advancement of both individuals and societies. Data from three rounds of the International Adult Literacy Survey (OECD & Statistics Canada, 2000) and the Adult Literacy and Lifeskills Survey (OECD & Statistics Canada, 2005) revealed that low skills are found in all countries and that those low skills pose problems for individuals trying to cope with work and life in modern societies.

11. The demands on literacy skills have always been increasing. As economies became increasingly industrialized, nations needed ever more educated populations and literacy skill took on greater importance. As nations progressed, the quantity and type of materials continued to expand and individuals found they were expected to use information from written materials in new and more complex ways.

12. Now, as information-based economies are succeeding industrial-based economies, literacy is again being transformed. A new form of text, digital text, makes increased demands on readers, and changes the ways in which text is used. The amount of information available and its uncensored nature emphasize the abilities needed to connect, evaluate and interpret information. In addition, computer technologies have added new dimensions related to the nonlinear, recursive, and interactive nature of these environments. Because individuals now often move through the material in their own ways when searching for information, they very often create their own "texts" in the sense that the total set of information that each individual encounters is unique. Collectively, the skills required to effectively use digital information are less well understood than traditional print skills but suggest that we will need to expand our definition of what it means to be literate.

13. A broadened view of literacy that includes skills and knowledge related to information and communication technologies is increasingly seen as an essential component of the knowledge, skills and attitudes that facilitate the creation of personal, social and economic well-being. (Kellner, 2002; Partnership for 21st Century Skills, 2003; Rainie & Horrigan, 2005; Senn-Breivik, 2005). In the global economy, individuals and nations with information and communication technology skills will most likely prosper while those lacking them will struggle to compete. It is therefore essential that we understand more about the ways in which we use information and communication technologies and the associated outcomes of those uses. According to Murnane and Levy (1996), these "new basic skills" are needed by everyone regardless of their aspirations, regardless of whether they are male or female, and regardless of their social and economic backgrounds.

14. At the same time that we see these increasing demands on individuals, research, especially in the United States and Canada, has shown had many individuals still have difficulty with the underlying skills, such as vocabulary and fluency, that are the building blocks necessary to developing the higher levels of literacy. To have a full picture of literacy in any society, it is necessary to have more information about these individuals because these groups are at the greatest risk of experiencing negative social, economic and labor market outcomes.

15. The Programme for the International Assessment of Adult Competencies is designed to assess the state of the skills of individuals and nations in this new information world. Clearly, one of those skills is literacy and the framework set out in this paper by the Literacy Expert Group for PIAAC reflects its understanding of what the literacy skills that are necessary in the 21st century are.

16. It is helpful to understand the process by which this document was developed. The Group met for three days in late May 2008 and reviewed the literacy framework that had guided the IALS and ALL assessments as well other reports and documents. General agreement was reached on expanding and reworking those frameworks to reflect new literacy demands. Stan Jones, as Chair of the Expert Group, prepared a draft which was then circulated to all members for comment. Mr. Jones then reworked the draft based on these comments to produce a version that was discussed at the June 2008 NPM meeting. After that NPM further drafts were developed and circulated for comments. The Literacy Expert Group met again in Valencia, Spain, in November, 2008 and reviewed a further draft. This final version was prepared following further discussion at an Expert Group meeting in March, 2009.

17. In its discussion of component reading skills, the Expert Group determined that developing a detailed framework and set of measures around these skills was also important. While these measures are an important part of the PIAAC Literacy Framework, their nature is somewhat different than the measures of literacy-in-use developed here. The role of the components in an overall assessment of literacy is included in this framework, but the complete discussion of the measures and their justification and development can be found in a separate document approved by the Literacy Expert Group.

18. The Group would like to express its considerable appreciation to Mary Lou Lennon for carefully and thoughtfully reviewing an early draft of the document and to Juliette Mendelovits and her item development team for their feedback. Our thanks also go to Jean-Pierre Jeantheau who served as a consultant at the first meeting and in the review.

19. Members of the Literacy Expert Group:

- Egil Gabrielsen - Norway
- Jan Hagston - Australia
- Stan Jones (Chair) – Canada
- Pirjo Linnakylä - Finland
- Hakima Megherbi - France
- John Sabatini - USA
- Monika Tröster - Germany
- Eduardo Vidal-Abarca - Spain

Defining literacy for PIAAC

20. In any large-scale assessment, the definition of the domain to be measured provides a foundation for the design of the assessment and sets the boundaries for what will be included. Such a definition needs to be situated in the wider context of adult literacy, but constrained by the goals of the

PIAAC project. When considering the construct definition for PIAAC, the Literacy Expert Group began by looking at previous assessments. Literacy assessments have been part of two major international surveys of adult skills: the International Adult Literacy Survey (IALS) in the 1990s and the Adult Literacy and Lifeskills Survey (ALL) in 2003 and 2006. These were, in turn, informed by several prior national surveys. However, during the 20 years that these assessments cover, adults have faced new literacy opportunities and demands, especially as the use of -mail and other digital media has grown. For PIAAC to appropriately assess adult literacy in the 21st century, it is necessary to broaden the construct to include new modes of text. PIAAC also provides an opportunity to deepen our understanding of the cognitive skills that underlie adult literacy and of the role that engagement plays in literacy. The Group noted that the PISA 2009 definition specifically references both electronic texts and engagement (OECD: Programme in International Student Assessment).

21. Given the opportunity to rethink literacy as a construct, the PIAAC Literacy Expert Group was guided by four principles:

a) The definition must be descriptive and not normative. Many definitions of literacy set out an ideal state, or a goal for literacy action. The underlying construct for PIAAC has to recognize that individuals vary in their ability to succeed across the full range of literacy tasks they face in modern society.

b) The definition should support an expanded conception of literacy, in these ways:

  - There should be an assessment of the underlying skills of those at the lower levels of literacy;
  - The range of texts to be considered should be broader than in previous assessments; in particular, the definition should include those texts often identified as electronic texts; and
  - The type literacy behaviors measured should go beyond simply using text for an immediate purpose, so as to enable a deeper understanding of literacy ability.

c) At the same time, the definition should support a link to the IALS and ALL assessments to support the analysis of trends.

d) Cognitive assessment alone is not sufficient to understanding the status of literacy in a particular society. It is also important to consider the engagement of individuals in literate activities.

22. The Literacy Expert Group considered a number of definitions that had been recently proposed, and found the following definition from a UNESCO expert group in 2003 to be useful:

> *Literacy is the ability to identify, understand, interpret, create, communicate and compute, using printed and written materials associated with varying contexts. Literacy involves a continuum of learning in enabling individuals to achieve their goals, to develop their knowledge and potential, and to participate fully in their community and wider society.*

23. Some elements of this definition were beyond the scope of the current PIAAC assessment (though, perhaps, not future ones), especially the productive elements of creating and communicating.

Computing was not appropriate for inclusion in the PIAAC literacy definition as it is in the domain of the Numeracy Expert Group. Other parts needed to be rephrased to better support an assessment.

24. In the end, the Literacy Expert Group found that an expanded and re-ordered version of the IALS/ALL[1] definition would meet both the descriptive, expansive, and linking criteria it wanted for PIAAC.

> *Literacy is understanding, evaluating, using and engaging with written texts to participate in society, to achieve one's goals, and to develop one's knowledge and potential.*

25. In the rest of this section, we set out what we understand by the key elements of this definition

Written text

26. Previous literacy assessments have focused primarily on informative texts of both continuous and non-continuous form. It is the intention of the new construct to expand the range of texts to include a greater variety of text types, such as narrative and interactive texts, and a greater variety of media. Until recently, most adult reading was of material printed on paper. Now, adults need to access and use text that is displayed on a screen of some kind, whether of a computer, a PDA, an ATM, or a Blackberry or iPhone. The PIAAC definition encompasses all these.

27. It is worth noting that including electronic text opens the assessment to new types of text and content. While one can find examples of similar texts in paper, they are much less common in that form. Some of these novel form/content combinations include interactive texts, such as exchanges in comments sections of blogs or in e-mail response threads, multiple texts, whether displayed at the same time on a screen or linked through hypertext, and expandable texts, where a summary can be linked to more detailed information if the user chooses.

Understanding

28. A basic task for the reader is constructing meaning, large and small, literal and implicit, from text. This can be as basic as understanding the meaning of the words, to as complex as comprehending the underlying theme of a lengthy argument or narrative. Certainly, evaluating or using a text implies some level of understanding and so provides an indirect measure of it, but it is the intent of the PIAAC assessment to include some more direct measure of it. The components framework provides a construct to support basic understanding, but the literacy assessment, itself, should also include tasks that explicitly tap more complex understanding, such as the relation(s) between different parts of the text, the gist of the text as a whole, and insight into the author's intent. Readers also have to understand the social function of each text and the way this influences structure and content.

Evaluating

29. Readers continually make judgments about a text they are approaching. They need to assess whether the text is appropriate for the task at hand, determining whether it will provide the information they need. They have to make judgments about the truthfulness and reliability of the content. They

---

[1] The IALS/ALL definition reads: "Literacy is defined as a particular capacity and mode of behaviour: the ability to understand and employ printed information in daily activities, at home, at work, and in the community – to achieve one's goals and to develop knowledge and potential."

need to account for any biases they find in the text. And, for some texts, they must make judgments about the quality of the text, both as a craft object and as a tool for acquiring information.

30. Because the provenance of many digital texts is obscure and because it is much easier to distribute digital texts widely and anonymously, such judgments are especially important for electronic texts. Sources for online information are more varied, ranging from authoritative sources to postings with unknown or uncertain authenticity. All information must be evaluated in terms of accuracy, reliability and timeliness, but this is particularly important with online material.

Using

31. Much adult reading is directed toward applying the information and ideas in a text to an immediate task or goal or to reinforce or change beliefs. Nearly all the tasks in previous international assessments have been of this kind. In some cases, using a text in this way requires just minimal understanding, getting the meaning of the words with some elementary recognition of structure (many menus, for example). In others, it requires using both syntactic and more complex structural understanding to extract the information. In all cases though, the reader approaches the text with a specific task in mind.

Engaging with

32. Many adults appear to read text only when some task requires them to do so. Others (sometimes) also read for the pleasure it brings them and for general interest. Some adults read only what others – employers, governments – make necessary, while others read things of their own choosing as well. That is, adults differ in how engaged they are with text and how much a role reading plays in their lives. Studies have found that engagement with (attitude toward and practice of) reading is an important correlate with the direct cognitive measures. As such it is necessary to understand these differences to get a full picture of adult literacy.

Participate in society

33. While earlier definitions referred to the role of literacy in "functioning" in society, the PIAAC use of "participating" is meant to focus on a more active role for the individual. Adults use text as a way to engage with their social surroundings, to learn about and to actively contribute to life in their community, close to home and more broadly. And for many adults, literacy is essential to their participation in the labor force. In this, we recognize the social aspect of literacy, seeing it as part of the interactions between and among individuals.

Achieve one's goals

34. Adults have a range of needs they must address, from basic survival to personal satisfaction, to professional and career development, to participation in society. Literacy is increasingly complicit in meeting those needs, whether simply finding one's way through shopping, or negotiating complex bureaucracies, whose rules are commonly available only in written texts. It is also important in meeting adult needs for sociability, for entertainment and leisure, for developing one's community and for work

Develop one's potential

35. Surveys suggest that many adults engage in some kind of learning throughout their life, much of it self-directed and informal. Much of this learning requires some use of text and as individuals want to improve their life, whether at work or outside, they need to understand, use, and engage with printed and electronic materials.

The literacy domain

36. While this definition gives us a broad picture of what PIAAC literacy proposes to measure, and what it does not, a more detailed discussion of the domain is needed to guide item development, to provide a basis for assessing the validity of the results, and to support analysis and reporting of the findings. In this we are able to build on the work from IALS and ALL, which has demonstrated the power of a few variables to account for much of the variance in individual scores. We do not propose to substitute a different type of model here, but to extend the old one to different types of text and new aspects of reading and to recast some elements to support a deeper understanding of adult reading.

37. One aspect of the literacy construct, engagement, is best assessed through background questions and not in the cognitive assessment. As such, it is covered in a separate section that follows. As well, basic understanding is assessed in a set of measures known as the components test; it, too, is discussed separately.

38. In creating items for an assessment the writer has two primary elements to manage. One is the text itself, with several important features that need to be considered. One of these features is the medium, format and type of text – its character as an object and another is the social setting in which the text is most naturally situated. The second primary element is the task the item writer sets for the assessment. A number of characteristics are known to play a role in the difficulty of these tasks and the writer needs to keep these in mind in constructing the items. These characteristics also play an important role in the analysis of the results and in descriptions of the different levels of literacy measured by the assessment. While we discuss these two elements separately, there is a close connection between them; in many ways a text determines the types of tasks that be asked using it and the item writer needs to balance the need for adequate coverage of text characteristics with the need for adequate coverage of task characteristics. Nonetheless, we discuss these two primary elements separately in this framework.

Texts

39. Any assessment of literacy requires texts to serve as the stimulus for the tasks that form the test. We have found it useful to organize the texts for PIAAC in a number of ways:

a) Medium (print and digital)

b) Format (continuous and non-continuous)

c) Type (rhetorical stance)

d) Physical layout (type of matrix organization)

e) Features unique to digital texts, and

f) Social context.

Medium

40. A major development of PIAAC over previous adult surveys is the inclusion of digital (or electronic) texts. We discuss below some key features of digital texts that distinguish them in practice from printed texts. We recognize that many texts that would previously be encountered in printed form are now just as likely to be accessed digitally. For example, consumer information about a product that was once printed in a brochure may now be available only in some electronic format. Indeed, many of

the respondents in PIAAC will encounter all the texts electronically, some of which will be simple copies of printed texts. We want to distinguish digital texts not simply by the medium in which they occur, but by the use they make use text navigation and display features that found only through digital devices. Any text that could appear on a printed page exactly as it appears on a screen will be considered a *print* text; any text that could not appear on a printed page with all its features intact will be considered a *digital* text[2].

Format

41. In IALS and ALL, and in PISA, texts were classified as continuous (prose) or non-continuous (document). This is an important distinction, as each format requires different text knowledge and a different approach to text processing. At the same time many actual texts involve some elements that are continuous and some that are non-continuous. Thus, the distinction is better made on the basis of what type(s) of text a task requires.

a) *Continuous.* This type of text is conventionally made up of sentences formed into paragraphs. Some continuous texts include typographic features, such as indenting and headings, that signal the organization of the text, but many do not. Examples of continuous texts include newspaper and magazine articles, brochures, manuals, e-mails and many web pages.

b) *Non-continuous.* This type of text uses explicit typographic features, rather than paragraphs, to organize information. While there may be full sentences in some non-continuous texts, most consist of words or phrases organized by some kind of matrix arrangement. Tables, graphs, charts and forms are all examples of non-continuous texts.

c) *Mixed.* This type of text has both continuous and non-continuous elements. Examples of mixed texts include web pages with a list of links, newspaper articles that incorporate line graphs or pie charts, and brochures with attached order forms.[3]

d) *Multiple.* Multiple texts consist of texts that have been generated and which make sense independent of each other. The texts are juxtaposed or loosely linked for a particular purpose. The relationships among the component texts need not be obvious. The texts may be contradictory or complementary. Such texts are common in digital settings, but are also found in print environments.

Text type (Rhetorical Stance of the Text)

42. The IALS / ALL framework classified continuous texts by their rhetorical stance, since all share much the same structure (sentence and paragraph, with or without headings). But non-continuous texts also share the same rhetorical stances. Therefore, we propose to identify the stance of all types of text using the six categories employed in the IALS / ALL assessments. We would note that although narrative texts have been included as a text type in the past, few narrative texts have actually appeared in previous assessments. Our intent is that such texts should form a part of the PIAAC assessment. We

---

[2] Although many pdf files are simply scanned copies of printed texts, the Acrobat reader application adds navigation, search and annotation features not available for the printed version. Thus, a pdf text would be a digital text by this description.

[3] There are texts, such as comics and graphic novels, which rely on graphic elements to carry important parts of the information. It seems to us that these are a special case of continuous text where the pictorial displays function to organise the sequence of ideas somewhat as the ordering of paragraphs does.

have eliminated one text type that had been listed for IALS / ALL: Hypertext, as it is not a rhetorical category, but a structural type which will be included under electronic text for PIAAC.

43. The point of having rhetorical stance as a variable is not that there is any evidence that difficulty is affected by the stance, but as a way ensuring that a variety of texts are included on the assessment.

44. The six types[4] of rhetorical stance for PIAAC are as follows:

a) *Description* is the type of text where the information refers to properties of objects *in space.* A page of a manual that identifies the parts of some device, such as a Cuisinart, is a description, as is a verbal depiction of a piece of art.

b) *Narration* is the type of text where the information refers to properties of objects *in time.* Stories recounted to make a point, such as fables, are narrations, as are texts about the steps what an individual took solve a problem.

c) *Exposition* is the type of text in which the information is presented as composite concepts or mental constructs, or those elements into which concepts or mental constructs can be analyzed. The text provides an explanation of how the component elements interrelate in a meaningful whole. A text that explains the nature of some health problem or one that talks about the effect of climate change would be an exposition.

d) *Argumentation* is the type of text that presents propositions as to the relationship among concepts or other propositions... An important subclassification of argument texts is persuasive texts. Newspaper editorials are one example, and advertisements are another.

e) *Instruction* (sometimes called *injunction*) is the type of text that provides directions on what to do. Most equipment manuals contain instruction texts, as do other guides, such as those about first-aid or some leisure activity.

f) *Records* are texts that are designed to standardize, present and conserve information without embedding in other stances. A table of standings in a sports league is an example of a record, as is a graph of the changes in oil prices. The minutes of a meeting constitute another type of record.

45. The existing items from ALL and IALS that have been selected as linking items for PIAAC cover these categories well, with the exception of narrative texts as noted previously. A goal for the new PIAAC literacy items is to include a range of texts that represents all six of these types.

Classifying non-continuous texts

46. In additional to their rhetoric, non-continuous texts differ in their structural organization. While all continuous texts have the same sentence-in-paragraph form, non-continuous texts differ in form and a major challenge to understanding their content is to understand the meaning of their forms. The IALS / ALL framework identified five types of non-continuous structures that we propose to adopt for PIAAC:

---

[4] These are taken almost directly from the IALS / ALL frameworks, as we saw no reason to change them.

a) *Matrix Documents*. This set of non-continuous text consists of four types of increasingly complex documents that have simple lists as their basic unit.

  - A simple list consists of a label and two or more items, where the label serves as the organizing category and the items all share at least one feature with the other items in the list. A basic shopping or 'to do' list is an example of this basic structure.

  - Next are combined lists, which consist of two or more simple lists. One list in a combined list is always primary and, as such, is ordered to facilitate looking up information within the list and then locating parallel information within the other lists. An email inbox is an example of the combined list structure with its related lists of sender names, subjects, dates and file sizes.

  - Intersected lists are the third type of matrix document and are composed of exactly three lists. Two of the lists form a row and column defining the cells which contain the third or intersected list. Television listings are a common example of an intersecting list with the channels and times defining the program content listed in each cell.

  - The fourth and most complex type of matrix document is the nested list. In order to economize on space, as well as to display comparative information, designers sometimes combine two or more intersecting lists to form a nested list. In a nested list, one type of information will be repeated in each of the intersecting lists. The intersecting list of unemployment rates, for example, may have separate entries under each month for males and females; in this case, gender would be nested under month.

b) *Graphic Documents*. A major function of graphic documents is to provide a succinct visual summary of quantitative information. Included in this group of texts are pie charts, bar charts, and line graphs. While these appear to be very different types of documents on the surface, they all derive from, or can be transformed into, either a combined, intersecting, or nested list.

c) *Locative Documents*. Like graphic documents, locative documents or maps portray information visually. Unlike graphic documents that display quantitative information, maps either portray the location of persons, places, or things in space, or depict characteristics of different geographic regions (e.g., types of vegetation or characteristics of a population).

d) *Entry Documents*. In matrix and graphic documents, the author provides the information that must be read and used. In contrast, entry documents, or forms, require the reader to provide information that can range from very simple to complex. For example, the reader may be asked to simply check a box; write a single word, number, or phrase; or construct a series of phrases or sentences. Generally speaking, forms provide the reader with a label or category for which the reader is asked to provide specifics.

e) *Combination Documents*. Some displays, especially graphic documents, rely on the use of other types of text for their interpretation. Maps and graphs, for instance, often include legends that display important information that must be read and understood. In addition, designers sometimes include more than one document for display or comparative purposes.

Digital text features

47. The new definition introduces digital texts as a type to be included in the assessment. It is important to note that digital texts use many of the same organizing principles and rhetoric as continuous and non-continuous texts. However, these texts have properties that function as navigation tools and features and distinguish them from paper and ink texts[5].

a) *Hypertext*. Electronic texts may provide direct links to other texts and it may be necessary to follow these links to gain a full understanding of a topic. The physical process is normally to click with a mouse on a highlighted word or phrase in a text to access another text with (additional) ideas and information relevant to the highlighted part of the passage. We would initially distinguish two main types:

    - *Index-like.* The initial text is a list of topics, from which the reader selects one or more for additional information. A common example is a news site on the web which lists headlines on which the reader clicks to view the full stories. An e-mail inbox is another example of an index hypertext; clicking on an entry takes the reader to the full e-mail. In many workplaces such index hypertexts are used to retrieve information for a specific task. The entry screen lists categories of information from which the reader selects and the appropriate information is then retrieved and displayed by the computer (or other device; PDAs are often used in this way). Schedules and electronic calendars are another example of an index-like hypertext as the user can usually click on an entry in a schedule for additional information about that entry.

    - *Text-embedded.* In this type of hypertext a link is embedded in a complete text and the reader is taken to a second text that expands on the immediate topic. A common example is Wikipedia, in which an entry includes many links to other entries or to other web sites. A second kind of text-embedded hypertext includes links to other parts of the same text. For example, footnotes in electronic texts can often be accessed by clicking on the footnote number in the text.

b) *Interactive*. In electronic text a reader often comes across texts that have been created by a series of authors. E-mail exchanges, where the sequence of messages is retained when replies are simply added to previous messages, are a common example of this interactive text. Comments sections of blogs or other web documents such as news sites that allow comments on stories are another instance. In these texts, later entries often cannot be understood without understanding prior contributions.

c) *Other Navigation Features*. Digital documents typically have navigation features that differ from printed texts. While it seems intuitive to experienced readers to turn a page in a multi-page text, it is something that must be learned. In the same way,

---

[5] We believe there are analogs in standard printed texts to both of these, but they are much more common in digital text and are easier to construct and use in their digital form. We should also note that the use of computer-assisted text layout has resulted in paper-and-ink texts that are more complex than previously. One only needs to compare a newspaper from the 1960s to one from today to see the impact of this change.

understanding that the scroll bar takes the reader to additional text, or that a digital text might require the reader to click on a next page (or previous page) button to move to new text are features that the digital reader must learn.

Social contexts

48. Adult reading normally is part of a social setting. Both the motivation to read and the interpretation of the content may be influenced by the context. As a result, a fair assessment must include material from a broad range of settings, so as to include some material that would familiar to any participant.

49. The following content areas have proven useful in previous assessments:

a) *Work and occupation* includes materials that deal in general with various occupations but not job-specific texts, finding employment, finance, and being on the job.

b) *Personal uses*

- *Home and family* includes materials dealing with interpersonal relationships, personal finance, housing, and insurance.
- *Health and safety* includes materials dealing with drugs and alcohol, disease prevention and treatment, safety and accident prevention, first aid, emergencies, and staying healthy.
- *Consumer economics* includes materials dealing with credit and banking, savings, and advertising, making purchases, and maintaining personal possessions.
- *Leisure and recreation* includes materials involving travel, recreational activities, and restaurants, as well as material read for leisure and recreation itself.

c) *Community and citizenship* includes materials dealing with community resources and staying informed.

d) *Education and training* includes materials that deal with opportunities for further learning.

Tasks

50. While any reading event is a complex cognitive operation, typically one aspect dominates any particular instance. A reader may be looking for a particular piece of information, but other aspects do come into play. The reader must judge whether the information source is reliable and whether it has the appropriate content. The reader must understand the semantic content of the text. But these are secondary to the main task of locating a specific piece of information. At other times, the reader may be primarily interested in understanding some phenomenon. Again judgments will have to be made and specific information will have to be considered, but the primary goal is broad understanding. The item writer attempts to capture this primary purpose in creating a task and any assessment must have tasks that focus on a variety of aspects of reading.

Aspects of tasks

51. Both IALS and ALL identified three broad aspects of tasks that readers were asked to carry out: those that require identification of pieces of information in the text, those that require connecting different parts of the text, and those that require some understanding of the text as a whole.

a) *Access and identify information in the text.* On some occasions adults are simply seeking specific items of information from a text. What time does the movie start? How many cups of flour are used for this cake? What does this candidate propose to do about the roads?[6] Sometimes finding the needed information is relatively simple, as it is directly and plainly stated in the text. However, identify tasks are not necessarily easy ones. For some tasks, inferences may be required and rhetorical understanding may have to be called upon. For example, using a text to find the reasons for a change of action by the local government may require an understanding of how reasons are presented in text. In addition, sometimes more than one piece of information is required. IALS / ALL called these access and identify tasks locating (when only one piece of information was required) and cycling (when more than one is required).

   Most tasks that require only identification would be classified as "using" tasks. But some evaluation tasks might require identifying several distinct pieces of information which might be compared for their relevance in a particular situation.

b) *Integrate and interpret (relate parts of text to each other).* Often tasks require the reader to understand the relation(s) between different parts of a text. These relations include problem-solution, cause-effect, category-example, equivalency, compare-contrast, and understanding whole-part relationships. To complete such tasks, the reader has to determine what the appropriate connection is. This may be explicitly signaled, as when the text states "the cause of X is Y", or may require inferencing by the reader. The parts to be related may be near each other in the text or may be in different paragraphs or even in different texts. In IALS / ALL such tasks were usually called integrate tasks.

   Compare and contrast, two basic evaluation steps, are examples of relating parts of a text to each other. Establishing what is the basis of a relationship between parts is a form of understanding at the sub-text level. Some parts of a text must be understood in the context of the rest of the text, as they take on essential elements of meaning from the larger text in which they occur.

   Readers are sometimes called upon to come to some understanding of a text as a whole. For example, the reader may need to determine the purpose of a text or comprehend its main theme. Again the text itself may make this explicit, as with a title or an introductory sentence or paragraph, but often it is something readers must discover on their own and produce a paraphrase or summary.

c) *Evaluate and reflect*. Evaluation and reflection involve drawing on knowledge, ideas or values external to the text. The reader has to approach the text from outside, assessing the relevance, credibility, truthfulness of the information or argumentation presented

---

[6] Note that this is a different task than answering should I vote for this candidate, but may be a step toward making that decision.

in the text. The reader may also evaluate the purposefulness, register, structure or reader awareness of the text, or how successfully the writer is using the evidence and language to argue or persuade a reader. Evaluation is, particularly, important in reading electronic texts where the reader needs to be more alert to the text's accuracy, reliability and timeliness.

Readers also need to be aware of how the author of a text is attempting to persuade them to a particular end. Such meta-textual awareness is part of evaluating and reflecting on a text. And a reader needs to determine whether they are the intended audience for the text.

Cognitive representations

52. In understanding a text, a reader has to create a mental representation from linguistic materials of the text. For continuous texts these materials are words, phrases, and sentences. For non-continuous texts, the materials also include the matrix and list relations underlying the arrangement of text elements. There is a substantial body of research[7] on how readers create these representations for continuous texts, but, despite the ubiquity of non-continuous texts in adult reading, there is little work on how such representations are created for them[8].

Factors that affect the construction of representations

a) *Transparency of the information.* An important factor in task difficulty is the transparency of the information in the text. When the question refers explicitly to the superficial information (literal information), it's easier to process. For some tasks, the needed information is explicitly signaled; a telephone number always has a particular form and may also be preceded by "Tel" in the text. The text may have a title, or the problem and solution may be directly labeled as such (explicit signal).

b) *Degree of complexity in* making inferences

Paraphrase. Readers have to process linguistic information by mobilizing their lexical and syntactic-semantic knowledge. Simple examples would be a task requiring readers to find information about the cost of an automobile in a table using the word 'car', or such as knowing that "ate" in a text indicates food.

High level text inference. In a problem-solution text, for example, neither the problem nor the solution need to be explicitly signaled; rather the reader may have to infer what the problem (and/or solution) is from the text itself. And the reader cannot necessarily assume that the problem statement will precede the solution.

Extra-textual inference. Some tasks require the reader to bring information from outside the text or from another text in order to understand parts of the text in question. For example, in a notice about local road repair projects, the reader may be expected to bring external knowledge about the types of roads in that area to understand the actual repair proposals.

---

[7] In particular, see Garrod & Sanford, 1994; Gernsbacher & Foertsch, 1999; Oakhill & Garnham, 1988; Kintsch, 1998; and Perfetti, 1994.

[8] There has been more interest in non-continuous texts among researchers concerned with digital texts. See, particularly, Rouet (2006), which is also useful on the processing differences of digital and print texts.

- *Semantic complexity and syntactic complexity.* Studies of both oral and written text have shown that the more concrete the information is the easier the task is. Tasks requiring the reader to identify persons, things or places tend to be easier than those that involve abstract properties, such as goals, conditions and purposes. The grammar structure of the sentence (question and text) could be more or less complex. For example, negative phrases are more complex than affirmative phrases. The presence of subordinate clauses (question or text) improves the complexity of syntactic processing.

- *Amount of information needed.* The more information the reader needs from the text to complete the task, the more difficult that task will be. The distinction between locate and cycle tasks in IALS/ALL was based on this (cycle tasks required the reader to make several independent identifications). As well, the amount of text that must be processed also plays a role in the difficulty of any task.

- *Prominence of the information.* If the information the reader needs is located in a prominent location in the text (in the first or last sentence of a paragraph, in a main rather than subordinate clause, at the top or bottom of a list) it will be easier to access.

- *Competing information*. The more potentially relevant information that the reader has to sift through to access the needed information, the more difficult the task will be. This is especially true if there is information in the text that might plausibly be appropriate, but is incorrect. For an obvious example, if a text has a telephone, a fax number and a mobile number, it will be more difficult to find the fax number, than if that were the only number in the text.

- *Text features*. The degree to which the reader has to construct relations among parts of the text affects difficulty. When there are large numbers of anaphoric references the reader must sort out and when text cohesion signals are absent, the reader will find the task more difficult.

Consequences for test development

53. The expanded definition of literacy has some important consequences for test development for PIAAC literacy. In the assessment structure proposed for PIAAC, there will be 48 literacy items on the computer based test and 24 on the paper on pencil test. Of these 29 of the computer and 19 of the paper and pencil items will be linking items drawn from the IALS / ALL pool. Only 19 computer and 5 paper and pencil items will be new. As the linking items test just part of the PIAAC framework, namely just the using written text portion, almost all item development needs to be directed to new parts of the framework.

54. In order to have the items necessary for the final test a larger pool of items will need to be developed. As Table 1 shows, a total of 42 items from ALL and IALS have been selected for development as linking items. All 42 of these items will be authored for computer-based delivery and a subset of 25 items will also be used for the paper and pencil version of the assessment. A total of 70 new items will be developed, 55 for the computer and 15 for paper and pencil. The 15 paper-and-pencil items may be subset of the 55 developed for the computer-based version of the assessment or they may be unique to the paper and pencil version.

55. The items prepared in the development phase will be reviewed by countries and the consortium. Forty-two linking items and up to 40 new items will be selected for the field test. Required numbers of items for the field test and main assessment are shown.

Table 1. Item development targets

	Development		**Field-test**		**Main assessment**	
	Linking Items	New	Linking items	New	Linking items	New
Paper & pencil	25	15	25	10	19	5
Computer-based	42	55	42	30	29	19

Item development by task characteristics

56. The Literacy Expert Group has specified overall targets in terms of the distribution of items across the three defined task characteristics: text type, context, and process. These distributions are shown in Tables 2 through 4 below.

Table 2. Distribution of items by medium

Medium	**Distribution of items (%)**
Print-based texts	
As used in previous studies: newspapers, magazine, books, brochures, manuals, announcements, letters, advertisements, etc.	60 – 70
New to PIAAC: texts which combine prose and document elements, texts + images, and diagrams.	10 – 15
Digital	
Including hypertext, interactive environments such as message boards and chat rooms, texts which combine prose and document elements, and texts + images	20 – 25

Note: Each category includes continuous, non-continuous and combined texts.

57. As Table 2 shows, the goal will be for electronic texts to comprise between 20 and 25 percent of the assessment. The bulk of the texts will be more traditional paper-based texts; however, within that category the expert group would like to include sets of tasks based on materials where prose and document or text and images are integrally related and respondents must use and/or integrate information from both sources.

58. The distribution by context area shown in Table 3 is included to ensure that materials for the assessment represent a broad range of settings. As the motivation to read and interpretation of content may be influenced by context, including a broad range is meant to help ensure that no group of respondents will be either advantaged or disadvantaged based on their familiarity with, or interest in, a particular context.

Table 3. Distribution by items by contexts

Contexts	**Distribution of items (%)**
Work/occupation	15
Personal • Home and family • Health and safety • Consumer economics • Leisure and recreation	40
Community and citizenship	30
Education and training	15

59. Table 4 shows the targeted distribution of items by task aspects. The operations shown apply to each of the components of the definition, except for engagement. That is, a reader can access and identify information, integrate and interpret, and evaluate and reflect in order to understand, to evaluate and to use text.

Table 4. Distribution by task aspects

Task aspects	**Distribution of items (%)**
Access and identify information in the text • Locate • Cycle	30 – 40
Integrate and interpret (relate parts of text to each other) • Cause/effect • Compare/contrast	40 – 50
Evaluate and reflect	15

60. Although PIAAC does not plan to provide separate scales for continuous (prose) and non-continuous (document) tasks, it is important that there be approximately an equal number of tasks of each kind, recognizing that the framework does call for some tasks that use both continuous and non-continuous sources. These latter items, called mixed items here, are an important new feature of the PIAAC framework and the assessment must include items of this type.

Engagement

61. The concept of reading engagement is an important one in adult literacy, referring to the degree of importance of reading to an individual and to the extent that reading plays a role in their daily life. Empirical studies with children and adults have shown that differences in engagement are systematically related to differences in performance on assessments. Engagement theory typically identifies 5 integrated aspects of the concept.

a) *Amount and variety of reading.* The more one reads and the more different types of reading (purposes, types of text) one uses, the greater one is engaged with reading.

b) Interest in reading. The more one seeks out reading as a means of obtaining information and for enjoyment, the greater one is engaged with reading.

c) Control. The more one feels in control of what one reads and is able to direct one's own reading, the greater the engagement.

d) Efficacy. The more an individual feels able to read well, especially the confidence to read successfully new texts, the greater one is engaged with reading.

e) Social interaction. The more one is interested in sharing reading experiences and seeks out others to talk about reading, the more one is engaged with reading.

62. Experience in IALS and ALL has demonstrated that aspects 1 and 2, and to some extent 6, can be reliably evaluated through self-reports, but that adults do not, indeed may not be able to, reliably self-report control and efficacy. Questions appropriate for assessing engagement have been recommended for inclusion in the background questionnaire.

Component assessment

63. In previous surveys, the information on the reading abilities of adults with poor skills was often insufficient to get a proper understanding of their difficulties. Because they could answer so few of the items on the test, there was little on which to build a description of their abilities. The literacy framework for PIAAC includes a component test intended to provide that information.

64. The Components Assessment is more fully described in a separate document. Because we regard it as a supplement to the main literacy assessment a summary of that framework is included here.

65. The components framework builds upon a basic principle of learning to read, now widely researched and accepted internationally. That is, the comprehension or 'meaning construction' processes of reading are built upon a foundation of component skills and knowledge of how one's writing system works. The evidence of this knowledge and skills can be captured in tasks that examine a reader's ability and efficiency in processing the elements of the written language – letters/characters, words (and nonwords), sentences, and larger, continuous text segments.

66. A second principle guiding the components design is that the main interest is in whether the adults surveyed can apply their existing language and comprehension skills to the processing of printed texts. The components tasks are not designed to separately assess the level of language skills in the target print literacy writing system. It is assumed that the adults surveyed will have basic oral vocabulary, syntactic/ grammatical, and listening comprehension skills in the target language. We provide a component measure of basic oral vocabulary as an indicator that individuals surveyed have a threshold level of language proficiency. However, independent measurement of language proficiency is not a basic feature of the component framework.

67. A third principle of this model of reading is that the level of proficiency, efficiency, and integration of component skills is indicative of level and learning potential in reading development. As skills and knowledge accumulate, the ease of processing familiar text-based information increases. Component efficiency is typically indexed by assessing speed or rate of processing, as well as

accuracy. As learners, we spend extra time, effort, and energy to solve problems that are novel. On familiar tasks, we can often respond, accurately, quickly, with seemingly little conscious effort. When the tasks are easy, we can spend more effort solving and learning from more complex problems and tasks.

68. Work with this model has typically identified five components:

- alphanumeric perceptual knowledge and familiarity,
- word recognition,
- word knowledge (vocabulary),
- sentence processing, and
- passage fluency.

69. In skilled reading, these components are integrated to support literacy performance. During acquisition, even with adults, they may be measured separately, with different profiles having implications for learning, instruction, and policy.

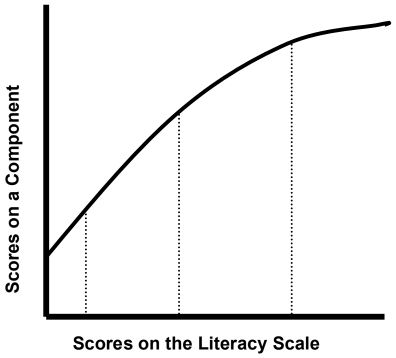
70. Two of these components, a) and b), are particular to the writing system of a language. Because some languages have more transparent writing systems than others, it is difficult to develop measures that could be used to compare different national populations.[9] As a consequence, the first two are offered as optional elements of the assessment.

71. The components are seen as an integral part of the overall literacy assessment. They are necessary, though not sufficient elements of skilled reading. In analysis, the component scores should be reflected on the literacy distribution, as in Figure 1, to show the relation between component acquisition and literacy ability.

---

[9] For example, a component study with adults in Canada found significant differences in the transparency of the French and English writing systems which made comparisons of results in the two languages difficult.

Figure 1. Proposed approach to reflecting components scores on the PIAAC literacy scale



72. Details of each of these five components are available in the full Components Assessment framework.

Developing the components assessment

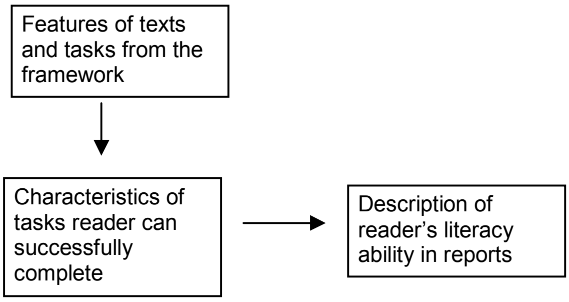
73. A prototype components assessment will be developed in English by the Consortium. Some parts may require nothing more than translation to be usable in other languages. Other elements may need to be developed specifically for some or all languages. For these, the English will serve as guidelines. The intent is to provide comparable assessments that will support international comparisons.

Analysis

74. It is anticipated that the scores on the PIAAC literacy assessment will be determined as they were for IALS and ALL and that the PIAAC scores will be equated to those on the previous surveys, using a 0-500 scale. Levels, similar to those used in IALS/ALL should also be computed and reporting should make use of these levels, of the 0-500 scale and of the components analysis.

75. Descriptions of the literacy ability of participants should be based on the characteristics of the texts and tasks they are successful with, using the features of this framework, as in Figure 2.

Figure 2. Relation of reporting framework to literacy framework



REFERENCES

Cain, K., Oakhill, J.V., Barnes, M.A., & Bryant, P.E. (2001). Comprehension skill, inference-making ability, and their relation to knowledge. *Memory and Cognition, 29,* 850-859.

Cain, K. & Oakhill, J.V., & Bryant, P.E. (2000). Investigating the causes of reading comprehension failure: The comprehension-age match design. *Reading and Writing, 12,* 31-40.

Garrod, S.C., & Sanford, A.J. (1994). Resolving sentences in a discourse context. How discourse representation affects language understanding. In M.A. Gernsbacher (Ed.), *Handbook of Psycholinguistics* (pp.675-698). New York: Academic Press.

Gernsbacher, M.A., & Foertsch, J.A. (1999). Three models of discourse comprehension. In S. Garrod, & M. Pickering (Eds.), *Language Processing* (pp. 283-300). East Sussex, UK: Psychology Press

Kellner, D. (2002). Technological transformation, multiple literacies, and the re-visioning of education. Retrieved August 8, 2005 at www.gseis.ucla.edu/faculty/kellner/essays/technolgicaltransformation.pdf

Kintsch, W., (1998). *Comprehension: A paradigm for cognition*. New york: Cambridge University Press.

Levy, F. & Murnane, R.J. (2004). *The new division of labor: How computers are creating the next job market.* Princeton, NJ: PrincetonUniversity Press.

Murnane, R. J., & Levy, F. (1996). *Teaching the new basic skills: Principles for educating children to thrive in a changing economy.* New York, The Free Press.

Partnership for 21st Century Skills (2003). Learning for the 21st Century. Retrieved August 8, 2005, at www.21stcenturyskills.org/downloads/P21_Report.pdf

Perfetti, C.A., (1994), Psycholinguistics and reading ability. In M.A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 849-894). New York, NY: Academic Press.

Perfetti, C.A., & Hart, L. (2001). The lexical bases of comprehension skill. In D.S. Gorfein (Ed.), *On the consequences of meaning selection: Perspectives on resolving lexical amgiguity. Decade of behavior* (pp. 67-86). Washington: American Psychological Association.

Oakhill, J.V., & Garnham, A. (1988). *Becoming a skilled reader*. Oxford: Basil Blackwell.

Rainie, L., & Horrigan, J. (2005). A decade of adoption: How theInternet has woven itself into American life. Pew Internet & American Life Project. Retrieved August 8, 2005, at www.perinternet.org/PPF/r/148/report_display.asp

Rouet, J.-F., Lowe, R., & Schnotz, W., eds. (2006). *The skills of document use.* New York: Springer.

Senn-Breivik, P. (2005, March). 21st century learning and information literacy. *Change*, 21-27.

ANNEX 2: ADULT NUMERACY- A CONCEPTUAL AND ASSESSMENT FRAMEWORK FOR PIAAC

Members of the Numeracy Expert Group: Iddo Gal (Israel; Chair), Dave Tout (Australia), Sean Close (Ireland), Myrna Manly (USA), Lene Johansen (Denmark), Terry Maguire (Ireland), Silvia Alatorre (Mexico), Jeff Evans (UK).

Introduction

76. This document presents a framework for conceptualizing numeracy and developing a scale for the direct assessment of adult numeracy as part of OECD's Programme for International Assessment of Adult Competencies (PIAAC). Numeracy as viewed here refers to adults' 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.

77. The conceptual framework and assessment issues discussed in this document were developed on the basis of several lines of work. This document builds on conceptual and assessment frameworks and cumulative wisdom developed in connection with prior surveys of adult skills, primarily the Adult Literacy and lifeskills project [ALL] and the International Adult Literacy Survey [IALS], but also surveys of school-age students (e.g., PISA, TIMSS). During preparation of this document by members of the PIAAC Numeracy Expert Group, a detailed review was conducted of professional and research literature in relevant areas such as on adult competencies, workplace skills, adult learning, or mathematics and statistics education. Ideas and feedback were also obtained from an international expert panel in 2006-2007 and all participating countries could react to drafts circulated for commentary during 2008.

78. This framework is organized in six parts which cover separate but interrelated issues:

- Part 1. A rationale for assessment of numeracy in PIAAC
- Part 2. Conceptual and theoretical foundations about adult numeracy
- Part 3. Facets of numeracy
- Part 4. Scale development: Principles, constraints, implementation
- Part 5. Differences between PIAAC's numeracy and related constructs and scales
- Part 6. Summary and further reflections

Part 1. A rationale for assessment of numeracy in PIAAC

Overview

79. PIAAC is a policy-driven initiative intended to provide policy makers and key stakeholders at the national and international levels with information that can inform policy-setting and planning of social interventions and programs. PIAAC's overarching goals (OECD, 2006) are twofold. First, to identify and measure differences within and across countries in "*literacy competencies for the information age – the interest, attitude, and ability of individuals to access, manage, integrate, and evaluate information, construct new knowledge, and communicate with others in order to participate*

effectively in the information age". Second, to assess the relationship of adult competencies with economic and social outcomes believed to underlie both personal and societal success (e.g., earnings, employment, educational attainment, participation in further learning) and optionally with additional outcomes or processes at the individual level (e.g., health, social capital) or workplace level, and with transitions at key points over the lifespan, such as school-to-work and possibly other stages.

80. The OECD planning paper states (OECD, 2006) that PIAAC is expected to provide reliable, valid, and valuable information to policy makers, differentiate the performance of low-scoring adults in each participating country, as well as include technology-based measures that tap into higher-order reading and thinking skills. As such, the design for the direct assessment in PIAAC incorporates computer-based measures of competencies, yet also use paper and pencil means for adults who are unfamiliar or uncomfortable with computers.

81. PIAAC is further expected to enable continuity with and links to the two previous international adult assessments, the International Adult Literacy Survey (IALS) and the Adult Literacy and Lifeskills (ALL) Study. Nineteen OECD countries[1] participated in either IALS or ALL, with nine participating in both. For these countries, it is paramount that PIAAC's direct assessment enables capitalization on their previous investments and provides an indication of how adults' competencies have changed since the previous measurement point(s). Equally important to maximizing countries' past investments in adult learning, is adding value to IALS and ALL by seeking innovation by broadening what skills are measured as well as how they are measured. As the intent of PIAAC is to have its results linked to previous international adult assessments, PIAAC has been designed with a specification by OECD that 60% of the literacy and numeracy tasks will come from item pools used in ALL and IALS. As a result, this conceptual framework for assessing numeracy in PIAAC maintains conceptual and pragmatic links to the numeracy framework developed for ALL.

82. Finally, OECD has expressed a desire that PIAAC's direct assessment will be conceptually compatible with PISA. Although no direct statistical links or common items are necessarily expected between the two assessment programs, such a linkage can provide policy-makers with information about the spectrum of competencies across different points of the lifespan, enable analyses of antecedents, consequences, and correlates of the distribution of competencies, and help to identify implications and relevant social interventions.

Rationale for assessing numeracy in PIAAC

83. 'Numeracy' is listed in the PIAAC overarching framework as a component of the broad set of 'literacy competencies'. Yet, numeracy should be viewed as a key competency which is *not* subsumed under 'literacy' as this later term has been traditionally defined regarding reading, writing, and comprehending the meaning of text or communicating through textual means. While literacy and numeracy in the traditional sense have some linkages, numeracy is a broad construct with a life of its own and it has a central and often quite distinct role in adults' lives. A later section reviews various perspectives which inform the conceptualization of numeracy, leading to a definition of numeracy for PIAAC and description of facets or dimensions of numerate behavior.

---

[1] Countries that participated in IALS include: Australia, Belgium (Flemish Community), Czech Republic, Canada, Denmark, Finland, Germany, Great Britain, Hungary, Ireland, Italy, Korea, Netherlands, New Zealand, Norway, Poland, Sweden, Switzerland, and the United States. Countries that participated in the first stage of ALL include: Bermuda, Canada, Italy, the Mexican State of Nuevo Leon, Norway, Switzerland, and the United States. ALL's second stage included Australia, Korea, Hungary, New Zealand, and the Netherlands.

84. This framework is founded on the assumption that a direct assessment of numeracy in PIAAC is an essential and worthwhile undertaking (Willms, 2006; Murray, 2006), for four separate but related reasons:

a. ***Numeracy is essential for adults and for the societies in which they live.*** Basic computational or mathematical knowledge has always been considered as part of the fundamental skills that adults need to possess to function well and be able to accomplish various goals in their everyday, work, and social life. Societies now present increasing amounts and wider range of information of a quantitative nature to citizens from all walks of life, in diverse contexts such as regarding health risk factors, school performance, or financial planning and insurance purchasing, to name just a few. As workplaces are becoming more concerned with involving all workers in improving efficiency and quality, the importance of numeracy skills is growing. Numeracy-related skills have been shown to be a key factor in labor market participation, sometimes even more so than literacy skills. Adults with lower skills in numeracy and literacy are more likely to be unemployed or require social assistance. Further, some numeracy skills are deemed essential for post-secondary education in many areas, including but not limited to hard sciences, engineering and technology. (Jones, 1995; Murnane, Willett & Levy, 1995; Hoyles, Wolf, Molyneux-Hodson, & Kent, 2002; Coulombe, Tremblay, & Marchand, 2004; Desjardins, Murray, Clermont & Werquin, 2005).

b. ***Public policy in most countries includes separate investments in literacy and numeracy.*** The separate acquisition of skills in these two fundamental areas is emphasized throughout both primary and secondary school systems, and in adult education or nonformal learning schemes. Countries expect that investment in literacy and numeracy will increase citizens' ability to act independently towards their own progress and income security, thereby reducing future social expenditures as well as contributing to citizens' participation in economic and social life in an information-laden society (European Commission, 1996). Numeracy has been shown to be associated with future learning which is important for re-training and for upgrading of skills as industries evolve (Marr & Hagston, 2007). The demands for higher numeracy performance will affect the employability of the labor force.

c. ***The policy and program responses are different for numeracy than for literacy.*** Efforts to improve literacy and numeracy levels of specific population groups are not implemented via the same mechanisms—they often require different experts, resources, and learning systems because of differences in the underlying knowledge components and learning trajectories. It is vital that nations have information about their workers' and citizens' numeracy, independently of other competency areas, in order to evaluate the human capital available for advancement, to plan school-based and lifelong learning opportunities, and to better understand the factors that affect citizens' acquisition and usage of numeracy (Johnston, & MacGuire, 2005).

d. ***Numeracy skill levels are not measured well by literacy measures.*** It is not possible to represent the *numeracy* levels in a population via people's performance on literacy measures that examine how well people read, process, and comprehend various types of texts and documents, or communicate about such texts. As

> explained later in more detail, numeracy involves, among other things, the handling of arithmetical processes, understanding of proportions and probabilistic ideas, understanding of numerical, geometric and graphical types and representations of quantitative information, critical interpretation of statistical or mathematical messages, ability to solve various types of quantitative problems, and other elements or processes that bear little relation to what is subsumed by literacy measures. (Coben, 2000; Gal, van Groenestijn, Manly, Schmitt, & Tout, 2005)

85. It follows that a direct assessment of numeracy in PIAAC can provide policy makers and other stakeholders with a sound basis for evaluating the distribution of the actual numeracy competence in the adult population.

Conceptual and theoretical foundations

86. The conceptualization of 'numeracy' in an international context is a challenging undertaking. Like literacy, the term numeracy has multiple meanings across countries and languages. In some countries the term numeracy relates to basic skills which school children are expected to acquire as a prerequisite to learning formal mathematics at higher grades. In other countries the term numeracy encompasses a broad range of skills, knowledge and dispositions that adults should possess but it does not necessarily relate to formal schooling (Baker & Street, 1994; NRDC, 2006). Finally, some countries do not even have a word such as numeracy; therefore, as part of educational or policy-oriented discourse in such countries, experts or translators either had to invent a special new word for it (e.g., 'Numeratie' in Canada, 'Numeralitet' in Denmark), or use other phrases such as "mathematical literacy", "functional mathematics", or terms equivalent to "computational ability". Such diversity in terminology, or the lack of an accepted term with which policy-makers feel comfortable, can complicate the communication with and among policy makers interested in PIAAC.

87. The range of meanings attached to the term numeracy and the lack of an equivalent term across languages may create miscommunications or gaps in expectations regarding what will be measured by a numeracy scale in PIAAC. This can affect the perceived policy relevance of a numeracy scale. Thus, attention has to be given to making sure that discussions regarding numeracy assessment in PIAAC are based on a consensus about the scope of the term and recognition of its centrality in a wide range of adult life circumstances.

88. However, it must be remembered that what will be measured by a numeracy assessment scale is *jointly* determined by two interrelated factors – by (1) a conceptual scheme describing numeracy and its elements, and (2) by an assessment scheme describing how the general conceptualization of numeracy is operationalized and manifested in the nature and range of tasks used in the assessment scale and the mode of administration and scoring. Of course, the conceptualization and assessment of adult numeracy involves many questions, such as: What are the key numeracy tasks which adults have to face in their lives? What facets or sub-domains are subsumed under 'adult numeracy'? What are the differences and commonalities between numeracy and related key constructs such as Quantitative Literacy which was assessed in IALS or Mathematical Literacy which is assessed in PISA?

89. This part is organized in three sections. In section 1, the notion of "competence" as defined by OECD is outlined. In section 2, the contexts and situations in adults' lives which require numeracy are reviewed. Section 3 examines perspectives on the meaning of numeracy and prior definitions and conceptualizations of numeracy, leading to a definition of numeracy for PIAAC. Part 3 which follows further examines the dimensions of numerate behavior, including contexts, expected responses, content areas of mathematical information and ideas, and representations, as a way of operationalizing the numeracy construct for scale development. It also discusses enabling processes, both cognitive and noncognitive or dispositional, which underlie numerate behavior. Subsequent parts outline principles

for assessment of numeracy in PIAAC, and comment on differences and commonalities between PIAAC's numeracy and related constructs assessed in IALS and PISA.

Numeracy as a competence

90. The conceptualization of numeracy in an assessment program which focuses on "literacy competencies for the information age" has to be congruent with the broader notion of "competence". Within OECD, prior work on the Definition and Selection of Competencies project (DeSeCo; see Rychen & Salganic, 2003) has defined competence as "the ability to meet individual or social demands successfully, or to carry out an activity or task". The DeSeCo view, which was adopted by OECD and also informed the design of assessment scales for PISA, places at the forefront how individuals function in the face of external demands that may stem from a personal or social context of action. DeSeCo (2002: 8-9) conceptualizes competencies as internal mental structures, i.e., abilities, capacities or dispositions embedded in the individual:

> *Each competence is built on a combination of interrelated cognitive and practical skills, knowledge (including tacit knowledge), motivation, value orientation, attitudes, emotions, and other social and behavioral components that together can be mobilized for effective action. Although cognitive skills and the knowledge base are critical elements, it is important not to restrict attention to these components of a competence, but to include other aspects such as motivation and value orientation.*

91. Further, DeSeCo (2002:7) argues that the terms "skills" and "competencies" are not synonyms. Skills designate an "ability to perform complex motor and/or cognitive acts with ease, precision, and adaptability to changing conditions", while competence designates "a complex action system encompassing cognitive skills, attitudes and other non-cognitive components".

92. The conceptualization of numeracy discussed below, which is based on a review of scholarly literature and research findings, operates on two levels. It relates to numeracy as a construct describing a competence as defined above, and to numerate *behavior* which is the way a person's numeracy is manifested in the face of situations or contexts which have mathematical elements or carry information of a quantitative nature. In this way, inferences about a person's numeracy are possible through analysis of performance on assessment tasks designed to elicit numerate behavior. In congruence with the above view of a competence, numeracy will be described as comprised both of cognitive elements (i.e., various knowledge bases and skills) as well as non-cognitive or semi-cognitive elements (i.e., attitudes, beliefs, habits of mind, and other dispositions) which together shape a person's numerate behavior.

Contexts and demands for numeracy

93. Once a view of numeracy as a competence as defined above is adopted, a discussion of what is encompassed by numeracy (and numerate behavior) has to start by identifying the nature of the contexts which contain mathematical[2] elements, or which include information of a quantitative nature,

---

[2] The term "mathematical" is used here as inclusive of situations where *statistical* or *probabilistic* information may appear or where statistical thinking or statistical literacy are required as well. Such usage is made for brevity and convenience only. It is acknowledged that statistics is not a branch of mathematics, and that statistical reasoning and statistical literacy have unique elements, concepts and processes which are not mathematical in nature (Moore & Cobb, 2000).

that adults face and which pose demands with which they have to cope. This in turn provides the basis for describing the knowledge elements and supporting processes which enable adults to cope with real-world numeracy tasks (Ginsburg, Manly & Schmitt, 2006), and can later help to form a road map which can guide the design and selection of tasks for inclusion in the numeracy assessment in PIAAC.

94. The literature pertaining to the uses of numeracy in the real world can be divided into three strands: literature on the roles of literacy and numeracy in adults' lives, on the mathematical demands of workplace and functional settings, and on educational perspectives on mathematical needs of school graduates and citizens. These areas are certainly intertwined but also offer complementary ideas, hence each is reviewed separately below.

The roles of literacy and numeracy in adults' lives.

95. The purposes served by adults' numeracy may parallel those served by adults' literacy, and further, people's numeracy may at times relate to or even depend in part on literacy skills or other lifeskills. Work to describe the purposes served by adults' literacy and numeracy skills has been conducted in several countries. In Australia, for example, Kindler et al., (1996) reported on four such purposes: *literacy for self-expression, literacy for practical purposes, literacy for knowledge, and literacy for public debate*. In the USA, the National Institute for Literacy has sponsored efforts to define critical skill areas. As part of its Equipped for the Future initiative, four broad types of purposes were identified (Stein, 1995): *Literacy for access and orientation in the world, literacy as voice to one's ideas and opinions, literacy for independent action, solving problems and making decisions as a parent, citizen and worker, and literacy as a bridge to further learning and to keep up with a rapidly changing world.*

96. Work has also been done in adult education contexts to identify different purposes and functions of using mathematical knowledge. In Australia, for example, one key project (Kindler et al. 1996), pointed to four broad categories regarding the uses of numeracy: *Numeracy for practical purposes* addresses aspects of the physical world that involve designing, making, and measuring. *Numeracy for interpreting society* relates to interpreting and reflecting on numerical and graphical information in public documents and texts. *Numeracy for personal organization* focuses on the numeracy requirements for personal organizational matters involving money, time and travel. *Numeracy for knowledge* describes the mathematical skills needed for further study in mathematics, or other subjects with mathematical underpinnings or assumptions.

97. A scheme developed by Steen (1990), a noted mathematics educator, outlines five dimensions of numeracy:

- *Practical*, focused on mathematical and statistical knowledge and skills that can be put to immediate use to cope with tasks in daily life
- *Professional*, focused on the mathematical skills required in specific jobs
- *Civic*, focused on benefits to society
- *Recreational*, related to the role of mathematical ideas and processes in games, puzzles, sports, lotteries, and other leisure activities
- *Cultural*, concerned with mathematics as a universal part of human culture (and related to appreciation of mathematical aspects such as in cultural or artistic artifacts)

98. Overall, the purposes regarding literacy and numeracy appear to agree and suggest that adults need to be able to apply their numeracy and literacy skills to tasks with a social or personal purpose in both informal and more formal contexts (NRDC, 2006). Such perspectives supplement Bishop's (1988) proposal that there are six modes of mathematical actions that are common in all cultures and pertain both to children and adults: counting, locating, measuring, designing, playing and explaining.

Numeracy in the workplace and in functional settings.

99. Mathematical and statistical skills that are important in adults' work have been described in large-scale efforts to define "core skills" or "key competencies" that workers should have, usually in response to the need to maintain economic competitiveness and improve employability of adults and school graduates. In addition, several projects looked specifically at the mathematical skills of workers in a range of occupational groups or workplace clusters.

100. Basic computational knowledge has always been considered as part of the fundamental skills that adults need to possess, but recent skills frameworks claim that workers need to possess a much broader range of mathematical skills. Examples exist in many countries and the following selective description from the United States is indicative of the nature of such efforts. Following earlier research by a task force of the American Society of Training and Development (Carnevale, Gainer, & Meltzer, 1990), the U.S. Secretary of Labor's Commission on Achieving Necessary Skills (SCANS) (Packer, 1997) has differentiated between mastery of basic arithmetical skills and much broader and flexible understanding of principles and underlying ideas subsumed under the notion of mathematical skills (SCANS, 1991, p. 83):

> *SCANS arithmetical skills:* Performs basic computations; uses basic numerical concepts such as whole numbers and percentages in practical situations; makes reasonable estimates and arithmetic results without a calculator; and uses tables, graphs, diagrams and charts to obtain or convey quantitative information.
>
> *SCANS mathematical skills:* Approaches practical problems by choosing appropriately from a variety of mathematical techniques; uses quantitative data to construct logical explanations for real world situations; expresses mathematical ideas and concepts orally and in writing; and understands the role of chance in the occurrence and prediction of events.

101. Based on a later survey of employers, industry trainers, and educators, among others, Forman & Steen (1999) similarly argued that quantitative skills desired by employers are much broader than mere facility with the mechanics of addition, subtraction, multiplication, and division and familiarity with basic number facts; they also include some knowledge of statistics, probability, mental computation strategies, some grasp of proportional reasoning or modeling relationships, and broad problem-solving and communication skills about quantitative issues.

102. Work on mathematical skills and their use in specific workplaces has been conducted over the last decade in both the manufacturing and service sectors in several countries such as the UK, Australia, USA, and others (e.g., Buckingham, 1997; Bessot & Ridgway, 2000; Hoyles et al., 2002; Fitzsimons, 2005; Skills Australia, 2005). Overall, these studies complement the SCANS study and suggest that employees need to possess a range of specific mathematical skills or knowledge, such as the following key (but not the only) examples:

- skills in both fast and accurate computations but also estimation, and knowing when each skill is required and why

- ability to deal with proportions and percents
- understanding measurement concepts and procedures
- working with or creating simple formulas
- a sense for the use of models and modeling in foreseeing future needs
- understanding of basic statistical concepts and displays

103. In addition, on a broader and less technical level, these studies argue that workers need to be able to make decisions in the face of uncertainty in real situations, prioritize actions and make choices regarding the approach to handling different tasks, depending on changing external demands. As well, there is a need for workers to be able to communicate with other workers or clients or understand written documentation (e.g., through text or with tables, charts, and graphs) about issues such as quantities, schedules, variation over time, results of quantitative projections, or analysis of different courses of action in this regard. Such findings echo the earlier distinctions made by the SCANS analysis between the need to attend both to basic arithmetical skills and more elaborate and complex mathematical skills in the workplace, but also highlight some areas where specific literacy and communication skills are intertwined with numeracy skills.

104. An important research literature has also accumulated over the last decades regarding the ways in which people use mathematical skills or cope with mathematical tasks in both formal (i.e., school-based) and informal (i.e., everyday, workplace) contexts (e.g., Rogoff & Lave, 1984; Resnick, 1987; Saxe, 1988; Carraher, Schliemann, & Carraher, 1988; Scribner & Sachs, 1991; Nunes, 1992; Presmeg, 2007). While too complex to discuss in detail here (see Greeno, 2003, for one of several reviews of this literature), among other things these studies highlight the situatedness of mathematical knowledge used in functional contexts and the need for actors in different contexts to develop situation-specific mathematical procedures and know-how. Further, numerous researchers (e.g., Straesser, 2003; Wedege, 2003; Williams & Wake, 2007) have argued, based on ethnographic analyses of workers' activities in diverse industries, that important portions of the mathematical activities at work are made "invisible" to occasional observers as well as to the workers themselves, or are disguised as nonmathematical. Various factors have been posited as causing this phenomenon, such as the encapsulation of many mathematical activities into routines or automated procedures; the use of tools and instruments or information technology (e.g., spreadsheets); the normative use of job-specific linguistic terms that are different than traditional school terms; or the division of labor among different workers.

105. Based on such and related findings, many projects have argued that mathematical skills as used in the workplace are often different and broader in scope than what is traditionally taught in school mathematics, but also take on different forms depending on the specific work context (Marr, & Hagston, 2007). Overall, the above suggests that what employers, training and employment specialists, and researchers know about the mathematical or statistical demands of different occupations may be incomplete. Further, projections of future skills demands usually focus on shifts in demand for *workers* with given skills, e.g., how many more engineers, technicians, or call-center operators will be required with a given mixture of currently-defined skills (Karoly, 2007), *not* on the changing future numeracy skills that workers will require.

106. Most sources discuss future skills required in occupational or job-market sectors, not to skills required in family, civic, or community contexts. Yet, with the accessibility of high-speed internet connections in homes, more adults are able to connect to, search, and make use of an increasing array of information sources, including many with quantitative components, such as regarding health,

personal finances, comparative shopping, sports, education, official statistics, and more. Many service organizations increasingly open more opportunities for customers not only to *access* information, but also *interact and take action*, through Internet websites and other technology-based devices (e.g., ATM machines, 'smart' cellular phones, GPS navigation systems), thus involving adults in new types of transactions and activities. These changes are important because they blur the distinction between "work" and "non-work" situations and their skill demands in the area of numeracy.

107. Given the above, the conceptualization of numeracy for PIAAC was derived with reference to the types of numeracy demands as depicted earlier in this subsection. Further, a working assumption has been made that it is not feasible to employ assessment items that are very workplace-specific (e.g., couched in the context of a single workplace or occupation) because mathematics or statistics as used in this context may not be visible or familiar to most other adults (Hoyles et al., 2002).

Educational perspectives on numeracy and informed civic participation.

108. A growing dialogue about the goals and impact of mathematics education in schools has intensified in recent years. This is in part due to economic pressures and industry expectations on the one hand, but also due to the realization that mathematical knowledge and skills serve multiple and separate gateway functions on the other hand. Specifically, mathematical competencies affect chances of entry into key occupational tracks (mainly in science, technology, and economics) and may affect employability and labor-force participation, underlie some important aspects of civic participation, and may impact on the possibilities of certain population groups for social equality and mobility. While the dialogue about these issues admittedly overlaps to some extent the points raised earlier in discussing the roles of literacy and numeracy in society, it is worth elaborating upon because it brings forward some additional points and broadens the understanding of contexts where demands on adults' numeracy exist.

109. Various arguments have been forwarded over the last few decades to support a broadening of the conceptions regarding the mathematical skills and knowledge that school graduates should possess, and the ways in which learned knowledge serves adults (Ernest, 2004). Educators working both with school students and adults increasingly aim to assist learners in developing mathematical concepts and skills in ways that are personally meaningful but also functional. Such approaches usually assume that there is often more than one right way to cope with a real-world functional task, and that adults require access to a repertoire of strategies for solving functional problems. Adults' personal methods of using mathematics are encouraged and valued. This is often a significant difference from traditional (pre-reform) school-based mathematics teaching, within which school students were often expected to solve a problem following the one correct method or algorithm, introduced by the teacher.

110. Several decades ago, ideas already began emerging in different countries that since mathematics is an essential aspect of society, mathematics education in schools should be derived from or prepare learners for broad real-life situations in family, work, community, and other contexts (National Council of Teachers of Mathematic, 2000; Willis, 1990), beyond employers' desire to focus mostly on practical or job-specific numeracy skills. Two early influential examples are the recommendations of the Cockcroft Committee in the UK (Department of Education and Science/Welsh Office, 1982), and Freudenthal's work in the Netherlands which has led to the Realistic Mathematics Education movement (Heuvel-Panhuizen & Gravemeijer, 1991). Over the last two decades, various countries (e.g., Australia, UK) have adopted adult education frameworks which give explicit attention to numeracy skills. For example, in the UK, the Adult Numeracy Core Curriculum aims to move learners through up to five levels of demand, where the expectation is that at the highest one (level 2), an adult:

> *Understands mathematical information used for different purposes and can independently select and compare relevant information from a variety of graphical, numerical and written material – e.g. compare data using mean and median, work out discounts as fractions and percentages of amounts, work out distances and lengths from scale drawings* (Gillespie, 2007: p.4).

111. Educators have also paid much attention to the importance of quantitative literacy in civic and social contexts, and argued that mathematical knowledge is a crucial part of a common fabric of communication indispensable for modern civilized society, in part because it is the language of science and technology. Thus, it has been claimed (National Research Council, 1989) that understanding of public discussions and reports about socially important topics such as health and environmental issues is impossible without using the language of mathematics.

112. Further, it has been claimed that in a society in which the media constantly present information in numerical or graphical form to all citizens, the ability to interpret quantitative and statistical messages is vital for all adults (Paulus, 1995; Steen, 1997). It is essential for all adults to possess the ability to critically reflect on quantitative information encountered in various media sources and documents (Frankenstein, 1989), and to understand how to be a careful or critical consumer of statistical arguments of various kinds (Gal, 2002; Utts, 2003; Watson & Callingham, 2003).

113. Indeed, the dialogue about the various demands on adults' knowledge has been reflected in part in the emphasis in PISA on the assessment of mathematical literacy and science literacy. Such constructs pertain, broadly speaking, to school students' readiness for entering adults' life contexts; it is indicative that they have been chosen to be the focus of assessment rather than more traditional notions of formal knowledge in mathematics or science areas which were assessed primarily in earlier studies.

More on numeracy situations and demands.

114. The discussion above suggests that numeracy is required so that people can effectively cope with or respond to a range of situations that are embedded in a life stream with real, personal meaning to them. The situations that call for activation of the numeracy competency can be situated in a hypothetical "numeracy task space", defined by dimensions such as the nature of the required response, the number and characteristics of the quantitative elements in the situation, or the extent and nature of literacy processes involved. Based on Gal (2000), below are described three key types of situations which illustrate the range of numeracy demands placed on adults.

1. ***Generative situations*** demand that people count, quantify, compute, or otherwise manipulate numbers, concrete objects, visual elements, and so forth, to create/generate new numbers or estimates. Examples are calculating the total price of products while shopping, finding the number of boxes in a crate, measuring the area of a room to be painted in order to calculate the amount of materials needed to do the job, reading a menu and computing the cost of a specified meal, filling out an order form for a product, figuring out travel times between train stations based on a timetable, and so forth. Generative situations include computational or quantitative literacy tasks (Kirsch et al, 1993), but certainly go beyond them, such as when measurements have to be made regarding length, volume, time, etc.

The numerical information in many types of generative situations may be evident in the situation itself (e.g., real objects to be arranged, sorted, counted, or measured; a graph on a computer display). Yet, numerical information may also be communicated through text or be embedded in different types of text; hence, such situations may also involve language skills to varying degrees. In generative situations, tools such as a hand-held calculator, a computer-based application, or a measuring tape or ruler may ease the mechanics of performing needed calculations or increase accuracy, although a person may choose not to use them. Even when such tools are used, a person still needs to know how to use them efficiently and effectively, and his or her operations are at times more likely to be correct or accurate if the person has alternative strategies (e.g., mental calculations) to check on accuracy and completeness of actions performed.

2. ***Interpretive situations*** demand that people make sense, and grasp the implications, of messages that contain information of a mathematical or statistical nature but that *do not involve direct manipulation of numbers*. An example is being faced, when reading the newspaper, with a report of results from a recent opinion poll based on a small sample, and having to decide if to take as valid a generalization made by the writer about differences between two populations. Paulus (1995) describes many mathematical statements made in the media that call for careful consideration of their validity; other examples can be added where references to proportions, averages, samples, bias, correlation, risk, or causality are discussed or implied, such as in the context of genetic or medical counseling, or understanding of statistical process control displays.

   In simple interpretive situations, such as when one has to read nutrition information or a drug label to decide if a product contains certain ingredients (e.g., sugar, allergens, contaminants) whose dosage exceeds an allowed limit, the response can be assessed as correct or incorrect. Yet in more complex interpretive situations, the response expected is usually the creation of an *opinion*, and to create this opinion one needs to invoke and answer a set of critical before the information or arguments presented are accepted as credible or valid. In such cases, the response, i.e., opinion, has to be judged in terms of its reasonableness and the quality of the arguments or evidence on which it is based. Thus an opinion cannot always be easily classified as “right/wrong” or "accurate/inaccurate", as with responses to many generative situations.

3. ***Decision situations*** demand that people locate and consider multiple pieces of information in order to determine a course of action, typically in the presence of conflicting goals, constraints, or uncertainty. Two key subtypes here are: *Optimization tasks*, which require the identification of optimal ways to use resources such as money or supplies, or schedule personnel or time (see SCANS, 1991); and *Choice tasks*, which require a choice among alternatives, such as which of several apartments to rent, which pension or

> health insurance plan to join, whether to undergo a surgical medical procedure which has known probabilities of certain side effects.
>
> It is important to note that optimization and choice tasks can be part of a broader *problem-solving process*, where alternatives have to be generated and then evaluated. Thus, what is being termed here a *decision* situation can at times also be viewed as a *problem-solving* situation.
>
> As with interpretive situations, a response to a decision situation has to be assessed in terms of its reasonableness, because the response may be based on multiple pieces of quantitative information (e.g., timetables, financial figures, statistical trends, event probabilities), the response may be created through a process that involves both generative and interpretive steps, and the response is shaped after a person evaluates the quality of the final decision or choice against external contextual criteria. An example is when a small business owner has to compare financial information from several banks to decide which loan schedule is the best or most manageable, given the current financial situation and anticipated costs and revenues of the business. The assessment of the reasonableness of a response in a decision situation may be further complicated, beyond what happens in an interpretive situation, because the response in a decision situation may also be shaped by a person's subjective preferences and value system, assumptions he or she makes about future trends or event probabilities, and other factors.

115. To be sure, the three types of numeracy situations described above are not mutually exclusive, and other cases may exist, possibly of a hybrid nature. Further, in considering the implications of these and other types of situations for the numeracy competency required of adults, it is important to keep in mind the impact of evolving technologies. As has been argued and documented by many sources (Expert Group on Future Skills Needs, 2007; Gatta, Appelbaum, & Boushey, 2007; Karoly, 2007), and summarized in the PIAAC planning documents, adults are presented with ever-increasing amounts of information of a quantitative nature through Internet-based or technology-based resources. More so than in prior decades, more types of quantitative information are more readily available, but this information has to be located, selected or filtered, interpreted, at times questioned and doubted, and analyzed for its relevance to the responses needed, whether generative, interpretative, or decision-oriented.

Towards a definition of numeracy for PIAAC

116. Reaching a consensus on a definition of numeracy that can fit an international program of assessment is a challenging undertaking. First, as noted above there are various country-specific connotations for numeracy, if such a term at all exists in a local language. Second, there are overlapping or competing constructs such as quantitative literacy, mathematical literacy, functional mathematics, and so forth (Hagedorn, Newlands, Blayney, & Bowles, 2003). Third, an attempt to discuss the definition and meaning of numeracy is complicated by the fact different stakeholders already view it from within a given lens imposed by the historical and cultural aspects, whether organizational, social, economic, or linguistic, of the systems within which they operate. For example, some of the existing conceptions of numeracy were developed by educators working in delivery systems for school children, while other stakeholders link the term numeracy only to adult-related competencies.

117. With the above in mind, and using the conceptions of competence and of the contexts for numeracy presented earlier as a backdrop, the remainder of this section is organized as follows: First, a review of some of the many perspectives on numeracy is presented so as to portray the key ideas that prior workers and scholars have addressed when discussing numeracy. Next, a definition of numeracy for PIAAC is presented, followed by a discussion of the facets of numerate behavior, including key content areas of mathematical knowledge and other cognitive and non-cognitive enabling processes and factors which take part in or affect numerate behavior. Finally, differences between numeracy and the constructs assessed in PISA and IALS are examined, in particular mathematical literacy and quantitative literacy.

Perspectives on [adult] numeracy.

118. Formulation of what numeracy encompasses have evolved since the term was introduced in the 1959 Crowther Report in England and Wales. Maguire and O'Donoghue (2003) have recently reviewed and organized conceptions of numeracy from several countries (Ireland, Canada, USA, UK, the Netherlands, Denmark and Australia) along a continuum of increasing levels of complexity or sophistication. *Formative* conceptions view numeracy as related to basic arithmetic skills. *Mathematical* conceptions consider numeracy in a contextualized way, as a broader set of mathematical knowledge and skills (beyond basic computations) of relevance in everyday life. Finally, *integrative* conceptions consider numeracy as a multifaceted, sophisticated construct incorporating not only mathematics but also communicative, cultural, social, emotional, and personal elements which interact and pertain to how different people function in their social contexts. (Coben, 2000; Condelli, Safford-Ramus, Sherman, Coben, Gal, & Hector-Mason, 2006).

119. At this time, formative conceptions which view numeracy as basic computational facility are often associated with how numeracy is viewed in connection with goals of primary schooling, and reflected in how numeracy is defined when classifying literacy/numeracy levels worldwide (UNESCO, 1997). Most extant conceptions which adult education, workplace training, and national and international assessments have adopted fall at different points across the mathematical and integrative phases described by Maguire and O'Donoghue. Below are four different but related views of numeracy, the first pair from the UK, and the second pair from Australia. These definitions illustrate that conceptions evolve over time and that variability can be noticed even within the same national system.

> *[numeracy is]…an 'at-homeness' with numbers and an ability to make use of mathematical skills which enables an individual to cope with the practical mathematical demands of his everyday life…[and] an ability to have some appreciation and understanding of information, which is presented in mathematical terms, for instance graphs, charts or tables or by reference to percentage increase or decrease* (From the Cockcroft report: Department of Education and Science/Welsh Office (1982), p. 11).
>
> *The ability to use mathematics at a level necessary to function at work and in society in general…understand and use mathematical information; calculate and manipulate mathematical information; interpret results and communicate mathematical information* (From the UK government's Skills for Life strategy to improve standards of adult literacy and numeracy in England, DfEE, 2001, p.3).

> *Numeracy is the mathematics for effective functioning in one's group and community, and the capacity to use these skills to further one's own development and of one's community* (Beazley committee, 1984, Australia).
>
> *Numeracy involves abilities that include interpreting, applying and communicating mathematical information in commonly encountered situations to enable full, critical and effective participation in a wide range of life roles* (Queensland Department of Education, 1994, Australia).

120. An interesting case study of defining numeracy is offered by Lindenskov and Wedege (2001). Based on their work in adult and mathematics education in Denmark, they have imported numeracy from English-speaking countries and introduced a new term, Numeralitet, with a conceptual framework that was later adopted by the Danish Ministry of Education. According to this perspective, it is essential to distinguish between what numeracy is, or ought to be, from the individual's and from society's points of view. Lindenskov and Wedege (2001) advocate a societal view, whereby numeracy is seen as a competence that involves a dynamic interaction between functional mathematical skills and conceptions and operations on the one hand, and a series of activities and various types of data and media on the other. They argue that this skill- and activity-based view should be coupled with the understanding that in principle all people need to have this competence, and that numeracy is a competence determined by society and technology and that it changes in time and space along with social change and technological development.

121. Other views of numeracy, usually developed by adult education experts, focus on the role of adults as reflective communicators and critical consumers of information in society who are involved in the exchange and interpretation of messages encountered in media or in political and community contexts (Frankenstein, 1989). Johnston (1994) argues:

> *"To be numerate is more than being able to manipulate numbers, or even being able to 'succeed' in school or university mathematics. Numeracy is a critical awareness which builds bridges between mathematics and the real-world, with all its diversity"* (Johnston, 1994).

122. The definition quoted from the UK's Cockcroft Committee (1982) has been quite influential in that its conception of numeracy implied it is an ability to cope with various functional tasks in real-world contexts as well as interpretive tasks, but also pointed to the centrality of underlying supporting noncognitive components. These key ideas are reflected, albeit with different terminologies and foci, in other views of numeracy. Another important commonality is the presence of mathematical elements or ideas in real situations, and the notion that these can be used or addressed by a person in a goal-oriented way, dependent on the needs of the individual within the given context, i.e., home, community, workplace, societal action, etc.

Numeracy-literacy connections.

123. Several scholars and projects have pointed to the need to consider *literacy* when discussing numeracy, as the two are related and can affect each other (Baker & Street, 1994). Examples are when quantities are described in words and not in numbers, or appear within surrounding text whose interpretation is essential to understanding what is being required in terms of computations, or when there is a need to understand the mathematical relationships described in simple phrases, e.g., realizing that "four more than" is a different relationship than "four times as much."

124. An important aspect of the literacy-numeracy linkage has been acknowledged early on by the Kirsch & Mosenthal construction of literacy as comprised of Prose, Document, and Quantitative dimensions (Kirsch, Jungblut, & Mosenthal, 1998). Accordingly, adults skills in dealing with arithmetical operations embedded in text have been assessed by the Quantitative Literacy scale in IALS and several prior national studies. However, other areas where literacy and numeracy are linked do exist and need to be recognized, such as in the context of interpretation of statistical arguments in media articles (Gal, 2002a), or comprehending financial information which pertains to planning one's retirement pensions or medical benefits. Thus, while it is possible to define numeracy in general terms without invoking literacy, as all the definitions quoted above have done, the structure of the tasks and demands in adults' lives shows that these areas cannot be considered as mutually exclusive. Mathematical or statistical information is carried by or embedded in text in some, but certainly not all, contexts in which adults have to function. To the extent this happens, one's performance on numeracy tasks will depend not only on formal mathematical or statistical knowledge but possibly also on literacy-related factors such as vocabulary, reading comprehension, reading strategies, or prior literacy experiences.

Numeracy and numerate behavior in the ALL survey.

125. The above ideas informed the conceptualization of numeracy for the Adult Literacy and Lifeskills (ALL) survey, which was developed in 1998-2000 by an international team (Gal, van Groenestijn, Manly, Schmitt, & Tout, 2005). This was the first time the construct of numeracy had to be defined in a comparative assessment context and not purely in an educational context. Cognizant of the complexity and multi-faceted nature of the numeracy construct, the ALL team developed a three-tier conceptualization which attempted to reflect key perspectives of numeracy on the one hand, but also enable operationalization of the construct in an assessment scale on the other (Tout, 2006). The three tiers are a brief definition of numeracy, a more elaborate definition of numerate behavior, both presented below, and a detailed listing of components of the facets of numerate behavior (see Gal et al., 2005).

> Numeracy *is the knowledge and skills required to effectively manage and respond to the mathematical demands of diverse situations.*
>
> Numerate behavior *is observed when people manage a situation or solve a problem in a real context; it involves responding to information about mathematical ideas that may be represented in a range of ways; it requires the activation of a range of enabling knowledge, factors, and processes.*

126. Both the brief and elaborate definitions shown above were seen by the ALL numeracy team to be required, given the needs of a comparative assessment. A brief definition is essential to simplify communication with various stakeholders, such as policy-makers and experts. However, as with most brief definitions of complex constructs, the language used is broad and abstract, hence the definition cannot be explicit about what a numerate person can do and what behavior to observe in an assessment. With this in mind, the more detailed definition of numerate behavior was developed as a way to emphasize four key facets or dimensions which were seen by the ALL numeracy team as underlying numerate behavior, as follows:

- *Contexts*: The range of external demands (e.g., work, home, etc.),
- *Responses*: What a person can do in response to the external demands (e.g., compute, interpret, communicate, etc.),

- *Mathematical ideas/content*: The informational content which the context carries or enables access to, which can be seen as mathematical (or statistical) in nature and hence of interest in the context of an assessment of numeracy (e.g., numbers, proportions, measurements, statistical concepts, etc).

- *Representations*: The different ways in which the mathematical (or statistical) information exists or is conveyed to the person in the given context (e.g., text, numbers, graphs, etc.)

127. The advantage of using a more elaborate definition of numerate behavior was that it is more explicit about what to examine in an assessment, and thus serves as a springboard for developing an actual specification for an assessment scale. It is important to also note that the definition of numerate behavior points to the presence of both cognitive and noncognitive factors which underlie or enable effective numerate behavior. Ideally, coverage of both cognitive and non-cognitive aspects of numerate behavior is essential in order to generate a full picture regarding the competence of numeracy.

Definition of numeracy for PIAAC.

128. The development of the conceptualization and definition of numeracy for PIAAC went through several stages of work and consultation. An expert panel appointed to develop the overall assessment design for PIAAC presented in summer 2006 tentative recommendations regarding all competencies to be assessed in PIAAC (OECD, 2006) and then proposed to define numeracy as: *"The ability to use, apply, and communicate mathematical information".* Various perspectives on numeracy and its assessment were later examined by participants at the Canada-OECD Expert Technical Workshop on Numeracy, which met in November 2006 in Ottawa; a tentative working definition of numeracy was then proposed for PIAAC and included in a draft framework circulated for external review (Gal, 2007). Further development of the numeracy framework has been undertaken by the Numeracy Expert Group for PIAAC appointed in April 2008, which released a revised framework for review by all participating countries in October 2008.

129. In general, work on the development of a numeracy framework for PIAAC, together with the assessment scale and related item pool, has been conducted with two somewhat conflicting objectives in mind. One objective is the need to maintain compatibility with the conceptualization of numeracy in the ALL survey, given the need for PIAAC to provide trend data related to ALL results. For this reason PIAAC was designed with a specification that 60% of the literacy and numeracy tasks that will be employed in the final assessment scale will come from item pools used in ALL and IALS. The other objective is the need to extend the ALL definition in light of PIAAC's overarching conceptualization of “literacy competencies in the information age", and consider new or emerging uses of numeracy in the adult world.

130. Taking all the above into consideration, numeracy has been defined for PIAAC 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*

131. This definition captures essential elements in numerous conceptualizations of numeracy in the extant literature; it is compatible with the definition used for ALL and appears to provide a solid basis from which to develop an assessment scale for PIAAC with its emphasis on competencies in the information age. The inclusion of "engage" in the definition signals that not only cognitive skills but also dispositional elements, i.e., beliefs and attitudes, are necessary for effective and active coping with numeracy situations. It is also important to note that while the definition of numeracy for PIAAC has been developed in the context of an assessment programme, it has been crafted so as to contribute to

public dialogue regarding the goal of educational and social interventions focused on developing adult competencies in general, and adult numeracy and related mathematical and statistical skills and dispositions in particular.

132. However, since numeracy is a broad, multifaceted construct referring to a complex competency, the definition of numeracy given above should not be considered by itself, but should be coupled with a more detailed definition of *numerate behavior* and with further specification of the facets of numerate behavior. This pairing is essential in order to enable operationalization of the construct of numeracy in an actual assessment, thereby contributing to the assessment's validity and interpretability, and in order to further broaden the understanding of key terms appearing in the definition itself. Consequently, a definition of numerate behavior similar in general terms to the one used for the ALL survey, but shorter, has been adopted for PIAAC: **Numerate Behavior** *involves managing a situation or solving a problem in a real context, by responding to mathematical content/information/ideas represented in multiple ways.*

Table 1: Numerate behavior – key facets and their components

Numerate behavior 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 / analyze
    - communicate
3. **to mathematical content/ information/ ideas:**
    - quantity & number
    - dimension & shape
    - pattern, relationships, change
    - data & chance
4. **represented in multiple ways :**
    - objects & pictures
    - numbers & mathematical symbols
    - formulae
    - diagrams & maps, graphs, tables
    - texts
    - technology-based displays

Numerate behavior 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 & attitudes
- numeracy-related practices and experience
- context/world knowledge

133. The definition of numerate behavior pertains to four facets: *Contexts, Responses, Mathematical content/information/ideas, Representations*. Table 1 lists the components of the four facets, and these components are explained in more detail the next section. Table 1 is based on the original description of the facets of numerate behavior developed for the ALL survey, but some changes have been implemented, such as the addition of "access" and of "evaluate/analyze" as possible responses, the merging of the content categories of "change" and "pattern and relationship", or the reference to "technology-based displays" as another representation mode.

134. It should be noted that the bottom part of Table 1 also lists several enabling factors and processes elaborated in section 3.5, whose activation underlies numerate behavior. Most of these enabling factors and processes appeared in the ALL conceptual framework, but some changes were introduced, such as the positioning of "adaptive reasoning and mathematical problem-solving" as a separate enabling factor. Overall, the definition of numerate behavior presented earlier, together with the details in Table 1 and the further explanations below, provided a roadmap for the development of a numeracy scale for PIAAC.

Facets of numeracy

Facets of numerate behavior

135. This section[3] elaborates on the facets of numerate behavior and their components, listed in Table 1. The discussion of the first facet, context, mainly revisits ideas which appeared in section 2.2 above regarding contexts and demands for numeracy. Elaborations on the other facets are based on materials and ideas in the bibliographic sources mentioned earlier, and the analysis of the components of adult numeracy by Ginsburg et al. (2006) which was based on an integrative review of multiple numeracy frameworks from several countries (see also Hagedorn et al, 2003). This section has also benefited from the positions presented in a report of the UK's National Research and Development Centre for Adult Literacy and Numeracy (NRDC, 2006), background papers prepared for the OECD-Canada Expert workshop on numeracy (Nov. 2006, Ottawa) and suggestions made by workshop participants, external reviews of earlier drafts of this framework, and professional perspectives of PIAAC's Numeracy Expert Group.

Facet 1: Contexts.

136. People try to manage or respond to a numeracy situation because they want to satisfy a purpose or reach a goal. Four types of contexts where demands on people's numeracy may appear are described below. These are not mutually exclusive and may involve the same underlying mathematical themes.

a. **Everyday life.** The numeracy tasks that occur in everyday situations are often encountered in personal and family life, or revolve around hobbies, personal development, and interests. Representative tasks are handling money and budgets, comparison shopping, personal time management, making decisions involving travel, planning holidays, mathematics involved in hobbies like quilting or wood-

---

[3] Selected text portions in section 3.1, as well as in subsections 3.1.3 – 3.1.5, were adapted and expanded from the numeracy framework developed in 1998-2000 for the ALL survey, and used here with permission of Statistics Canada. The final version of the ALL numeracy framework is presented by Gal, van Groenestijn, Manly, Schmitt, & Tout (2005).

working, playing games of chance, understanding sports scoring and statistics, reading maps, and using measurements in home situations such as cooking or home repairs.

b. **Work-related.** At work, one is confronted with quantitative situations that often are more specialized than those seen in everyday life. In this context, people may develop good skills in managing situations that might be narrower in their application of mathematical themes. Representative tasks are completing purchase orders, totaling receipts, calculating change, managing schedules, budgets, and project resources, using spreadsheets, organizing and packing different shaped goods, completing and interpreting control charts, making and recording measurements, reading blueprints, tracking expenditures, predicting costs, and applying formulas.

c. **Societal or community.** Adults need to know about trends and processes happening in the world around them (e.g., regarding crime, health issues, wages, pollution) and may have to take part in social events or community action. This requires that adults can read and interpret quantitative information presented in the media, including statistical messages and graphs. Also, they may have to manage situations like organizing a fund-raiser, realizing the fiscal effect of community programs, or interpreting the results of a study about a health issue.

d. **Further learning.** A numeracy competency may enable a person to participate in further study, whether for academic purposes or as part of vocational training. In either case, it is important to be able to know some of the more formal aspects of mathematics that involve symbols, rules, and formulas and to understand some of the conventions used to apply mathematical rules and principles.

137. It should be emphasized that performance in all of the above contexts is based on a combination of cognitive and non-cognitive elements, and thus requires that we think of numeracy as a competence as defined earlier, not just as possession of a set of technical skills or know-how. For example, engagement in further learning of mathematical topics, whether in formal or informal contexts, involves willingness to start such learning in the first place as well as perseverance in such learning. For such engagement to occur, an adult is required to have positive beliefs and attitudes about mathematics and about oneself as a person capable to cope with mathematical tasks.

Facet 2: Responses.

138. In different types of real-life situations, people may have to react with diverse types of responses, grouped below under three broad headings: Identify, locate, or access; Act upon or use; and Interpret, evaluate/analyze, communicate. It should be noted that while these types of responses are described separately, in real life they may co-occur in a dynamic fashion, and may vary from simple to more complex. The extent to which response types will vary in complexity or co-occur in an integrated way depends on various aspects of the situation at hand discussed later, such as the density of the information available, presence of distracting information, transparency of the task, literacy demands, number of steps and iterations involved. Further, responses are shaped by the interaction between

situational demands on the one hand and the goals, skills, dispositions, and prior practices and experiences of the person on the other hand.

a. **Identify, locate, or access**. In virtually all situations, people have to identify, locate or access some mathematical information present in the task or situation confronting them that is relevant to their purpose or goal. When it exists alone, this response type often requires only low level mathematical understanding or application of simple arithmetic skills. Usually, however, this response type is subsumed or co-occurred with the other types of responses listed below.

b. **Act upon or use.** In situations described earlier (see 3.2.4) as "generative", people have to perform actions on the mathematical information which can be identified in the situation, or use known mathematical procedures and rules. Acting upon or using encompasses arithmetical operations such as when counting, doing calculations "in the head", with pen and paper or with a calculator. Acting upon or using may also involve ordering or sorting, estimating, figuring out an area or volume of a certain object in an approximate way, or using various measuring devices to generate needed mathematical information of a more exact nature. Finally, acting upon may involve using (or developing) a formula which serves as a model of a situation or a process.

c. **Interpret, evaluate/analyze, communicate.** This response type encompasses three separate but related responses, described below in more detail:

Interpret. Some situations do not demand any direct manipulation or action on available quantitative information, but the interpretation of the meaning and implications of given information of a mathematical or statistical nature. Further, in such situations, described earlier as "interpretive", the person in the situation may need to not only interpret mathematical or statistical information but also make a judgment or create an opinion, such as about trends, changes, or differences described in a graph or in a text appearing in a newspaper article or advertisement. It should be emphasized that interpretive responses may be in reaction not to information that is numerical (i.e., figures or statistical data), but to broader mathematical or statistical concepts possibly expressed in oral, textual, or visual manner, including ideas such as rate of change, proportions, shape of distribution, samples, bias, correlation, probability or risk, or causality.

Evaluate/analyze. This response category is in part an extension of the Interpret response type described above, which was used in ALL. It accommodates responses that may be more likely in situations requiring a person to analyze a problem and in so doing evaluate the quality of the solution against some criteria or contextual demands, and if needed cycle again through the interpretation, analysis and evaluation stages. Such situations may be encountered in various contexts, including in dynamic or information-rich technology environments, or those termed earlier 'decision situations'. Examples are when, as part of developing a solution to demands of a given situation, an adult has to process raw quantitative information through technology-enabled channels (e.g., sift through a website of a national statistical office), or retrieve and integrate information from multiple sources after evaluating their relevance to the task at hand (e.g., compare information from different sources regarding costs of competing courses of action).

A common feature that pertains both to "interpret" and "evaluate/analyze" responses is that the judgments or opinions expected in the situation may need to be critical in nature. Examples are when a person needs to question the validity of the data or information presented, identify gaps between the available information and the conclusions presented by a source (e.g., a journalist or politician), or reflect about the proposed implications of the data, both for himself or herself as an individual or for the wider community. As a result, these two responses types have an inherent overlap between them.

Communicate. In addition to the responses listed above, a person may have to represent and communicate about the mathematical information given, describe the results of one’s actions or interpretations to someone else, or explain and justify the logic of one’s analysis or evaluation. This can be done via oral or written means (ranging from presenting a simple number, a word, all the way to a detailed explanation), through a drawing (a diagram, map, graph) or by generating a computer-based display (e.g., by referencing a spreadsheet-based chart showing the results of “what if” scenarios), and various combinations of these and other modes of communication and illustration.

139. Problem-solving is not seen as a separate response type, but rather assumed to be part of the demands set forth by the external situation - as is implied in table 1, the goal of numerate behavior is managing a situation involving a numeracy task or solving a numeracy-related problem. Hence, several response types described above may be called upon and co-occur when people have to solve numeracy-related problems, especially novel ones. Such responses may be aided or organized by more generalized skills of adaptive reasoning and problem-solving, examined later on as part of "enabling processes" that underlie numerate behavior.

140. Note about assessment of certain response types. The ideas above describe key ways in which people may respond to mathematical/statistical tasks embedded in a range of real-life situations. However, one needs to distinguish between a conceptual framework (discussed in this section) and an assessment framework (discussed later on). Not all real life numeracy tasks can necessarily be simulated well in a specific assessment. Further, the ability of an assessment to actually *capture, evaluate, and score* responses associated with numerate behavior ultimately depends on the technical aspects of that assessment. While the computer-based assessment platform chosen for PIAAC offers many advantages, the design specifications adopted by participating countries pose limitations on skills assessments in PIAAC, due to the computer-based environment employed for the assessment, the need for immediate scoring of responses given the adaptive testing process necessary for efficient ability estimation, and restrictions on testing time per respondent which are typical in large scale household surveys. These realities necessitate the use of short separate tasks, exclude extended problem tasks, and prohibit the use of most types of numeracy tasks that require respondents to communicate via free-form text input. Specifically, tasks requiring communication-based responses, such as when adults have to explain interpretations of given information, or describe their evaluation or analysis of a situation *or their thinking about that situation*, could hardly be used in the direct assessment of all skills targeted by PIAAC. Such tasks do comprise an important, inseparable part of the landscape of adult numeracy situations and are an inherent part of the conceptual framework of adult numeracy, yet very few could be included in the item pool for the first cycle of PIAAC.

Mathematical content/information/ideas.

141. Mathematical information can be classified in several ways and on different levels of abstraction. One approach is to refer to fundamental "big ideas" in mathematics. Steen (1990), for example, identified six broad categories: Quantity, Dimension, Pattern, Shape, Uncertainty, and Change. Rutherford & Ahlgren (1990) described networks of related ideas: Numbers, Shapes, Uncertainty, Summarizing data, Sampling, and Reasoning. Dossey (1997) categorized the mathematical behaviors of quantitative literacy as: Data representation and interpretation, Number and operation sense, Measurement, Variables and relations, Geometric shapes and spatial visualization, and Chance. More broadly, many curriculum frameworks around the world in one way or another refer to these key areas, albeit using somewhat different terminologies and with somewhat different groupings (e.g., NCTM, 2000).

142. Based on such and related classifications, the ALL numeracy framework (Gal et al., 2005) defined five areas of mathematical content and ideas that characterize the mathematical demands faced by adults: Quantity & number; Dimension & shape; Data & chance; Pattern, functions, & relationships; Change. In PIAAC the first three areas were retained, while the last two areas listed for ALL were united to create a single area called "Patterns, relationships and change". Such a change was deemed sensible because these two areas are sufficiently related. Indeed, Ginsburg et al, 2006, in their analysis of the components of numeracy, subsumed "change" as part of a unified domain termed "patterns, functions, and algebra." Uniting the two areas has to advantages: it enables PIAAC to focus on domains that are more distinct from each other, as well as maintains better conceptual compatibility with PISA, an issue discussed in Section 5.

143. With the above in mind, four key areas of mathematical content, information and ideas are covered by the numeracy assessment in PIAAC and are briefly summarized below.

a. **Quantity and Number.** *Quantity* is described by Fey (1990) as an outgrowth of people's need to quantify the world around us, using attributes such as: numbers of features or items; costs and charges for goods and services; size (e.g. length, area, and volume); temperature, humidity, and pressure of our atmosphere; populations and growth rates of species; revenues or profits of companies, etc. *Number* is fundamental to quantification and different types of number constrain quantification in various ways: whole numbers can serve as counters or estimators; fractions, decimals and percents as expressions of greater precision, parts or comparisons; and positive and negative numbers as directional indicators. In addition to quantification, numbers are used to put things in order and as identifiers (e.g., telephone numbers or zip codes). There is also the requirement to operate on such quantities and numbers (the four main operations of +, –, x, ÷ and others such as squaring). Facility with quantity, number, and operation on number requires a good "sense" of magnitude. Contextual judgment comes into play when deciding how precise one should be or which tool (calculator, mental math, a computer) to use. Money and time management, the ubiquitous mathematics that is part of every adult's life, depends on a good sense of number and quantity. A basic level numeracy task might be figuring out the cost of one can of soup, given the cost of 4 for $2.00; a task with a higher cognitive demand could involve more complex numbers such as when figuring out the cost when buying 0.283 kg of cheese at 12.95 Euros per kg.

b. **Dimension and shape**. *Dimension* includes “big ideas” related to one, two, and three dimensions of “things” (using spatial and numerical descriptions), projections, lengths, perimeters, areas, planes, surfaces, location, etc. Facility with each dimension requires a sense of "benchmarks" and estimation, direct measurement and derived measurement skills. *Shape* is a category describing real images and entities that can be visualized (e.g., houses and buildings, designs in art and craft, safety signs, packaging, snowflakes, knots, crystals, shadows and plants),in both two and three dimensions . Direction and location are fundamental qualities called upon when reading, interpreting or sketching maps and diagrams. This content area requires an understanding of units and systems of measurement, both informal and standardised such as the Metric and Imperial systems. A basic numeracy task in this fundamental aspect could be shape identification whereas a complex task might involve describing the change in the capacity of an object when one dimension is changed.

c. **Pattern, relationships, and change.** It is frequently written that mathematics is the study of patterns and relationships. *Pattern* is seen as a wide-ranging concept that covers patterns encountered all around us, such as those in musical forms, nature, traffic patterns, etc. It is argued by Senechal (1990) that our ability to recognize, interpret, and create patterns is the key to dealing with the world around us. The human capacity for analyzing and identifying patterns and relationships undergirds much mathematical thinking. *Relationships* and *change* relate to the mathematics of how things in the world are associated or develop. Individual organisms grow, populations vary over time, prices fluctuate, and objects traveling speed up and slow down. Some characteristics or values can change directly in proportion or relation to another change, whilst other characteristics may change in the opposite direction or in a different way. Change and rates of change help provide a narration of the world as time marches on. The ability to generalize and to characterize relationships between variables is a crucial gateway to understanding basic economic, political or social analyses. This domain includes the ability to develop and/or use a mathematical formula between the different variables involved in a situation, alongside the need to be able to understand, use and apply a sense of proportional reasoning. A lower level numeracy task may ask someone to use a familiar formula such as that for calculating the area of a square or rectangle. More demanding tasks involving relationships and change may require using formulae such as for calculating compound interest or one’s BMI (Body Mass Index). Or, tasks could require using an electronic spreadsheet or a Web-based dedicated calculator (applet) for exploring "what if" scenarios related to different interest rates, or to different levels of weight loss or weight gain, and their impact on one's long-range savings or health risk levels, respectively.

d. **Data and chance.** Data and chance encompass two related but separate topics. *Data* covers "big ideas" such as variability, sampling, error, or prediction, and related statistical topics such as data collection, data displays, and graphs. Modern society demands that adults interpret and produce organizers of data such as frequency tables, pie charts, graphs and to sort out relevant from irrelevant data. *Chance* covers "big ideas" related to probability, subjective probability, and relevant statistical methods. Few things in the world are 100% certain; thus the ability to attach a number that represents the likelihood of an instance is a valuable tool whether it has to do with the weather, the stock-market, or the decision to board a plane. In this mathematical category, a simple numeracy skill might be the interpretation of a simple pie chart; a more complex task would be to infer the likelihood of an occurrence, such as predicting the weather, based upon past information.

Representations of mathematical information.

144. Mathematical information in a situation may be available or represented in many forms. It may appear as concrete objects to be counted (e.g., people, buildings, cars, etc.) or as pictures of such things. It may be conveyed through symbolic notation (e.g., numerals, letters, and operation or relationship signs). Sometimes, mathematical information will be conveyed by formulae, which are a model of relationships between entities or variables. Mathematical information may be encoded in visual displays such as a diagram or chart; graphs and tables may be used to display aggregate statistical or quantitative information (by displaying objects, counting data, etc.). Similarly, a map of a real entity (e.g., of a city or a project plan) may contain information that can be quantified or mathematized. Last but not least, textual elements may carry much mathematical information or affect the interpretation of mathematical (and statistical) information, as explained further below.

145. A person may have to extract mathematical information from various types of *texts*, either in prose or in documents with specific formats (such as in tax forms). Two different kinds of text may be encountered in numeracy tasks. The first involves mathematical information represented in textual form, i.e., with words or phrases that carry mathematical meaning. Examples are the use of number words (e.g., "five" instead of "5), basic mathematical terms (e.g., fraction, multiplication, percent, average, proportion), or more complex phrases (e.g., "crime rate increased by half") which require interpretation, or coping with double meanings (or with differences in mathematical and everyday meanings of the same terms). The second involves cases where mathematical information is expressed in regular notations or symbols (e.g., numbers, plus or minus signs, symbols for units of measure, etc.), but is surrounded by text that despite its non-mathematical nature also has to be interpreted in order to provide additional information and context. An example is a bank deposit slip with some text and instructions in which numbers describing monetary amounts are embedded, or a parking ticket specifying an amount of money that has to be paid by a certain date due to a parking violation, but also explaining penalties and further legal steps that will be enacted if the fine is not paid by a certain date.

Enabling processes: cognitive and non-cognitive

146. People's numeracy competence is revealed through their responses (i.e., identifying, interpreting, acting upon, evaluating, and communicating) to the mathematical information or ideas that may be represented in a situation or that can be applied to the situation at hand. It is clear that numerate behavior will involve an attempt to engage with a task and not delegate it to others or deal with it by intentionally ignoring its mathematical content. Numerate behavior, however, depends not only on cognitive skills or knowledge bases, but also on several enabling factors and processes listed in Table 1 (NRDC, 2006; Tout, 2006).

147. Specifically, the enabling processes involve integration of mathematical knowledge and conceptual understanding with broader reasoning, problem-solving skills, and literacy skills. Further, numerate behavior and autonomous engagement with numeracy tasks depend on the dispositions (beliefs, attitudes, habits of minds, etc), and prior experiences and practices that an adult brings to each situation. These are briefly discussed below. Most of these enabling factors and processes have also been described by Kilpatrick (2001) as part of his analysis of the construct of mathematical literacy, and further examined and deemed relevance for description of adult numeracy in a recent analysis by Ginsburg et al. (2006).

Mathematical knowledge and conceptual understanding.

148. The notion of conceptual understanding refers to *an integrated and functional grasp of mathematical ideas* (Kilpatrick, Swafford & Findell, 2001: 118). Ginsburg et al (2006) suggest that the two aspects of conceptual understanding, i.e., it being integrated and functional, frame the ability to think and act numerately and effectively, and that across different numeracy frameworks in different countries, equivalent terms are used such as “meaning making,” “relationships,” “model,” and “understanding.” Conceptual understanding can help learners produce reasonable estimates that can help them catch computational errors, or realize that an exact product is not necessary, but an estimate is enough for the purpose. Ginsburg et al (2006) further explain that conceptual understanding permits one to be free from relying on memory for all methods and procedures, i.e., an adult can think about the meaning of the task and “construct or reconstruct” a representation that both illustrates what it means and suggests a method for solution. As an example they state that a fundamental conceptual understandings include interpreting and visualizing 23 x 13 as the repeated addition of 13 objects, 23 times (one could arrive at an accurate answer by adding groups), or as a 23 by 13 rectangular array (one could count the elements in the array).

Adaptive reasoning and problem-solving skills.

149. Throughout life, adults develop or apply diverse strategies to manage their quantitative situations. Some strategies may be based on prior formal learning, while others may be self-invented or adapted to fit the situation at hand. To solve computational problems or to manage certain quantitative tasks, people have to re-construct reality in a mathematical way, for example, model or mathematize. They can do so either on their own or in discussion with other people. Problem-solving strategies may include, e.g., extracting relevant information from the task/activity; rewriting/restating the task; drawing pictures, diagrams or sketches; guessing and checking; making a table; and/or generating a concrete model or representation (Kilpatrick, 2001; Ginsburg et al., 2006).

Literacy skills.

150. The ability to read, write, and talk are important skills in undertaking a numeracy task or activity or communicating the outcomes of working on such tasks. In cases where “mathematical representations” involve text, one's performance on numeracy tasks will depend not only on formal

mathematical or statistical knowledge but also on reading comprehension and literacy skills, reading strategies, and prior literacy experiences. For example, following a computational procedure described in text (such as the instructions for computing shipping charges or adding taxes on an order form) may require special reading strategies, as text is very concise and structured. Likewise, analyzing the mathematical relationships described in words requires specific interpretive skills, as in the simple case of recognizing the similarity of "the price doubled" and "the priced was twice as high", but the different meanings in "production levels were constant over the last five years" and "production levels constantly increased over the last five years".

Context/World knowledge.

151. Proper interpretation of mathematical information or quantitative messages by adults depends on their ability to place messages in a context and access their world knowledge, as well as rely on their personal experiences and practices, noted further below. World knowledge also supports general literacy processes and is critical to enable "sense-making" of any message. For example, adults' ability to make sense of statistical claims or media-based graphs will depend on information they can glean from the message about the background of the study or data being discussed. When interpreting statistical claims made by journalists, advertisers and the like, context knowledge is the main determinant of the reader's familiarity with sources for variation and error, helps to imagine why a difference between groups can occur (as in a medical or educational experiment), or what alternative interpretations may exist for reported findings about an association or correlation between certain variables. Likewise, world knowledge is a prerequisite for enabling critical reflection about statistical messages and for understanding the implications of the reported findings.

Beliefs and attitudes.

152. Research literature suggests that the ways in which a person responds to a numeracy task, including overt actions as well as internal thought processes and the adoption of a critical stance, depend not only on knowledge and skills but also on negative attitudes towards mathematics, beliefs about one's mathematical skills, habits of mind, and prior experiences involving tasks with mathematical content (Lave, 1988; Schliemann & Acioly, 1989; Saxe, 1991). In some cultures, some adults, including highly educated ones, decide that they are not "good with numbers" or have other sentiments or self-perceptions usually attributed to negative prior experiences they have had as pupils of mathematics (Tobias, 1993). Such attitudes and beliefs stand in contrast to the desired sense of "at-homeness with numbers" (Cockcroft, 1982) and can interfere with one's motivation to develop new mathematical skills or to tackle math-related tasks, and may also affect test performance (McLeod, 1992).

153. In real-world contexts, adults with a negative mathematical self-concept may elect to avoid a problem with quantitative elements, address only a portion of it, or prefer to delegate a problem, e.g., by asking a family member or a salesperson for help. Such decisions or actions can serve to reduce both mental and emotional load (Gal, 2000). Yet, such actions may fall short of autonomous engagement with the mathematical demands of real-world tasks, carrying negative consequences, e.g., not being able to fully achieve one's goals.

Numeracy-related practices and experiences.

154. Research suggests that, for adults as well as for children, mathematical knowledge develops both in and out of school (e.g., Schliemann & Acioly, 1989; Saxe, 1991; Lave, 1998). Saxe and his colleagues have written about the importance of cultural practice in the development of mathematical thinking and how such practices profoundly influence an individual's cognitive constructions and mathematical ideas, depending, e.g., on the artifacts or tools they use, the nature of the measurement

systems in their culture, the counting or calculating devices (abacus, calculator) they use, the distribution of work among family members, or general patterns and types of social activity. Further, the frequency of engaging with mathematical tasks or of exposure to mathematical or statistical information or displays, whether at work, home, when shopping, or in other contexts, is of much interest. Engagements or practices in this regard can be both the result of a certain skill level, but also the cause of observed skill levels, or at a minimum a factor influencing observed skill level apart from prior formal schooling.

155. The ideas above suggest that numerate behavior does not rely only on mathematical knowledge or related reasoning and problem-solving skills acquired as part of formal learning in a school context. Both attitudes and beliefs as well as numeracy-related practices and world knowledge are important enabling processes and may influence adults' ability to act in a numerate way. Therefore, scales assessing selected attitudes and beliefs about mathematics, and numeracy-related practices in work, everyday, and other settings, have been developed for PIAAC's Background Questionnaire (BQ). Information collected by such scales can help to explain differences in performance among adults, further inform our understanding of factors that affect skill acquisition and retention or motivation for further learning, and explain the links between numeracy and covariates such as participation in further learning or employment/unemployment status.

Scale development: Principles, constraints, implementation

156. The operationalization of the construct of numeracy in a large-scale assessment scale is affected by many factors which shape the extent to which the theoretical construct can be fully addressed by the actual collection of items used in the direct assessment. This part first describes general expectations in assessing adult numeracy gleaned from prior work on assessing adults' mathematical skills and the theoretical foundations reviewed above, followed by an outline of design constraints that affect the development of a scale for direct assessment of cognitive skills in PIAAC. Based on these foundations, an outline is presented of design principles that guide the assessment of numeracy in PIAAC, and further details on a supporting scheme regarding factors that affect task complexity (or item difficulty) which is of importance both for task design as well as interpretation of results regarding numeracy in PIAAC.

General ideas about shaping tasks for assessing adult numeracy

Task authenticity and realism.

157. Numerous authors have highlighted the need to retain in assessments of adults' numeracy the authenticity of assessment tasks and make them as similar as possible to the way adults encounter mathematics in different life contexts. Ginsburg et al (2006: 9), for example, claim:

> *"If one accepts the premise that 'realistic' is not the same as 'real', a serious question is raised about the extent to which 'efficient', short-response standardized test items are valid measures of a person's numeracy when the items are not structured to elicit the practices an adult actually employs in a real situation".*

158. It follows that differential performances can occur when assessment is divorced from, as opposed to contextualized in, realistic settings (Lave, Murtagh, & de la Rocha, 1984). Problem-solving in contextualized real life and work activities may differ from solving school-like problems (Resnick, 1987; Greeno, 2003). Thus, assessments of adult numeracy have to aim for a high degree of realism and authenticity in both stimuli and tasks presented to respondents. The desire to retain authenticity, however, may at times be at odds with the need to establish cultural appropriateness of tasks and stimuli and reduce context effects. Tasks deemed as authentic and valid in the context of one country

or culture may be unfamiliar to a smaller or larger degree in another cultural context. This is a traditional problem in cross-cultural testing that has challenged generations of test specialists.

159. Arguably the problem of authenticity and cultural appropriateness is lessened when testing pupils in schools, such as in PISA, because test designers can use conventional mathematical terminology, formulae, symbols, and so forth; this helps school-age assessments to standardize the demands from respondents by conveying the mathematical information embodied in different situations in consistent ways regardless of the cultural context. However, testing of adults' numeracy presents more challenges because many will not remember formal school-based notations or terminology. In countries where a sizable proportion of the population are immigrants or speak multiple home-based languages, the gaps between mother tongues and school-based mathematical linguistic conventions may further affect performance on some numeracy tasks. Thus, attention has to be given to linguistic and cultural factors when adapting items for adult assessments.

Task format and coding.

160. Another aspect of importance in designing assessment of adult numeracy is task format, i.e., forced-choice (or multiple choice) format versus a constructed-response format where respondents communicate in their own words the answers to tasks or questions given as part of the direct assessment, or otherwise are free to choose how to respond and are not limited to a specific and small set of given responses as in multiple-choice tests. Some of the key arguments for using constructed-response formats in adult numeracy assessment are that in most real-life situations, adults have flexibility in how they choose to respond to given tasks. Many real life tasks call for approximate answers or estimates rather than for accurate results, or for opinions or judgments that adults have to express in their own words. Further, there is long-standing awareness regarding the limited ability of forced-choice items to reflect reasoning or problem-solving processes and arguments that underlie the choice of a particular response. Thus, the use of multiple-choice items undermines the ability to assess the extent to which adults can "communicate mathematically". On the other hand, the coding of the constructed responses as being correct or incorrect can be more complicated and require much further training of coders, while items where respondents have to choose from among a limited set of possible answers sometimes (but certainly not always) offer advantages in terms of cost, speed, and reliability of coding.

Usage of calculators and other tools or objects.

161. The assessment of numeracy, whether by paper and pencil tasks or their computer-based equivalents, has to take into account that the practice of numeracy in everyday or work situations also involves the use of certain objects and artifacts. First we should examine the use of hand-held calculators, which by now are inexpensive and would be widely available to adults from all walks of life in many countries. Calculators have been included for quite some years now in school curricula so should be familiar to many adults. Thus, hand-held calculators are tools which are part of the fabric of numeracy life in many cultures. Increasingly, respondents in large scale tests are allowed, sometimes even expected, to use calculators. However, we still see discussions as to whether an assessment should be conducted with, or without, allowing respondents to use calculators or other technological tools.

162. It follows that adults should be given access to a calculator as part of an assessment of numeracy skills, and they can then choose if and how to use it. That said, when a calculator is made available, it is not possible to know what exactly the respondent does with it or for what it is being used in each task, e.g., does the respondent use it to compute a result which is then given as an answer to the task, or use it to verify results that were first obtained by mental or manual (written) calculation? It is also difficult to document problems in using a hand-held calculator without having an examiner looking all the time "over the shoulder" of the respondent and intruding into the respondents' work

process. Thus, while making a calculator available during testing is paramount, collecting information about its usage presents many challenges. Yet, without information about the purpose of usage, it is difficult to analyze whether the usage of a calculator helps adults cope with certain numeracy tasks, or to conclude what might be the educational or policy implications.

163. In addition to a calculator, other tools or objects could be used in certain assessment tasks. The use of a ruler or measuring tape, whether in a metric or imperial (inches) system are part of contexts where adult numeracy competence is manifested, both in certain work setting and parallel home settings (e.g., carpentry, construction, home remodeling projects). Further, the use of objects that can be counted or manipulated (coins, beans) can shed more light on the ability of low-ability or low-literacy individuals to handle certain everyday situations involving simple quantitative information. (The use of other more sophisticated objects, such as a computer spreadsheet, of course can also fit under the assessment of numeracy, but in PIAAC is taken under the framework of the Problem Solving domain).

164. The use of a calculator, ruler, or objects such as coins is in principle desired in an assessment of adult numeracy skills. Yet, actual implementation in a large-scale assessment carries both psychometric, operational, and cost implications when there is a need to test thousands of adults in their homes in multiple countries in a standard and efficient manner. For these reasons, in PIAAC's first cycle respondents are able to use hand-held calculators as well as paper (printed) rulers that have both metric and imperial measurements. However, it is not possible to use other types of objects such as country-specific coins due to the heterogeneity in this regard.

PIAAC approach to assessment

The PIAAC assessment design involves using a household survey methodology which assumes that overall testing time per respondent is around 60-80 minutes. During this time, there is a need to administer a short core test (screener), present direct assessment items in one or more of the different competency domains (Literacy, Numeracy, Problem-solving), and collect information about the respondents' background and various correlates of interest via the background questionnaire (BQ). The design for the direct assessment in PIAAC incorporates computer-based measures of competencies when possible, yet also use written ("paper and pencil") assessment booklets for adults who are unfamiliar or uncomfortable with computers.

Adaptive testing.

165. To increase assessment efficiencies, the direct assessment is administered to the majority of respondents via a computer platform, TAO, which uses a computer-based adaptive testing process. The adaptive testing process means that tasks (i.e., stimuli and questions about them) are shown on a computer screen, the respondents answer on the computer, and their *answers are automatically* (immediately) *scored as correct or incorrect*, without human judges or coders being involved in interpretation of responses. This automatic scoring is essential because adaptive testing is based on the cumulative performance on tasks; at various points during the assessment TAO decides, based on decision rules stored by the computer program, what additional assessment tasks (at higher or lower difficulty levels) to select from a pool of assessment items for presentation to the respondent.

166. The key advantage of an adaptive testing scheme is that is can achieve the best estimate of each respondent's ability level, using a smaller number of assessment items than in a traditional test design where respondents have to answer all questions included in the test, from easiest to most difficult. Thus, adaptive testing can enable deeper and more accurate assessment of respondents ability level, while reducing response burden and the chance respondents will face many tasks which are above their ability level and hence cause frustration.

167. However, the assessment of numeracy in the first cycle of PIAAC is constrained in several ways because of the assessment design. Firstly, the overall testing time per respondent does not allow inclusion of extended problems or lengthy simulations of complex authentic numeracy tasks, although it is recognized that ability to solve complex or extended numeracy problems is an inherent part of the numeracy competency. In order to cover all facets of the numeracy construct in the limited time available, the use of a larger number of short tasks is prescribed.

168. Secondly, the need to score all responses automatically limits the type of assessment tasks that can be used. While the TAO system allows respondents to provide an answer in several different modes (e.g., numeric entry, clicking on an area of the screen, choosing from pull-down menus), in its present stage of development it cannot accept most types of free-form text-based answers because of the huge possible diversity in how respondents may enter their answers. The limitations stem from the difficulty to automatically code (i.e., designate an answer as correct or incorrect) *free-form* responses in dozens of languages while accommodating various grammatical and syntactical structures, as well as overcoming typing mistakes which are naturally expected when people type text into a computer. Examples are when respondents:

- write number ranges or estimates which have multiple mathematically-equivalent representations, such as "a quarter", "0.25", "1/4", "1 in 4", or “around five to six”, "1.00 to 6.00",
- provide explanations of how a certain result was reached (“subtracted six from the sum of 30", "I did 30 - 6"),
- describe their interpretation of given information such as in a simulated media statement
- write justifications for their answers, or list arguments supporting their conclusions.

169. As a result of the restrictions discussed above, certain types of numeracy tasks, especially those involving interpretation or evaluation/analysis with communication responses, receive only partial or slight coverage in the first cycle of PIAAC. With this in mind, as part of the search for ways to circumvent somewhat the limitation on text-processing in the computer-based testing environment, in a few numeracy questions respondent may be asked to provide an explanation for a response by choosing from pre-designed encapsulated texts, so as to simulate the way a person provides a justification for an answer in real life. However, such experimental solutions are partial at best and have their own limitations. It is thus hoped that in future cycles of PIAAC, some of the current technical limitations will be resolved, allowing for broader coverage of more aspects of the numeracy construct. In addition, it should be noted that respondents uncomfortable or unfamiliar with using computers are directed to the paper-based branch of the assessment. While the tasks they encounter in that portion of the test are mostly duplicates of computer-based items, a few do require respondents to communicate interpretations or explanations of their reasoning about some tasks via free-form written answers, thus helping to expand the coverage of the numeracy construct slightly beyond what is possible in the computer-based assessment alone.

Item pools and scaling.

170. The intent of PIAAC is to have its results linked to previous international adult assessments. Therefore, the general PIAAC design requires that 60 percent of the literacy and numeracy tasks will come from item pools used in ALL and IALS. These former items serve as linking items, and in addition new items were developed for PIAAC that can fit the computer-based adaptive testing requirements and constraints. Overall, the items for numeracy assessment are expected to enable reporting of respondents' performance in a manner similar to the one used in ALL and IALS, which

scaled raw ability scores in the range 0-500, but mainly focused on reporting performance on five ability levels with the following tentative boundaries:

- Level 1: raw score of 0 – 225 (lowest level)
- Level 2: raw score 226 – 275
- Level 3: raw score 276 – 325
- Level 4: raw score 326 – 375
- Level 5: raw score 376 – 500 (highest level)

Principles for assessing numeracy in PIAAC

171. The development of numeracy assessment for PIAAC has been based on a number of general principles or guidelines listed below. These principles reflect the cumulative literature on large-scale assessment of mathematical skills and adult numeracy (Gal et al., 2005; Gillespie, 2004; Murat, 2005), and various background documents and positions prepared as part of the planning of PIAAC (e.g., Gal, 2006; Jones, 2006; Murray, 2006; Tout, 2006), the general ideas listed earlier in this section, as well as the known technical limitations in the first cycle of PIAAC:

a. *Items should cover as many aspects as possible within each of the four facets of the numeracy competency.* Items should require the activation of a broad range of skills and knowledge included in the construct of numeracy, as portrayed in the conceptual framework depicted in Table 1. Specifically, all four areas of mathematical content, information, or ideas (Facet 3 in Table 1) should be covered, with relative proportions as follows:
    - 30% Quantity & Number
    - 25% Dimension & Space
    - 20% Patterns, Relationships, Change
    - 25% Data & Chance

b. *Items should aspire to maximal authenticity and cultural appropriateness.* Tasks should be derived from real-life stimuli and pertain to all types of contexts or situations (i.e., everyday life, work, societal, further learning) that can be expected to be of importance or relevant in the countries participating in PIAAC. Item content and questions should appear purposeful to respondents across cultures, although it must be acknowledged that in a large-scale assessment such as PIAAC, not all items and contexts can be personally familiar to all adults within any one country, let alone across all countries.

c. *Items should have a free-response format, to the extent feasible by the computer platform used for administering the direct assessments in PIAAC.* Items should be structured to include a stimulus (e.g., a picture, drawing, visual display) and one or more questions, the answers to which the respondent communicates via the modes available within TAO, primarily: numeric entry, click, highlight a region of the stimulus, usage of various pull-down menus. (Text entry is limited to very specific words or sometimes a simple number due to the concerns listed above regarding the inability to score text entries with keying/typing errors, and the presence of multiple ways to express the same mathematical entities in words and/or numbers). In addition, items allowing a free-form response will be used in the paper-and-pencil portion of PIAAC, which some respondents will take, allowing for some expansion of possible responses, beyond those presently afforded within the computer platform.

d. *Items should spread over different levels of ability.* Items should span the range of ability levels anticipated within PIAAC participants, from low-skilled individuals (which are of interest in countries where policies and educational programs may be earmarked for low-skill populations), all the way to those with advanced competencies.

That said, it should be recognized that the need to reduce the number of items to be administered in any one domain has led designers of past assessments (IALS and ALL for adults, PISA for school students), as well as in PIAAC, to include few very easy items (i.e., items at level 1) and few very hard items (i.e., items at Level 5). Instead, respondents will be classified as at Level 1 if they could not do well on Level 2 tasks. Likewise, those classified at Level 5 will be those who performed well on Level 4 items and on the few real Level 5 items. It follows that a more detailed assessment of the specific skills that Level 1 respondents have requires a separate diagnostic assessment, such as the assessment of component literacy skills planned for PIAAC.

Given the above, to enable the adaptive testing process reach a an efficient estimation of respondents' ability levels, the following distribution of items at the different difficulty levels is likely to be sought for constructing the numeracy item pool for the Main assessment, based on the results of the field-test (pilot) in 2010:

- Level 1: 10% of the numeracy items
- Level 2: 25% of the numeracy items
- Level 3: 30% of the numeracy items
- Level 4: 25% of the numeracy items
- Level 5: 10% of the numeracy items

d. *Items should represent the different response types*. However as mentioned already, certain types of numeracy response types, especially those requiring the use of interpretation, evaluation, analysis and communication, will receive only partial or slight coverage in the first cycle of PIAAC due the the computer based assessment platform and its constraints at this stage. Therefore for PIAAC the response types of Interpret, Analyse/evaluate and Communicate have been collapsed into a single response type. It is hoped that in future cycles of PIAAC, some of the current technical limitations will be resolved, allowing for better coverage of more aspects of the numeracy response facet. Given the above, the following distribution of items requiring the different types of response types will be sought for constructing the numeracy item pool for the main PIAAC direct assessment:

- 10% Identify, locate or access
- 50% Act upon, use: order, count, estimate, compute, measure, model
- 40% Interpret, evaluate/analyze, communicate

e. *Items should vary in the degree to which the task is embedded in text.* Some items should be embedded in or include relatively rich texts, while others should use little or no text. This distribution aims to reflect the different levels of text involvement in real-world numeracy tasks, as well as reduce overlap with the literacy scale.

f. *Items should be efficient.* To allow for coverage of many key facets of the numeracy competency, the inclusion of a large number of diverse stimuli and questions will be needed. However, in light of testing time constraints, the use of short tasks is necessitated, precluding items that can simulate extended problem-solving processes or that require a lengthy open-ended response.

g. *Items should be adaptable to unit systems across participating countries.* Items should be designed so that their underlying mathematical demands are as consistent as possible across countries, regarding language and mathematical conventions. For example, items should be designed so that different currency systems or different systems of measurement (metric or Imperial) could be applied to the numbers or figures used. Items should retain equivalency with respect to their mathematical or cognitive demands after being translated.

Factors explaining item/task complexity

172. In planning an assessment, it is of course desirable to be able to understand what it measures. Assessment designers assume that when engaged with the assessment items (including tasks, questions, stimuli, etc), respondents activate cognitive processes and rely on stored knowledge and learned skills which are part of the construct being measured. Thus, differential performance levels can be accounted for by the underlying cognitive knowledge bases and other enabling processes. It follows that it is useful to have a theoretical model or set of assumptions regarding what factors cause certain tasks to be harder or more complex than others, so that the assessment results can be correctly interpreted. A model or scheme of factors affecting task complexity can also help when linking the assessment results to possible social (or educational) interventions, i.e., point to the skills that are lacking and have to be further developed in the population (Brooks, Heath, & Pollard, 2005).

173. Prior seminal work by Kirsch and Mosenthal (e.g., Kirsch, 2001) and earlier projects has pointed to several key factors which account for task difficulty when considering arithmetic items or items involving text comprehension, primarily readability, type of match, plausibility of distractors, operation specificity ('transparency'), and type of calculation and number of steps. The Kirsch & Mosenthal work has informed the design of assessment tasks for IALS and other surveys, and the interpretation of their results. In designing the ALL numeracy scale, the ALL Numeracy team has attempted to advance the Kirsch and Mosenthal complexity scheme and develop tentative assumptions regarding factors which affect difficulty of multiple types of new tasks introduced to measure the numeracy construct which were beyond those encompassed by the more focused construct of Quantitative Literacy in IALS. Examples are items involving percents, knowledge of measurement and spatial reasoning, statistical concepts, and so forth.

174. The developers of the Mathematical Literacy scale for PISA (2006) also recognized multiple factors affecting item difficulty, such as the kind and degree of interpretation and reflection required by the problem, the kind of representation skills required, or the kind and level of mathematical skill required, e.g., single-step vs. multi-step problems, or more advanced mathematical knowledge, complex decision-making, and problem solving and modeling skills, or the kind and degree of mathematical argumentation required. Further factors that are assumed to affect difficulty both in PISA, ALL and other surveys relate to the degree of familiarity with the context, and the extent to which tasks require reproduction of known procedures and steps or present novel situations requiring non-routine and perhaps more creative responses. It should be noted that the PISA description of complexity factors seems quite compatible with that of ALL, although some of the terminology is

different, and published PISA reports do not explain in detail how it was used to guide the design of specific items.

175. The complexity scheme for numeracy used in ALL (Gal et al., 2005) has been instrumental for the item development and scale construction stages of that study, especially in that it helped to evaluate in advance if items will span different difficulty levels. Given that PIAAC's numeracy assessment is founded on the principles developed for ALL and that the PIAAC numeracy assessment scale uses over two dozen linking items used in ALL, the ALL complexity scheme has been adopted as an analytic tool for item development and interpretation for PIAAC as well. Further details about this scheme are provided in Appendix 1, which is adapted from Gal et al. (2005).

Differences between PIAAC's numeracy and related scales

176. To gain a better understanding of what is measured in the numeracy domain in PIAAC, it is important to discuss the differences between numeracy and related constructs targeted in international assessments, such as quantitative literacy and mathematical literacy. As will be seen, the differences are more a matter of degree rather than these constructs being totally different from each other - after all in one way or another they all pertain to some aspects of people's mathematical knowledge. Further, it must be pointed out that the differences emerge in more clarity when looking not at definitions (i.e., the conceptual level) but at their *operationalization* (i.e., the assessment scale design, constraints on assessment, and the actual assessment tasks).

Adult assessments

177. Let us first examine some conceptions developed in international surveys of adult skills. A framework developed by Kirsch and Mosenthal (see Kirsch, Jungblut, & Mosenthal, 1998) to describe adults' literacy skills, including aspects of adult's quantitative skills, has been widely implemented in multiple national and international assessment projects, most recently the International Adult Literacy Survey (IALS; see Statistics Canada and OECD, 1996, 1997). The IALS framework made use of three literacy scales—Prose Literacy, Document Literacy, and Quantitative Literacy—to operationalize its conception of literacy. The PIAAC domain of numeracy is most closely related to the Document Literacy (DL) and Quantitative Literacy (QL) scales, defined as follows.

> DL: *The knowledge and skills required to locate and use information contained in various formats (including job applications, payroll forms, transportation schedules, maps, tables, and graphics).*
>
> QL: *The knowledge and skills required to apply arithmetic operations, either alone or sequentially, to numbers embedded in printed materials (such as balancing a check book, figuring out a tip, completing an order form, or determining the amount of interest on a loan).*

178. QL tasks as well as some DL tasks have addressed important aspects of people’s mathematical knowledge and skills. For example, DL tasks required respondents to identify, understand, and interpret information given in various lists, tables, charts and displays; this information sometimes included quantitative information, such as numbers or percents. QL tasks required respondents to apply *arithmetical* operations learned mostly in elementary grades; these tasks did not require respondents to cope with other types of mathematical information (e.g., measurements, shapes) or with information whose processing does not require comprehension of text. In addition, tasks used in both QL and DL scales called for a limited range of responses, i.e., exact computations or specific

types of interpretations. Such tasks and responses are important by themselves, yet they represent only a subset of the much wider range of tasks and responses that are typical of many everyday and work tasks, such as sorting, measuring, estimating, conjecturing, or using models (e.g., formulas). Thus it can be concluded that QL and DL cover a subset of the dimensions and ideas captured by Numeracy for PIAAC.

School-age assessments

179. In the context of international assessments of school-age students, a central construct is mathematical literacy. This term first appeared as part of the second TIMSS, where students in their final year of secondary schooling (usually 12th grade) were assessed not only on mathematical knowledge but also on "mathematics literacy" (Mullis et al, 1998:43), defined as follows:

> *Mathematics literacy items address number sense, including fractions, percentages, and proportionality. Algebraic sense, measurement, and estimation are also covered, as are data representation and analysis. Several of the items emphasize reasoning and social utility. A general criterion in selecting the items was that they should involve the types of mathematics questions that could arise in real-life situations and that they are contextualized accordingly.*

180. This definition illustrates the connection between conceptualization and operationalization as two building blocks of the construct. But it can also be seen that the construct of mathematics literacy was defined through its facets, without there being a general definition.

181. PISA (2006) defines mathematical literacy as follows:

> *Mathematical literacy is an individual's capacity to identify and understand the role that mathematics plays in the world, to make well-founded judgments and to use and engage with mathematics in ways that meet the needs of that individual's life as a constructive, concerned and reflective citizen.*

182. This definition shows some overlap and consistency with the conception of numeracy used in the present framework as well as with broader conceptions of *literacy* as adopted by IALS, ALL, and PIAAC.

Table 2: Mathematical content areas covered by PIAAC and PISA

PIAAC	PISA
- Quantity & number	- Quantity
- Dimension & shape	- Space and shape
- Data & chance	- Uncertainty
- Pattern, relationships, & change	- Change & relationships

183. Table 2 lists the four key content areas covered by the numeracy assessment in PIAAC, in comparison with the four content areas of mathematical literacy covered in PISA. While using

somewhat different terminologies, the two schemes of mathematical content refer to quite similar domains overall and point to conceptual compatibility between PIAAC and PISA.

184. There are several additional similarities between PISA's Mathematical Literacy assessment of school students of age 15 and PIAAC's Numeracy assessment of adults, such as:

- PISA examines how students of age 15 cope with tasks embedded in four contexts: personal, educational/occupational, public and scientific, which are quite similar to those contexts discussed in section 2.4.1 above regarding adult numeracy

- The PISA assessment framework for mathematical literacy highlights PISA's focus on real-world problems, moving beyond the kinds of situations and problems typically encountered in school classrooms, and demanding "the ability to apply those skills in a less structured context, where the directions are not so clear, and where the student must make decisions about what knowledge may be relevant and how it might usefully be applied" (PISA, 2006; p 72).

- The description of the four types of responses (i.e., identifying, locating or accessing; acting upon, using; interpreting; and communicating), bear some resemblance to the five steps in the mathematization cycle posited in the PISA framework as involved in solving real-life problems. These steps involve, among other things: identifying mathematical information in a situation after trimming away non-essential elements in reality, solving a mathematical problem while switching between representations and using formal operations or technical language as needed, linking mathematical solutions and making sense of them in light of the real situation, or explaining results.

- The enabling processes that support numerate behavior, discussed above, relate to the eight key competencies described in the PISA framework, such as thinking, reasoning, modeling, problem-posing and problem-solving, communication, representation, and using symbolic and technical language.

185. While there are numerous similarities between the framework for PIAAC numeracy and the PISA framework for mathematical literacy, differences also emerge, in part due to the different environments for which these frameworks have been developed. PISA relates to school-based populations, and while it is interested in students' performance on real life problem, a basic underlying assumption is that the performance is to be based on skills and dispositions acquired in a schooling context. As a result, descriptions of students' desired actions or underlying cognitive processes in operation are couched in a school-based environment. Indeed, an arguable cursory examination of published PISA items shows that there is room to further examine the extent to which items use realistic contexts and stimuli. Some items use formal symbolism that reflect an expectation for formal knowledge of what was taught in schools, yet such knowledge is less (or not) available to adults who have been out of formal school environment for years.

186. It can be assumed that adults, much more so than 15-year olds, have personal experiences and ways of coping with everyday situations which are different than those of school-age students. Hence, the types of responses envisioned of adults tested in PIAAC, and the explanations for underlying enabling or causative factors (such as "literacy skills") are not couched in a "mathematical problem-solving" culture. For this reason, task realism in PIAAC assessment may play a somewhat different role than in PISA. As a result, the numeracy framework for PIAAC, while being informed by the established literature on school mathematics, certainly goes beyond it, and at times uses different terms and ideas that are based on additional literatures.

187. Note should also be taken of the role of *literacy* (in the narrow, technical sense of reading and writing) in PISA's mathematical literacy. Despite the inclusion of the term “literacy” in “Mathematical Literacy”, the PISA mathematical assessment does *not* seem explicitly interested either in tasks where mathematical information is embedded in text, or in the influence of literacy skills on mathematical performance, described earlier in this framework. In creating the PISA Mathematical Literacy scale, relatively little effort appears to have been taken to control the literacy content of tasks; as a result, very high correlation were obtained between reading literacy and mathematical literacy scale scores; this in turn was one (but not the only) contributing factor towards a situation whereby countries scoring high on one scale usually (but not always) scored high on other scales. In contrast, the designers of the ALL numeracy scale sought to reduce the literacy demands of at least some of the numeracy items. This in turn contributed to numeracy scores having lower correlations with Document Literacy scores in ALL, compared to the relatively high correlations between Quantitative and Document Literacy scores in IALS.

Further issues in comparing large-scale assessments

188. The comparison of assessment frameworks, even in a single domain such as mathematics, can be a complex undertaking. This was demonstrated by a recent project of the National Center on Education Statistics (NCES) in the United States, which aimed to compare the mathematics frameworks and items for three large scale assessments, the National Assessment of Education Progress, TIMSS, and PISA (Neidorf, Binkley, Gattis & Nohara, 2006). The project involved comparisons of frameworks along multiple dimensions or topics, such as the mathematics content and process skills to be assessed, the main content areas included and the set of subtopics covered in each, calculator use policy, and so forth. The analysis of the commonalities and differences between the three assessments also included comparisons of hundreds of items, in terms of the mathematics content covered, performance expectations for different grade levels, the complexity of different tasks (e.g., the extent to which they require application of routinized versus novel approaches), cognitive processes underlying different items, item formats, and item contexts. This project required the work of a panel of a dozen experts over several days.

189. One of the general specifications for PIAAC's numeracy is that there would be conceptual continuity with Mathematical literacy as viewed by PISA. While the two constructs are related they should *not* be viewed as identical, for reasons explained above. Each construct has somewhat different operational implications at the level of assessment design and item content. For example, PISA items are somewhat different than PIAAC items due to differences in authenticity of tasks, the contexts from which tasks are drawn (and the role of technology and tools in them), and in particular the increased use of formal mathematical symbolizations, which many adults are normally not familiar with after leaving school. Further, numeracy implies the need for greater attention to dispositional aspects of the competence.

190. It follows from the above that the PIAAC numeracy scale is compatible and shares much common ground with PISA’s mathematical literacy scale, as well as with ALL’s numeracy and to some extent with IALS QL and DL scales. Yet, the nature of the commonalities and differences cannot be fully analyzed at this stage without further investment in a more fine-grained analysis and some systematization of terminology, thereby creating simpler bridges between PIAAC and PISA. Eventually, the meanings that can be attached to assessment results from PIAAC in the area of numeracy, and the degree of overlap between PIAAC’s numeracy and PISA’s mathematical literacy, depend not only on the conceptualization of numeracy in PIAAC, but probably more so on a host of other factors discussed earlier. It must be kept in mind that an analysis of commonalities and differences between PIAAC's and PISA's mathematical scales cannot be conducted only on the conceptual level - it needs to also consider the characteristics of the actual assessment tasks. Such characteristics include task realism, density of text in stimuli, and so forth, as well as the constraints on

assessment (i.e., what questions can be asked, what coding is possible, etc) imposed by the features of the computer platform for implementing PIAAC's direct assessments.

Summary and further reflections

191. Given the increasing need for adults to continuously adapt to changing citizenship and workplace demands (European Commission, 1996; Coben, O'Donoghue & FitzSimons, 2000), the assessment of numeracy is essential so that countries have a solid basis from which to design social interventions and effective lifelong learning opportunities that can improve competencies (OECD, 2006). Accordingly, numeracy has been conceptualized in this document as a broad construct that pertains to adults' level of coping with a diverse range of numeracy tasks couched in real-world contexts.

192. This document has reviewed literature and research regarding competencies (part 1), perspectives on the complex meanings associated with numeracy and related constructs, and the numeracy demands faced by adults in various contexts (part 2). Based on this background, later subsections in part 2 presented a definition of numeracy as a basis for developing the assessment scale for PIAAC. These theoretical foundations then served as the basis for discussing (part 3) facets of numerate behavior related to *Contexts, Responses, Mathematical content/information/ideas, and Representations*, each with several components. This document has also emphasized the importance of assessing dispositions and practices as an integral part of the numeracy competence, given that they affect performance on numeracy tasks and can correlate with various variables of interest. Later parts elaborated on assessment principles and design constraints inherent in PIAAC's assessment plan and computer-based platform (part 4), and examined commonalities and differences between PIAAC and other scales (part 5) in order to help readers understand how to interpret PIAAC results pertaining to numeracy and connect them with results regarding mathematical literacy (PISA) or Quantitative literacy (IALS).

193. At the *conceptual* level, the definition of numeracy in this framework is in general compatible with the ALL conceptualization, yet introduces some advances of a modest scope that go beyond what existed in ALL. These changes bring the conceptualization of numeracy presented in this document closer to the conception of "literacies in the information age" employed by PIAAC, while allowing for more compatibility with the PISA definition of mathematical literacy, as desired by OECD. The changes in the conceptualization of numeracy were also introduced with a long-range view in mind, to enable the accommodation of new types of numeracy-related tasks and demands faced by adults in the information age in future cycles of PIAAC, while maintaining a common conceptual definition across assessment cycles.

194. It should be recognized that some modest overlap exists between key constructs measured in PIAAC, i.e., Document literacy, Numeracy, and Problem-Solving in Technology-Rich Environments, on the conceptual level *but* also separately at the scale or item level. In particular, numeracy includes as one of multiple sub-areas the ability to read and interpret quantitative information in graphs and tables, yet this is also subsumed as one part of Document Literacy. The existence of such an overlap is sensible and expected. After all, many real-world tasks, such as interpretive tasks which do not require the manipulation of numbers, or tasks involving quantitative statements embedded in text, require adults to *integrate* the use of numeracy and literacy skills (Kirsch et al., 1993; Gal, 2002a). Likewise, numeracy involves the ability to solve multi-step or extended problems couched in a technology context, not just short simple tasks, and some numeracy-related content is encountered when adults handle problems involving financial, scheduling, or other everyday tasks. Hence tasks chosen for the Problem-solving scale may also touch on numeracy topics. The presence of overlap between scales, however, should not be a cause for concern or a reason for restricting the conceptual definition of numeracy. The construct of numeracy stand on its own and reflects an authentic part of the adult world.

In addition, one must also consider that "assessment drives instruction": Excluding some areas from the *conceptualization* of numeracy to reduce overlap or inter-scale correlations may limit what delivery systems in, e.g., educational or workplace training contexts, will target when trying to develop desired competencies.

195. As emphasized earlier, what is measured by a numeracy assessment scale (direct assessment) and the associated BQ items (regarding numeracy-related practices, attitudes, beliefs, and so forth) is determined not only by a conceptual framework describing numeracy and its facets and enabling factors - it is also determined by an *assessment* framework. Such a framework describes how the general conceptualization of numeracy is operationalized and manifested in the nature and range of tasks used in the actual assessment, and specifies what limitations may be created due to the mode of task administration and of scoring.

196. Overall, an assessment is a complex dynamic system that combines conceptual and technical elements. Eventually, the outcomes of an assessment, and the reliability, validity, and usefulness of the findings and their interpretations, are influenced by criteria and values placed by the assessment designers, and by the many choices they make about task design and methodology. The implementation of the numeracy conceptual framework in the first assessment cycle of PIAAC is influenced by practical limitations of the computer-based platform, which prohibited the use of certain types of open-ended, communication and evaluative type items. In addition, the need to cover a broad construct within the short timeframe available for assessment, a typical situation in large-scale household-based assessments, has forced the use of short tasks with machine scorable response formats, further limiting the use of extended problems. It is hoped that these limitations will gradually lessen as greater sophistication in computer platforms will enable more flexibility in task design, question-posing and response formats in future cycles of PIAAC.

References

American College Testing (1997). *Work keys: Preliminary technical handbook*. Iowa City, IA: Author.

Australian Education Council (The Mayer Committee). (1992). *The key competencies report: Putting general education to work.* Canberra: The Australian Education Council and Ministers for Vocational Education, Employment, and Training.

Baker, D., & Street, B. (1994). Literacy and numeracy: Concepts and definitions. In T. Husen & E. A. Postlethwaite (Eds.), *Encyclopedia of education*. New York: Pergamon Press.

Beazley, K. (1984). Education in Western Australia: *Report of the Committee of Inquiry into Education in Western Australia.* Education Department of Western Australia.

Benn, R. (1997). *Adults count too: mathematics for empowerment*, Leicester, UK: National Institute of Adult Continuing Education (NIACE)

Bessot, A., & Ridgway, L. (Eds.) (2000). *Education for Mathematics in the Workplace.* New York: Springer.

Bishop, A. (1988). *Mathematical enculturation: a cultural perspective in mathematics education*. Dordrecht, Holland: D. Reidel Publishing Co.

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]

Buckingham, E. (1997). *Specific and generic numeracies of the workplace*. Melbourne: Deakin University.

Carnevale, A. P., Gainer, L. J., & Meltzer, A. S. (1990). *Workplace basics: The essential skills employers want*. San Francisco: Jossey-Bass.

Carraher, T. N., Schliemann, A. D., & Carraher, D. W. (1988). Mathematical concepts in everyday life. In G. B. Saxe & M. Gearhart (Eds.), *Children's Mathematics* (pp. 71-88). San Francisco: Jossey-Bass.

Coates et al. (1995). *National Reporting System*. Australia: Commonwealth of Australia and the Australian National Training Authority.

Coben, D. (2000). Numeracy, mathematics, and adult learning. In I. Gal (Ed.), *Adult numeracy development: Theory, research, practice* (pp. 33-50). Cresskill, NJ: Hampton Press.

Coben, D., O'Donoghue, J., & FitzSimons, G. (Eds.)(2000), *Perspectives on Adults Learning Mathematics*. London: Kluwer Academic Publishers.

Coben, D., Colwell, D., Macrae, S., Boaler, J., Brown, M., & Rhodes, V. (2003). *Adult numeracy: Review of research and related literature.* London: National Research and Development Centre for Adult Literacy and Numeracy (NRDC).

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 U.S. Department of Education, Office of Vocational and Adult Education). [Online: www.ed.gov/about/offices/list/ovae/pi/AdultEd/math.html].

Coulombe, S., Tramblay, J. F. & Marchand, S. (2004). *Literacy scores, human capital and growth across fourteen OECD countries.* Ottawa: Statistics Canada.

Department of Education and Science/Welsh Office (The Cockroft Report). (1982). *Mathematics counts: Report of the committee of inquiry into the teaching of mathematics in schools*. London: Her Majesty's Printing Office.

Desjardins, R., Murray, S., Clermont, Y., & Werquin, P. (2005). *Learning a living: First results of the Adult Literacy and Life Skills Survey.* Ottawa: Statistics Canada.

Dossey, J.A. (1997). Defining and measuring quantitative literacy. In L.A. Steen (Ed.), *Why numbers count: Quantitative literacy for tomorrow's America*. New York: College Entrance Examination Board.

Ernest, P. (2004). Relevance versus utility: Some ideas on what it means to know mathematics. In Clarke, B., Clarke, D. M., Emanuelsson, G., Johansson, B., Lambdin, D. V., Lester, F. K., Wallby, A., & Wallby, K. *International Perspectives on Learning and Teaching Mathematics*, Göteborg, The Netherlands: National Centre for Mathematics Education, Göteborg University.

European Commission (1996). *White paper on education and training: Teaching and learning towards the learning society*. Luxembourg: Office for Official Publications of the European Commission.

Expert Group on Future Skills Needs (2007). Responding to Ireland's growing skill needs: the second report of the expert group on future skills needs. Belfast: Author.[online: www.skillsireland.ie/press/reports/pdf/egfsn0003_2nd_skills_report.pdf]

Fey, J. T. (1990). Quantity. In L. A. Steen (Ed.) *On the shoulders of giants: New approaches to numeracy*. Washington, DC: National Academy Press.

Fitzsimons, G. (2005). Numeracy and Australian workplaces: Findings and implications. *Australian Senior Mathematics Journal, 19*(2), 27-42.

Forman, S. L. & Steen, L. A. (1999). *Beyond eight grade: Functional mathematics for life and work*. University of California, Berkeley: National Center for Research in Vocational Education.

Frankenstein, M. (1989). *Relearning mathematics: A different third 'R' – Radical maths*. London: Free Association Books.

Gal, I. (1997). Numeracy: Imperatives of a forgotten goal. In L. A. Steen (Ed.), *Why numbers count: quantitative literacy for tomorrow's America* (pp. 36-44). New York: The College Board.

Gal, I. (1999). Links between literacy and numeracy. In D. A. Wagner, B. Street, and R. L. Venezky (Eds.), *Literacy: an international handbook*. Boulder, Colorado: Westview Press.

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

Gal, I. (2002b). Dispositional aspects of coping with interpretive numeracy tasks. *Literacy and Numeracy Studies*, 11(2), 47-61.

Gal, I. (2007). Assessment of adult numeracy in PIAAC: A conceptual and development framework. (Unpublished manuscript prepared for OECD). Haifa, Israel: University of Haifa.

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: Statistics Canada.

Gatta, M., Appelbaum, E., Boushey, H. (2007). High-touch and here-to-stay: Future skills demands in low wage service occupations. (Paper commissioned for a workshop on research evidence related to future skills demands). Washington, DC: National Academies Center for Education. [online: www7.nationalacademies.org/CFE/Future_Skill_Demands]

Gillespie, J. (July 2004). The "Skills for Life" national survey of adult numeracy in England. What does it tell us? What further questions does it prompt? Paper presented at ICME-10, the 10th International Congress on Mathematics Education (Copenhagen, Denmark).

Gillespie, J. (2007). National surveys of adult numeracy: Contributions from computer-based task settings and modes of response. (Discussion paper following the Skills For Life National Survey in England). Unpublished manuscript. Author.

Ginsburg, L., Manly, M., & Schmitt, M. J. (December 2006). *The components of numeracy.* (NCSALL Occasional Paper). Harvard Graduate School of Education, Cambridge, MA: National Center for the Study of Adult Learning and Literacy. [online: www.ncsall.net]

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.

Hagedorn, L., Newlands, J., Blayney, I., & Bowles, A. (2003). *Frameworks for adult numeracy education: A survey and discussion*. Ontario, Canada: National Literacy Secretariat. [online: http://www.nald.ca/fulltext/framewrk/framewrk.pdf]

Heuvel-Panhuizen, M. van den, and Gravemeijer, K. P. E. (1991). Tests are not all that bad: An attempt to change the appearance of written tests in mathematics instruction at the primary school level. In L. Streefland (Ed.), *Realistic mathematics education in primary school*. The Netherlands: Utrecht.

Hoyles, C., Wolf, A., Molyneux-Hodson, S., & Kent, P. (2002). *Mathematical skills in the workplace. Final report to the Science, Technology and Mathematics Council.* Institute of Education, University of London. [www.basic-skills.co.uk/site/page.php?p=735]

Johnston, B. (1994, Summer). Critical numeracy? *Fine print*, Vol.16 (4), 7-12.

Johnston, B., & MacGuire, T. (2005). *Adult numeracy: policy and practice in global contexts of lifelong learning*. Melbourne: Adult Literacy and Numeracy Australian Research Consortium. [online: www.staff.vu.edu.au/alnarc/revamp/publications]

Jones, S. (1995). The distribution of literacy. In *Literacy, economy and society*. Ottawa: Statistics Canada.

Jones, S. (2006). *Designing Numeracy in PIAAC.* (Background paper prepared for the OECD-Canada Expert Technical Workshop on Numeracy, Ottawa, November 10, 2006). Ottawa, Canada.

Karoly, L. A. (Feb. 2007). *Forces shaping the future U.S. workforce and workplace: implications for 21st century work.* (Testimony presented before the House Education and Labor Committee). Report CT-273. Arlington, VA: Rand Corporation [online: www.rand.org/pdfrd/pubs/testimonies/ct273/]

Kilpatrick, J. (2001). Understanding mathematical literacy: The contribution of research. *Educational Studies in Mathematics 47*(1), 101–116.

Kilpatrick, J., Swafford, J., Findell, B. (Eds.)(2001). Mathematics Learning Study Committee, National Research Council: Conclusions and recommendations. In *Adding It Up: Helping Children Learn Mathematics* (pp. 407-432). Washington, DC: National Academies Press.

Kindler, Jan, et al. (1996). *Certificates in General Education for Adults*. Victoria, Melbourne: Adult, Community and Further Education Board.

Kirsch, I. S., Jungblut, A., & Mosenthal, P. B (1998). The measurement of adult Literacy. In S. T. Murray, I. S. Kirsch, & L. B. Jenkins (Eds.), *Adult literacy in OECD countries: Technical report on the first International Adult Literacy Survey* (pp. 105-134). Washington, DC: National Center for Education Statistics, U.S. Department of Education.

Lave, J. (1998). Cognition in practice: Mind, mathematics, and culture in everyday life. Cambridge, UK: Cambridge University Press.

Lester F. K. (Ed.)(2007). *Second handbook of research on mathematics teaching and learning (A project of the National Council of Teachers of Mathematics)*. NY: Information Age Publishers.

Lindenskov, L. & Wedege, T. (2001). *Numeracy as an Analytical Tool in Mathematics Education and Research*. Roskilde, Denmark: Centre for Research in Learning Mathematics. [online: http://www.statvoks.no/emma/index_numeracy_tine.htm]

Maguire, T., & O'Donoghue, J. (2003). Numeracy concept sophistication - an organizing framework, a useful thinking tool. In J. Maass & W. Schlöglmann (Eds.), *Learning mathematics to live and work in our world. ALM-10. Proceedings of the 10th international conference on Adults Learning Mathematics, Strobl, Austria* (pp. 154-161). Linz, Austria: ALM and Johannes Kepler Universität Linz.

Marr, B., & Hagston, J. (2007). *Thinking beyond numbers: Learning numeracy for the future workplace*. Adelaide, South Australia: National Centre for Vocational Education Research (NCVER). [online: www.ncver.edu.au/publications/1795.html]

McLeod, D. B. (1992). Research on affect in mathematics education: A reconceptualization. In Grouws, D.A. (Ed.), Handbook of research on mathematics teaching and learning (pp. 575-596). New York: Macmillan.

Moore, D. S., & Cobb, G. W. (2000). Statistics and mathematics: Tension and cooperation. *American Mathematical Monthly, 107*(7), 615-630.

Mullis, I. V. S., Martin, M. O., Beaton, A. E., Gonzalez, E. J., Kelly, D. L., & Smith, T. A. (1998). *Mathematics and science achievement in the final year of secondary school: IEA’s Third International Mathematics and Science Study (TIMSS)*. Boston: Center for the Study of Testing, Evaluation, and Educational Policy.

Murat, F. (October 2005). *Les compétences des adultes à l’écrit, en calcul et en compréhension orale*. Report #1044). Paris: INSEE. [online: http://www.insee.fr]

Murnane, R., Willett, J., and Levy, F. (1995, March). *The growing importance of cognitive skills in wage determination* (Working Paper No. 50-76). Cambridge, MA: National Bureau of Economic Research.

Murray, T. S. (2006). *Reflections on the rationale for, and measurement of, numeracy in PIAAC.* (Background paper prepared for the OECD-Canada Expert Technical Workshop on Numeracy, Ottawa, November 10, 2006). Ottawa: Statistics Canada.

National Council of Teachers of Mathematics (2000). *Principles and standards for school mathematics*. Reston, VA: Author.

National Research Council (1989). *Everybody counts: A report to the nation on the future of mathematics education*. Washington, DC: National Academy Press.

Neidorf, T. S., Binkley, M., Gattis, K., & Nohara, D. (2006). *Comparing Mathematics Content in the National Assessment of Educational Progress (NAEP), Trends in International Mathematics and Science Study (TIMSS), and Program for International Student Assessment (PISA) 2003*

Assessments. Technical Report. NCES 2006-029. Washington, D.C. National Center for Education Statistics. [online: http://nces.ed.gov/pubsearch].

NRCD (2006). *Programme for the International Assessment of Adult Competencies: An adult numeracy assessment instrument for the UK*. (Background paper prepared for the OECD-Canada Expert Technical Workshop on Numeracy, November 10, 2006, Ottawa). London: National Research and Development Centre for Adult Literacy and Numeracy.

Nunes, T. (1992). Ethnomathematics and everyday cognition. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 557-574). New York: Macmillan.

Oates, T. (1992). Developing and piloting the NCVQ core skill units: An outline of method and a summary of findings (Report 16). London: National Council on Vocational Qualifications (NCVQ).

OECD (2002). Definition and selection of competencies: conceptual and theoretical foundations strategy paper. (Report DEELSA/ED/CERI/CD(2002)9). Paris: Author.

OECD (Nov. 2006). Planning for the direct assessment of literacy competencies in the information age for PIAAC. (Report EDU/EC/CERI(2006)25). Paris, Author.

Paulos, J. A. (1995). *A mathematician reads the newspaper*. New York: Anchor Books/Doubleday.

PISA (2006). *Assessing scientific, reading and mathematical Literacy: A framework for PISA 2006*. Paris: Author. [www.pisa.oecd.org]

Presmeg, N. G. (2007). The role of culture in teaching and learning mathematics. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning*. NY: Information Age Publishers.

Queensland Department of Education (1994). Unpublished manuscript. Australia.

Resnick, L. B. (1987). Learning in school and out. *Educational Researcher*, *16,* 13-20.

Robitaille, D. F., and Travers, K. J. (1992). International studies of achievement in mathematics. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 687-709). New York: Macmillan.

Rogoff, B., & Lave, J. (Eds.)(1984). *Everyday cognition: Its development in social context*. Cambridge, MA: Harvard University Press.

Rutherford, F. J. & Ahlgren, A. (1990). *Science for all Americans*. New York: Oxford University Press.

Rychen, D. S., & Salganic, L. H. (Eds) (2003). Key competencies for a successful life and a well-functioning society. Gottingen, Germany: Hogrefe & Huber.

Saxe, G. B. (1992). Culture and cognitive development: Studies in mathematical understanding. Hillsdale, NJ: Lawrence Erlbaum Associates.

Saxe, G. B., Dawson, V., Fall, R., & Howard, S. (1996). Culture and children's mathematical thinking. In Sternberg, R, and Ben-Zeev, T. (Eds.), *The nature of mathematical thinking* (pp. 119-144). Hillsdale, NJ: Lawrence Erlbaum Associates.

Schliemann, A. D., and Acioly, N. M. (1989). Mathematical knowledge developed at work: The contribution of practice versus the contribution of schooling. *Cognition and Instruction*, 6, 185-221.

Secretary's Commission on Achieving Necessary Skills (1991). *What work requires of schools: A SCANS Report for America 2000*. Washington, DC: U.S. Dept. of Labor.

Senechal, M. (1990). Shape. In L.A. Steen (Ed.) On the shoulders of giants: New approaches to numeracy. Washington, DC: National Academy Press.

Service Skills Australia (2005). Profiles of Communication and Maths (Literacy and Numeracy) skills for industries and employees covered by the Service Skills Council. Canberra, Author.

Statistics Canada and Organisation for Economic Cooperation and Development (1996). *Literacy, economy, and society: First results from the International Adult Literacy Survey.* Ottawa: Author.

Statistics Canada and Organisation for Economic Cooperation and Development (1997). *Literacy skills for the knowledge society: Further results from the International Adult Literacy Survey*. Ottawa: Author.

Steen, L. A. (2001). *Mathematics and democracy: The case for quantitative literacy.* Washington, DC: Woodrow Wilson National Fellowship Foundation.

Steen, L. A. (Ed)(1990). *On the shoulders of giants: New approaches to numeracy*. Washington, D.C.: National Research Council.

Steen, L. A. (Ed)(1997). *Why numbers count: quantitative literacy for tomorrow's America*. New York: The College Board.

Stein, S. G. (1995). *Equipped for the future: A customer-driven vision for adult literacy and lifelong learning.* Washington, D.C.: National Institute for Literacy.

Straesser, R. (July 2003). Mathematics at work: Adults and artefacts. In J. Maasz & W. Schloeglmann (eds)., Learning mathematics to live and work in our world: Proceedings of the 10th international conference on adults learning mathematics (pp. 30-37). Linz, Austria: Johannes Kepler Universitat Linz.

Tobias, S. (1993). *Overcoming math anxiety*. New York: Norton.

Tout, D. (2006). *Review of Numeracy component of PIAAC.* (Background paper prepared for the OECD-Canada Expert Technical Workshop on Numeracy, Ottawa, November 10, 2006). Melbourne, Australia: Centre for Adult Education.

UNESCO (1997). *International standard classification of education: ISCED*. Paris: Author.

Utts, J. (2003). What educated citizens should know about statistics and probability. *The American Statistician*, 57(2), 74-79.

Watson, J. M., & Callingham, R. (2003). Statistical literacy: A complex hierarchical construct. *Statistics Education Research Journal, 2*(2), 3-46.

Wedege, T. (July 2003). Sociomathematics: Researching adults' mathematics at work. In J. Maasz & W. Schloeglmann (eds)., *Learning mathematics to live and work in our world: Proceedings of the 10th international conference on adults learning mathematics* (pp. 38-48). Linz, Austria: Johannes Kepler Universitat Linz.

Williams, J. & Wake, G. (2007). Black boxes in workplace mathematics. *Educational Studies in Mathematics*, 64, 317–343.

Willis, S. (1990). *Being numerate: What counts?* Victoria: Australian Council for Educational Research (ACER).

Willms, J. D. (2006). Policy imperatives for the international measurement of literacy *and numeracy competencies in working-age populations*. (Background paper prepared for the OECD-Canada Expert Technical Workshop on Numeracy, Ottawa, November 10, 2006). Canada: University of New Brunswick.

FACTORS AFFECTING COMPLEXITY OF NUMERACY ITEMS

197. This appendix describes a scheme of factors that account for the difficulty of different numeracy assessment tasks. The scheme was developed for the Adult Literacy and Lifeskills survey by the ALL Numeracy team. The following text is copied (with permission) from the ALL Technical Report, see Gal et al., 2005. This scheme was found useful to inform item development, i.e., help in the creation of items that spread over a range of difficulty levels. Results from the ALL pilot study showed that predicted difficulty of items used by the scheme described below was highly correlated with observed difficulty (r = 0.79). Because of the recursive nature of the testing of this scheme (e.g., the same individuals wrote the scheme and rated the complexity of items), caution should be exercised in further interpretive use of the present version. While further validation is needed, the scheme in its current state nonetheless appears to also be a possible useful tool for interpretation of testing results.

Previous research on task complexity

198. In IALS, three factors were found to be the principal components of task difficulty (regarding literacy or text-based tasks): plausibility of distractors, type of match required, and type of information required. The difficulty of the Quantitative Literacy tasks appeared to be a function of several other factors:

1. The particular arithmetic operation required to complete the task
2. The number of operations needed to perform the task
3. The extent to which the numbers are embedded in printed materials
4. The extent to which an inference must be made to identify the type of operation to be performed (i.e. problem transparency; see below)

199. The IALS QL difficulty factors overall fit those used in large-scale assessments of mathematical skills (with children), which often make use of three or four factors:

1. *The mathematical concepts involved:* number systems and number sense, spatial and geometrical topics, functions and algebra, chance/statistics topics, etc. Concepts that are related to topics taught in lower grades are considered easier.
2. *The complexity of operations*: addition, subtraction, multiplication, and division, as well as dealing with whole numbers, with decimals, and with percents. Operations that are related to topics taught in lower grades are considered easier.
3. *The number of operations*: one-step problems are considered easier than multi-step problems.
4. *Problem transparency*: This factor is sometimes relevant; it refers to the extent to which the problem situation includes clearly identified numbers or entities and the extent to which it is clear what operations or actions to perform. To the extent that

these are not clear or transparent, respondents have to extract needed information by applying comprehension and inference strategies, making the task more complex.

200. There are other adult-related assessment projects on which to draw to develop the levels of complexity. Both the Essential Skills Research Project and the Applied Numeracy sub-test of the Work Keys test battery (American College Testing, 1997) use a two-factor model of complexity in their description of numeracy levels. The first factor, "operations required;" is seemingly straightforward and refers to the difficulty of operations called for. However, this is complicated by the level of difficulty of the numbers being manipulated: computations that include fractions and decimals are usually more difficult than those with whole numbers.

201. The Essential Skills model spells out two sequences of complexity on this factor: *Operations* and *Translation* of information (sometimes called 'problem transparency').

Operations:

1. Only the simplest operations are required and the operations to be used are clearly specified. Only one type of mathematical operation is used in the task.
2. Only relatively simple operations are required. The specific operations to be performed may not be clearly specified. Tasks involve one or two types of mathematical operation. Few steps of calculations are required.
3. Task may require a combination of operations or multiple-applications of a single operation. Several steps of calculation are required. (More advanced operations may call for multiplication or division.)
4. Tasks involve multiple steps of calculation.
5. Tasks involve multiple steps of calculation. Advanced mathematical techniques may be required (e.g., percents, ratios, proportions).

Translation (Problem Transparency)

1. Only minimal translation is required to turn the task into a mathematical operation. All the information required is provided.
2. Some translation may be required or the numbers needed for the solution may need to be collected from several sources. Simple formulae may be used.
3. Some translation is required but the problem is well defined.
4. Considerable translation is required.
5. Numbers needed for calculations may need to be derived or estimated; approximations may need to be created in cases of uncertainty and ambiguity. Complex formulae, equations or functions may be used.

202. Two considerations prompted us to question the appropriateness of using mathematics-related frameworks (from Essential Skills or elsewhere) as the sole source for development of a complexity scheme for items assessing *adults'* ability to cope with real-world numeracy tasks. First, effective coping with many real-world quantitative problems depends upon people's ability to make sense of and

interact with different types of texts. This is hardly recognized by the Essential Skills model. Hence, it was essential to add difficulty factors that acknowledge the inherent links between literacy and numeracy, quite similar to those used in IALS.

203. Another, albeit a more restricted consideration, is that the ordering of complexity of tasks by the type of operation performed may not be as clear with adults as it may be with children. Such ordering in school-based assessments is predicated on traditional school curricula, where more advanced topics are learned at higher grades. However, adults are known to use a lot of invented strategies, perhaps more so, and more efficiently so, than children. Multiplication or division problems, which can prove relatively hard for some young people, may be solved by seemingly simpler strategies, such as by repeated addition or repeated subtraction; complex numbers may be broken down in ways that ease mental load, and so forth. In addition, adults' familiarity with everyday contexts, such as with monetary entities, facilitates their performance with some seemingly advanced concepts. For example, specific benchmark values of fractions and percents, such as 1/2, 1/4, 50%, or 25%, are familiar to many people; as a result, they may be easier to manage than expected, violating curriculum-based ordering of difficulty. Hence, an *overall complexity level* has to be used, in order to weight these "inconsistencies" in ordering of difficulty levels proposed in other schemes.

Complexity factors in the ALL survey

204. The above literature review suggests that a framework of factors affecting the complexity of numeracy tasks should not only address factors related to the numerical and textual aspects of tasks, but should also address other issues. It should treat separately the number of operations and the type of operations from the *type of mathematical (or statistical) information to be processed*, which may involve numbers explicitly but also other types of mathematical information. In so doing, the desired framework of complexity factors should take into account the broad scope of the definition of numeracy, i.e., reflect the variation within contexts, the range of mathematical ideas/content, the types of possible responses, and the types of representations that cut across adult life contexts.

Table 2: Complexity Factors—Overview

Aspects	Category	Range
Textual aspects	1. Type of match/problem transparency	Obvious/explicit to embedded/hidden
	2. Plausibility of distractors	No distractors to several distractors
Mathematical aspects	3. Complexity of Mathematical information/data	Concrete/simple to abstract/complex
	4. Type of operation/skill	Simple to complex
	5. Expected number of operations	One to many

205. With the above considerations in mind, five key factors have been identified that are predicted to affect, separately and in interaction, the difficulty level of numeracy tasks to be used in the ALL survey. These five "complexity factors" are outlined in Table 2 and are organized in two sets: two factors that address mainly textual aspects of tasks, and three factors that address the mathematical aspects of tasks. These five factors are listed separately for clarity of presentation, but in actuality are *not* independent of each other and do interact in complex ways. Each factor is examined in some detail below, followed by a later subsection that describes the calculation of an overall complexity level for each item, taking into account all five factors.

Type of Match/Problem Transparency.

206. This is a combination of the factor of Problem Transparency outlined above, and of an IALS factor called Type of Match. Problem Transparency is a function of how well the mathematical information and tasks are specified and includes aspects such as how apparently the procedure is set out, how explicitly the values are stated, etc. Type of Match refers to the process that a respondent has to use to relate the requested action in the question to the information in the task or text, which can range from a simple action of locating or matching to more complex actions that require the respondent to perform a number of searches through the information given. This measure of complexity for a numeracy task incorporates the degree of text embeddedness of the mathematical information.

207. In easy tasks, the type of information (e.g., numerical values) and the operations needed are apparent and obvious from the way the situation is organized. In more difficult ones, the values must be located or derived from other values; the operations needed may have to be discovered by the performer, depending on his or her interpretation of the context and of the kind of response expected. As well, numeracy situations may involve text to varying degrees, and this text may be of different degrees of importance. There may be a situation where there is little or no text. Some situations may involve pure quantitative information that is to be interpreted or acted upon with virtually no text or linguistic input. In other words, the performer derives all the information needed to respond from the objects present in the situation or from direct numerical displays.

208. At a higher level, some textual or verbal information may be present alongside the mathematical information. The text can provide background information about the problem situation, or some instructions. For example, a bus schedule, cooking instructions, and a typical school-type word problem all involve some text and some numbers. Still other situations would be heavily text-based or may not involve any numbers or mathematical symbols at all, just plain text. The task will contain mathematical or statistical information that a person needs to understand and, in some cases, act upon,

but it will be much less transparent. It may be heavily embedded in dense text or may require using information from a number of sources within or even outside the text/task.

209. This factor requires that a task will be analyzed in terms of the questions: How difficult is it to identify and decide what action to take?, and How many literacy skills are required?

Plausibility of distractors.

210. This variable is literacy related, even though it can involve mathematical components. In general, literacy tasks are easiest to process when there are no plausible distractors in the text, that is, there is no other information in the text that meets any of the requirements of the task. At higher levels of difficulty, tasks can involve irrelevant information both within the question as well as within the text. In terms of mathematical information, a low level of plausible distractors would mean that no other mathematical information was present apart from that requested, making the numbers or data required easy to identify. At a higher level, there may be either some other mathematical information in the task (or its text) that could be a distractor, or the mathematical information given or requested could occur in more than one place. A higher level of complexity could also mean that outside information (e.g. the knowledge of a formula) may be needed to answer the question.

211. This factor requires that a task will be analyzed in terms of the questions: How many other pieces of mathematical information are present?, and Is all the necessary information there?

Complexity of Mathematical Information.

212. Some situations present a person with simple mathematical information, such as concrete objects (to be counted), simple whole numbers, or simple shapes or graphs. At lower skill levels, the information will be more familiar, whereas at higher levels, the information may be less familiar. Situations will be more difficult to manage if they involve more abstract or complex information, such as very large or very small numbers, unfamiliar decimals or percents, information about rates, or dense visual information, as in a diagram or complex table.

213. This factor requires that a task will be analyzed in terms of the question: How complex is the mathematical information that needs to be manipulated or managed?

Type of Operation/Skill.

214. Some situations require simple operations, such as addition or subtraction, or simple measurement (e.g., finding the length of a shelf), or recognition of shape. These are usually easier to analyze mathematically than situations that require multiplication or division, and than situations that require using exponents. While the difficulty of recognizing and carrying out the operation implied by a situation (be it additive, multiplicative, etc.) has direct bearing on task complexity, there may be exceptions that occur when alternative approaches are obvious. There are some tasks that combine both interpretive and generative skills and may involve a deeper conceptual understanding than merely carrying out a procedure. Other more complex tasks may involve an explanation of one's reasoning. The interpretation of information appearing in graphs, for example, becomes more complex if comparisons, conjecturing, or "reading beyond the information given" is required.

215. This factor requires that a task will be analyzed in terms of the question: *How complex is the mathematical action that is required?*

2.5 Expected Number of Operations.

216. Tasks that require acting upon the mathematical information given may call for one application (step) of an operation, or for one action (e.g., literal reading of information in a table, or measurement). More complex tasks will demand more than one operation, which may be the same or similar to one another, such as the steps involved in multiple passes on the data or text. Still more complex tasks are those that involve the integration of several different operations.

217. This factor requires that a task will be analyzed in terms of the question: *How many steps and types of steps are required?*

Overall Complexity Level

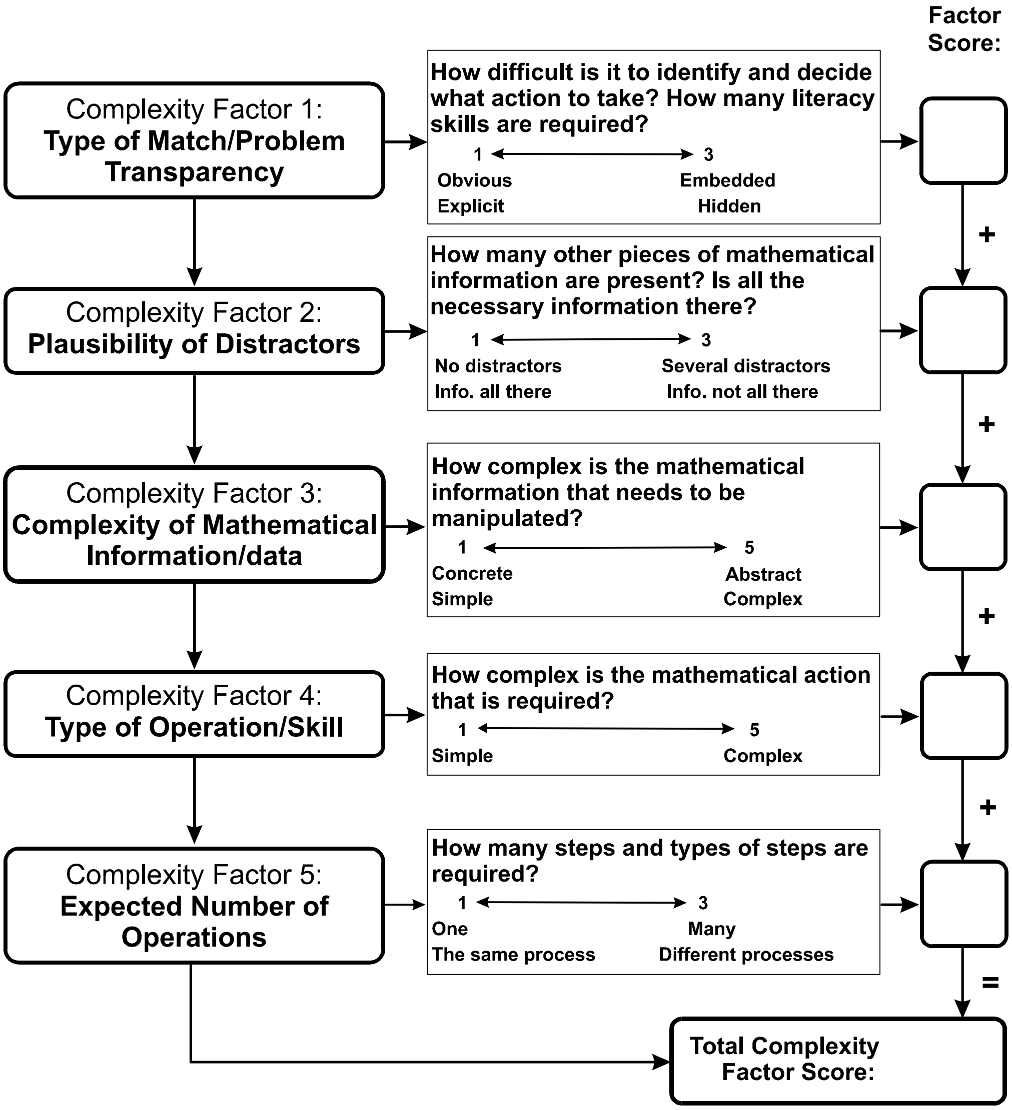
218. It is possible to estimate the overall difficulty level of a specific item by first scoring the item on each of the five factors of complexity, according to the levels described in Appendix 2, and then summing together the scores for each factor. Figure 1 below explains the process; Appendix 2 describes each level of the five factors in detail. The total summary score can range between 5 (easiest) and 19 (most difficult).

219. The estimation process outlined in Figure 1 suggests that each factor has a separate contribution to an item's overall difficulty or complexity. However, it can be hypothesized that as tasks become more complex, actual performance on items may increasingly depend not only on each factor by itself, but also on the interplay or interaction between them. Hence, the computational process suggested in Figure 1 can provide only approximate information about an item's anticipated difficulty level.

220. Further, the difficulty of a task cannot in some cases be predicted without taking into account characteristics of the person who interacts with the task. The same task may be more difficult for some individuals and less difficult for other individuals, depending on factors such as their familiarity with the context in which a task is situated, knowledge of formal mathematical notations, background world knowledge, as well as general literacy, problem-solving, and reasoning skills. For example, it could be predicted that a task that involves the composition of a fertilizer would be more difficult for an urban apartment dweller than for a rural farmer whereas a task that uses a bus schedule would be more difficult for the farmer. For the above reasons, the prediction of the difficulty of a task in isolation of detailed knowledge about the respondent himself can only be an estimate.

221. Despite the above limitations, the scheme of complexity factors developed for numeracy assessment in ALL comprises a theoretical contribution. It provides a conceptual basis for predicting the different levels of complexity of a broader range of items well beyond those involving arithmetic operations only. Indeed, this scheme was highly correlated with observed difficulty (r = 0.79). Because of the recursive nature of the testing of this scheme (e.g., the same individuals wrote the scheme and rated the complexity of items), caution should be exercised in further interpretive use of the present version. While further validation is needed, the scheme in its current state nonetheless appears to also be a possible useful tool for interpretation of testing results.

Figure 1. Complexity Flow chart



Scoring for each of the Complexity Factors

Complexity Factor 1. Type of match/Problem transparency How difficult is it to identify and decide what action to take? How many literacy skills are required?		
Score 1	**Score 2**	**Score 3**
In the question and the 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	In the question and the stimulus, the information, activity or operation required: - is given using clear, simple sentences and/or visualizations where some translation or interpretation is required - is located within a number of sources within the text/activity. - fairly closed question	In the question and the stimulus, the information, activity or operation required: - is embedded in text where considerable translation or interpretation is required and/or - may need to be derived or estimated from a number of sources within or outside the text/activity and/or - the information or action required is not explicit or specified - more complex, open-ended task

Complexity Factor 2. Plausibility of distractors How many other pieces of mathematical information are present? Is all the necessary information there?		
Score 1	**Score 2**	**Score 3**
- no other mathematical information is present apart from that requested—no distractors	- there is some other mathematical information in the task that could be a distractor - the mathematical information given or requested can occur in more than one place - may need to bring to the problem simple information or knowledge from outside the problem.	- other irrelevant mathematical information appears - mathematical information given or requested appears in several places. - necessary information or knowledge is missing, so outside information or knowledge needs to be brought in

Complexity Factor 5. Expected number of operations How many steps and types of steps are required?		
score 1	**score 2**	**score 3**
one operation, action or process	application of two or three steps, the same or similar operation, action or process	integration of several steps covering more than one different operation, action or process

Complexity Factor 3. Complexity of mathematical information/answer required

How complex is the mathematical information that needs to be manipulated?

score 1	score 2	score 3	score 4	score 5
Context Based on very concrete, real life activities, familiar to most in daily life.	Based on common, real life activities.	Based on real life activities, but less often encountered.	Based on real life activities but unfamiliar to most	Based on abstract ideas or unfamiliar activity in a context new to most.
Quantity Whole numbers to 1,000 Fractions, decimals, percents - benchmark fractions ($^{1}/_{2}$, $^{1}/_{4}$, $^{3}/_{4}$) - decimal fraction for a half only (0.5) and equivalent as a percentage (50%)	- large whole numbers including millions - other benchmark fractions, like $^{1}/_{3}$ and $^{1}/_{10}$ - common decimals, like 0.1, 0.25 to 2 decimal places - common whole number percents, like 25% and 10%.	- large whole numbers including billions - other fractions - decimals to 3 decimal places (other than money) - all whole number percents	- negative integers - all remaining fractions, decimals and percentages	- all remaining types of rational (and some irrational) numbers including directed numbers
Pattern and relationship - very simple whole number relations and patterns	- simple whole number rates and ratios - whole number relations and patterns	- rates and ratios - relations and patterns including written everyday generalizations	- complex ratios, relations, patterns - simple formula	- formal mathematical information such as more complex formulae, knowledge of relationships between dimensions or variables, etc
Measures/ Dimension/Space - standard monetary values - common everyday measures for length (whole units) - time (dates, hours, minutes) - simple, common 2D shapes - simple localised maps or plans (no scales)	- everyday standard measures for length, weight, volume , including common fraction and decimal units - common 3D shapes and their representation via diagrams or photos - common types of maps or plans with visual scale indicators	- other everyday measures (area included) including fraction and decimal values - more complex 2D and 3D shapes, or a combination of 2 shapes - area and volume formulae - common types of maps or plans with ratio type scales	- all kinds of measurement scales - complex shapes or combinations of shapes	
Chance/Data - simple graphs, tables, charts with few parameters and whole number values - simple whole number data or statistical information in text	- graphs, tables, charts with common data including whole number percents—whole number scales in 1s, 2s, 5s or 10s - data or statistical information including whole number percents	- graphs, tables, charts with more complex data (not grouped data) - more complex data or statistical information including common average, chance and probability values - scales: more complex whole number, fractional or decimal	- complex graphs, tables or charts including grouped data - complex data or statistical information including probabilities, measures of central tendency and spread	

Complexity Factor 4. Complexity of Type of operation/skill

How complex is the mathematical action that is required?

score 1	score 2	score 3	score 4	score 5
Communicate no explanation - a single simple response required (orally, or in writing)	- no explanation - a simple response required (orally, or in writing	- simple explanation of a (level 1 or 2) mathematical process required (orally, or in writing) -	- explanation of a (level 3) mathematical process required (orally, or in writing)	- complex, abstract and generative reasoning or explanation required
Compute - a simple arithmetical operation (+, -, x, ÷) with whole numbers or money	- calculating common fraction, decimal fraction and percentages of values - using common rates (e.g. $/lb.); time calculations; etc - changing between common equivalent fraction, decimal and percent values, including for measurements e.g. $^1/_4$ kg = 0.250kg	- more complex applications of the normal arithmetical operations such as calculating with fractions and more complex rates, ratios, decimals, percentages, or variables - simple probability calculations	- applications of other mathematical operations such as squares, square roots, etc	- more advanced mathematical techniques and skills e.g. trigonometry
Estimate	- estimating and rounding off (when requested) to whole number values or monetary units	- estimating and rounding off to requested number of decimal places	- making a contextual judgment re whether a found answer is realistic or not and changing the answer to the appropriate correct rounded (but not necessarily mathematically correct) answer.	
Use formula/ model	- evaluating a given formula involving common operations (+, -, x, ÷)		- developing/creating and using straight forward formulae - using strategies such as working backwards or backtracking (e.g. 15% of ? = $255)	- generative reasoning - using and interpreting standard algebraic and graphical conventions and techniques
Measure - knowing common straight forward measures - naming, counting, comparing or sorting values or shapes	- visualizing and describing shapes, objects or geometric patterns or relationships - making and interpreting standard measurements using common measuring instruments	- using angle properties and symmetry to describe shapes or objects - estimating, making and interpreting measurements including interpolating values between gradations on scales - converting between standard measurement units within the same system	- calculating measures of central tendency and spread for non-grouped data - converting between non-standard measurement units within the same system - counting permutations or combinations	- converting between measurements across different systems

Interpret - locating/identifying data in texts, graphs and tables - orientating oneself to maps and directions such as right, left, etc	- reading and interpreting data from texts, graphs and tables - following or giving straight forward directions	- interpolating data on graphs - calculating distances from scales on maps	- generating, organising, graphing non-grouped data - extrapolating data - reading and interpreting trends and patterns in data on graphs, including slope/gradient	- graphing grouped data - calculating measures of central tendency and spread for grouped data

ANNEX 3: PROBLEM SOLVING IN TECHNOLOGY RICH ENVIRONMENTS

Introduction

Structure of the present document

222. This document introduces the domain of problem solving in technology rich environments in PIAAC and explains how the domain will be assessed. The document contains four parts. Part 1 discusses the development of information technologies and their uses in today's societies. Part 1 also introduces various constructs of problem solving, especially as they apply in technology rich environments. Part 2 provides a definition of problem solving in technology rich environments and elaborates on some of the assumptions underlying this definition. Part 3 presents the core dimensions that will be represented in the tasks to be included in the assessment. Part 3 also contrasts the new problem solving in technology rich environments (PS-TRE) domain with the other domains assessed in PIAAC, as well as with other constructs of technology knowledge and use. Finally, Part 4 reviews some practical aspects of the assessment, especially regarding the evaluation of problem solving strategies in the context of digital technology.

Impact of digital technologies on instruction and assessment

223. During the past decades, digital technologies have deeply transformed the way individuals learn, communicate, work, and more generally the way they function in societies. Shopping, traveling, and interacting with administrations and services, to cite only a few examples, more and more routinely involve the use of digital technologies. Over just two decades, microcomputers, laptops, mobile phones, and the Internet have provided users with powerful tools to search for and make use of immense repertories of information and services. Increasingly versatile mobile technologies allow users to stay connected almost regardless of where they are and what they are doing. And the integration of digital tools in homes, cars and appliances potentially increases the safety, flexibility, and effectiveness of many activities of everyday life. Navigating in unfamiliar environments, preventing domestic accidents, or encouraging the rational use of energy and water resources are just a few concrete benefits that can be expected from technological advances.

224. Thus, early visionaries' dreams about a universal "information society" are, to some extent, coming true. But as computer technologies were pervading most areas of human activities, their actual impact on the development of human communities was found to be more ambiguous than initially expected (Forester, 1992). To start with, the spread of information technologies across the world is far from homogeneous. Despite two-digit growth rates reported in many countries at the dawn of the 21st century, the distribution of digital technology uses across geographic and socioeconomic borders remains very uneven. Far from the rapid progress toward universal access envisioned by some a decade ago, as of May 2008, only one human in five had ever used the Internet (http://www.internetworldstats.com/). Personal access to computers and the Internet ranged from about zero in the poorest countries to 70% in the wealthiest ones. But even in those latter countries, access was still unevenly distributed across the socioeconomic spectrum, contributing to the so-called "digital divide" (Norris, 2001).

225. Furthermore, access to computers and high-speed connectivity does not necessarily result in more social participation and well being. As surveys have revealed, heavy Internet use sometimes leads to a greater feeling of loneliness and less genuine interactions with others (Kraut et al., 1998). And, regardless of prior experience with technology, demographic factors such as age, level of income or education often correlate with users' ability to access and make effective use of computers in purposeful contexts (Marquié & Baccarat, 1997; Sweets & Meats, 2004).

226. Using computers or other digital devices to perform personal or work-related activities often presents a challenge for the everyday user. People often have trouble installing, setting up, and learning how to use new digital devices and software applications. Users often confine themselves to a few basic, but ineffective, procedures. Then, even routine computer use for mundane tasks is often prone to errors, delays and incidents. To take just one example, information searches using web sites often present complex problem solving situations for laypersons. Among other reasons, the nonlinear linking of information in hypertexts is often experienced as cumbersome or unpractical, yielding a sense of disorientation and cognitive overload (Rouet, 2006). Studies of information evaluation and integration also suggest that laypersons' source evaluation skills are challenged - and often overwhelmed.

227. More experience on the part of users and improved design practices on the part of developers have been pointed out as important conditions for digital technologies to fulfill their potential. Though necessary, these are probably not sufficient conditions to make the online world accessible to all. There is a growing body of evidence that mere exposure to technology is not sufficient for people to achieve a satisfactory level of skill in purposeful technology-based tasks. So-called computer literacy skills must be integrated with deeper and more abstract problem solving skills (Lazonder & Rouet, 2008). This is because the most relevant uses of computers are not easily handled through the mere application of simple routines (e.g., launching an application and clicking on a set of buttons or links). Instead, to successfully complete computer-based tasks, people must be able to analyze the various requirements of the task, set up appropriate goals and plans, and monitor their progress through the task until the task purposes are achieved. This process seems best captured by the cognitive psychology construct of "problem solving", a construct that has been the subject of extensive research over the past forty years, in traditional as well as technological contexts.

228. Analyzing the problem solving skills involved in the uses of digital technologies for various purposes and properly assessing the distribution of such skills in the general public appear to be two important conditions for the advancement of technology-rich societies. The data can be turned into high return education and training investments. These efforts, however, should be targeted towards relevant portions of the public, relevant educational levels, and relevant skills. The present framework aims to contribute to these important undertakings.

Defining problem solving in technology rich environments

229. In this section, we introduce a definition of problem solving in technology rich environments. We first discuss the notion of problems and problem solving, focusing in particular on "information-rich" problems. Then we provide and explain the definition of PS-TRE to be used in PIAAC.

Problems and problem solving

230. A problem is usually defined as a situation where a person cannot immediately and routinely achieve his or her goals due to some kind of obstacle or difficulty. The ability to solve problems is considered one of the most complex and sophisticated aspects of human cognition (Newell & Simon, 1972). In order to solve a problem, individuals have to first become aware of a difference between the current state of affairs and the state of affairs that would correspond to the satisfaction of their goals. In

other words, they have to come to an understanding of the nature of the problem. This is also called "problem finding". Individuals then have to engage in a series of thinking processes and concrete actions in order to (a) define a set of sub-goals and steps through which the problem may be solved (also called planning or "problem shaping"), and (b) perform the actions required to reach those sub-goals until the situation reaches a satisfactory state. Throughout the problem solving activity, individuals have to monitor their progress and, where necessary, reconsider their goals and actions. For instance, individuals may face an unexpected outcome or find themselves at an impasse. In such cases, they may have to reconsider their understanding of the problem or the actions they have decided to take in order to solve the problem.

231. Problem solving also normally requires a range of tools and information resources. In traditional areas, people may use source documents, paper and pencil, calculator and other devices. In the context of technology-rich environments, people may use Internet-based services such as search engines and Web pages, but also desktop software such as spreadsheets, email, or file management systems in order to solve their problem. Tools and technologies are normally meant to facilitate the resolution of the problem. They may, however, also contribute to making a problem more difficult, especially when a person has limited knowledge and experience with the use of those tools and technologies.

232. In concrete, everyday situations, problems and problem solving often involve other persons. For instance, people may be asked to solve a problem for another person, they may need to get information or advice from another person, or they may want to communicate the solution to another person. The actions needed in order to solve the problem may then include spoken or written interpersonal communication (e.g., comprehending instructions, asking questions or explaining). Thus, communication skills must be considered a factor in assessing problem solving skills. In the context of technology-rich environments several powerful tools for rapid (e.g. mail and chat software) and broad (e.g. blogs, shared applications) ways of communication are available, thereby enabling collaborative problem solving activities by people being at different places. Such tools require special skills for computer-mediated communication (Bromme, Hesse & Spada, 2005).

233. From a cognitive perspective, problem solving involves a complex hierarchy of processes and skills. The core characteristic of problem solving is that it is impossible for a person to achieve the goal through routine actions. In problem solving, one has to reflect on the situation in order to identify the proper arrangement of decisions and actions that may lead to a solution. Thus, the status of problems is conditional upon a person's familiarity with the problem or category of problems. With learning and practice, some activities that were initially experienced as problem solving may become routine activities. Examples include fundamental skills such as reading and performing mental calculations, as well as everyday tasks such as tying one's shoes, replacing a broken light bulb, or installing new software on a personal computer.

234. Regardless of a person's ability level, some problems are intrinsically more complex than others (Funke & Frensch, 2007). Dimensions of problem complexity include the clarity of the initial situation; the number of subgoals and steps needed to solve the problem; the amount of information to be considered; and the pragmatic constraints that surround the person's activity (e.g., time constraints, level of stakes or hazard, probability of unexpected events or outcomes). The complexity of a problem also varies as a function of the arrangement of informational and other resources in the problem-solving environment (i.e., extrinsic cognitive load, Sweller, Chandler, Tierney, & Cooper, 1990). For instance, reducing the distance between two pieces of information that need to be combined (e.g., a diagram and a legend) can make it easier for a person to solve a problem.

235. Research on problem solving has also established distinctions between various types of problems. One important distinction is between closed vs. open problems. In closed problems the amount of resources (e.g., objects, tools) available and the range of possible actions is limited. An example is a

chess game where moves are limited by the size of the chessboard and the rules of the game. In other problems, the potential resources and possible actions are, in principle, unlimited (Goel & Pirolli, 1992). For instance, finding one's way in an unfamiliar city or designing a new kitchen may be considered open problems.

236. Another important distinction is between well-defined and ill-defined problems (Voss & Post, 1988). Well-defined problems come with a set of circumstances that clearly let the person know what they are supposed to do. For instance, "*Using the attached schedule, find a train to go from Paris to Amsterdam on Tuesday, October 15th leaving no earlier than 11 a.m. and arriving no later than 9 p.m.*" would be considered a well-defined problem. In contrast, "*Find a way to go to Amsterdam on Tuesday, October 15th*" is an example of a less well-defined problem. It is important to note, however, that there is no straightforward link between the definition of a problem and its absolute level of difficulty. Sometimes ill-defined problems are easier to solve because they allow several solution paths. But ill-defined problems also require the problem solver to set up appropriate subgoals and operators, and to select appropriate resources, which may increase the difficulty of the problem.

Information problems

237. The distinction between information-rich and information-lean problems deserves particular attention in the context of PIAAC. Because digital technologies are primarily aimed at storing, processing, representing, and communicating symbolic information, the types of problems that will be used in the PIAAC problem solving assessment clearly belong in the first category. This contrasts, for instance, with logical or mathematical problems where the complexity lies in the computational reasoning, not so much in the information to be accessed and used. The PS-TRE domain of PIAAC is intended to cover the specific class of problems people deal with when using information and communication technologies (ICT). Those problems share the following characteristics.

2. The existence of the problem is primarily a consequence of the availability of new technologies. One example relates to the vast amount of information now available on the World Wide Web. The Web has enlarged the access to specialist knowledge for laypersons. This gives rise to problems related to locating and evaluating information for quality and credibility, e.g., when seeking advice about legal issues or medical conditions (Stadtler & Bromme, 2007). Evaluation and critical judgment of information are core aspects of literate Internet use (Gilster, 1997) and will be one focus of the PS-TRE assessment. Other examples include the increasing capacity of electronic storage devices, with the subsequent problems of organizing and sorting large numbers of files; or the growing practice of social communication on the Web, with the subsequent problem of learning and making use of new social norms as regards private vs. public information.

3. The problem solution requires the use of computer-based artifacts (tools, representational formats, computational procedures) that were not available previously or at least not available to the general public. An example is the management of personal finance using spreadsheets, statistical packages and graphical tools. Here the problem itself may not be new (i.e., keeping spending in balance with income) but the new artifacts modify the distribution of work across social agents (professional vs. laypersons) and they deeply transform the procedures and steps required to solve the problem.

4. The problems are related to the handling and maintenance of technology rich environments themselves (e.g., how to operate a computer, how to fix a settings problem, how to use the Internet browser in a technical sense).

238. Most of the problems that correspond to those broad characteristics require one to handle vast amounts of symbolic information. Therefore, they require an ability to deal with semantic content or meaning. Examples include understanding command names in drop down menus, the naming of files and folders, hits in a search engine, or links in a web page. Furthermore, many problems require the person to read and understand electronic texts, graphics and numerical data. Therefore, understanding and evaluating meaningful information available in technology rich environments is central in the construct of PS-TRE.

Definition

239. In the context of the PIAAC survey, problem solving in technology rich environments is defined as follows:

> *"Problem solving in technology rich environments involves using digital technology, communication tools and networks to acquire and evaluate information, communicate with others and perform practical tasks. The first PIAAC problem solving survey will focus on the abilities to solve problems for personal, work and civic purposes by setting up appropriate goals and plans, accessing and making use of information through computers and computer networks."*

240. The two sentences in the definition each serve a specific purpose. The first sentence is aimed at providing a broad basis for the first as well as subsequent surveys of PS-TRE. The second sentence acknowledges some constraints that limit the scope of the first survey. We provide below a series of more specific comments on the words and phrases used in this definition.

"using digital technology, communication tools and networks"

241. PIAAC focuses on problems that are specifically related to the use of ICT. The problem solving context means that routine or basic ICT skills will not be central to the framework. Instead, PS-TRE will focus on situations that involve the active construction of goals and strategies on the part of the user. We also acknowledge the increasing diversity and versatility of digital technologies, and we emphasize that a proper assessment of PS-TRE should not be limited to traditional desktop computing. Instead, we envision that mobile and integrated technologies may be involved in new types of problem solving that will need to be represented in future assessments.

"to acquire and evaluate information"

242. This phrase acknowledges that most uses of digital technologies involve the use of symbolic information, such as texts, graphics, links, and commands. Symbolic information is used as part of human-computer interfaces (e.g., icons, commands) and it constitutes the primary content of most computer applications (e.g., word processor, spreadsheet, Internet Browser, and email applications).

"communicate with others"

243. An important role of digital technologies is to provide powerful and flexible means for people to communicate with each other. Examples include email, chats, short message systems, and IP audio-visual communication. Digital communication may take place in the context of purposeful, problem-like situations and therefore it is an integral part of the PIAAC PS-TRE construct.

"and perform practical tasks"

244. The ability to solve problems with digital technologies is tightly related to the achievement of personal, civic and work-related purposes, which, in turn, take the form of concrete, practical tasks. Examples include shopping, learning about laws and regulations, and organizing teamwork through online agendas and reservation systems. The problems assessed in PIAAC will use authentic, meaningful scenarios based on surveys of computer uses and input from participating countries.

"The first PIAAC problem solving survey".

245. This is the first attempt to assess PS-TRE on a large scale and as a single dimension. This creates many challenges as regards the definition of tasks and the practical collection of data. Furthermore, digital technologies keep evolving at a rapid pace, as do the personal, social, and work-related uses of those technologies. While setting the stage for further rounds of surveys, the present framework will take a perspective on PS-TRE that takes into consideration feasibility issues as well as possible evolutions of technology and technology uses.

"will focus on the abilities to solve problems for personal, work and civic purposes"

246. In order to reflect the pervasiveness of ICT in the society, PIAAC PS-TRE will assess problem solving ability based on scenarios that pertain to these three important contexts.

"by setting up appropriate goals and plans,"

247. An assessment of problem solving capacity should focus on situations where test takers cannot immediately reach their goal based on routine, mechanistic sets of actions. Instead, we focus on tasks that require test takers to actively construct a solution based on the resources available in the assessment environment.

"accessing and making use of information"

248. Again, this phrase emphasizes a specific aspect of PS-TRE, namely that these are often information-rich problems that require individuals to access, interpret and integrate multiple sources of information.

"through computers and computer networks".

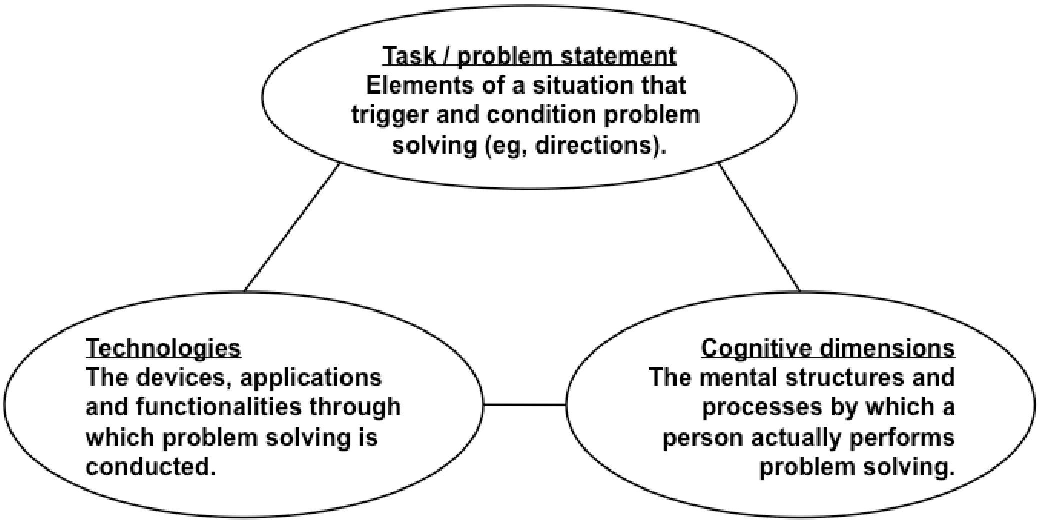
249. There is more to "technology rich environments" than merely personal computers. A full assessment of PS in TRE would require a range of devices that mimic the diversity and versatility of digital technologies of today's world. However, for feasibility reasons, this first survey will be limited to problems requiring the use of computers and Internet-based services.

Organization

Core dimensions of problem solving in technology rich environments

250. The domain of problem solving in technology-rich environments (PS-TRE) may be organized along three key dimensions (Figure 1).

Figure 1. Three core dimensions of problem solving in technology-rich environments



251. "Cognitive dimensions" involve the mental structures and processes by which a person actually performs problem solving. These include goal setting and monitoring progress; planning; locating, selecting and evaluating information; and organizing and transforming information.

252. "Technologies" are the devices, applications and functionalities through which problem solving is conducted. These include hardware devices (laptop computers for PIAAC), simulated software applications, commands and functions, and representations (text, graphics, and so forth).

253. "Tasks" are the circumstances that trigger a person's awareness and understanding of the problem, and that determine the actions to be taken in order to solve the problem. In naturalistic situations, a wide range of conditions can initiate problem solving. For instance, a computer user may realize that their mailbox is crowded and that they need a new categorization scheme; or they may need to reflect on a complex issue (such as finding out more about a medical treatment) and decide to look for information on the Web. In test-taking contexts, tasks are more explicitly assigned to participants. They include the question and task instructions presented to test takers, as well as the specific materials and time constraints associated with the test.

Cognitive dimensions

254. Table 1 summarizes the cognitive dimensions of problem solving that will be assessed in PIAAC. These are goal setting and progress monitoring, planning and self-organizing, acquiring and evaluating information, and making use of information.

Table 1. Cognitive dimensions in PS-TRE

Dimension	Examples
Goal setting and progress monitoring.	Identifying one's needs or purposes, given the explicit and implicit constraints of a situation Establishing and applying criteria for constraint satisfaction and achievement of a solution Monitoring progress Detecting and interpreting unexpected events, impasses and breakdowns
Planning, self-organizing	Setting up adequate plans, procedures, and strategies (operators) Selecting appropriate devices, tools or categories of information
Acquiring and evaluating information	Orienting and focusing one's attention Selecting information Assessing reliability, relevance, adequacy, comprehensibility Reasoning about sources and contents
Making use of information	Organizing information, integrating across potentially inconsistent texts and across formats, making informed decisions Transforming information through writing, from text to table, from table to graph, etc. Communicating with relevant parties

Technology dimensions

255. Table 2 summarizes the technology dimensions that are taken into account in PIAAC. Hardware devices include artifacts that rely on digital technologies, such as desktop or laptop computers, mobile phones, and so forth. These devices are increasingly part of other devices, such as cars or homes, hence the phrase "integrated digital technologies". It is important to note that, regarding hardware devices, only laptop computers with simulated software applications will be included in the first cycle of this assessment. In addition, for operational reasons, sound, animations and videos will not be included. However, the general definition above provides for the inclusion of other digital devices in future cycles.

256. In addition to artifacts, technology rich environments involve the use of software applications. In turn, these applications rely on commands, functions and representations of information. We list commands and functions independent from applications because some commands and functions may be found across a broad range of applications. Therefore, it is unclear if the knowledge of these commands is linked to that of particular applications where they may be found. Examples include "sort" or "find"

commands. Similarly, texts, graphics and other representations are quite independent from the specific application where they may be found.

Table 2. Technology dimensions of PS-TRE

Dimension	Examples
Hardware devices	Desktop or laptop computers, mobile phones, personal assistants, geographical information systems, integrated digital devices
Software applications	File management, Web browser, Email, Spreadsheet
Commands, functions	Buttons, Links, Textboxes, Copy/Cut-Paste, Sort, Find
Representations	Texts, Sound, Numbers, Graphics (fixed or animated), Video

Note. Only laptop devices, a few simulated software applications and a restricted range of representations will be included in the first cycle of PIAAC.

Task dimensions

257. Table 3 summarizes the dimensions of the tasks that will be assessed in PIAAC PS-TRE. These include the purpose and context in which the task is performed, the intrinsic complexity of the problem, and the explicitness of the problem statement and task directions given to the test taker.

Table 3: Task dimensions in PS-TRE

Dimension	Examples
Task purposes (contexts)	Personal, Work/occupation, Civic purposes.
Intrinsic complexity	Minimal number of steps required to solve the problem Number of options or alternatives at various stages in the problem space Diversity of operators required, complexity of computation/transformation Likelihood of impasses or unexpected outcomes Number of constraints to be satisfied Amount of transformation required to communicate a solution
Explicitness of problem statement	Ill-defined (implicit, unspecified) vs. well-defined (explicit, described in detail)

258. It is important to note that the "intrinsic complexity" of a problem is not a simple straightforward dimension. Intrinsic complexity may be characterized through a set of more specific dimensions, which include: the minimal number of steps or actions requested to solve the problem, the number of options at each stage, the diversity of operators and the complexity of mental reasoning and/or computation, the probability of impasses or unexpected outcomes, the number of constraints to be satisfied, and the amount of composition or transformation required in order to communicate a solution.

Number of steps or sub-goals required to reach a solution

259. Tasks which present a problem with a single goal and few required steps are likely easier than those with multiple goals or sub-goals which require a number of steps to reach a solution.

Probability of impasses or unexpected outcomes

260. Tasks that include unanticipated impasses or outcomes are expected to be more difficult. One advantage of a computer-delivered assessment of problem solving is the possibility of building tasks in which additional constraints or outcomes can be introduced as a test taker works through a task. For example, at a defined point in a task, an unexpected email might display, adding new information that test takers must take into consideration while working towards a problem solution.

Amount of transformation and generation required to communicate a solution

261. Tasks requiring test takers to represent or compose information to convey a solution would likely be more difficult than tasks with more defined responses. Examples of transformation and generation tasks include constructing a table, re-representing text in a graph, or writing a justification. In order to keep the PS-TRE framework distinct from numeracy, we do not include the production of statistical graphs in the tasks. Neither do we include the writing of lengthy open justifications, because of difficult scoring issues. We may, however, ask participants to evaluate the communicative effectiveness of a graph or to select among several possible justifications.

Specificity and explicitness of task constraints

262. It is expected that tasks which explicitly define the problem to be solved and the steps required to reach a solution will be easier than tasks which present ill-defined problems. A problem situation that requires the selection of operators, sub-goals, or to define successful achievement of a goal makes the problem more difficult.

Problem solving in technology rich environments vs. other domains of PIAAC

263. The constructs of literacy, numeracy, and PS-TRE rely on the same “core” cognitive processes. For instance, the ability to decode printed symbols and at least a minimal working memory capacity are required for tasks in any of these domains. PS-TRE, however, will assess a set of competencies that are distinct from the other two constructs. Aspects that distinguish PS-TRE from the other domains include the following.

- Problem solving specifically assesses goal setting, monitoring, and planning in technology rich environments. Therefore PS-TRE tasks will emphasize the processes of problem finding and problem shaping that are typical in these environments. Problem solving tasks will include selecting an appropriate software application, deciding on a strategy among several possible, making use of adequate functionalities in a context-sensitive manner, interpreting ill-structured texts and making use of online forms.

- Problem solving tasks will be carried out in environments that involve multiple, complex sources of information. Some of the tasks will require the test taker to use multiple environments and to shift across environments. PS-TRE will therefore assess decision-making with regard to information sources to be used (for example, the act of choosing which environment to use or whether or not to go to another web site.) Evaluation will be included as a critical underlying part of problem solving. Additionally, selecting appropriate devices or tools will take a prominent role for this domain.

264. In terms of information processing, problem solving is a specific construct in that:

- It focuses on the pragmatic evaluation of sources in terms of reliability and the adequacy of information relative to the problem statement as opposed to mere topical relevance, which is more applicable for literacy.

- It focuses on the integration of information across sources, especially in cases where the sources provide inconsistent information.

265. PS-TRE tasks will be kept as simple as low as possible on numeracy and literacy demands in order to increase the specificity and validity of the construct.

Problem solving in technology rich environments vs. other related constructs

266. Problem solving in technology-rich environments is related to several other constructs that have been designed either as part of international surveys or in other contexts. We point out the similarities and differences between PS-TRE and some other constructs below.

267. The expert group addressed the issue of how the problem-solving domain is differentiated from the general ICT domain. ICT skills may be broadly defined as "the interest, attitude, and ability of individuals to appropriately use digital technology and communication tools (…)" (Lennon et al., 2003). It is acknowledged that, just like literacy and numeracy, ICT skills underlie PS-TRE. However, the PS-TRE construct aims to go beyond purely instrumental skills related to the knowledge and use of digital technologies. The cognitive dimensions of problem solving are considered the central object of the assessment, with the use of ICT as secondary.

268. Distinctions between PS-TRE and the Big6™ Skills also exist. The Big6™ skills model is defined as an information and technology literacy model and curriculum (Eisenberg, 2008). It rests on the acknowledgement that information problems involve a series of core steps: namely task definition, source selection, location and access of content information, extraction of content information, synthesis, and evaluation. Although logically sequential, the steps do not necessarily follow each other linearly during complex information problem solving. The Big6™ model is meant to describe a wide range of information problems, among which are problems requiring the use of ICT. The Big6™ model, however, does not specifically focus on tasks that are novel or complex (i.e., actual *problems*), as the PS-TRE framework does.

Operational aspects

269. Part 4 reviews some practical aspects of the assessment, especially regarding the evaluation of problem solving strategies in the context of digital technology.

Prerequisite skills

270. Achievement of PS-TRE tasks presupposes the mastery of foundational ICT skills. These include skills associated with manipulating input and output devices (e.g., the mouse, keyboard, and digital displays), awareness of concepts including files and folders, and an understanding of basic file management operations such as save, open, close, delete, move, and rename. In addition, test takers should be at least minimally familiar with simple graphical interface features, such as the iconic representation of files and folders, hyperlinks, scrollbars, and different types of menus and buttons.

Task development

271. Test developers will use the defined task characteristics, which include cognitive dimensions, technology environments, and contexts, to build and code tasks for the problem solving assessment. Assuming the development of 25 tasks for consideration for the field test, it is recommended that the proportion of tasks be as shown in Table 4. Additionally, the distribution across contexts is recommended to be 40% personal, 30% occupational and 30% civic.

Table 4: Distribution of tasks as a function of environment and cognitive dimensions

	Web	Spreadsheet	Email	Multiple
	9	4	6	6
Goal setting and monitoring progress	2	1	1	1
Planning	2	2	2	4
Acquiring and evaluating information	3	0	0	0
Making use of information	2	1	3	1

272. In addition, the task dimensions of intrinsic complexity and explicitness of the problem definition, as listed in Table 3, will be included as variables during development as they are expected to influence the difficulty of items in the problem solving assessment.

Item development goals

273. The test design for PIAAC specifies the number of scenarios needed for the problem solving assessment, as shown in Table 5.

Table 5: Problem solving development targets

	Development Pool	Field Test	Main Assessment
Total	25 task proposals	16 scenarios (5, 10, or 15 minutes)	8 scenarios (5, 10 or 15 minutes)

274. It is anticipated that the three types of scenarios listed will vary in terms of specific characteristics including: the number of required cognitive processes (e.g., goal setting and monitoring; planning; acquiring and evaluating information; making use of information), the number and kind of steps required to complete a task, the inclusion of unexpected outcomes or impasses to which a test taker must respond, and the extent to which tasks are open-ended or explicitly broken down into a series of defined steps. More specifically:

- Five-minute scenarios may range from easy to more difficult but they are envisioned as the most directed and least complex of the tasks in the assessment. For example, a test taker may be given a simulated page of hits from an Internet search and be asked to evaluate the choices and select one that meets a short set of specified criteria.

- Ten minute-scenarios will likely involve multiple steps and, in some cases, multiple technology environments. For example, test takers might be presented with a problem which requires them to locate an email message, open an attachment, and then use information in the attachment to create a brief table that will represent the information for a specific purpose.

- Fifteen-minute scenarios will be designed to simulate real-life problem-solving activities that are recursive and more exploratory in nature. The tasks will often cross multiple environments, and require test takers to employ several, if not all, of the cognitive components. A sample complex scenario would be one in which a test taker has to conduct a search in the simulated web environment, integrate and evaluate information across a number of sites and then use the information to generate a summary to be shared as part of a community presentation.

Capturing and scoring performance indicators

275. Skilled problem solving includes the ability to effectively reach problem-solving goals, and to do so using the most efficient combination of means and actions. Therefore, the assessment of PS-TRE skills requires both a measure of problem-solving performance and a measure of strategy effectiveness.

276. The cognitive components that underlie PS-TRE (Table 1) all contribute to skilled performance. The competent problem solver is presumably good at setting goals, planning, acquiring and making use of information. The components may nevertheless be related to different underlying cognitive abilities. For instance, goal setting may depend on a person's reasoning ability, whereas locating information may depend on the person's visual scanning and reading ability. A fine-grain assessment of problem solving skills should be based on a broad set of indicators of the underlying cognitive components. The constraints of a large-scale international assessment, however, do not allow the application of the relevant fine-grain measurement procedures. Therefore, only overall indicators of problem solving performance and strategy will be collected in PIAAC. The manipulation of task and environment characteristics across items will ensure that all the required cognitive components will be assessed.

277. A critical aspect of the test development process will be defining the actions to be captured by the software and determining construct-based scoring criteria. The test delivery software will be able to collect more than the output or product of a participant’s response to a task. For any given task, the computer can collect a variety of information including time spent, actions taken, and the sequence in which actions are completed. This information provides direct evidence of the processes and strategies participants use to solve the presented problems, and therefore it improves the inferences we can make about their knowledge and skills. Given the possibility to capture process and strategy information, it is anticipated that the survey will yield an extensive amount of raw data. It is therefore important to use the framework variables, the hypotheses about task characteristics that will impact performance, and

information from the cognitive labs to define those behaviors that provide the best evidence about performance and are therefore important to capture and score.

278. It is anticipated that the scoring models developed around each scenario and task will allow for a range of data to be aggregated into variables that can be used to understand performance across the problem solving scale. This will include those aspects of performance that are associated with fulfilling the task demands as well as those related more to strategies and processes that are undertaken. Subject matter experts including task designers will provide initial input into the creation of these scoring models. These models will be reviewed and changes made as appropriate based on feedback from reviewers, the cognitive labs and field test data. The goal is to use the raw data that is captured to create a set of variables that can reliably capture and distinguish performance along the problem solving scale.

References

Bromme, R. Hesse, F.W. & Spada, H. (Eds.) (2005). *Barriers and biases in computer-mediated knowledge communication - and how they may be overcome*. New York: Springer.

Eisenberg, M.B. (2008). Information literacy: essential skills for the information age. *Journal of Library and Information Technology*, *28*, 39-47.

Forester, T. (1992). Megatrends or Megamistakes? What Ever Happened to the Information Society? *The Information Society*, *8*,133-146.

Funke, J., & Frensch, P. A. (2007). Complex problem solving: The European perspective - 10 years after. In D. H. Jonassen (Ed.), *Learning to solve complex scientific problems* (pp. 25-47). New York: Lawrence Erlbaum Associates.

Gilster, P. (1997). *Digital literacy*. New York: Wiley.

Goel, V. & Pirolli, P. (1992). The structure of design problem spaces. Cognitive Science, 16, 395-429.

Kraut, R., Lundmark, V., Patterson, M., Kiesler, S., Mukopadhyay, T., & Scherlis, W. (1998). Internet paradox: a social technology that reduces social involvement and psychological well-being? American Psychologist, 53, 1017-1031.

Lazonder, A.W., & Rouet, J.-F. (2008). Information problem solving instruction: some cognitive and metacognitive issues. Computers in Human Behavior, 24, 753-765.

Lennon, M., Kirsch, I., Von Davier, M., Wagner, M., & Yamamoto, K. (2003). Feasibility Study for the PISA ICT Literacy Assessment. Princeton, NJ: Educational Testing Services.

Marquié, J.-C., & Baracat, B. (2001). Being over 45 in an ever-changing technological context. in J.C. Marquié, D. Paumes Cau-Bareille, and S. Volkoff (Ed.) Working with age (pp. 273-284). London: Taylor and Francis.

Moore, P. (1995). Information problem-solving: A wider view of library skills. Contemporary Educational Psychology, 20, 1-31.

Newell, A. & Simon, H.A. (1972). Human Problem Solving. Englewood Cliffts, NJ: Prenctice Hall.

Norris, P. (2001) Digital Divide: Civic Engagement, Information Poverty, and the Internet Worldwide, Cambridge University Press, Cambridge.

Rouet, J.-F. (2006). The skills of document use: from text comprehension to Web-based learning. Mahwah, NJ: Lawrence Erlbaum Associates.

Stadtler, M. & Bromme, R. (2007). Dealing with multiple documents on the WWW: The role of metacognition in the formation of documents models. *International Journal of Computer Supported Collaborative Learning, 2*, 191-210

Sweets, R., & Meates, A. (2004). ICT and low achievers: What does PISA tell us? In Karpati, A. (Ed.) Promoting equity through ICT in education: Projects, problems, prospects. Budapest, Hungarian Ministry of Education and OECD.

Sweller, J., Chandler, P., Tierney, P. & Cooper, M. (1990). Cognitive load as a factor in the structuring of technical material. Journal of Experimental Psychology: General, 119, 176-192.

Voss, J.F. & Post, T.A. (1988). On the solving of ill-structured problems. in M.T.H. Chi, R. Glaser & M.J. Farr (Eds.) The nature of expertise (261-285). Hillsdale, NJ: Lawrence Erlbaum Associates.

ANNEX 4: PIAAC READING COMPONENTS - A CONCEPTUAL FRAMEWORK

Introduction

279. The Components Assessment Framework described in this report builds upon the basic principle that comprehension, that is, the 'meaning construction' processes of reading, are built upon a foundation of knowledge of how one's language is represented in one's writing system, that is, component print skills. This basic principle of learning to read has now been widely researched and accepted internationally (Curtis, 1980; Oakhill, Cain, & Bryant, 2003; Perfetti, 1985, 2003; Sabatini, 2003; Strucker, Yamamoto, & Kirsch, 2004). Evidence of an individual's level of print skill can be captured in tasks that examine a reader's ability and efficiency in processing the elements of the written language – letters/characters, words, sentences, and larger, continuous text segments.

280. A second principle guiding the components design is that the main interest is in whether the adults surveyed can apply their existing language and comprehension skills to the processing of printed texts. The components tasks are not designed to separately assess the level of language skills in the target print literacy writing system, nor the literacy skills assessed in the main literacy survey. If the adults surveyed are non-native speakers of the target language, and do not have basic oral vocabulary, syntactic/grammatical, and linguistic comprehension skills, then that will result in poor performance on component reading tasks. We cannot differentiate low language skills from low literacy skills in the component tasks.

281. A third principle of this model of reading is that the level of proficiency, efficiency, and integration of component skills is indicative of level and learning potential in reading development. As skills and knowledge accumulate, the ease of processing of familiar, text-based print increases. Component efficiency is typically indexed by assessing speed or rate of processing, as well as accuracy. As learners, we spend extra time, effort, and energy to solve problems that are novel. On familiar tasks, we can often respond accurately, quickly, with seemingly little conscious effort. When the tasks are easy, we can spend more effort solving and learning from more complex problems and tasks. Speed or rate can be approximated by recording the time it takes to complete certain tasks or by setting a time limit and observing how many items are completed in the time frame allotted.

282. Finally, two guiding assumptions of this assessment framework are made. The first is that the adults to be sampled are in the low end of the continuum of reading ability (as evidenced by low performance on the screening instrument). The model of reading acquisition, development, and choice of item types and difficulties described below holds most strongly in this range of non- and developing readers. Different assumptions about component inter-relationships may hold for a population of more skilled readers.

283. The second assumption is that the set of component items administered in each country will reflect the linguistic characteristics of the language of assessment. As the relationship of the language to the writing system may be very different in different languages, the nature of the items used to assess the

components will need to be adapted based on consideration of those differences. This will best ensure comparability across languages. A base set of components items in English and the guidelines for translation and adaptation are provided in a separate document.

Measuring component skills

284. The primary goal of the component skills battery is to help us better understand the "reading" profiles of adults at the low end of the literacy spectrum. In designing these measures in English, we can adopt the assumptions of the simple view of reading to organize our assessments to maximize useful profile information. As described by Hoover & Tunmer (1993): "the simple view makes two claims: first, that reading consists of word recognition and linguistic comprehension; and second, that while each of these components is necessary for reading, neither being sufficient in itself." (p.3) Word recognition is a stronger predictor of reading level in the early years of reading development. As word recognition becomes more fluent and automatized, listening comprehension becomes a stronger predictor of reading ability, though word recognition continues to contribute significant variance even in skilled readers (Gough & Walsh, 1991; Cunningham, Stanovich, & Wilson, 1990; McCormick, 1994). Strucker, Yamamoto, & Kirsch (2003) use a similar component framework when they describe *print components* (e.g., decoding accuracy and fluency) and *meaning components* (e.g., oral vocabulary).

285. In skilled reading, these components are integrated to support literacy performance. Even during acquisition of reading skill, the components do not strictly develop hierarchically. One learns to understand basic sentences and paragraphs, even as one learns individual words and how to decode. One does not wait to 'master' decoding skills before learning to construct meaning. However, during acquisition, the components may be measured separately, with different profiles having implications for learning, instruction, and policy.

286. Nonetheless, a foundation of *decoding* and *word recognition* skills is necessary (albeit not sufficient) to enable the growth in proficiency of meaning/comprehension level skills. However, the decoding/word recognition components are highly dependent on the precise nature of each language to its writing system. Aspects that affect difficulty in development of learner proficiency include whether the writing system is alphabetic, syllabic, logographic, or some combination; the degree of regularity of the relationship between the print and oral language forms; and how morphological and grammatical/syntactical features of the language are encoded in words. For these reasons, it is difficult to ensure cross-language comparability, as this requires evaluating how to match the sources of difficulty in acquiring these print skills for each language, and balancing them across stimuli and tasks. While we do not provide a base set of decoding/word recognition items in English, we discuss how to develop such items and tasks below, so that each country interested in assessing these foundational print skills can design them.

287. *Vocabulary*, *sentence*, and *basic passage* understanding in print comprise the meaning based component skills we will assess. We discuss each component in more detail in the following sections, along with how they depend on and may be linked back to the more basic print skills of decoding and word recognition as appropriate.

Framework of reading components

288. Conceptually, we begin this discussion of the reading components framework with the basic print skills (alphanumeric perception and efficiency and word recognition/decoding), but it should be noted that these are optional parts of the components assessment. Participating countries that wish to assess alphanumeric perception and word recognition/decoding will find development guidelines in a separate

document. The base set of components to be assessed are word meaning (print vocabulary); sentence processing; and basic passage comprehension.

Alphanumeric

289. Recognizing the alphabet is a core prerequisite of reading ability development in alphabetic writing systems. It remains a significant predictor of early reading acquisition in the U.S. (Adams, 1990). Part of this is explained by the obvious need of visual recognition of the letters to understand oral instruction. If an instructor asks the learner to find the word that begins with the letter 'bee', then the learner must identify the visual symbol from the auditory label. This is the most basic step of sight-to-sound correspondence – matching the letter name to the printed symbol and vice versa. Even observing that not all the letter names correspond to letter sounds (e.g., the letter name of 'w' is pronounced 'double-you') in English and that different languages have different names for letters (e.g., in German 'b' is pronounced 'bay') does not change the fundamental value of knowing this sometimes arbitrary set of associations.

290. However, just accurately being able to puzzle out the names of letters is not as indicative as having automatized this important symbol system. This latter skill serves as a foundation for a) benefiting from oral instructional settings, and b) focusing attention on higher level skills. In the context of a broad survey it is indicative of individuals' experience with the writing system either through schooling or attempts at reading print. From data collected in the U.S. and elsewhere, the rate of rapid naming of alphanumeric symbols remains moderately correlated with overall reading skill across developmental and adult levels (e.g., van den Bos, Zijlstra, & Spelberg, 2002; Sabatini, 2002)

291. Rapid naming of alphanumeric lists can be used as basic reading measures, as well as covariates for better understanding profiles or eliminating extraneous variance from inferences we might wish to make about subgroups. The main types of information provided by these tasks are as follows.

- Index of familiarity with basic perceptual codes of the writing system. (Numbers and letters will be over-learned symbol systems and frequent exposure to them should result in efficient perceptual identification.)

- Index of baseline pronunciation rates.

292. Rapid naming of letters and numbers have typically been administered as separate tasks. Letters are a slightly stronger predictor of reading than numbers, but both tasks are generally more strongly correlated with each other than with overall reading ability. This pattern from the research literature holds because frequent exposure to printed texts in rich text settings (e.g., in schools, workplaces) typically involves exposure to both letters and numbers. However, it is possible that in some settings, such as communities with minimal printed materials available, exposure and knowledge of printed digits is higher than for letters or vice versa. This highlights the value of measuring both independently, rather than assuming one can be a proxy for the other. Therefore, we recommend both letter and digit recognition tasks when this assessment is used.

Decoding and visual word recognition

293. Most models of reading development recognize the centrality of rapid, automatic visual word recognition to reading ability (Abadzi, 2003; Adams, 1990; Perfetti, 1985). The visually presented printed real word (a spelling or orthographic representation in alphabetic languages) is transformed by the perceptual-cognitive system for processing into semantic (meaning) and phonological (sound-based) code

systems. It is widely documented that the sound-based code is used in phonological working memory during the process of meaning construction or comprehension (Gathercole & Baddeley, 1993).

294. The semantic and phonological systems described for reading are the same cognitive systems used in language comprehension more generally. This forms the basis for the claim that visual word recognition 'feeds' the more general language processing system that also is used when listening to language or internal speech (Perfetti, 2003). Put another way, the goal of word recognition is to permit the individual to use the full extent of their language skills to construct meaning or comprehend as early in the cognitive processing of print as possible.

295. Without going into great detail on the mechanisms of word recognition (which are still under study in the psychological sciences), there are two basic behavioral skills that are indicative of proficiency in word recognition. The first is the accumulation of sight word knowledge of real words in the language. In English, one can identify a relatively smaller set of words that appear frequently in everyday texts, as compared to all words in the language that one might find in a dictionary. Most of these frequent printed words are words most skilled speakers of a language have in their speaking/listening lexicon/vocabulary.[13] Accurate and rapid recognition of frequent words is a strong index of word recognition efficiency and proficiency.

296. The second, more fundamental skill is decoding (also referred to as *word attack* or *cyphering*). This skill enables the generation of plausible pronunciations of printed words and conversely, plausible phonetic spellings of heard words. Decoding has been described as the fundamental word learning mechanism in alphabetic languages (Share, 1997), and therefore an essential component to measure directly. In alphabetic systems, decoding requires knowledge and skills in how lexical and sublexical sight-to-sound correspondences represent words in the language. Acquiring mastery of this skill is somewhat easier in languages in which the sight-to-sound correspondences are highly regular and predictable (e.g., German, Serbo-Croatian, Spanish, Turkish). With only a modest input of instruction, learners in these languages can often generate pronunciations for novel printed words and produce the correct pronunciation (i.e., the pronunciation that matches the typical spoken form in the language).

297. In languages with less regular correspondences, there are many alternate pronunciations for any given spelling (and vice versa), so more learning and instructional effort may be required to achieve proficiency. For example, the 'ou' vowel sound is pronounced differently in the English words 'could', 'though', 'thought', 'found'. In contrast, 'word', 'bird', 'heard', 'curd', 'nerd' all rhyme when pronounced, but the vowel sound is represented visually by different letters.

298. As noted, sight recognition of frequent printed words is a direct index of the accumulated knowledge of word reading. Several sources are available for getting an approximate list of these frequent words in English (e.g., Kucera and Francis, 1967). The ability of an individual to read a selected sample of such frequent, well known words without the benefit of passage context is a useful index of how many words an individual can recognize in print. However, one cannot tell based solely on accuracy whether the words were processed more like sight words or decodable words. The distinction is one of degree as much as kind. A sight word is a printed word that has been seen often enough by the individual that it is recognized 'by sight', in contrast to a novel or pseudoword in which one must apply one's decoding knowledge of sight-to-sound correspondences to generate a pronunciation.

---

[13] This line of reasoning begins to get more complicated as one tries to categorize grammatical and morphological features in determining what counts as a word. For example, a dictionary will not list every verb tense as a separate meaning, though visually and auditorily they are different. In general, morphological, grammatical and syntactic variations across languages interact with word recognition in different degrees as well.

299. In skilled reading, both skills are necessary and applied rapidly, automatically, and strategically as needed. If one only measured decoding skill, one might have an estimate of the growth potential for learning real words, but under or overestimate accumulated knowledge of sight word knowledge directly necessary to reading and understanding printed texts. If one only measured sight word knowledge, then one might under or overestimate the growth potential. For example, low literate adults in the U.S. have been found to have sight word knowledge of frequent English words gleaned from years of formal schooling and exposure to printed text that overestimates the decoding knowledge and skills they can apply to learning novel words. Though they have some functional literacy ability, their reading growth seems to be stunted by their slow progress in learning new sight words (Davidson & Strucker, 2002; Greenberg, Ehri, & Perin, 1997, 2003; Sabatini, 2003).

300. Therefore, in the components measures for word recognition in English and other alphabetic languages that are less regular, sight word and decoding tasks can be separated. However, when the spelling system is highly consistent, predictable, and regular, it may be more efficient to use one or the other type assessment.

Word meaning (print vocabulary)

301. Very simply, a barrier to understanding what one reads is not knowing the meaning of the printed words. One can infer meanings of unknown words from context (while reading or listening), but this typically produces provisional, uncertain, and incomplete word meanings – the understanding of which must be separately verified (e.g., checking definition in a dictionary).

302. In the component skills framework, we seek to determine whether individuals can identify in print, words in the everyday listening lexicon of average adult speakers of the language. That is, the emphasis is on the everyday words of the language, rather than specialized technical or academic words that may be known by some but not most of the population. This would be the language used in the neighborhood or market. It would be the language of popular media such as newspaper, radio, and television. This is the most cross-country, comparable vocabulary. The purpose of the vocabulary component measure of this survey, then, is not so much to measure the full extent of individual's vocabulary knowledge, so much as determining whether individuals could understand words when reading print that they could otherwise understand when listening to those words.

303. How do decoding and word recognition affect print vocabulary knowledge? As reading skill develops, one would expect a greater facility in learning new words from print whether or not one hears them used in oral language contexts. *Decoding* skill is critical to this word learning function of reading (for generating plausible pronunciations and storing memory traces), as well as strong reading comprehension skills (for inferring meaning from context). One would also expect that high-frequency, familiar words in the language will be recognized with ease and automaticity in print as one's reading skill increases. When we use the term *word recognition*, we refer to the memory trace of words that one sees in print frequently. So, both basic decoding and word recognition are pre-requisite to print vocabulary skill, though they are not the same as knowing the word meanings. As each vocabulary task will ask adults the meaning of printed words, the primary construct is vocabulary, with decoding and word recognition skills serving as necessary moderators of performance.

304. Assessing vocabulary knowledge can be slippery. Words have multiple meanings. Individuals can know partial meanings of words or know only a specific meaning of a word when the word is used in a specific context. For assessment purposes, many of these difficulties can be circumvented by using words that are concrete and visualizeable. Such words can be made into picture vocabulary items. Respondents are shown common words in the language, then must select from several line drawings depicting common things (e.g., book, chair, cat). Or, one can show a picture and have individuals select from a set of words,

the word that best matches the picture.[14] Care should be taken to select items that are expected to be well known by most adults in the population. Rare and infrequent words are useful in assessing the breadth of vocabulary across the entire population, but are less helpful in making claims about the reading vocabulary of adults with low literacy skills. Care must also be taken to select items that are known cross-culturally. For example, a 'raccoon,' which is indigenous to parts of North America, but not necessarily world wide, would not be a good candidate item for a multi-country survey, whereas the 'moon,' which is known in all countries, is a good candidate.

Sentence processing

305. A variety of psychological studies of reading show that the sentence is a natural breakpoint in the reading of continuous text (e.g., Kintsch, 1998). A skilled reader will generally pause at the end of each sentence. A variety of operations are typically performed including encoding the propositions of the sentence, making anaphoric inferences, relating meaning units to background knowledge and to previous memory of the passage as it unfolds, and deciding which meaning elements to hold in working memory. Thus, each sentence requires some syntactic and semantic processing.

306. By controlling the difficulty of the vocabulary in a sentence (i.e., use simple words the individual can recognize and knows the meaning of), one can vary the sentence complexity to get an indicator of the individual's proficiency at basic semantic meaning construction processing of print. Several measurement focal points are possible, depending on the assumptions about the population and claims one is interested in making.

- If one can assume that the population has a basic command of the grammar and syntax of the language, then the emphasis will be on whether they can apply their language skills in the context of printed text.

- If one cannot, then one may also put emphasis on assessing their basic command of the grammar and syntax of the language.

307. In the components framework, we emphasize the first point. Given a goal of cross-country comparability, varying grammatical/syntactic complexity may not be an ideal strategy, as it is difficult to create items that would be judged as equally difficult (grammatically) in different languages. The recommended strategy, as demonstrated in the examples that follow, is to vary the length of sentences within a basic grammatical structure, and to vary the logical relationships that comprise meaning. By adding more phrases, clauses, conditions, and relations, one will increase length and processing demands, minimizing processing of more complex grammatical structures that may be specialized in languages (e.g., past perfect subjunctive).

308. Thus, in the component measure of sentence processing, sentence length or complexity is varied. The individual is asked to make a judgment whether the sentence makes sense based on the content of the sentence, either in relation to common knowledge about the world (see Example 1) or based on the internal

---

[14] A distinction is typically made between receptive versus expressive or productive vocabulary. In expressive or productive tasks, the examinee sees stimuli such as a picture and must produce the correct word. In receptive tasks, they choose the correct response from among alternatives. We have adopted the latter for several reasons. First, receptive vocabulary is generally larger than productive, so it is better suited to low ability populations. Second, constructed responses are more difficult to score objectively, as respondents may give correct answers unanticipated by the test developer, which should be credited. With choices, only one response is the best answer, even if there are other correct answers as well, disambiguating the score interpretation.

logic of the sentence (Example 2). This task demand is consistent with the 'evaluation' goal of reading in the PIACC framework. Even at the most basic reading level, as one constructs meaning from a sentence, one needs to evaluate that meaning against ones knowledge of the world to judge its veracity. That is, one cannot always believe what one reads. This item type thus measures a combination of basic sentence meaning processing, as well as basic comprehension monitoring and evaluation.

309. For example, one could write items such as:

- Example 1: "The sky is green." YES or **NO**
- Example 2: "If a house is taller than a person, then the person is shorter than the house." **YES** or NO

310. The primary emphasis is on whether respondents can apply their existing language skills in a reading literacy context, not on higher level vocabulary and background knowledge, syntactic/grammatical knowledge, or reasoning skills. These are critical skills for all levels of reading ability, but in the sentence component task we seek to minimize their influence. They are more robustly embedded in the tasks on the main literacy survey. At the basic reading level, we are interested in a more fundamental construction of meaning as a building block to higher level comprehension skill.

311. The simple judgment of whether a sentence is sensible is designed to focus the assessment on basic literal comprehension. However, the logical truth or falsity of basic facts in the empirical world can be slippery. One can imagine exceptions to most absolute statements. Also, language is often used figuratively and metaphorically (e.g., "The sky is grey or black" may conjure images of a storm or night time; 'The sky is pink' conjures images of a sunset.) In the item design of sentences for this component, an attempt should be made to minimize ambiguity of meaning as much as is possible.

Basic passage comprehension

312. Skilled reading is rapid, efficient, and fluent (silent or aloud). Theoretical discussions and definitions of reading fluency generally refer to three aspects of fluency: accuracy, rate, and prosody/expressiveness (Fuchs, Fuchs, Hosp, & Jenkins, 2001; Kame'enui & Simmons, 2001; NRP, 2000; Rasinski, 2006). At a minimum, reading fluency is indicating that visual word identification processes are efficiently feeding language processing systems (e.g. working memory) to produce outputs. The outputs do not necessarily imply the construction of meaning or comprehension as we commonly imagine it. Skilled readers can read familiar texts somewhat fluently aloud without attending to the meaning. However, when oral reading fluency has been operationalized as relatively error free reading of a simple passage aloud at a normal speaking rate, it has reliably served as a solid indicator of the integration of some basic component skills (e.g., Daane, Campbell, Grigg, Goodman & Oranje, 2005; Wayman, Wallace, Wiley, Ticha, & Espin, 2007). On the other hand, breakdowns in accuracy, rate, or both, suggest difficulties in other subcomponents.

313. In recent research, a silent reading assessment task design has gained empirical support as an indicator of basic reading fluency and comprehension. The design uses a forced-choice cloze paradigm, that is, a choice is given between a word that correctly completes a sentence in a passage and an option that is incorrect. The incorrect item is meant to be obviously wrong to a reader with some basic comprehension skills. Distractors may be grammatically or semantically wrong. By giving the participant only a fixed amount of time to do the task, a measure of reading efficiency and fluency is assessed. In this component assessment participants must focus attention on comprehension as they read (Samuels, 2006). Thus, the integration of decoding, word recognition, vocabulary, and sentence processing is required to construct the

basic meaning of a short passage. Fluent, efficient performance on such a basic, integrated reading task is a building block for handling longer, more complex literacy texts and tasks.

314. For cross country comparability, it may be best to allow individuals as much time as necessary to complete each passage, then record total time required to complete. This is because average reading rates of skilled readers may vary from country to country, primarily as a function of language, writing system, and cultural variables. For most low skill adults, the accuracy score will be sufficient to estimate their basic comprehension ability. For the very low skilled beginning reader, the measurement falls more heavily on basic reading comprehension and the time to complete will add very little additional information about their skills.

315. However, near the top of the low ability scale, adults may reach ceiling level total correct scores. By taking into consideration total time to complete the task, we can estimate their basic reading efficiency. A skilled reader would be able to choose all correct responses quickly, without much effort, and continue on reading at a normal rate. By collecting performances from a sub-sample of skilled readers in each country, we will have a benchmark for relative scaling efficiency/fluency of these low-moderate skilled adults across countries. That is, low ability adults with high accuracy scores on this task possess some basic comprehension skills, but still differ in their efficiency in processing continuous text relative to skilled adult readers.

Conclusion

316. Components assessment tasks are designed to inform our understanding of the basic reading skills that underlay proficient literacy performance levels. They help us describe what low literate adults can do and therefore form a basis for learning, instruction, and policy with respect to helping low literate adults achieve higher literacy levels. In this framework, we have focused attention on those component skills that show the greatest promise for cross-country comparability, specifically reading vocabulary, sentence comprehension, and basic passage comprehension and fluency. We have also described how decoding, word recognition, and familiarity with the basic print codes of a language are related and necessary to achieving higher levels of literacy skill. Noting how the language-specific nature of these latter skills makes it difficult to construct items and tasks that are easily comparable across languages, we encourage their use nonetheless as beneficial to within country insights into learning, instruction, and policies directed at improving the achievement levels of low literate adults.

References

Abadzi, H. (2003). *Improving adult literacy outcomes: Lessons from cognitive research for developing countries.* Washington, DC: The World Bank.

Adams, M.J. (1990). *Beginning to Read: Thinking and learning about print.* Cambridge, MA: MIT Press.

Cunningham, A.E., Stanovich, K.E., & Wilson, M.R. (1990). Cognitive variation in adult students differing in reading ability: Associations with vocabulary, general knowledge, and spelling. Journal *of Educational Psychology*, *83*, 264-274.

Curtis, M.E. (1980). Development of components of reading skill. Journal of Educational Psychology, 72, 656-669.

Daane, M. C., Campbell, J. R., Grigg, W. S., Goodman, M. J., & Oranje, A. (2005). *Fourth-grade students reading aloud: NAEP 2002 special study of oral reading* (No. NCES 2006-469). Washington, DC: U. S. Department of Education, Institution of Education Sciences, National Center for Educational Statistics.

Davidson, R.K. & Strucker, J. (2002). Patterns of word recognition errors among adult basic education native and nonnative speakers of English. *Scientific Studies of Reading*, *6*(3), 299-316.

Fuchs, L. S., Fuchs, D., Hosp, M. K., & Jenkins, J. R. (2001). Oral reading fluency as an indicator of reading competence: A theoretical, empirical, and historical analysis. *Scientific Studies of Reading*, *5*(3), 239-256.

Gathercole, S.E. & Baddeley, A.D. (1993). *Working memory and language*. Hillsdale, NJ: Erlbaum.

Gough & Walsh (1991). Chinese, Phoenecians, and the orthographic cipher of English. In S. Brady & D. Shankweiler (Eds.), *Phonological processes in literacy* (pp. 199-209). Hillsdale, NJ: Erlbaum.

Greenberg, D., Ehri, L. C., & Perin, D. (1997). Are word-reading processes the same or different in adult literacy students and third-fifth graders matched for reading level? *Journal of Educational Psychology*, *89*(2), 262-75.

Greenberg, D., Ehri, L. C., & Perin, D. (2002). Do adult literacy students make the same word-reading and spelling errors as children matched for word-reading age? *Scientific Studies of Reading*, *6*(3), 221-241.

Hoover, W.A. & Tunmer, W.E. (1993). The components of reading. In G.G. Thompson, W.E. Tunmer, and T. Nicholson (Eds.), *Reading acquisition processes* (pp. 1-19). Philadelphia: Multilingual Matters.

Kame'enui, E. J., & Simmons, D. C. (2001). Introduction to this special issue: The DNA of reading fluency. *Scientific Studies of Reading, 5*(3), 203-210.

Kintsch, W. 1998 *Comprehension: A paradigm for cognition*. Cambridge, UK: Cambridge University Press.

Kucera, H. & Francis, W.N. (1967). *Computational analysis of present-day American English*. Providence, RI: Brown University Press.

McCormick, S. (1994). A nonreader becomes a reader: A case study of literacy acquisition by a severely disabled reader. *Reading Research Quarterly, 29*, 156-177.

National Institute of Child Health and Human Development. (2000). *Report of the National Reading Panel. Teaching children to read: an evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: Reports of the subgroups* (NIH Publication No. 00-4754). Washington, DC: U.S. Government Printing Office.

Oakhill, J., Cain, K., & Bryant, P. (2003). The dissociation of word reading and text comprehension: Evidence from component skills. *Language and Cognitive Processes*, 18(4), 443-468.

Perfetti, C. A. (1985). *Reading Ability*. New York: Oxford University Press.

Perfetti, C.A. (1992). The representation problem in reading acquisition. In P.B. Gough, L.C. Ehri, & R. Treiman (Eds.), *Reading acquisition* (pp.145-174). Hillsdale, NJ: Erlbaum.

Perfetti, C. (2003). The universal grammar of reading. *Scientific Studies of Reading*, 7(1), 3-24.

Rasinski, T. (2006).Reading fluency instruction: Moving beyond accuracy, automaticity, and prosody. The Reading Teacher, *59*(7), 704-706.

Sabatini, J.P. (2002). Efficiency in word reading of adults: Ability group comparisons. *Scientific Studies of Reading*, *6*(3), 267-298.

Sabatini, J. P. (2003). Word reading processes in adult learners. In E. Assink & D. Sandra (Eds.), *Reading Complex Words: Cross-Language studies.* London: Kluwer Academic.

Samuels, S. J. (2006). Toward a model of reading fluency. In S. J. Samuels & A. E. Farstrup (Eds.), *What research has to say about fluency instruction* (pp. 24-46). Newark, DE: International Reading Association.

Share, D.L. (1997). Understanding the Significance of Phonological Deficits in Dyslexia. *English Teacher's Journal*, 51, 50-54.

Strucker, J., Yamamoto, K., & Kirsch, I. (2003). The relationship of the component skills of reading to IALS performance: Tipping points and five classes of adult literacy learners

(NCSALL Reports #29). Cambridge, MA. National Center for the Study of Adult Learning and Literacy.

van den Bos, K. P,. Zijlstra, B. J., & Spelberg, H.C. (2002). Life-Span data on continuous naming speeds of numbers, letters, colors, and pictured objects, and word-reading speed. *Scientific Studies of Reading*, *6*(1), 25-49.

Wayman, M. M., Wallace, T., Wiley, H. I., Ticha, R., & Espin, C. A. (2007). Literature synthesis on curriculum-based measurement in reading. *The Journal of Special Education, 41*(2), 85-120.