Review of the PIAAC Numeracy Assessment Framework:

Final Report

Dave Tout, Diana Coben, Vince Geiger, Lynda Ginsburg, Kees Hoogland, Terry Maguire, Sue Thomson, & Ross Turner

January 2017

Australian Council *for* Educational Research

ACER

This work has been produced by the Australian Council *for* Educational Research (ACER) on behalf of the OECD. Funding was provided through the OECD. The views and opinions expressed in this document are those of the authors/project team and do not necessarily reflect the official views of the OECD or of its member countries.

Published by the Australian Council for Educational Research
19 Prospect Hill Road
Camberwell, Victoria, 3124, Australia
Ph: +61 3 9277 5555

ISBN: 978-1-74286-516-4

Please cite the work as follows:
Tout, D., Coben, D., Geiger, V., Ginsburg, L., Hoogland, K., et al. (2017). Review of the PIAAC Numeracy Assessment Framework: Final Report. Camberwell, Australia: Australian Council for Educational Research.

ABSTRACT

The Programme for the International Assessment of Adult Competencies (PIAAC) is an international assessment of the proficiency of adults (aged 16-65 years) in key information processing skills (reading, numeracy and Problem Solving in Technology-Rich Environments). The PIAAC Survey of Adult Skills has revealed that a considerable number of adults in OECD countries possess only limited literacy and numeracy skills. The OECD is currently reviewing the content of the frameworks and cognitive assessment instruments for the 2nd cycle of PIAAC ready for delivery in 2021-22.

This report is the result of a review of the numeracy construct and assessment in PIAAC. It recommends a range of areas for potential improvements and enhancements, including of the definition and elaborations of adult numeracy used in the framework, and the assessment content. Many of the suggestions arise out of the concern that the existing framework and assessment do not reflect some of the realities of the skills and knowledge adults now need to succeed in work, life, and citizenship in the 21st century.

The report also recommends the development of a PIAAC numeracy components assessment, which would have parallel aims to the existing reading components assessment, and provide insights into the skills and knowledge of the significant number of adults with low levels of numeracy.

CONTENTS

INTRODUCTION

The aim of this ACER-led OECD project was to prepare a paper reviewing the framework that guided the assessment of numeracy in the 1st cycle of the Programme for the International Assessment of Adult Competencies (PIAAC). The review aimed to evaluate the extent to which the framework reflects current understandings of adult numeracy and continues to be an appropriate basis for the assessment of the capacity of adults to undertake successfully the range of numeracy tasks that they will face in their everyday and working lives in the third decade of the 21st century.

Background

The Programme for the International Assessment of Adult Competencies (PIAAC) is an international assessment of the proficiency of adults (aged 16-65 years) in key information processing skills (reading, numeracy and Problem Solving in Technology-Rich Environments). The first cycle of the study involved three waves of data collection: the first in 2011-12 (24 countries), the second in 2014-15 (nine countries) with a third wave planned for 2017-18 (six countries).

The assessment was undertaken using a household survey methodology. Respondents undertook the assessment in their own homes or in some other agreed location under the supervision of trained interviewers. The assessment was undertaken either as a computer-based assessment (CBA) on a laptop computer or, in the case of adults with little or no familiarity with ICT (or who refused to undertake the test on computer), in a paper and pencil format.

The OECD has always conceived PIAAC as a repeated cross-sectional study, and is currently in the process of developing the 2nd cycle of the assessment, with data collection planned for 2021-22. As part of the development phase, the OECD has instituted reviews of the cognitive assessment instruments and background questionnaires, and work is being undertaken on the development of possible new content.

The Numeracy Assessment

An assessment framework guides the assessment of numeracy in PIAAC, defines the construct of numeracy and steers the development of the items and the interpretation of results. The framework for the assessment of numeracy in the first cycle of PIAAC is described in OECD (2012) and, in more detail, in PIAAC Numeracy Expert Group (2009) and in Gal et al (2014). The PIAAC framework drew heavily on the numeracy assessment framework of the Adult Literacy and Life Skills Survey (ALL) (see Murray et al, 2005). The assessment framework provides a definition of the domain and the features of the construct in terms of the dimensions of content, cognitive strategies and range of applications.

PIAAC defines numeracy as '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'. The definition of numeracy is complemented by a definition of 'numerate behaviour': 'Numerate Behaviour involves managing a situation or solving a problem in a real context, by responding to mathematical content/information/ideas represented in multiple ways'.

It is important to note that the PIAAC numeracy assessment describes the full range of numeracy capability (as defined above) in the adult population. This covers at one extreme, adults who have university level training and, at the other, adults who have very low levels of education (e.g. who left school at or before the age of 15). At the same time, it covers both young adults still in education and adults who completed their formal education 30-50 years prior to undertaking the assessment.

The assessment of numeracy in the first cycle of PIAAC involved the use of 57 test items. As the assessment linked to ALL, 19 items were linking items that had been previously used in ALL. Examples of numeracy items can be found in OECD (2012, 2013c) and at http://www.oecd.org/skills/ESonline-assessment/takethetest/. Access to the full set of PIAAC numeracy items were provided to the consultants engaged to undertake the review of the framework.

The assessment of numeracy in the 2nd cycle of PIAAC will link to the assessment used in the 1st cycle in order to evaluate change in the numeracy proficiency of the adult populations in participating countries over time. At the same time, it will be important that the framework identifies a construct of numeracy that is relevant to that the realities of the third decade of the 21st century as well as reflecting contemporary understandings of adult numeracy and incorporates relevant developments in testing practice and makes the best use of the available testing technologies.

Issues

In particular, the review needed to address the following issues as specified in the briefing paper provided by the OECD:

- **Theoretical developments**: identify any theoretical developments in the understanding and conceptualisation of adult numeracy that are relevant for the assessment of numeracy in PIAAC;
- **21st century digital implications**: discuss how to ensure that the assessment reflects the importance of digital information, representations, devices and applications as realities that adults have to manage in dealing with the numerical demands of everyday life;
- **Assessment developments**: identify any developments in the assessment of numeracy (particularly of adults) that could be relevant for PIAAC (e.g., item types and formats, use of animation, and modelling);
- **Relationship with PISA**: discuss how the relationship between the PIAAC numeracy framework and the PISA mathematical literacy framework and assessment should be conceived, developed (if appropriate) and presented;
- **Numeracy or mathematical components**: evaluate the utility and feasibility of the implementation of an equivalent to the PIAAC reading components and numeracy assessments; and
- **Recommendations**: make recommendations regarding the definition of the construct of numeracy and the priorities for development of the assessment framework for numeracy in the second cycle of PIAAC.

The Review Team

Dave Tout, Senior Research Fellow, Numeracy and Mathematics, Vocational, Adult and Workplace Education, Australian Council for Educational Research, Australia.

Professor Diana Coben, PhD, Director of the National Centre of Literacy and Numeracy for Adults at the University of Waikato, New Zealand and Emeritus Professor of Adult Numeracy at King's College London.

Professor Vince Geiger, Research Fellow, Learning Sciences Institute of Australia, Australian Catholic University, Australia.

Dr Lynda Ginsburg, Senior Research Associate for Mathematics Education, Center for Math, Science and Computer Education, Rutgers University, NJ, United States.

Dr Kees Hoogland, Researcher and Curriculum Developer, SLO - Netherlands Institute for Curriculum Development, The Netherlands.

Dr Terry Maguire, Director, National Forum for the Enhancement of Teaching and Learning in Higher Education, Ireland.

Dr Sue Thomson, Director, Educational Monitoring and Research Division and Research Director of the Australian Surveys research program, Australian Council for Educational Research, Australia.

Ross Turner, Principal Research Fellow, Centre for Global Education Monitoring, Australian Council for Educational Research, Australia.

Structure of the report

This report has separate sections relating to each of the different issues needing consideration and review. The final section synthesises the research and review into a set of recommendations. The sections are:

1. Theoretical developments
2. 21st century digital implications
3. Assessment developments
4. Relationship with PISA
5. Numeracy or mathematical components
6. Recommendations.

Under each of the above issues, suggestions for reviewing the existing PIAAC Framework or assessment have been organised, where relevant, under a number of different aspects of the framework:

1. Framework definition and description
2. Mathematical content
3. Contexts
4. Responses/actions
5. Item formats
6. Representations, reading demands, and authenticity-related issues
7. Other factors.

Each section does not address all aspects.

PIAAC Numeracy Framework document

There are in effect two existing PIAAC Numeracy Framework documents: the full version used as the basis for developing the assessment: *PIAAC Numeracy: A Conceptual Framework* (PIAAC Numeracy Expert Group, 2009) and the abbreviated version published in *Literacy, Numeracy and Problem Solving in Technology-Rich Environments: Framework for the OECD Survey of Adult Skills* OECD (2012). The former, full version of the framework is the one referred to throughout this report.

EXECUTIVE SUMMARY

This report is the result of a review of the numeracy construct and assessment in PIAAC. It recommends a range of potential improvements and enhancements, including of the definition and elaborations of adult numeracy used in the framework, and the assessment content. Many of the suggestions arise out of the concern that the existing framework and assessment do not reflect some of the realities of the skills and knowledge adults now need to succeed in work, life, and citizenship in the 21st century.

PIAAC is an international assessment of the proficiency of adults (aged 16-65 years) in key information processing skills (Reading, Numeracy and Problem Solving in Technology-Rich Environments). The PIAAC Survey of Adult Skills has revealed that a considerable number of adults in OECD countries possess only limited literacy and numeracy skills. In almost all countries, a sizeable proportion of adults (18.5% on average) has poor reading skills (Level 1 or below) and even poorer numeracy skills (22.7% on average at Level 1 or below).

Given that PIAAC is a repeated cross-sectional study, and that the OECD is currently in the process of developing the 2nd cycle of the assessment ready for delivery in 2021-22, the content of the frameworks and cognitive assessment instruments are being reviewed. PIAAC's numeracy framework defines the construct of numeracy and guides the development of the items and the interpretation of results. It should also be noted that the existing PIAAC framework drew heavily on the numeracy assessment framework of the Adult Literacy and Life Skills Survey (ALL), written in the late 1990s, and the ALL assessment implemented in the mid-2000s.

The PIAAC numeracy assessment framework that has defined the construct of numeracy and guided the development of the items and the interpretation of results has been robust and has stood the test of time since its initial development. It is opportune, however, to reflect on its status and to review how the existing framework and assessment can be enhanced and refined prior to the next delivery of PIAAC. It is therefore important that the revised framework should identify a construct of numeracy that is relevant to the realities of the third decade of the 21st century, reflects contemporary understandings of adult numeracy, and makes the best use of available testing technologies.

Being numerate in the 21st century means being able to cope with the aspects of the world as we encounter it, which includes digital and technological aspects. The reality that techno-mathematical aspects of the workplace and society are now ubiquitous must be acknowledged.

The review found that the existing pool of assessment items are not taking advantage of current and available assessment opportunities, and that better utilisation of 21st century assessment developments should be implemented. For an assessment of numeracy in PIAAC, the assessment stimuli and items should be, as far as possible, authentic and representative of what adults might meet in their lives. A recommendation is that the next iteration of PIAAC harness the potential of technology and digital tools to support a more effective and representative 21st century assessment. Use of different technology, media and associated supports has the potential to transcend not only some of the challenges with literacy and/or language that may impede some adults' ability to demonstrate their numeracy capability, but also to make the assessment more relevant to the 21st century. However, a balance needs to be kept between numeracy tasks embedded in digital and technological environments versus those embedded in other, non-digital environments. The results from the Problem Solving in Technology-Rich Environments component of PIAAC showing a high proportion of adults with no or extremely limited ICT skills also need to be taken into account in reviewing the technological and digital aspects of the PIAAC Numeracy framework and the associated assessment.

The report also recommends the development of a PIAAC Numeracy Components assessment. This new assessment would have parallel aims to the existing Reading Components assessment: to develop a set of "fine-grained tasks" so that "at least some of these adults would demonstrate some level of **numeracy** knowledge and skills", and would provide insights into the skills and knowledge of low level numerate adults.

SECTION I. THEORETICAL DEVELOPMENTS

In this section, the Review Team identifies any theoretical developments in the understanding and conceptualisation of adult numeracy that are relevant for reviewing the framework and assessment of numeracy in PIAAC.

The PIAAC numeracy conceptual framework and its associated assessment construct were published in 2009 (PIAAC Numeracy Expert Group, 2009). The numeracy framework and assessment in PIAAC is based in large part on the conceptual framework and item pool developed for the Adult Literacy and Lifeskills survey (ALL) (Statistics Canada and OECD, 2005). Work on the numeracy domain of ALL started in 1998. As such, most research considered by the Review Team for this review dates from the period after the PIAAC Numeracy framework was developed and written—2008 or later.

General comments

Most recent reviews indicate that there continues to be a shortage of any empirical or theoretical developments in research on adult numeracy (see e.g., Carpentieri et al., 2010; Condelli et al. 2006; Geiger et al., 2015; Windisch, 2015). Despite this, the Review Team's own research (which included recent reports about the teaching, learning and descriptions of adult numeracy practices; e.g., Chisman, 2011; Griffiths et al., 2013; NIACE, 2011) found that a number of issues should be considered and addressed in the review and rewriting of the PIAAC numeracy framework for the next cycle of PIAAC.

Issues

Framework definition and description

The review has raised a number of issues to address in the revision of the PIAAC numeracy framework and its descriptions.

Terminology - *competency* versus *skill*

The two terms *Competency* and *skill* are not the same but both are used (without a clear differentiation) in the current PIAAC numeracy framework document and in the titles of the reports and surveys (e.g. compare The Survey of Adult Skills with the Programme for the International Assessment of Adult Competencies).

Competencies can mean different things in different situations, and in different cultures. For example, a reductive notion of competence is used as a synonym for skill in some adult education settings within Australia. Competencies can however also mean the combination of skills and aspects of higher order thinking such as strategic planning. The latter is reflected in the OECD's Definition and Selection of Competencies project (DeSeCo) (OECD, 2005), which was developed to provide a framework informing the identification of key competencies in international surveys measuring the competence level of young people and adults (see Rychen, 2004, p. 321). In *Chapter 7. The Survey of Adult Skills (PIAAC) and "Key Competencies"* of the PIAAC Readers Companion (OECD, 2013d), the OECD discusses this issue of terminology and the use of the terms *skill* and *competency* and concludes that there is much overlap in their understanding and use.

The Review team recommends therefore, that in the revision of the PIAAC numeracy framework, where possible, the use of the terms 'competence' and 'skill' be made as explicit and as consistent as is possible.

Definition of numeracy and numerate behaviour

A number of issues exist in the existing definition of numeracy and associated description of numerate behaviour in PIAAC, which drives the assessment content and its parameters. Presently:

> *PIAAC defines numeracy as 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. (PIAAC Numeracy Expert Group, 2009, p. 21)*

The PIAAC framework further describes a more detailed definition of numerate behaviour with specification of a number of different facets:

> *Numerate Behaviour involves managing a situation or solving a problem in a real context, by responding to mathematical content/information/ideas represented in multiple ways. (PIAAC Numeracy Expert Group, 2009, p. 21)*

The Four facets described are Contexts, Responses, Mathematical content/information/ideas, and Representations. Further, several enabling factors and processes whose activation underlies numerate behaviour are also listed.

There are four related issues arising from the research into theoretical developments to consider in reviewing the existing PIAAC framework definition and elaboration of numerate behaviour:

- disposition to use mathematics
- seeing mathematics in a numeracy situation
- critical reflection and action
- degree of accuracy.

Disposition to use mathematics

There is the need to consider the issue of a person's judgement on how to use mathematics (or not) in solving a numeracy problem. The issue of choice or disposition when engaging with and solving a numeracy problem is an important factor to consider in an adult's use and application of mathematics in a real world situation (Geiger et al., 2014; Goos et al, 2014). Are individuals able to choose to use mathematics when it is relevant and appropriate? This relates to mathematics anxiety and a significant number of individual's negative disposition to mathematics and their decision to avoid using mathematics, even when appropriate. Research about mathematics anxiety is well documented and demonstrates that it can have a significant impact on performance in mathematics (e.g., see: Buckley, 2013; Ma, 1999; Tobias, 1993).

There are three potentially related aspects behind this issue of disposition in relation to solving a numeracy problem where an adult is expected to use and apply some form of mathematical knowledge in a real world situation:

- using other means than mathematics to solve a problem when mathematics should have been the obvious and most sensible approach;
- using formal mathematics when other sense-making methods would be more efficient; or
- avoiding doing anything at all and not attempting to solve the numeracy problem at hand.

To some extent, the existing PIAAC framework discusses this issue:

> *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. (PIAAC Numeracy Expert Group, 2009, p. 21)*

The Review Team recommends that this issue of disposition be expanded on in the framework, and addressed either within the definition itself and/or in the elaboration under numerate behaviour and its listed enabling factors and processes:

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

Seeing mathematics in a numeracy situation

Research indicates that an important aspect of a person's numeracy or numerate behaviour is their capability to "see" or notice when mathematics is embedded in a real world situation—how to recognise the mathematics and to potentially take the next step and act on it. The ability to see the mathematics that surrounds adults in their everyday life is an important issue in relation to being numerate—to link the mathematics they learned in school with mathematics embedded in a real world situation (Maguire & Smith, 2016; Roth, 2012). This issue is also identified as important in research about workplace numeracy, for example in calculating medication dosage (Weeks et al., 2013). This relates to the first process described as part of PISA's problem solving cycle for mathematical literacy: *formulating*. PISA describes *formulating* as identifying how to apply and use mathematics to the problem being posed in the real world; it includes being able to take a situation and transform it into a form amenable to mathematical treatment (OECD, 2013a).

Critical reflection and action

While the current framework mentions the notion of critical reflection under the facet *Responses* in its elaboration of numerate behaviour, having a critical orientation or reflection are aspects of numeracy that could be emphasised and described further. It is important for individuals in their lives as citizens and workers to critically review the mathematics used and the outcomes obtained to reflect on and question real-world implications, to be capable of following up with appropriate actions, and to make decisions and judgements. A critical orientation is also about supporting an argument or position with mathematical evidence or challenging the argument or position of another person or organisation.

This capability to reflect critically and to act is named and described explicitly in some other models and frameworks (e.g., Geiger et al., 2014; Goos et al, 2014). The third process in the PISA mathematical literacy problem solving cycle, which is called *Interpreting and evaluating*, also covers elements of critical reflection and describes reflecting upon mathematical solutions or results and interpreting or evaluating them back in the context of the initial problem, including critiquing and identifying the limits of the model used to solve a problem (OECD, 2013a).

Degree of accuracy

The present PIAAC numeracy framework does not explicitly address the issue of the degree of precision or accuracy that may be required in the solution of a numeracy problem. It is expected that a numerate person would use estimation and other skills to check the outcomes and decide on the appropriate degree of accuracy required when solving a problem. This is particularly true within a workplace environment, where precision, accuracy and working within specified tolerances can be critical. On the other hand in other situations and applications, there are instances in being numerate where accuracy is not a critical component (e.g., in relation to some spatial skills, in graphical/data interpretation and analysis, or in estimating quantities, where an order of magnitude estimate can often suffice.

Summary regarding the descriptions of numeracy and numerate behaviour

The Review Team recommends these four issues be accommodated both by addressing them explicitly in the background information and related discussions in the framework and, where appropriate, by reworking the definition of numeracy itself and/or in expanding on the framework's elaboration and description of numerate behaviour through its set of facets. Any changes made regarding these four issues would have consequences for the assessment construct and the variables and parameters assessed across the item pool. For instance, an increased focus on degree of accuracy could require new test items be developed to cover that aspect of the framework.

Mathematical content

The issue of the mathematical content in PIAAC is discussed later under *Issue IV: Relationship with PISA*.

Contexts

The issue of context in PIAAC is discussed later under *Issue IV: Relationship with PISA*.

Responses/actions

Reviewing and enhancing the Responses/actions facet of the elaboration of numeracy in the existing PIAAC framework can address and resolve a number of the above issues. The current framework groups those responses under three broad headings: *Identify, locate, or access*; *Act upon or use*; and *Interpret, evaluate/analyse, communicate* and the set of test items need to provide opportunities to demonstrate each of these behaviours. These three parameters and their descriptions should be reviewed to better address and cover the above issues, and potentially the list of Responses/actions may need to be reworked and changed.

The first issue from above (Disposition to use mathematics) should be addressed and incorporated into the description of this facet of Responses/actions. The other three issues (Seeing mathematics in a numeracy situation; Critical reflection and action; and Degree of accuracy) all suit themselves to being made more explicit as part of the description and elaboration of the different Responses/actions and as well new test items may need to be developed to provide opportunities for demonstrating these aspects of numerate behaviour.

Mathematical representations, reading demands, and authenticity-related issues

In their comparison of the PISA and PIAAC frameworks and assessments, Gal and Tout addressed this issue and commented there were three issues to address:

> *…how information is represented within assessment tasks, the amount and difficulty level of the text components of items, and how it is perceived by the respondents because of its degree of familiarity or authenticity. (Gal & Tout, 2014, p. 38)*

In the existing PIAAC numeracy framework, there is discussion of the fact that mathematical information in a real-life situation can be represented in many forms, including: concrete objects to be counted or pictures of such things; through symbolic notation; formulae; diagrams and charts; graphs and tables. Mathematical information can also be embedded in various types of texts, either in prose or in documents with specific formats with words or phrases that carry mathematical meaning, or expressed in notations or symbols. There is little mention of digital representations, which is an issue that should be addressed. The next iteration of PIAAC should build on the fact that adults are living and working in a far more digitally rich world than that in which earlier PIAAC conceptualisations were developed. This issue is addressed more fully in *Issue II: 21st Century Digital Implications*.

In relation to textual components and reading demands, there are a range of issues in relation to the intersection between literacy and numeracy skills and the role that reading literacy aspects take in solving a numeracy problem. A goal in the item development process is often to make the wording of numeracy items as simple and direct as possible, in order to help minimise the reading literacy demands. In PIAAC's complexity scheme, two of the factors (type of match/problem transparency, and plausibility of distractors) relate to literacy aspects, and these descriptions and classifications assisted in addressing and minimising the reading literacy demands of stimuli and questions. However, it is clearly acknowledged in the description of numeracy and its elaborations, and then reflected in the PIAAC Complexity scheme, that reading literacy is an integral and important aspect of numeracy. Certainly, in society and workplaces that adults occupy, tasks and challenges do not neatly divide into, or present as, discreet 'literacy' and 'numeracy' tasks. Real world situations and demands cross those kinds of educationally defined boundaries.

The reality is that in using authentic situations as the basis for the numeracy assessment tasks where the mathematics is embedded in a real world setting, the associated information and data can be very complicated, unfamiliar and involve a heavy reading load. This can create challenges in trying to focus the assessment on the mathematics and numeracy skills and knowledge. In relation to writing PISA mathematical literacy test items, Tout and Spithill wrote that:

> *One approach was to start with a real world context and develop it into a unit. Often the mathematics was too highly embedded in the context and to extract the*

> *mathematical model required too much reading and understanding of the situation, which would block many students from solving the problem. Another issue was that the mathematics and the required values or numerical information to be manipulated in the real world context were also complex and so calculation would be time consuming and too open to arithmetical errors. The challenge was to therefore simplify the real world context, the related stimulus and its embedded mathematical information to make it accessible whilst still maintaining its authentic aspect. (Tout & Spithill, 2015, p. 154)*

Recent research that systematically compares descriptive assessment tasks with more depictive representations of problem situations through using illustrations and photos and minimising the use of words (Hoogland, 2016), gives an indication that even the use of simple supporting illustrations and images could improve performance by a small margin.

However, are there possible consequences that flow from minimising the text or literacy demands of numeracy items? Through doing this, are the PIAAC (and PISA) items made comparably easier than the real life problems adults actually face in their lives, and hence less authentic? In relation to PISA and PIAAC numeracy items, Tout and Gal (2015) stated that:

> *The ability to read and understand texts that include mathematical content, and to speak and write about their mathematical aspects, is crucial for coping with tasks that are likely to appear in the adult world, as text is sometimes a gateway to understanding the context or to gaining valuable [mathematical] information. Yet, in both PISA and PIAAC the reading demands of many items were minimised through reduction of the complexity of the text, alongside the use of supporting illustrations. (Tout & Gal, 2015, p. 704)*

This is intrinsically linked to the issue of authenticity. Based on the description and understanding of numeracy in PIAAC, there were deliberate attempts to avoid what are traditionally seen as school, curriculum-based word problems that are often contrived and have little real world relevance or authenticity (Palm 2006, 2008a; Stacey, 2015). Although briefly mentioned in the PIAAC numeracy framework this issue of authenticity was neither highlighted nor discussed in any depth. However, the existing PIAAC numeracy items were in fact developed based on identifying situations and tasks from different countries that provide authentic stimuli and then writing sets of questions using the information in the stimulus.

Authenticity of assessment tasks has now been further researched and documented (Palm, 2006, 2008a & 2008b; Stacey, 2015). This research and associated frameworks should be considered in future iterations of PIAAC and in the development of any new PIAAC numeracy items.

Summary

This section raises a significant number of challenges and points to the need for careful consideration in the revisions to the PIAAC numeracy framework and in the development of any new assessment items. The Review Team recommends that there are four related issues to be explicitly addressed in updating and refining the existing PIAAC framework definition and description:

- disposition to use mathematics
- the ability to see mathematics in a numeracy situation
- critical reflection
- degree of accuracy.

These four issues should be addressed in the background information and related discussions in the framework and, where appropriate, by reworking the definition of numeracy itself and/or in expanding on the framework's elaboration and description of numerate behaviour through its set of facets.

Research related to the issue of authenticity of assessment tasks should also be considered in future iterations of PIAAC and in the development of any new PIAAC numeracy items. Based on the review in relation to *Issue III. Assessment Developments,* the Review Team recommends that PIAAC could better harness the potential of technology to transcend some of the challenges with literacy and/or language that may impede some adults' ability to demonstrate their numeracy capability. However, as stated above in relation to authenticity, literacy demands often go hand-in-hand with the numeracy challenges that adults meet in the real world, so this consideration needs balancing across all items, in order to ensure that items are as far as possible authentic and representative of what adults might meet in their lives.

SECTION II. 21ST CENTURY DIGITAL IMPLICATIONS

This section addresses the issue of ensuring that the PIAAC numeracy assessment better reflects the importance of information in digital formats, representations, devices, and applications as realities that adults have to manage in dealing with the mathematical demands of everyday life in the 21st century.

General comments

There is acknowledgement across sectors—education, government and business—that the skills and knowledge adults now need to succeed in work, life, and citizenship have dramatically changed in the 21st century, often driven by technological advances. The OECD summary describes how:

> *The technological revolution that began in the last decades of the 20th century has affected nearly every aspect of life in the 21st: from how we "talk" with our family and friends, to how we shop, to how and where we work. …. New means of communication and types of services have changed the way individuals interact with governments, service suppliers and each other. These social and economic transformations have, in turn, changed the demand for skills as well. While there are many factors responsible for these changes, … technological developments, particularly information and communications technologies, … have profoundly altered what are considered to be the "key information-processing skills" that individuals need as economies and societies evolve in the 21st century. […]*
>
> *With manufacturing and other low-skill tasks in the services sector becoming increasingly automated, the need for routine cognitive and craft skills is declining, while the demand for information-processing skills and other high-level cognitive and interpersonal skills is growing. In addition to mastering occupation-specific skills, workers in the 21st century must also have a stock of information-processing skills, including literacy, numeracy and problem solving, and "generic" skills, such as interpersonal communication, self-management, and the ability to learn, to help them weather the uncertainties of a rapidly changing labour market. (OECD, 2013c, p.46).*

While this is widely acknowledged, there appears to be little research about the role and impact of 21st century skills in relation to adult numeracy or mathematical literacy skills in general. Most research focusses on numeracy and mathematics practices in the workplace (see e.g., AAMT & AiGroup, 2014; Hoyles et al, 2002; Hoyles et al, 2010; Geiger et al., 2015; Kent et al, 2011; Straesser, 2015; Wake, 2015; Weeks et al, 2013; Zevenbergen, 2004). One of the key outcomes of the research is that because of the impact of technology and digital tools and processes, the mathematics or numeracy tasks that people undertake at work involve more than basic calculation skills or 'by hand' skills and straightforward procedural competence. These practices involve more sophisticated mathematical problem solving skills and understandings, and entail the ability to recognise and engage with the mathematics that is fully embedded within messy workplace settings.

Much 21st century workplace mathematics and numeracy practice is integrated with technology. The skills required in the 21st century include a range of mathematical capabilities such as understanding and interpreting graphical information, interpreting measures in terms of what the data are saying about a manufacturing process, making use of spreadsheets, interpreting visual, computer-generated 3D representations or virtual images, and more. Hoyles et al. (2010) argue that this requirement for mathematical capabilities will be driven by the need to improve production processes and productivity; that is, there will be greater demand for what they call techno-mathematical literacies:

> *We therefore decided to introduce the term Techno-mathematical Literacies, developing from the idea of mathematical literacy that was used in our previous*

> *research … This literacy involves a language that is not mathematical but 'techno-mathematical', where the mathematics is expressed through technological artefacts. (Hoyles et al, 2010, p. 14)*

In relation to technology at work, a recent Australian study about the use of mathematics in the workplace found similar connections and entanglements between mathematics and technology:

> *Many people in the workplace are engaged with technology, particularly in using spreadsheets and graphical outputs. There is an inter-dependency of mathematical skills and the use of technology in the workplace in ways that are not commonly reflected in current teaching practice. The perception is that technology is transforming workplace practices and the use of technology has changed the mathematical skills required – while not reducing the need for mathematics. Through technological change, mathematics has become more important and more embedded in the role of the modern worker. (AAMT & AiGroup, 2014, p. 2)*

Much workplace numeracy and mathematics research also comments on the disconnect between mathematics learnt at school and how mathematics is used in the workplace. The above Australian study about the use of mathematics in the workplace summed up much of this research:

> *Although the skills observed appear to be fundamental, it is their use and application in work contexts that is not straightforward. (AAMT & AiGroup, 2014, p. 1)*

This report went on to describe more fully the differences between school mathematics and workplace mathematics use:

> *Mathematics is applied in both routine and complex tasks requiring sophisticated use of fundamental mathematical skills and 'judgement' or 'problem-solving' procedures. Workplace mathematics is performed differently to school mathematics. Mathematical demands may be present implicitly in the workplace tasks, often through tasks that are not obviously mathematical." (AAMT & AiGroup, 2014, p.2)*

This is consistent with earlier research by Steen in the United States:

> *"Mathematics in the workplace makes sophisticated use of elementary mathematics rather than, as in the classroom, elementary use of sophisticated mathematics. Work-related mathematics is rich in data, interspersed with conjecture, dependent on technology, and tied to useful applications. Work contexts often require multi-step solutions to open-ended problems, a high degree of accuracy, and proper regard for required tolerances. None of these features are found in typical classroom exercises." (Steen 2004, p.55)*

It needs to be emphasised that sense of number still underpins much of the mathematical thinking required—including fluency and flexibility in mental calculations and estimations.

However, it is not simply the demands of those who control the means of production that is driving the use of digital technologies in the workplace; workers themselves also use technology to support their *thinking*. That is, it is not just about the use of digital tools to replace traditional physical or cognitive skills. In particular, digital tools increasingly mediate young workers' ways of reasoning, acting, and working (Zevenbergen 2004; Jorgensen 2011). At the same time, these new ways of thinking and acting are reshaping the structuring practices and deployment of skills in workplaces. Zevenbergen argues that this allows young workers to solve problems in often more inventive ways than their more experienced co-workers do.

More than just workplace practices

Also worth remembering is that PIAAC is a survey of adult competencies across all aspects of adult life—not just about workplace and employment, and not just about engaging with numeracy and mathematics within technologically rich environments. Therefore, the focus must be on both life as an individual, and as part of society and citizenship. Being numerate in the 21st century means being able to cope with the aspects of the world as we encounter it, and that includes with the digital aspects of information and society— the reality is that technology is now ubiquitous throughout all aspects of many societies. Services, interactions and communications outside the workplace have all changed in the 21st century, often driven by technological advances. This includes online processes such as banking, purchases, bookings, reviewing information (health, housing, etc.) and making decisions based on that information. It includes functioning in the bureaucratic world (applying for permits, social security applications and processes, applying for jobs, managing insurances, etc.), use of different media (e.g., the internet, online news, Facebook, podcasts, videos, etc.), use of different aspects of communication (email, SMS, apps, Twitter, Instagram, etc.), and of a range of software and technology at home and in the community. Technology has meant greater market penetration and influence. The influences of social and mass media has implications for informed and participatory citizenship, and hence for citizens to be critical consumers of the mass and social media.

Digital technologies change situations and environments through the way they mediate reasoning and activity and so affect the structuring of practices. Questions to address include in what situations do digital tools provide an advantage? All situations; or are there times when numerate responses are not assisted by use of digital tools? Is the engagement with and use of digital tools and processes also a matter related to user dispositions?

Non-technologically based environments

Many numeracy-related tasks undertaken by adults, however, still relate to non-technologically based environments. Are the skills different when, for example, adults interpret their finances and bank statements using paper-based statements as opposed to using online banking? Is it not how the information is interpreted (and the adult's ability to do that) regardless of how it is presented, that is important? What are the different or unique skills and abilities in play in numeracy actions within the different environments?

In reviewing the PIAAC numeracy framework and assessment, a balance needs to be maintained between numeracy and mathematics actions embedded in digital and technological environments, and numeracy and mathematics embedded in other modes. This should be considered across the four PIAAC Contexts of *Everyday life*, *Work-related*, *Societal or community* and *Further learning.*

Responses to the variable presence of technology in the workplace and in other places in which adults exist must also recognise the ICT-related skills and abilities of the broader adult population that the PIAAC survey assesses. Non-digital test items need to be maintained as part of PIAAC as a proportion of the population has low or no familiarity with ICT devices such as computers. PIAAC should continue to be delivered in both a computer-based and a paper-based version, with filtering processes in place to direct respondents to the most appropriate delivery mode.

Problem solving in technology-rich environments in PIAAC

The PIAAC Survey of Adult Skills also includes an assessment of Problem Solving in Technology-Rich Environments (PSTRE). This is defined as "using digital technology, communication tools and networks to acquire and evaluate information, communicate with others and perform practical tasks". It focuses on "the abilities to solve problems for personal, work and civic purposes by setting up appropriate goals and plans, and accessing and making use of information through computers and computer networks" (OECD, 2012).

The PSTRE results should be considered in the review and implementation of the numeracy construct in the next delivery of PIAAC. Are there constraints due to the ICT-related capabilities of adults across the PIAAC age-range of 16-65 years that will affect what can be asked of the general adult population, in relation to numeracy skills interacting with technology (techno-mathematical literacies)?

The PIAAC survey provides two different pieces of information regarding the capacity of adults to manage information in technology-rich environments: the proportion of adults who have sufficient familiarity with computers to use them for PIAAC tasks, and the assessment of the ability of adults with at least some ICT skills to solve the PSTRE tasks.

The PIAAC PSTRE assessment results show that there are adults with no or extremely limited ICT skills in all of the participating countries. The Survey of Adult Skills found that:

> *From around 7% to 27% of the adult population reported having no experience in the use of computers or lacked the most elementary computer skills, such as the ability to use a mouse. In addition, there are also adults who appear to lack confidence in their ability to use computers, primarily because they use them infrequently. Of the adults undertaking the assessment, most were proficient at Level 1, which involves the use of familiar applications to solve problems that involved few steps and explicit criteria, such as sorting e-mails into pre-existing folders. As would be expected, young adults are less likely than their older compatriots to lack computer skills or to have low proficiency in problem solving in technology-rich environments. At the same time, there are several countries in which the proportion of young adults who can effectively solve more complex problems in computer environments is surprisingly low. (OECD, 2012, p. 98)*

These results from the Problem Solving in Technology-Rich Environments component of the Survey of Adult Skills with its warnings on the high proportion of adults with no or extremely limited ICT skills need to be taken into account in reviewing the technological and digital aspects of the PIAAC Numeracy framework and the associated assessment.

Summary

21st century digital technologies provide tools and processes that mediate thinking as well as action and are not just devices that can be used to complete manual, hands-on tasks more efficiently. Many technologies are thinking tools, which in turn have a range of implications for life, citizenship and the workplace. These tools and processes can often change the numeracy task itself and so transform practices within adults' lives and within the workplace. The use and application of a range of techno-mathematical literacies underpins much of this. These 21st century aspects of numeracy practices need to be reflected in a revised framework and in the associated assessment tasks, but a balance needs to be kept between numeracy and mathematics actions in digital and technological environments versus those embedded in other, non-digital ways.

Issues

Framework definition and description

The new PIAAC numeracy framework, including the assessment content, should consider the available research about the competencies and knowledge adults need to succeed in work, life and citizenship in the 21st century and the impacts of technological advances. This should include considering the implications of research into workplace numeracy and techno-mathematical literacies, and how this impacts on society and the individual in general. These issues should be addressed in both the definitions and the elaborations in the facets described as part of numerate behaviour. This also has implications for the assessment construct and assessment items.

Some of these issues are discussed below under the different parameters of the PIAAC numeracy framework.

Mathematical content

Despite predictions that the availability of digital tools and technologically advanced processes would change the nature of the mathematics we teach and use in the real world, there is little evidence that this has been the case. There is still little connection between the world of school mathematics and the world of mathematics outside the classroom.

Revisions of mathematical curricula across the globe have often undertaken to be futures-oriented at their inception, yet have failed to deliver any transformational changes. The reality is that what is taught across the globe at all levels of education and what teachers work with is very much the same mathematical content with differences being generally superficial; e.g., the way we organise, package and label mathematics for appropriation by students and teachers. What has changed is the emphasis placed on different types of mathematics. As pointed out above, there is less demand for highly developed 'by hand' skills because of the impact of digital tools, although adults still need strong number sense and understandings of basic mathematical operations (e.g., Bakker, 2014; FitzSimons & Coben, 2009; Geisinger, 2016; Marr & Hagston, 2007; Smith, 2004).

Therefore, the review of the existing descriptions of the four content facets in the PIAAC numeracy framework (quantity and number; dimension and shape; pattern, relationships, change; and data and chance) should be reviewed in light of the above research about 21st digital information, representations, devices and applications including of techno-mathematical literacies.

Note: the issue of the mathematical content descriptions in PIAAC is also further discussed under *Issue IV: Relationship with PISA*.

Purposes and demands for numeracy use

The PIAAC Numeracy framework argues that young people and adults need to manage or respond to a situation involving numeracy because they want to satisfy a purpose or reach a goal. The framework defines numeracy and numerate behaviour, including describing the different settings and situations (named as Contexts in the framework) where mathematical demands can be made.

> *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 elements, or which include information of a quantitative nature, that adults face and which pose demands with which they have to cope. (PIAAC Numeracy Expert Group, 2009, p. 11)*

The current framework document proposes and discusses three different aspects related to the potential different purposes and demands for numeracy skills and capabilities:

> *… 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. (PIAAC Numeracy Expert Group, 2009, p. 11)*

In addressing the first aspect, literacy and numeracy in adults' lives, the current framework refers to Kindler's (1996) work that looked at different purposes and functions of using mathematical knowledge. This framework proposed four broad categories regarding the uses of numeracy—these same categories are also used in another numeracy framework (VCAA, 2008). The four categories are *Numeracy for practical purposes*; *Numeracy for interpreting society*; *Numeracy for personal organisation*; and *Numeracy for knowledge*. These categories can be used to reflect on how these different purposes and uses might interact with digital information and technology. Table 1 below shows some possible connections.

Category	**Related to**	**Connections with digital information and technology**
Numeracy for practical purposes	Aspects of the physical world that involve designing, making, and measuring	e.g., many aspects of measuring are now digital – theodolites, inclinometers, medical equipment/monitors, etc. e.g., design aspects are now available digitally, via software such as Computer-aided design (CAD) or online design software for kitchen/house planning
Numeracy for interpreting society	Interpreting and reflecting on numerical and graphical information in public documents and texts	e.g., much digital information is presented in digital and graphical formats, often dynamic in nature, including the use of spreadsheets for analysis. Even common software such as Word has sophisticated graphic and data options available. e.g. use of data, statistics and quantitative information through social and mass media for advertising, news and political information dissemination, etc.
Numeracy for personal organisation	Numeracy requirements for personal organisational matters involving money, time and travel	e.g., digital diaries, online banking, online shopping and planning, GPS and Google maps
Numeracy for knowledge	Mathematical skills needed for further study in mathematics, or other subjects with mathematical underpinnings or assumptions	The degree of technology inclusion is dependent on the programs of study—some are technology intensive, others less so. But often it is expected to be able to use and work with sophisticated digital and technological tools, including calculators, software, etc.

Table 1. Four categories of numeracy use and their connections with technology

The skills and competencies adults now need to succeed in work, life and citizenship have dramatically changed in the 21st century, often driven by technological advances. The above reflection about numeracy use indicates a strong connection and entanglement of digital information and technology with literacy and numeracy use in adults' lives. The ubiquitous presence of social and mass media also carries implications for informed and participatory citizenship, particularly the need for citizens to be critical consumers of such media. This issue of the connection of numeracy with digital information and technology receives no attention under this discussion in the existing framework document, and given that this connection is likely increase, this section and aspect of the framework should be revised and updated to address the issue.

The second aspect of numeracy use discussed in the current framework relates to the mathematical demands of workplace and functional settings. This section is comprehensive but, as mentioned earlier, it is somewhat dated with little reference to digital information and technological advances. The focus is generally on purely mathematical skills and types of mathematics in formal and informal settings. There is very little reference to the role of digital tools and processes. There is acknowledgement that "broader" strategic and critical capabilities are needed—but again little about how digital tools and processes can mediate this sort of thinking. Paragraph 29, pp. 13 and 31 (PIAAC Numeracy Expert Group, 2009) makes reference to technology but this is not discussed in depth and makes no reference to the use of mobile technologies or social media. The roles that digital tools and processes play and whether they enhance or inhibit adults' capabilities to engage with, interpret and resolve numeracy based problem situations should be addressed, especially in relation to the techno-mathematical literacies workplace research discussed earlier. Again, an update and rethink is needed here in light of technological advances, while remembering, as stated previously, that many tasks as an individual, citizen or worker still relate to non-technology-based environments.

The third aspect of mathematics use discussed in the current framework relates to the educational perspectives on mathematical needs of school graduates and citizens. This section recognises the importance of broadening the mathematical skills school graduates possess because of the need to apply these in civic and social situations and the relationship of these skills to the language of science and technology. This relates to the earlier discussions about the disconnect between school mathematics and the world outside the mathematics classroom. Paragraph 37, pp. 15-31 (PIAAC Numeracy Expert Group, 2009) also draws attention to the need for a critical orientation towards the interpretation of mathematical information presented in numerical or graphical form, but again does not mention digital information or related formats.

In summary, there is little discussion of the connections between technology, the different settings and situations in which mathematics is embedded and the demands for numeracy across this section of the current framework document. The impact of the increasingly globalised and technologised world, where digital tools are now near ubiquitous (most situations and environments are now entangled with digital aspects or elements), needs to be reflected in the revised PIAAC Numeracy framework.

Responses/actions

In section 3.1.2 Facet 2: Responses (PIAAC Numeracy Expert Group, 2009, pp. 24-26), the current framework describes and elaborates how in different real-life situations, adults may have to react to a numeracy problem with different types of responses. The current framework groups those under three broad headings: *Identify, locate, or access*; *Act upon or use*; and *Interpret, evaluate/analyse, communicate*.

Integrated into this discussion about these responses are some examples of 21st century digital information and technological scenarios, although there is no discussion, similar to issues mentioned above, about whether such digital information, tools, processes and applications affect the sorts of responses that could be expected of respondents or not. Do digital technologies affect the way people need to respond and interact, through the way they mediate reasoning and activity, and hence on the outcomes and the way responses are structured? For example, the manipulation of virtual three-dimensional objects; the ability to sort and analyse datasets; utilising online calculators (currency exchanges or online purchases), or the ability to interpret the consequences of manipulating relationships or formulae based on real-life variables represented graphically. This section could further develop the work on techno-mathematical literacies, to revise and update these responses to see if they need extending or refining. How does all this relate to and impact on the existing list of actions: *Identify, locate, or access*; *Act upon or use*; and *Interpret, evaluate/analyse, communicate*?

This issue about responses and actions relates to another challenge raised in the current framework. This concerns the limitations and restrictions on conducting the actual PIAAC assessments, including the computer-based assessment platform and the need for immediate scoring of responses given the adaptive testing process. Whilst this is discussed further in *Issue III. Assessment Developments*, there are some relevant points to make regarding the responses aspect and the issue with regard to 21st century digital information and technological advances. Here are some comments from the existing PIAAC Numeracy framework document:

> *… one needs to distinguish between a conceptual framework … and an assessment framework ... 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 behaviour ultimately depends on the technical aspects of that assessment. …*
>
> *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. … 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. (PIAAC Numeracy Expert Group, 2009, p. 26)*

The implications of these technical, implementation restrictions are that they potentially hinder the ability to adequately represent the range of 21st century digital information and technological situations and the consequent actions and responses expected. The existing PIAAC numeracy items generally expect a static response, with little interactivity available either in the stimulus or in the response types.

In the context of the 21st century, especially within a workplace setting, the kinds of numeracy problems that are solved are complex, multidisciplinary and often open-ended. The answers and responses do not come in a multiple choice format with a single right answer—the next PIAAC assessment needs to better reflect 21st Century digital and technological actions and responses.

Representations, reading demands, and authenticity-related issues

Under *Facet 4: Representations of mathematical information*, the PIAAC Numeracy framework states that mathematical content/information/ideas can be represented in multiple ways: objects and pictures; numbers and mathematical symbols; formulae; diagrams and maps, graphs, tables; texts; and finally, technology-based displays. None of these is expanded in much detail (PIAAC Numeracy Expert Group, 2009, p. 28). The issue of different forms of representation of information is raised but digital formats are not addressed.

The nature of information graphics (e.g., Lowrie & Diezmann, 2009) is only now being unpacked within the field of mathematics education. Societies are becoming increasingly reliant on representing information both diagrammatically and graphically. The new, more dynamic representation of data and information needs to be addressed. It is now no longer a matter of interpreting static images, as in the existing PIAAC item pool, but also how new scenarios and different problems can be posed by interpreting and manipulating dynamic representations.

As developments in the 21st century impact on how mathematical and numerical information is represented, the PIAAC framework should be updated to reflect these changes. The Review Team recommends that these could be reorganised under new categories such as physical, representational, and digital.

As discussed in *Issue I. Theoretical Developments* and in *Issue III. Assessment Developments*, both authenticity and literacy-related demands are elements of the numeracy challenges that adults meet in the real world. The assessment items should be, as far as possible, authentic and representative of what adults might meet in their lives. This is true in relation to digital and technological aspects and the need to include such representations in the next iteration of PIAAC.

The Review Team recommends that the next iteration of PIAAC harness the potential of technology to support a more effective and representative 21st century assessment, through greater use of non-text based media in assessment items. Use of different technology, media and associated supports has the potential to transcend not only some of the challenges with literacy and / or language that may impede some adults' ability to demonstrate their numeracy capability, but also to make the assessment more relevant to the 21st Century.

It should be noted that it will be important to address the issue of instructions required for the respondent to understand and manipulate any such interactive stimuli and responses, and how this can be facilitated and supported in order to actualise the effective use of such dynamic 21st Century representations.

The review and updating of the current framework in terms of its representations, reading demands, and authenticity against 21st century digital developments is consistent with the need to update other aspects and facets of the PIAAC numeracy framework and assessment. However, this needs to be balanced with the risks of using more sophisticated technology, to guard against confounding numeracy skills and competencies.

Item formats

The way in which 21st century digital information, representations, devices and applications affects the item types utilised in PIAAC is more related to *Issue III. Assessment Developments* and is discussed more fully there. However, there are specific aspects that are relevant where it is important to realise some of the complex techno-mathematical literacy related skills might require new and different item types. In a report about 21st century skills and how they are assessed, Geisinger concluded that it is:

> *…possible to assess a number of the very complex knowledge and skills considered a part of what is being termed 21st century skills. ... It must be noted that all of the latent constructs measured in these studies are complex. They cannot easily be evaluated via simple paper-and-pencil or multiple-choice assessments. (Geisinger, 2016, p. 248)*

Much of the existing PIAAC item pool is based around static images and are similar to paper-based assessments transferred onto a computer. This does not seem to reflect how many numeracy tasks and actions are now situated and practised in the 21st century. This will require a rethink of the available item assessment formats for the next delivery of PIAAC.

As Gal and Tout said in their comparison of the PIAAC and PISA assessments:

> *the PIAAC computer-based assessment is considerably more limited in terms of its range of item types and responses available compared to PISA. … This enables PISA to offer potentially richer descriptions of students' abilities to reason and communicate their mathematical skills and knowledge. (Gal & Tout, 2014, p. 37)*

As described further in *Issue III. Assessment Developments*, the challenge is how test developers can write computer-based items that take all the above on board and can utilise both 21st century assessment advancements and reflect the use of 21st century digital and technological tools and processes in adults' lives.

The Review Team recommends that the next delivery of the PIAAC numeracy assessment should build on assessment developments and lessons such as those learnt from CBAM in PISA 2012 and including the range of other assessment developments described in *Issue III. Assessment developments*, to introduce new item formats that better reflect 21st century related information, stimuli and numeracy tasks and assessment responses.

The other key challenge in such an international assessment as PIAAC is how to assess some of these 21st century digital and technological aspects of numeracy use in a generic way, so that it is accessible to most of the population across a spread of countries. A balance needs to be kept between numeracy and mathematics tasks and actions embedded in 21st century digital and technological environments with those embedded in modes that are more traditional.

It needs to be remembered that the assessment of numeracy in the first cycle of PIAAC involved the use of 57 test items, which included 19 linking items that had been previously used in ALL. A good proportion of these items will continue to be used in the next iteration of PIAAC as linking items. As this item pool is considered relatively static and not representative of how many numeracy tasks and actions are now situated and practised in the 21st century, it would therefore be opportune that any new test items written to complement the existing pool should take on board potential innovations due to 21st century requirements and enhancements.

SECTION III. ASSESSMENT DEVELOPMENTS

In this section, the Review Team identifies developments in the assessment of numeracy (particularly of adults) that could be relevant for PIAAC – e.g., item types and formats, use of animation, and modelling.

General comments

It must be emphasised that the definition of Numeracy in PIAAC and its associated description of numerate behaviour drives the assessment content and its parameters. Presently:

> *PIAAC defines numeracy as 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. (PIAAC Numeracy Expert Group, 2009, p. 21)*

The PIAAC framework further describes a more detailed definition of numerate behaviour with specification of a number of different facets of numerate behaviour:

> *Numerate Behaviour involves managing a situation or solving a problem in a real context, by responding to mathematical content/information/ideas represented in multiple ways. (PIAAC Numeracy Expert Group, 2009, p. 21)*

Any developments, refinements, and enhancements to assessment processes and capabilities identified in this review may also inform revisions to the definition and elaboration of numeracy, not only to the methods of conducting the assessment. The multifaceted definition of numeracy implies multimodal ways of assessing this concept, and because of new developments in technology and communication, new assessment developments could inform revisions to the definition of numeracy.

It also needs to be acknowledged up front that the existing PIAAC items are based around static images and associated responses and are more like paper-based assessments transferred onto a computer, partly due to the transfer of many of the paper-based items from the previous ALL assessment to the computer-based assessment in PIAAC. As well, the platform used for PIAAC was quite restrictive in terms of modalities and interactions that were available to house the stimulus and for the responses that could be automatically scored, and this platform is now quite outdated.

Assessment developments

There have been technologically-driven advancements in the educational measurement and assessment field in the 21st century, some of them based around the need to assess 21st century skills. Education and educational assessment have moved on from an emphasis on memory and routine problem solving. There is much research about such developments (e.g., see Bennett, 2015; Geisinger, 2016; Parshall et al, 2002; Shute et al, 2016).

At the beginning of the 21st century, computer-based assessments were starting to be accepted as a way forward:

> *Computer-based assessments reflect a relatively modest deployment of technology for testing purposes at the present time. Many computer-based assessments, including adaptive tests, capitalize on innovations along key dimensions: item format, response action, media inclusion, level of interactivity, scoring, and communication of test results (Parshall et al, 2002, p. 40)*

Bennett (2015), who has been researching and mapping educational assessment for a considerable period, describes three generations of assessment. He described first-generation technology-based testing as largely an infrastructure-building activity, laying the foundation for tests to be delivered in a new medium, where much of the testing closely resembled traditional tests. In his description of second-generation tests, he argued that qualitative change and efficiency improvement become the driving goals (Bennett, 1998, 2010), and where the tests use less traditional item formats, moved towards new multimodal formats and where there were attempts to measure new constructs. The driving force was often the technology. Bennett describes a third generation assessment as one of reinvention occurring on multiple fronts where these assessments were able to serve both institutional and individual-learning purposes. They are designed from both cognitive principles and theory-based domains, and where the assessments utilise "complex simulations and other interactive performance tasks that replicate important features of real environments, allow more natural interaction with computers, and assess new skills in more sophisticated ways" (Bennett, 2015, p. 372).

A more recent summary of 21st century advancements endorsed the potential for enhancing assessment opportunities:

> *This wave of innovation, ushered in by advances in the learning sciences and technology, has revolutionized the science of assessment, permitting greater ecological validity and feedback to students related to the breadth and depth of knowledge and skills learned in-situ, including so-called 21st century skills (e.g., critical thinking, creativity, collaboration, and problem solving). That is, advances in technologies and their integration with assessment systems have allowed for the assessment of multidimensional learner characteristics (cognitive, metacognitive and affective) using authentic digital tasks (e.g., games and simulations). (Shute et al, 2016, p. 36)*

Augmented Reality and Virtual Reality

Some of the 21st century technologically driven developments included in the above advancements that can be considered in relation to assessment relate to Augmented Reality and Virtual Reality (e.g., see Bower, Howe, McCredie, Robinson, & Grover, 2014; Sommerauer & Müller, 2014). This diagram in Figure 1 below shows the mixed reality spectrum that could be a model for future assessment development.

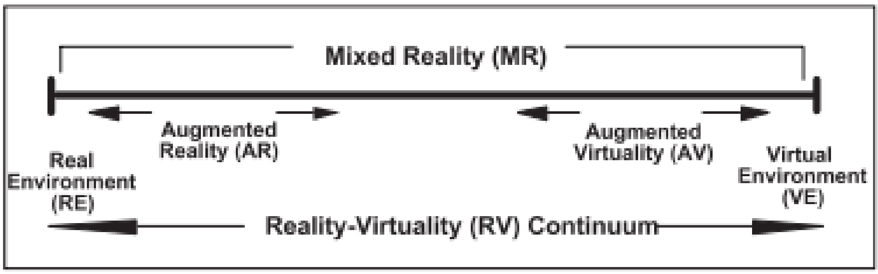

Figure 1. Mixed reality spectrum (Milgram & Colquhoun. 1999, p.9)

In relation to Augmented Reality, Bower et al (2014) stated:

> *It is important for educators to anticipate developments in Augmented Reality so that they can prepare for what is to come. Future developments in Augmented Reality will undoubtedly include new trigger types (sound, temperature, smell, voice recognition, etc.), more intelligent input recognition (e.g., more accurate gesture detection as is already being evidenced by products such as Leap Motion) and increased sophistication of expression types (for instance, vibration, more complex 3D interactive models, scripts to networked devices such as printers and lights as outlined by "Internet of Things" proponents). (Bower et al, 2014, p.12-13)*

In considering the future development and potential enhancements for the assessment of numeracy in PIAAC, there are therefore many different computer-based models and options available to inform how numeracy might be more effectively assessed in future iterations of PIAAC.

Computer-based assessment of mathematics and numeracy

The literature specifically on computer-based assessment of mathematics (CBAM) and multimedia learning of mathematics (e.g., Atkinson, 2005) mostly focuses on the multimodal representation of mathematical concepts: calculating, graphs, diagrams, computer algebra systems, spreadsheets, statistical programs, etc. However, in the computer-based assessment of numeracy another focal point is also of importance, namely the role of the representation of the problem (the situations and settings in which the mathematics is embedded). More general research on representations of real-life situations in education and assessment should also be considered (e.g., see Schnotz, 2002, 2005; Schnotz, Baadte, Müller, & Rasch, 2010; Schnotz & Bannert, 2003; Schnotz & Kürschner, 2008), as well as research on more general multimedia learning (Mayer, 2005, 2009), while being aware of the cognitive load discussion (Sweller, 2005, 2010; Van Gog, Paas, & Sweller, 2010).

Computer-based assessment of mathematical literacy in PISA 2012

The existing PIAAC numeracy item pool is mainly based around static images and the items are more like paper-based assessments transferred onto a computer, which does not seem to reflect how many numeracy tasks and actions are now situated and practised in the 21st century (see Issues I and II). The Review Team recommends that the next delivery of the PIAAC numeracy assessment builds on assessment developments as described above and also learns from the experiences in the development and implementation of, for example, CBAM in PISA 2012.

In an analysis and review of the Computer-based assessment of mathematics items developed for PISA 2012, the ACER test developers created a list of features that benefited and advantaged CBAM test items over traditional paper-based assessments of mathematics:

- Their appeal to students' interactive learning styles increases the engagement of students with the task.
- Items are less dependent on text and reading skills, which means students can access an item from visual cues, and then use the text to confirm the required response details.
- Response modes are more flexible and less daunting. Students can easily edit a response, so they are more inclined to "have a go".
- Relevant calculations can be automated, which means answers are correct, and less time is taken. This allows items to address higher-order mathematical reasoning.
- Items can assess spatial and visual skills using accurate simulations and manipulatives that are not readily available in pencil-and-paper formats.
- Items can test ability to use a wider range of problem solving strategies, such as observation of patterns and trends, and of the effect of manipulations and actions.
- Items can simulate computer-based processes, such as spreadsheets, drawing tools and graphing tools, and handling information in an online environment.
- Systems can collect data about what the student did within an item, such as the time taken, number of clicks, processes followed and the final state. (PISA Mathematics Expert Group, 2009)

A useful classification of the CBAM item types in PISA 2012 was also developed by the ACER test development team for the set of items (PISA Mathematics Expert Group, 2009; Tout & Spithill, 2015). The labels and descriptions for the different categories developed, which are not mutually exclusive, are:

AC Automatic calculation, where calculation could be automated "behind the scenes" to support assessment of deeper mathematical skills and understanding.

AnM Animation, and/or manipulation.

DSV Drawing, spatial, visual cues and/or responses, such as representations of three-dimensional objects that could be manipulated.

ICT Simulation of computer applications (e.g., using the data sorting capability of an 'imitation' spreadsheet).

IGr Interactive graphing allowing automatic mathematical function graphing and statistical graphing.

Web Simulation of web-based applications or contexts, with or without computer-based interactivity (e.g., buying goods online).

Nil A small number of items require no specific CBAM application.

The responses and item types available and that were used and could be **automatically** scored in CBAM in PISA 2012 included, amongst others:

- Selected response formats (e.g., Multiple Choice, Complex Multiple Choice)
- Short numeric responses
- Click-on and hot-spots
- Drag-and-drop
- Pull down menu
- Matching and ordering
- Manipulation of images to a correct, final position and solution.

This range of CBAM item types and related functionality has already been developed and tested on items used in an international assessment; they were developed, trialled and psychometrically checked for validity, reliability, and fairness. This means that they are technically viable for use within PIAAC. One media that the 2012 PISA CBAM test developers had wanted to use and implement was the use of videos as part of the stimulus for items, but this was not able be easily implemented for 2012, but could be considered in future test development for PIAAC.

PISA 2012 also had a significant proportion of open response type questions that are human-marked. A challenge facing PIAAC is that items with an open response format often appear to be more realistic, taking more of an adult approach and closer to problem solving situations in real life. However, this will be an issue if it is required, as in the last cycle of PIAAC, the assessment is adaptive and all items need to be autoscored. An advantage of considering the use of open response type questions that are human-marked, is that responses become a rich source for future item development, and for potential use in auto-scored versions. Although human marking is more expensive and time consuming, it could be worth reconsidering, at least for any field trial stages, in order to do justice to the multifaceted concepts being assessed in PIAAC.

In some PIAAC items, PISA CBAM items, and now in other assessments, one option for auto-scored items is to have Constructed Response Auto-coded items. This involves respondents keying in a relatively short response (e.g., range of numerical values or short set of words) and then an algorithm is applied to score the item. The challenge is to write tightly structured and framed test questions that can enable such restricted open responses, where correct answers can be predicted, and that can be then automatically scored. Having a comprehensive field trial process enables the collection of all of the responses and then refinement and updating of the correct answers.

It is worth noting that through the potential future use of virtual or augmented reality, there may be new ways that the problem situation can be described, depicted, animated, or shown, scored, and marked.

One of the challenges however, no matter the item type, is how to write and develop suitable test items utilising the potential multimodal features of 21st century assessment techniques and capabilities. As Tout and Spithill said in writing about test development for the computer-based assessment of mathematics (CBAM) in PISA:

> *The challenge posed to both the test developers and the IT platform development team and programmers was not to make CBAM a version of the paper-based assessment transferred onto a computer. The intention was to develop items that both reflected the use and application of mathematics within a computer-based environment, but also used the ability to ask different types of questions hence*

> *assessing different aspects of mathematical literacy that were not possible with the paper-based assessment. (Tout & Spithill, 2015, pp. 151-153)*

As such, it is recommended as stated earlier, that the next delivery of the PIAAC numeracy assessment builds on assessment developments to introduce new assessment content, representations and item formats that better reflect 21st century related digital representations, stimuli and numeracy tasks and assessment responses. The 2012 PISA CBAM model is a known, relevant and sound starting point.

Summary

It is important to make a distinction and achieve a balance between the drive that stems strongly from development of technologies, that can be used in assessments (technology-driven) and the drive to design an assessment that is closely related to the concepts that are designed around the construct of numeracy (concept-driven).

In the current situation of the last PIAAC assessment, there is a gap between the sophistication of the concept of numeracy used and the design of the assessment. As acknowledged above, the assessments that exist in the current item pool are relatively simple and one-dimensional. This analysis is corroborated in the literature (e.g., Bennett, 2015).

More sophisticated assessments utilising some of the possibilities outlined above are not necessarily aimed at more complex or higher order skills, but focus more on the multifaceted and multimodal nature of numeracy problem situations encountered in real life. To assess a sophisticated concept of numeracy there is the need for multimodal options to better represent reality, in which the respondents can show their competence (or not).

Based on the above research and discussions, the following section outlines a possible process and structure by which the development of the numeracy assessment could move forward for the next iteration of PIAAC.

Assessment development enhancements to numeracy in PIAAC

Specifically, there is the need to frame any assessment development enhancements to PIAAC numeracy around two underpinning aspects of the PIAAC numeracy construct:

1. that PIAAC is based on a multifaceted concept of numeracy and has an associated multimodal assessment; and
2. that PIAAC is an assessment of how well individuals can use their mathematical knowledge and skills to solve problems stemming from pragmatic (i.e., real-world) situations, needs or demands.

Multifaceted concept of numeracy and multimodal assessment

The PIAAC definition and description of numeracy falls into the category that Maguire & O'Donoghue (2002) called the "integrative phase". They classified the development of definitions and ideas about numeracy into three phases: formative, mathematical, and integrative. In this "integrative" classification, as with PIAAC, numeracy is viewed as a complex, multifaceted and sophisticated construct incorporating each individual's mathematics, communication, cultural, social, emotional, and personal aspects in a real world situation. Numeracy, as with mathematical literacy in PISA, is seen as a sophisticated capability requiring more than just calculations and basic mathematics. These more integrative approaches to numeracy have become influential over the last few decades, as illustrated by projects that define numeracy instructional content standards and assessment frameworks such as in PISA, ALL, and PIAAC and national adult curriculum frameworks/standards (e.g., see DoE, 2013; NALA, 2015; McLean et al 2012; Tertiary Education Commission, 2008). The assessment of such a multifaceted phenomenon therefore requires a multifaceted and multimodal set of assessment items.

Assessing a multifaceted concept with simple assessment tasks has two negative implications:

1. The capabilities of individuals to cope with complex and multifaceted mathematics problems from real life are not assessed in full when the items are too straightforward and one-dimensional.

2. The rollback effect of an international assessment of adult competencies on adult numeracy education practices is not to be underestimated. There is a responsibility for assessment developers such as those in PISA and PIAAC that their framework and assessment items are in sync with and reflect the complex, multifaceted constructs and concepts being assessed.

Moreover, the plea for multimodal assessment of literacies is broader than only in mathematical literacy or numeracy. Compare for instance with Hung, Chiu, & Yeh (2013) in their article: *Multimodal assessment of and for learning: A theory-driven design rubric*:

> *Given the changing nature of literacy, there is an urgent need to develop alternative ways of assessment in support of students' new literacy practices in the digital age. While emergent models of multimodal assessment are being developed in theoretical contexts, the study reported in this paper illustrates how multimodal theories can be realized in classroom practice. (Hung et al, 2013, p. 400)*

Future iterations of PIAAC should harness the potential of ongoing development of multimodal assessments.

Challenges of 21st Century assessment developments

Given the assessment developments described above, especially the enhanced ability to have assessment content based in more realistic, in-situ applications, the assessment of a multifaceted and broad construct of numeracy as defined in PIAAC should endeavour to take advantage of such developments. Enhancing the existing relatively static set of numeracy items and moving towards a more realistic and authentic assessment of adults' numeracy capabilities is recommended.

However, on a pragmatic level, some of the innovations and developments arising from these technologically-driven advancements in educational measurement and assessment in the 21st century will need to be carefully reviewed and considered as to the feasibility of their use and practical implementation for PIAAC in 2021-22. Some issues that would need to be considered include:

- The costs—some technologies, media and tools would potentially be expensive to use and implement—both from a content development point of view (the production of videos/animations/etc.) and from a delivery and implementation perspective (conducting such assessments in people's homes).
- The time available for testing—would the use of such innovations in assessment take substantially more time for the delivery of the PIAAC assessment?
- The feasibility of producing and using any animations, simulations, video or audio support in potentially 30 different languages, which is not necessarily easy, may be costly and would require substantial quality assurance processes of translations.
- The performance of the use of any such innovations, especially the use of simulations (the use of games is highly unlikely to be relevant in such an adult assessment at this point in time), and hence the cost of the mandatory trialling and psychometric checking for performance, reliability and validity.

It would make sense that as part of the revision of the numeracy framework and the associated development of new assessment items a staged, incremental process documenting what is feasible in the immediate future for the next delivery of PIAAC and what could be considered as part of longer term, horizon planning be developed. A timeframe for implementation and possible delivery in 2021-22 would need to be included so that sufficient time was allowed for implementation.

The Review team recommends that a staged expansion of the scope of PIAAC's assessment options be carefully considered. That is, identify what is desirable, balanced by what is feasible, and build towards the goal over a period of time. And it needs to be remembered that there will always be a substantial set of link items from previous assessments and that the Field Trial can check how any new items work compared with existing items. The Review team is not arguing that all new items should be of this enhanced, multimodal type, as other changes to the framework will require other items types to be developed and utilised.

Therefore, to help both describe and identify the possibilities, the next section sketches a possible dimension or spectrum of assessment options and modalities (see Figure 3). The aim is that it will help in developing assessment items that span from what is available now (closed questions and simple contextual problems) to more complex, multimodal assessment items. The latter are either available already, or will become available in the near future (including more complex and depictive representations of real world problems and more ways to interact with the testing tools).

The availability of new multimodal approaches to assessment prior to the next phase of item development provides a developmental spectrum on which a decision can be made as to which interval or spread of this dimension should be utilised for the assessment in the next PIAAC release. The final decisions of the spread of items across the spectrum would be based on assessment, practical and policy perspectives.

Real-world contextualised tasks

In PIAAC the definition and description of numeracy and numerate behaviour is framed from the perspective of the ability of individuals to cope with tasks that are likely to appear in the real world, and that contain mathematical or quantitative information, or that require mathematical or statistical skills and knowhow. This ability can only be assessed when the mathematics is embedded in contexts and situations that are linked to problems that exist in the real life of individuals and that are authentic (or that are at least imaginable). The issue of authenticity of assessment tasks was discussed in some detail in *Issue I. Theoretical developments* and it was recommended that the issue of authenticity should be taken into account in future iterations of PIAAC and in the development of any new items.

The representation of a problem situation and a question together make a so-called contextual mathematical (or numeracy) problem that is the basis of an assessment item. It is useful to distinguish between the various elements of such a numeracy assessment. The more “authentic”, “realistic”, or “imaginable” these problem situations are, the more valuable and valid (i.e. does it measure what it should measure?) is the numeracy assessment.

Below in Figure 2. are two possible simple diagrams of an assessment item, with a problem situation area, the question to be posed, and then the response area/space. How these areas are organised will depend on a number of factors, including the mode of delivery, the response type and the level of interactivity available. In more interactive assessments, the response space can also be integrated with the problem situation space.

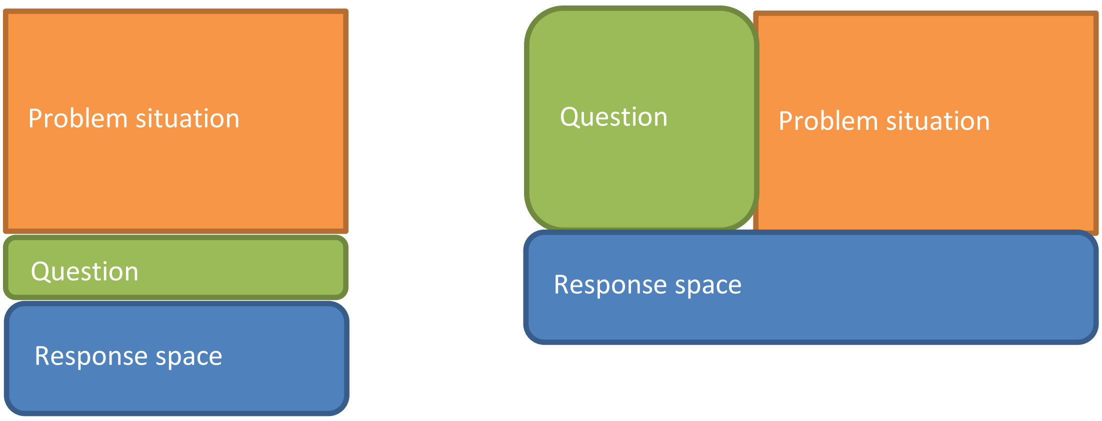

Figure 2. Basic layouts of a contextual mathematical problem

Considering this in its broadest sense, the problem stimulus and situation can be described in words only, depicted through images and illustrations, animated, or shown through virtual or augmented reality, or it might be a simulation of a real situation, or be a real situation.

Similarly, the question can be described, depicted, delivered in text, with voice-over or through a video-question, or otherwise.

The response or answering space can range from the most simple ticking of the box (traditional multiple choice), to numerical fields, to keyboard interface fields. Further, more open and flexible multimodal spaces are now available: matching, click-on, graphing, drag-and-drop, highlighting, etc. Even on-line problem solving environments can be made available and used (spreadsheet, drawing or modelling software, and more). This could lead to a framework with several dimensions, with greater or lesser "sophistication", in the sense of more or less authentic or imaginable problem situations.

The dimension of assessment possibilities

The Review Team sketches (Figure 3 below) a dimension or spectrum of assessment options and modalities for developing assessment items that span from what has been available in the past to assessment items we could see in the future, including some still in the pioneering phase. This dimension deals primarily with the way the problem situation is communicated to the respondent but also includes reference to the response type.

The assumption of this dimension is that respondents have a device available and can access applications developed for PIAAC.

Representing the dimension in Figure 3 there are four classifications included:

- Our categorisation of possible representations from A through to G.
- Possible delivery mechanisms: pen-and-paper or computer-based assessment (CBA)
- The classification of the CBAM interactive item types from PISA 2012 described above.
- Suggestions for possible response types that can be automatically scored.

The recommendation is to potentially use this or a similar spectrum or description of delivery and assessment modalities, as a set of parameters to classify items and to create a spread across these categories to show the multimodality of problem situations and the multifaceted nature of numeracy practices and the PIAAC numeracy construct. This would be similar to how the PISA classification of the CBAM item types described above was used to monitor the range of items selected for PISA 2012.

The challenge of authentic questions that meet requirements for validity and reliability

In the theoretical and practical development of adult numeracy education (and other mathematics education sectors) and related assessment practices there is often the tendency to start from the left in Figure 3 below (decontextualised maths problems) and move to right (more sophisticated, multimodal contextual problems) on the sketched dimension(s), but sometimes the move to the right is limited in its extent. The extent of what is "doable" and feasible, and within the comfort zone of teachers, are limiting factors in this tendency.

Many assessment programs (local, regional, national and international) have tended to take the safest path by continuing to employ established and known assessment practices, hence limiting exploration of new possibilities. That conservative approach has also affected the extent to which experimentation occurs in school-level assessment practices. Additionally, psychometric discussions on validity, reliability, and usability in varied cultural settings has frequently encouraged use of one-dimensional, simple, preferably multiple-choice problems and answers.

However, it does need to be remembered that, for example, the OECD pursued a different direction in the range of CBAM item types used in PISA 2012. A range of item types was developed, trialled and used successfully in that international assessment across different countries and cultures and delivered in more than 30 different languages. They were psychometrically checked for performance, validity, reliability and fairness and found to work successfully and were therefore implemented in PISA 2012. This means that this delivery and associated item types are technically viable for use within PIAAC.

From both the above analysis and what is currently available for the delivery of assessments, categories A – D are commonly available and can be used straight away. For E and F, field trials and piloting are necessary, and consequent psychometric analysis needs to be undertaken to confirm their performance (as would still need to happen for any new items developed across categories A – D). For

G, some pioneering work and extensive piloting and psychometric analysis would be required (e.g., see Bower et al, 2014; Sommerauer & Müller, 2014). The evidence from the use of the PISA 2012 CBAM items demonstrates that at least up to category F is achievable.

The Review Team recommends that this dimension of possibilities, or some similar schema, should be taken into account in the development of the next set of PIAAC numeracy items. Having a spread of items developed and selected from across this spectrum would enable PIAAC numeracy to be better representative of the PIAAC numeracy framework and construct, add to the issue of authenticity and hence better assess adults' capabilities and competencies around numeracy practices in the 21st Century.

As mentioned previously, the existing item pool is considered relatively static and not representative of how many numeracy tasks and actions are now situated and practised in the 21st century. Given that a proportion of these items will continue to be used as link items, it would therefore be opportune that any new test items developed to complement the existing item pool should take on board potential enhancements and innovations due to 21st century requirements and assessment capabilities and have new items that are to the right on the spectrum of possibilities.

Challenges and Issues

First, as mentioned in the discussions about 21st Century Digital Implications, the results from the Problem Solving in Technology-Rich Environments (PSTRE) assessment in PIAAC need to be considered in relation to the review and implementation of the PIAAC numeracy construct. Are there constraints, due to the ICT-related capabilities of adults across the PIAAC age-range of 16-65 years, which will affect what can be asked of the general adult population in relation to numeracy skills interacting with technology? How could this be best reflected by selecting a range of item types and representations from across the proposed dimension of possibilities? Furthermore, as the Review Team has stated previously, this needs to be balanced against the risks of using too much sophisticated technology; there needs to be a balance between an individuals' ability to utilise digital tools to answer questions and their ability to solve the numeracy problem. The technology should not significantly, negatively impact on numeracy skills and competencies. Technology, on the other hand, now facilitates multimodal assessment capabilities, and may support numeracy tasks and questions being both more authentic and more accessible, with less reliance of reading texts.

Second, On a pragmatic level, some of the innovations and developments arising from technologically-driven advancements in educational measurement and assessment in the 21st century will need to be carefully reviewed and considered as to the feasibility of their use and practical implementation for PIAAC in 2021-22. The issues identified included:

- the costs—some technologies, media and tools would potentially be expensive to use and implement
- the time available for testing
- the translation issues of producing materials in potentially 30 different languages
- the suitability and performance of the use of any innovations.

It would make sense that as part of the revision of the numeracy framework and the associated development of new assessment items a staged, incremental process documenting what is feasible in the immediate future for the next delivery of PIAAC and what could be considered as part of longer term, horizon planning be developed.

Cultural fairness and applicability is also an issue in international assessments and brings its own set of challenges. Again, the standard practice is that a set of items will be developed, trialled, and psychometrically checked for validity, reliability, and fairness, including across countries. Similarly fairness and bias in relation to other characteristics such as gender would need to be checked and monitored.

The dimension of assessment possibilities

The figure represents a spectrum of options and possibilities for the delivery of assessments of numeracy where the scoring can be done **automatically**.

Category:	**A**	**B**	**C**	**D**	**E**	**F**	**G**
Type of representation	Decontextualised maths problem 2 x 3 = …	Simple contextual word-based problems You have 25 sheep. Three are stolen. How many do you have left?	Sophisticated contextual problems with images and descriptive representations (but no interactivity in response space)	Sophisticated contextual problems … with depictive representations and interactivity in response space, but no interactivity in **situation** space	Sophisticated contextual problems Short video clips or animations as representation with interactivity in response space, but no interactivity in **situation** space)	More sophisticated **multimodal** contextual problems with interactivity in **both** the **situation** and **response** spaces	Content of all previous categories, with augmented or virtual reality, with simulation of real situations or in real situations. Full interactivity.
Possible delivery	Pen-and-paper CBA	Pen-and-paper CBA	Pen-and-paper CBA	CBA	CBA	CBA	CBA
PISA interactivity classification	Nil	Web (static only) Nil	Web (static only) Nil	Web (static only) Nil	AnM Web (static only) Nil	AC AnM DSV ICT IGr Web	AC AnM DSV ICT IGr Web …
Possible response types (automatically scored)	Multiple Choice Numerical field	Multiple Choice Numerical field	Multiple Choice Numerical field	Multiple Choice Numerical field, Click on, drag and drop, pull down menu, matching, ordering, etc.	Multiple Choice Numerical field Click on, drag and drop, pull down menu, matching, ordering, etc.	Click on, drag and drop, pull down menu, matching, ordering, manipulating fields to create a correct solution (e.g., spreadsheet,), digital working space with digital tools	Click on, drag and drop, pull down menu, matching, ordering, manipulating fields to create a correct solution, digital working space with digital tools, physical actions in simulations (e.g., choosing an object by grabbing it)

Figure 3. Framework on increasing sophistication of representation of reality in contextual mathematical problems.

SECTION IV. RELATIONSHIP WITH PISA

In this section, the Review team discusses how the relationship between the PIAAC numeracy framework and the PISA mathematical literacy framework and assessment should be conceived, developed (if appropriate), and presented.

Note: in this section the references used are as those written and documented in the original, full PIAAC Numeracy framework (PIAAC Numeracy Expert Group, 2009). Similarly, the references to PISA are mainly to the PISA 2012 mathematical literacy framework and its descriptions (OECD, 2013a). This is because in 2012 mathematical literacy was the major domain for PISA, with the relevant framework revised and updated. It remains the mathematical literacy framework for PISA 2015.

General comments

Based on detailed comparisons of the two numeracy frameworks for PISA and PIAAC by the Review team, by Gal and Tout (2014), and from two of the Review Team who are familiar with both the full sets of PISA and PIAAC items and not just the publicly released items, it is apparent that both assessments describe and cover very similar territories.

On the conceptual level, numeracy and mathematical literacy are closely related constructs in terms of their core, underlying ideas. In relation to the definitions and descriptions of the constructs and what they are assessing, Gal and Tout, in their comparison of the two programs, summarised the similarities:

- *Both constructs refer to the ability of individuals to cope with tasks that are likely to appear in the real world, and that contain mathematical or quantitative information, or that require mathematical or statistical skills and knowhow.*
- *Both constructs focus on how well individuals can use their mathematical knowledge and skill to solve problems stemming from pragmatic (i.e., real-world) needs or demands, and to 'engage', manage, and understand various tasks in the world around them—rather than addressing decontextualised mathematical tasks.*
- *Both PISA and PIAAC describe mathematical literacy or numeracy as not synonymous with a minimal or low level of mathematical knowledge and skills. That is, both assessments view the constructs as describing competencies lying on a continuum, i.e., individuals could be placed on a scale from low levels to high levels. (Gal & Tout, 2014, pp. 47-48)*

They concluded that both the PIAAC and PISA frameworks and assessments have substantial conceptual similarities and also practical commonalities in their test items and design principles, as well as in the range of content areas and skills they cover (Gal & Tout, 2014). However, there are some differences between the two assessments, related to the diversity in the experiential backgrounds and the distances from schooling for adults compared to children. As Gal and Tout wrote:

> *Because many adults may not remember more formal school-based representations or technical language, the design of PIAAC items has taken into account from the outset the need to establish authenticity while reducing the use of formal notations and 'school-like' appearance. (Gal & Tout, 2014, p. 39)*

An examination of the item sets of both PISA and PIAAC shows that PISA is more interested in the ability of 15-year-olds to use and apply curriculum-based mathematical skills and knowledge embedded in a real world situation. On the other hand, the focus of PIAAC is somewhat less interested in how respondents use formal mathematical skills when solving a real life type mathematical problem. For example, in some of the PISA mathematical literacy items 15-year-olds are asked to use information from a real life situation to calculate and identify specific formal characteristics of linear equation graphs, such as the gradient and the y-intercept. This type of more formal mathematical knowledge is not assessed in PIAAC, as generally PIAAC respondents are not required to show evidence of their knowledge of the use and understanding of formal school-based mathematical notations, which are often forgotten from not having been in current or recent use.

Drivers/indicators of mathematical proficiency

One of the important features of both the PISA and PIAAC frameworks is the way that each has independently developed a schema that describes aspects of test items that drive item difficulty, and which indicate the mathematical proficiency of tested individuals and populations.

PIAAC does this in considerable detail in the Annex to the framework (PIAAC Numeracy Expert Group, 2009, pp. 44-56). As well as classifying test items according to the mathematical content knowledge required to complete each item, the Annex presents a detailed scheme designed to show the complexity of test items. Five 'complexity factors' are defined, and a scheme is presented for rating mathematical tasks according to the extent each factor is present in the test items. Examples are given to show how the rating scheme would be applied, and the assumption is that a total score across the factors for an individual item (20 score points could be generated) would be strongly related to item difficulty; and by implication, successful completion of particular items can be used as indicators of levels of mathematical proficiency.

The PISA framework specifies a set of 'fundamental mathematical capabilities', the activation of which is assumed to collectively provide indicators of mathematical literacy. A scheme designed to rate individual items based on the extent to which each of those capabilities is needed to respond to PISA questions has been developed as part of research activities documented by Turner, Blum and Niss (2015). This research has shown that the scheme predicts the difficulty of PISA test items. Evidence of activation of the capabilities is fundamental to PISA's descriptions of growing mathematical literacy.

Members of the Review Team familiar with the two schemes have been able to use and apply each scheme to the other program's test items. Given the richness and the research backing up these two schemes, it is recommended that they might profitably learn from each other as work on both PISA and PIAAC continues. Further research could be usefully undertaken to establish the degree to which the PISA capabilities could enrich the PIAAC complexity scheme, and vice versa, noting of course that the PIAAC scheme focusses essentially on characteristics of tasks, whereas the PISA scheme has its immediate focus on characteristics of people.

Summary

The Review Team agrees that given the general consistency in what is being assessed by both PIAAC and PISA, and based on their review and knowledge of the two frameworks and item pools, there is much to gain from making significant and more explicit connections and links between the PIAAC numeracy framework and the PISA mathematical literacy framework and assessment.

The first recommendation is that the relationship between the two scales should be established empirically. This would make analysis and comparison across the two assessments stronger and more useful for research purposes, both within countries and internationally. It would also enable research into how items are approached differently by those in school versus those that are not, and support provision of stronger data to enable research into progress from school into adult life. One way of achieving this would be that PIAAC includes a selection of appropriate PISA items so that the two scales can be anchored against each other and the two scales aligned psychometrically. The consequences, however, would be that this would require the use of probably a minimum of 15 items from PISA and either add considerably to the total number of items and/or reduce the ability to also include a range of new items as recommended earlier in relation to both 21st century skills and assessment developments. This could be seen as a one-off occurrence to PIAAC to establish the link with PISA. Alternatively, without affecting the delivery of PIAAC, the relationship between the scales could be established through an appropriately designed linking study. This would require between 15-20 items from each of PISA and PIAAC to be included and delivered as a separate, one-off assessment to suitable sample populations.

The following sections look at how strengthening the links between PISA and PIAAC could occur across a number of the features and parameters of PIAAC.

Issues

Framework definitions and descriptions

> *PIAAC defines numeracy as 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. (PIAAC Numeracy Expert Group, 2009, p. 21)*

PIAAC has a supporting description of numerate behaviour that specifies a number of different facets:

> *Numerate Behaviour involves managing a situation or solving a problem in a real context, by responding to mathematical content/information/ideas represented in multiple ways. (PIAAC Numeracy Expert Group, 2009, p. 21)*

The four facets that are described are: *Contexts*, *Responses*, *Mathematical content/information/ideas*, and *Representations*. Furthermore, the elaboration also lists and describes several enabling factors and processes, whose activation underlies numerate behaviour.

Comparing this to the PISA definitions and elaboration, mathematical literacy is defined in PISA 2012 and PISA 2015 as:

> *… an individual's capacity to formulate, employ, and interpret mathematics in a variety of contexts. It includes reasoning mathematically and using mathematical concepts, procedures, facts, and tools to describe, explain, and predict phenomena. It assists individuals to recognise the role that mathematics plays in the world and to make the well-founded judgments and decisions needed by constructive, engaged and reflective citizens. (OECD, 2016, p. 13)*

One core idea behind the PISA description of mathematical literacy is modelling, which assumes that when individuals use mathematics and mathematical tools to solve problems set in a real world situation, they work their way through a series of stages as depicted in Figure 4 (OECD, 2013a, p. 26).

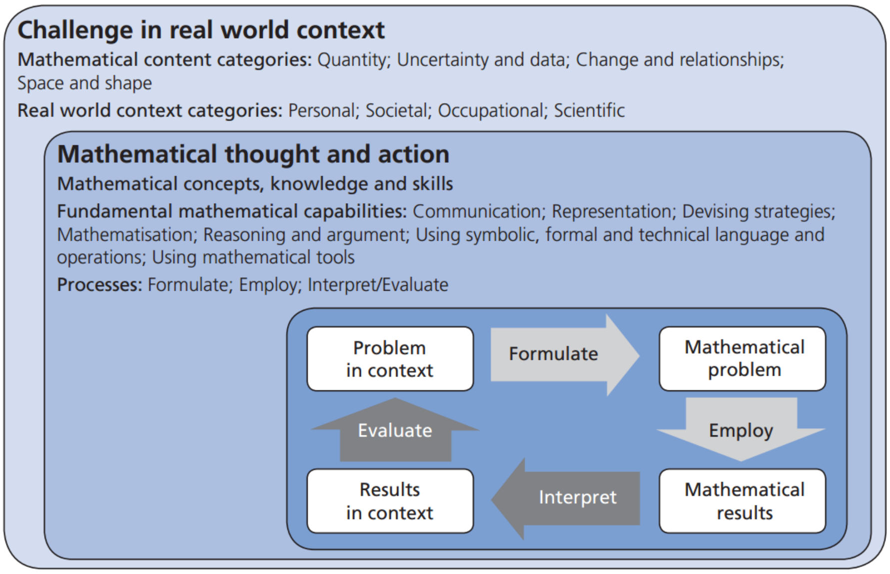

Figure 4. The PISA model of mathematical literacy in practice

The *formulating*, *employing*, *interpreting, and evaluating* processes are key components of the mathematical modelling cycle that has underpinned the mathematical literacy construct in PISA since its beginnings, and has been elaborated and enhanced in the description and definition of mathematical literacy for PISA 2012 and continuing for PISA 2015. These processes each draw on the problem solver's fundamental mathematical capabilities and knowledge in four overarching Content areas: Quantity and number, Uncertainty and data, Change and relationships, and Space and shape. PISA assessment items collectively (though not each individual item) are expected to reflect all the stages in the mathematical modelling process cycle illustrated in Figure 4.

The PISA definition and description is somewhat more sophisticated than the current PIAAC version. The PIAAC definition is very reactive as it refers to responding to situations (interpret and communicate) while PISA has elements of being proactive (explain and predict). Additionally the PISA definition points in the direction of being critical (making well-founded judgements and decisions).

The relationships between the two framework aspects of mathematical Content and Contexts and how they are named and described in PIAAC and PISA are discussed in the following two sections.

Another existing model for numeracy in the twenty-first century is illustrated in Figure 5 below, and has some consistency with both the PISA model and PIAAC's framework. This model incorporates four dimensions of settings/contexts, mathematical knowledge, tools, and dispositions that are embedded in a critical orientation to using mathematics (Goos et al, 2014, p. 84). These dimensions are described more fully in other publications (e.g., Geiger et al. 2014; Goos et al, 2014).

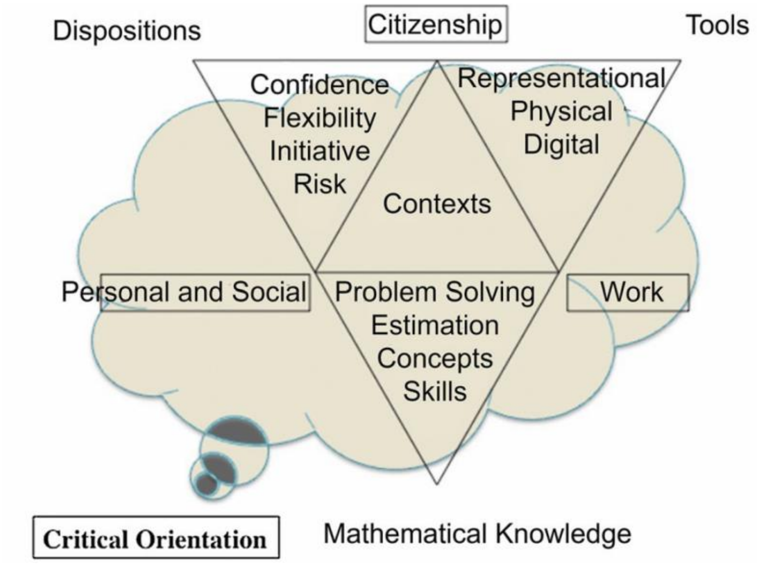

Figure 5. Model for numeracy in the twenty-first century

It is recommended that both the PISA problem solving model and its set of processes and this model for numeracy in the twenty-first century be considered in reviewing and refining PIAAC's definition of numeracy and particularly in how it conceives and describes numerate behaviour through its facets. This would then influence the assessment construct and parameters for item development and coverage.

One of the issues recommended for consideration in *Issue I. Theoretical Developments* was the notion of critical reflection in relation to PIAAC's elaboration and descriptions of numerate behaviour, especially in relation to *Facet 2: Responses*. Critical reflection or having a critical orientation is an aspect that could be emphasised and described further, and is named and described explicitly in the above model for numeracy in the twenty-first century. As well, this aspect was incorporated within the third process in the PISA mathematical literacy problem solving cycle, *Interpreting and evaluating*. Both should be considered in the review of the PIAAC numeracy definition and descriptions.

Mathematical content

Both numeracy in PIAAC and mathematical literacy in PISA use a non-school-curriculum focused approach to naming and describing the content areas covered in their assessments. The purpose behind both frameworks is describing mathematics for use and application outside of the classroom, and so the organisational structure for mathematical content knowledge is based on how mathematical phenomena are encountered in situations in the outside world. While the PISA and PIAAC frameworks were developed by independent teams, they use very similar descriptors for their content classifications, introducing and describing these in terms of the *big ideas* behind mathematics.

Table 2. below compares the terms used for the content areas in PISA 2012 with those used in PIAAC, and the recommended proportions of items in each category.

PIAAC	**PISA 2012**
Quantity and number (30%)	Quantity (25%)
Dimension and shape (25%)	Space and shape (25%)
Pattern, relationships and change (20%)	Change and relationships (25%)
Data and chance (25%)	Uncertainty and data (25%)

Table 2. Comparison of content terms in PISA and in PIAAC.

The two frameworks appear highly consistent in terms of their descriptions and structures of the mathematical content covered in their assessments. There are very similar spreads across each content area; PIAAC places a slightly higher emphasis on quantity and number than on pattern, relationships and change, most likely on the assumption these areas are slightly less prevalent in adult life. The PISA assessment indicated an equal percentage breakdown of each classification suggesting a slightly stronger emphasis on the more formal aspects of mathematical knowledge.

In light of the above and the comparative papers by Tout and Gal (Gal & Tout, 2014; Tout & Gal, 2015), the Review Team recommends that it would be appropriate to bring consistency and coherence to the labelling of what are essentially the same content areas in PISA and PIAAC. There is a lot to be gained by using the same language in both surveys.

Contexts

An important, common aspect of both the PIAAC and PISA frameworks is their emphasis on the use and application of mathematics in a real-life or realistic setting. PIAAC seeks to identify the opportunities that exist for adults to use mathematics in the various life situations they experience (including work-related, as well as those related to their personal life and in their interactions with the social and commercial worlds). PISA also seeks to identify situations in which students might confront an opportunity to use their mathematical knowledge, and uses these to devise problem situations in which mathematics is embedded.

Context is the parameter or term used in both PISA and PIAAC for naming and classifying these settings or situations where people use and apply their mathematical knowledge to solve a realistic problem. The main purpose behind the use of the chosen Context categories in both PISA and PIAAC is to ensure a mixture or blend across the four categories to help guarantee some degree of balance in the assessment, with no particular Context overwhelming the others (and therefore advantaging or disadvantaging respondents with greater or lesser daily interaction with some settings/contexts).

The names used for Contexts within PIAAC

Firstly, however, there are a number of different labels used for naming the Contexts within different existing PIAAC documents, and this should be addressed. In the original, full PIAAC Numeracy framework (PIAAC Numeracy Expert Group, 2009), which is the focus of this review, four Contexts are named and described as one of the facets of numerate behaviour: *Everyday life*, *Work-related*, *Societal or community* and *Further learning* (PIAAC Numeracy Expert Group, 2009, pp. 23-24).

In the Literacy, Numeracy and Problem Solving in Technology-Rich Environments: Framework for the OECD Survey of Adult Skills document (OECD, 2012), the four Contexts are labelled: Everyday life; Work; Society; and Further learning. In the OECD Skills Outlook 2013: First results from the Survey of Adult Skills document (OECD, 2013c) the Contexts for numeracy are: Personal, Work-related, Society and community, and Education and training. These four labels are consistent with the literacy Contexts used and named in that latter report. The descriptions are consistent across the different documents, so it appears that it is mainly the labels that have changed. The first recommendation therefore is that the labels used for naming the Contexts in the different PIAAC documents are made consistent for the next iteration of PIAAC.

Comparison of PISA and PIAAC Contexts

Note that in the review below the labels used are as they appear in the original, full PIAAC Numeracy framework (PIAAC Numeracy Expert Group, 2009).

Table 3. below compares the terms used for the Contexts in PISA 2012 (and 2015) with those used in PIAAC, and the recommended proportions of items in each category.

PIAAC	**PISA 2012**
Everyday life (25%)	Personal (25%)
Work-related (25%)	Occupational (25%)
Societal or community (25%)	Societal (25%)
Further learning (25%)	Scientific (25%)

Table 3. Comparison of Context terms in PISA and in PIAAC.

The sets of descriptors used in both frameworks regarding the first three Contexts (Everyday life/ Personal; Work-related/Occupational; Societal or community/Societal) are highly consistent with each other, with a similar spread of items. Any differences are due to the age of the two target groups, with some of the PIAAC example situations and situations described being more relevant to adults, and some of the PISA situations being more appropriate for 15-year-olds.

One of the Review Team's recommendations is that the PISA label *Personal* is preferable to the PIAAC label of *Everyday*. "Everyday" suggests some "sameness" in what people do which is not particularly challenging, whereas the term *Personal* aims to indicate that the issue at hand bears most directly just on that individual.

Another suggestion relates to *Work-related* versus *Occupational*. Here is the description from PISA for *Occupational*:

> *Problems classified in the occupational context category are centred on the world of work. Items categorised as occupational may involve (but are not limited to) such things as measuring, costing and ordering materials for building, payroll/accounting, quality control, scheduling/inventory, design/architecture and job-related decision making. Occupational contexts may relate to any level of the workforce, from unskilled work to the highest levels of professional work, although items in the PISA survey must be accessible to 15-year-old students. (OECD, 2013a, p. 37)*

For comparison, here is the description from PIAAC for *Work-related*:

> *At work, one is confronted with quantitative situations that often are more specialised 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, totalling 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. (PIAAC Numeracy Expert Group, 2009, p. 23).*

While there is a lot of consistency, there are a number of reasons why there is a need to compare the two categories and refresh the description of the *Work-related* Context in PIAAC. Firstly, the earlier discussions in *Issue II: 21st Century Digital Implications* relating to techno-mathematical literacies in workplaces needs to be better reflected in the PIAAC framework and assessment. Secondly, the current description starts by referring to “quantitative situations”—work-related settings cover a broad range of mathematical skills and knowledge, not just quantitative, and this should be broadened and described. Third, the phrase in the PISA description “Occupational Contexts may relate to any level of the workforce, from unskilled work to the highest levels of professional work” (PISA, 2013a, p. 37) is an aspect that could be better reflected in the PIAAC description. The Review Team feels that the PISA label *Occupational* is a better term to use, and would make the two surveys consistent in their terminology.

The fourth Context, *Further learning* in PIAAC and *Scientific* in PISA, is the Context that needs significant thought and consideration. In PIAAC this Context is described as related to adults needing to solve problems that may arise when participating in further study, whether for academic purposes or for vocational training, and is explicitly related to knowing about the more formal aspects of mathematics, including the conventions used to apply mathematical rules and principles. This has some similarity and consistency with the term intra-mathematical that PISA refers to within its description of *Scientific*:

> *… Particular contexts might include (but are not limited to) such areas as weather or climate, ecology, medicine, space science, genetics, measurement and the world of mathematics itself. … Items that are intramathematical, where all the elements involved belong in the world of mathematics, fall within the scientific context. (OECD, 2013a, p. 37)*

In OECD Skills Outlook 2013: First results from the Survey of Adult Skills (OECD, 2013c), this Context was named as Education and training. The sample item from this Context that is discussed in this document was the item “Candles” (OECD, 2013c, p.77). However, this item could also have been classified as Everyday life, Work-related, or even Societal or community. Is Education and training a better title than Further learning, or does Education and training overlap too much with Work-related or Occupational?

The Review Team suggests that the PISA classification and description of *Scientific* can be used to reflect on the *Further learning* Context of PIAAC*.* However, note that the PIAAC description of *Societal or community* includes reference to situations such as pollution or interpreting the results of a study about a health issue. Another possibility is the name *Technical*, which would possibly work as an alternative. Therefore, the review of the PIAAC Framework needs to address what this fourth Context of *Further learning* covers and includes, and the question of what the most appropriate name for it is.

Responses/actions

In section 3.1.2 Facet 2: Responses (PIAAC Numeracy Expert Group, 2009, pp. 24-26), the current framework describes and elaborates how in different real-life situations, adults may have to react to a numeracy problem with different types of responses or actions. The current framework groups those under three broad headings: *Identify, locate, or access*; *Act upon or use*; and *Interpret, evaluate/analyse, communicate*. PISA uses the three processes of “formulate”, “employ” and “interpret”, as its structure for organising the mathematical processes that describe what individuals do to connect the setting and application of a problem with the mathematics and thus solve the problem.

Table 4. below compares the terms used for the response-related descriptions in PIAAC with the three processes used in PISA 2012, and the recommended proportions of items in each category.

PIAAC	PISA 2012
Identify/locate/access (10 %)	Formulating situations mathematically (25 %)
Act on/use (order, count, estimate, compute, measure, model) (50 %)	Employing mathematical concepts, facts, procedures, and reasoning (50 %)
Interpret/evaluate (40 %)	Interpreting, applying and evaluating mathematical outcomes (25 %)

Table 4. Comparison of response descriptions used in PISA and in PIAAC.

Unlike a number of the other facets of numerate behaviour in PIAAC and their related descriptions, this facet of responses or actions has the least consistency between PISA and PIAAC. Tout and Gal concluded in their comparison between PISA and PIAAC 2012 that:

> *…there is a high level of consistency in their third category, and some consistency between the second pair of categorisations. However, there is a difference in that the PIAAC category of Act upon or use also includes processes that fit more closely to some of the PISA classification of Formulating, e.g., the aspects relating to modelling and developing a formula. Similarly, the first category has some commonalities, but in PIAAC is narrower in its spread. (Tout & Gal, 2015, p. 698)*

This aspect has also already been addressed in a number of the earlier Issues as requiring revision and enhancement. It would appear that this is an aspect of the PIAAC framework that probably needs major review, and PIAAC could potentially learn from the processes described in PISA.

Representations, reading demands, and authenticity-related issues

This issue of representations, reading demands, and authenticity in PIAAC has been discussed in a number of the earlier sections.

Item formats

In their review and comparison of the two numeracy frameworks for PISA and PIAAC, Gal and Tout (2014) concluded in relation to the issue of item formats that:

> *PISA 2012, with its more comprehensive range of item types and more interactive computer-based assessment, will enable richer and extended descriptions of sub-components of mathematical literacy compared to the information that can be generated by the numeracy assessment in PIAAC. (Gal & Tout, 2014, p. 52)*

Furthermore, as discussed in *Issue II. 21st Century Digital Implications* and *Issue III: Assessment Developments*, much of the existing PIAAC item pool is based around static images and is more like paper-based assessments transferred onto a computer, and this also does not seem to reflect the way numeracy tasks and actions are now situated and practised in the 21st century. As stated in these earlier sections, the Review Team recommends that the next delivery of the PIAAC numeracy assessment should build on assessment developments and lessons such as those learnt from CBAM in PISA 2012 to introduce new item formats.

SECTION V. NUMERACY OR MATHEMATICAL COMPONENTS

The utility and feasibility of implementing an equivalent to the reading components assessment in PIAAC's numeracy assessment is discussed in this section.

General comments

PIAAC has included for the first time an assessment of 'reading components' to evaluate how well individuals with low levels of proficiency master the basic building blocks of reading.

Brief background to the reading components

In discussing the reading components, Grotlüschen et al (2016) stated:

> *The reason for introducing reading components into the PIAAC literacy framework and surveys was to better understand what literacy skills adults below Level 1 possessed or lacked. These were individuals who in previous surveys essentially could not answer any of the literacy items correctly and therefore were assigned a score at the bottom of the literacy scale. Were these individuals entirely nonliterate? … The research literature suggested that, given the opportunity to engage in fine-grained skills tasks, at least some of these adults would demonstrate some level of literacy knowledge and skills (e.g., Baer, Kutner, and Sabatini, 2009; Jeantheau, 2006). These fine-grained skills are what we call reading components. … The foundational skills adults possess may be overlooked if only the performance on relatively complex, applied, real world tasks is observed. (Grotlüschen et al., 2016, p. 60)*

The delivery of the reading components included oral support by the test administrator and each component was timed in order to be able to get an estimate of fluency and automaticity:

> *Reading components were only administered to adults who took the paper-based version of the PIAAC survey. The paper booklet began with simple instructions that were read to the adult by the survey administrator. The three task sets always appeared in the same order – Vocabulary, Sentences, and Passages. … The administrator started the timer when the adult turned the page for the first item, then stopped it when the adult completed the final item. (Grotlüschen et al., 2016, p. 67)*

The purpose of the numeracy components assessment would have the same aims: to develop a set of "fine-grained tasks" so that "at least some of these adults would demonstrate some level of **numeracy** knowledge and skills".

PIAAC performance results at Pre level 1

The overall performance in PIAAC is that 5% of adults surveyed in the first wave of 24 countries were at Pre-Level 1. When including the second wave of countries, the results showed 6.7% were at Pre-Level 1 across 33 countries. This compares with a performance in reading literacy of 3.3% across the original 24 countries, and 4.5% for the second wave of 33 countries.

The percentages of adults performing at the lowest level in numeracy are significantly higher when compared with literacy (OECD, 2013c, 2016). Therefore, there is a very strong argument for developing an equivalent to the reading components assessment in the PIAAC numeracy assessment based on the higher numbers of adults performing at that level compared with reading.

Research

In numeracy, such "component skills" for adults have not been researched, theorised or examined quantitatively compared with how they have been in literacy (as briefly described above for the reading components in PIAAC). Therefore, for a numeracy components assessment in PIAAC, considerable work would need to be undertaken to establish what such components are and how they relate to the existing PIAAC below level 1 items and their descriptions. A range of potential sources of content would need to be researched that could be used as the basis of developing and trialling a numeracy components assessment.

In the sections below the Review Team considers what relevant research and information exists that could form the basis of such a development.

Issues

Delivery constraints

One critical issue to address in the development of a numeracy components assessment is that of the constraints imposed by the practicalities of delivering such an assessment. Factors to consider include: the time available; the number and range of items in terms of content areas, difficulty levels, and amount of text; the delivery and item types (oral instructions and support by administrator; pen-and-paper versus online delivery; interactive or not) and more. A number of these issues are discussed below, but the final solution that can be implemented will be dependent on the decisions made about the delivery issues mentioned here.

Given the likelihood of an interaction between low numeracy and low literacy skill levels, delivery of a numeracy components assessment should take special account of the reading demands of the assessment.

Framework definition and description

Once the specifications and details of the content and structure of potential components of numeracy is developed and described, the PIAAC definition of numeracy and the associated description of numerate behaviour would need to be reviewed in order to see if it needed to be changed to accommodate the numeracy components assessment.

Mathematical content and embedded nature of the items

There are two major challenges to consider in the development of a numeracy components assessment. One is the breadth and the level of the mathematical content that should be covered, and second, whether the use of real-world problems embedded in authentic situations is feasible in an assessment of the components of numeracy, or at least what considerations need to be accommodated in order for the components assessment to work.

As with the reading components, the aim is to better understand what numeracy and mathematical skills adults scoring below Level 1 possess or lack. These will be the individuals who in previous surveys essentially could not answer any or many of the numeracy items correctly.

Mathematical content

The current, lowest level in PIAAC is pre-Level 1, and the description of this level of numeracy performance in PIAAC is:

> *Tasks at this level require the respondents to carry out simple processes such as counting, sorting, performing basic arithmetic operations with whole numbers or money, or recognising common spatial representations in concrete, familiar contexts where the mathematical content is explicit with little or no text or distractors. (OECD, 2013c, p. 76)*

The existing three Pre-Level 1 PIAAC numeracy items are: the counting or estimating of the number of bottles (48) shown in a photo of two layers of 4 x 6 bottles; adding up three whole numbers in a short text where the total is just over 200; and identifying the item that was packed first from four supermarket price tags, each of which include the date packed. Therefore, the skills that need to be assessed in a numeracy components assessment would mainly need to be at a lower level than those three questions and their texts.

In a recent review of options for developing a low level assessment of numeracy for adults in low- and middle-income countries (UNESCO, 2016), the authors said:

> *It is therefore necessary to distinguish between people with no formal skills (those who have relatively few mental calculation skills beyond counting simple quantities and who cannot understand the meaning of written digits) and with low formal skills (those who can engage in some mental calculations using indigenous number systems or measurement techniques but know few print-based or formal numeracy symbols and systems, even if they may be able to complete very simply written math problems). (UNESCO, 2016, p. 284)*

These issues are at the heart of the development of a numeracy components assessment.

There are existing adult numeracy standards and frameworks in different countries that have described relatively low levels of numeracy competence, and these could be used as starting points for descriptions of possible numeracy components questions and tasks. One challenge is that many such frameworks, as with PIAAC Numeracy, do not detail or describe a level below PIAAC's existing Pre-Level 1.

What is common at the lowest levels of existing adult numeracy frameworks is that they describe mathematical content across a number of content areas, as with PIAAC's four content areas of

- Quantity and number;
- Dimension and shape;
- Pattern, relationships and change; and
- Data and chance.

For example, Ireland has five areas described:

- Quantity and Number;
- Data Handling;
- Pattern and Relationship;
- Problem Solving; and
- Shape and Space.

New Zealand has three areas described:

- Make Sense of Number to Solve Problems;
- Reason Statistically; and
- Measure and Interpret Shape and Space.

The Netherlands has described an entrance level for adults around four domains: numbers, proportions, measurement & geometry, and relations, stressing the concrete nature of the content with a few data, a minimum of text, rounded numbers and problems taken directly from everyday life and the work environment.

As examples of what is described at the lower levels approximating Pre-level 1 of PIAAC or lower, Table 5 below includes some sample statements from a number of different national **adult** curriculum frameworks/standards, organised against the PIAAC content areas (excerpts from: QQI, 2016; McLean et al 2012; Tertiary Education Commission, 2008).

Quantity and number
• Recognise the relationship between numerical value and groups of objects, up to and including 10. • Recognise the language of mathematics in everyday situations using elementary language, e.g., greater than, less than, bigger than, farther than. • Solve addition and subtraction problems by counting all of the objects. • Solve addition and subtraction problems by counting on or counting back, using ones and tens. • Solve multiplication problems by skip-counting, often in conjunction with one-to-one counting and often keeping track of the repeated counts by using materials (e.g., fingers) or mental images.

• Read and write personally relevant numbers, e.g., street number. • Recognise and write money as symbols (e.g., $12.50) up to $100. • Recognise and use ordinal numbers from first to tenth.
Dimension and shape • Identify key characteristics of shapes and forms, e.g., number of sides, corners and curves. • Use the language of measurement in relation to shape and form, e.g., longer, shorter, wider, narrower. • Sort and describe objects by their shape attributes. • Describe, name and interpret relative positions in space. • Compare and order objects directly, using attributes of length, area, volume and capacity, weight, angle, temperature and time intervals in order to understand the attributes. • Read digital time (not including concept of am/pm). • Identify dates in a calendar. • Recognise common time sequences; e.g., the order of the days of the week. • Identifies differences and similarities between common 2 dimensional (2D) shapes
Pattern, relationships and change • Make a pattern; e.g., a sequence of images, symbols or sounds with two variables (different colour, same shape, etc.).
Data and chance • Identify the use of data in everyday life; e.g., the numbers of people who want tea/coffee. • Sort objects according to their attributes, organise data about the objects and represent data, using concrete objects or pictures. • Identify all possible outcomes in situations involving simple (single-stage) chance. • Compare information and data within highly familiar simple texts, lists, charts, diagrams and tables.

*Table 5. Sample statements from national **adult** curriculum frameworks/standards, organised against the PIAAC content areas.*

However, the Review Team acknowledges that there is a potential problem with using national **adult** curriculum frameworks/standards directly, because some national adult numeracy frameworks and standards have been either developed formally to align with established, hierarchical levels in child-focused curricula or are at least built on notions of children's learning. This can be illustrated in a number of ways, for example, by simplistic, bounded statements such as 'can count to 20'; by specific, school based terminology such as the 'place values of digits in whole numbers up to 100'; or where percentages are not named and included until higher levels of performance. Such statements don't acknowledge the empirical data that exists from PIAAC as it does not match the knowledge of adults nor represent the day-to-day tasks that many adults can in fact successfully undertake, but who may nonetheless be performing at pre-Level 1 Numeracy in PIAAC.

With the above caution in mind, the other potential source of content that could be used as the basis of developing and trialling a numeracy components assessment would be from research and frameworks based on the work with the early learning of mathematics by children. Again, as representative examples only, below are a few samples of beginning level numeracy and mathematics frameworks that could be used as part of the research base.

In the UK, the Early Years Foundation Stage describes these developing core skills in mathematics:

- **Numbers:** children count reliably with numbers from 1 to 20, place them in order and say which number is one more or one less than a given number. Using quantities and objects, they add and subtract two single-digit numbers and count on or back to find the answer. They solve problems, including doubling, halving and sharing.
- **Shape, space and measures**: children use everyday language to talk about size, weight, capacity, position, distance, time and money to compare quantities and objects and to solve problems. They recognise, create and describe patterns. They explore characteristics of everyday objects and

> shapes and use mathematical language to describe them. (Department for Education, 2014, pp. 11-12)

Consistent with other frameworks, the Early Numeracy Research Project (ENRP) in Australia, a three-year Prep to Year 2 research project, developed a comprehensive framework of early mathematics learning in Number and Measurement and Space (Department of Education and Training, 2013). This project developed an empirically-based assessment instrument, which describes a hierarchy of mathematics skills across these content strands.

A common theme that comes up in relation to children's learning of mathematics is what is called *number sense* and related skills such as mental calculations and multiplicative thinking, and these areas are often argued as a critical foundation for successful further learning in mathematics (see for example, McIntosh et al 1992, McIntosh 2005, Siemon et al, 2010). Here is a description and structure of the different stages of the McIntosh et al 1992 framework of number sense:

> *Number sense: A propensity for and an ability to use number and quantitative methods as a means of communicating, processing and interpreting information. It results in an expectation that numbers are useful and that mathematics has a certain regularity (makes sense).*
> *Stage 1: Knowledge and facility with NUMBERS*
> *1.1 Sense of orderliness of numbers*
> *1.2 Multiple representations for numbers*
> *1.3 Sense of relative and absolute magnitude of numbers*
> *1.4 System of benchmarks*
> *Stage 2: Knowledge and facility with OPERATIONS*
> *2.1 Understanding the effect of operations*
> *2.2 Understanding mathematical properties*
> *2.3 Understanding the relationship between operations*
> *Stage 3: Applying Knowledge and facility with numbers and operations to COMPUTATIONAL SETTINGS*
> *3.1 Understanding the relationship between problem context and the necessary computation*
> *3.2 Awareness that multiple strategies exist*
> *3.3 Inclination to utilise an efficient representation or method*
> *3.4 Inclination to review data and result for sensibility.* (McIntosh, 1992, p. 4)

However, often the school focussed frameworks and descriptions, especially in relation to number sense, focus on the skills in isolation—they are not embedded in a real-life situation and depend on the formal knowledge of the school classroom. This issue of context-based versus non-context-based tasks is addressed in the next section.

The reality is that there is much consistency across the skills and knowledge described in both the adult and child frameworks, although most of the adult frameworks are not empirically based, and the detail in the developmental stages and levels can vary. However, there are enough available descriptions for a structure and potential range of mathematical skills to be described and covered in a numeracy components assessment, once decisions are made concerning the constraints of the testing delivery and time available.

As discussed above, a critical issue that needs to be decided in the development of a numeracy components assessment is which content areas of mathematics could be covered within the time available and that would provide insights into the skills and knowledge of low level numerate adults. While much of the research with children indicates that number sense is seen as a common area of focus, the issue with adults and mathematics in their lives is that other areas such as understanding of space, position and time are seen as at least as important in a range of situations and settings (such as those associated with the workplace). Increasingly there is an expectation that adults are graphically and digitally literate.

The challenge will be to develop assessment tasks that are pitched at the appropriate, mathematically low cognitive levels, and that are accessible to the wide range of adults undertaking the numeracy components assessment.

In discussing this issue, the Review Team considers the following as a beginning list of tasks across different content areas to research and consider in the development of a numeracy components assessment:

- sorting tasks (sort according to attributes – shape, colour, size, …)
- ordering and comparing objects or collections (e.g., smallest to largest; counters; area; volume)
- identifying and recognising different representations for numbers, including orally
- counting, basic calculating and estimating with numbers from 1 to 20, 10-fold and 100-fold numbers
- identifying and recognising dates and times
- interpreting data in a simple pictograph or table of values
- interpreting relational words (above, next to, …)
- identifying a location on a diagram, or a point on a simple graph or chart.

Given that a numeracy components assessment for adults will almost be starting from scratch and has a number of challenges, considerable work would need to be undertaken to establish what such components are and how they relate to the existing PIAAC below level 1 items and their descriptions. Models for both the content and delivery would need to be researched and developed, and then used as the basis of substantial testing and trialling prior to their implementation as part of a numeracy components assessment in the next iteration of PIAAC.

Formal versus informal mathematics and embedded or not?

Another significant issue that needs to be addressed in developing a numeracy components assessment, however, is that of the embedded nature of the mathematics in real world settings and situations and the role that this plays. This is often called context-based versus non-context-based tasks (or contextualised versus decontextualised). There needs to be an acknowledgment that formal, or school-based, mathematics is not the only mathematics—individuals acquire mathematical knowledge through both formal and informal learning, and that informal learning is as valuable as formal, school-based learning. The field of ethnomathematics richly documents this issue of “street maths versus school maths” and as this components assessment will often target adults with little formal schooling but who are functioning as adults in society, this issue needs to be taken on board and addressed. For example, D’Ambrosio (1985) theorised the concept of ethnomathematics. Carraher et al’s (1985) research with street children in Brazil found they could operate in quite sophisticated ways when using mathematics to survive in a commercial sense, although they had been previously adjudged as being incapable of doing mathematics in schools. For other related ethno-mathematics research see: Nunes et al, 1993; Harris, 1991; Powell et al, 1997.

Matthijsse (2000) specifically addressed the issue of how adults cope with mathematical knowledge in practical daily situations and the gap between school mathematics and its formal algorithms and the mathematics that adults use in their daily lives. He looked at the informal methods adults used in daily life, and found they were often anchored and embedded in familiar knowledge and real-life settings and situations. Although his focus was on instructional methods to use with learners, his research, like the other ethnomathematical research, indicates that this proposed PIAAC numeracy components assessment cannot be constrained by only offering non-context-based tasks with the mathematics being like formal, school-based questions. However, a significant risk, and challenge, exists with regard to cultural and the possible national specificity of particular rule of thumb or informal methods, and how these differences could be overcome in an international assessment. Given this, a low-level components assessment could aim to find out about adults’ informal/common sense ways of doing mathematics—what mental models and processes do adults use when solving a numeracy problem? In addition, can data and information be collected about the connections (or non-connections) between the school ways of doing mathematics (and the use of algorithms) versus the way adults solve such problems in everyday life?

Different people will have very different settings and applications in which they may comfortably and more confidently use their mathematical knowledge. Finding the right problem situation or setting for an individual so that they can demonstrate their understanding of mathematics concepts will be a challenge. At this more basic level of mathematical knowledge, the familiarity of the setting and situation could be critical. A potential solution could be to use a form of adaptive delivery to allow respondents to be able to select from a range of settings and situations where the same content and level of mathematics content is embedded.

In relation to the four named PIAAC numeracy Contexts (*Everyday life*, *Work-related*, *Societal or community* and *Further learning*), it would make sense for the numeracy components to work with the more common, generic and familiar settings and Contexts which would appear to be *Everyday life* and *Society*. The three existing Pre-Level 1 numeracy items are located within those Contexts. Again, a challenge exists in how to use work-related situations, given that research shows that adults with poor formal skills are often able to function 'perfectly well' in particular jobs where they have learned rule of thumb or other methods that enable them to get by.

One other issue is that it would seem obvious that some of these numeracy components assessment items should be based around money, which appears to have the advantage of being (a) number-based, and (b) important in most adults' lives, and hence relatively familiar. Money is, however, highly country-dependent: its very familiarity is grounded in its localisation in a particular set of relationships, financial and otherwise, and these are not necessarily consistent across countries. Monetary systems across participating countries vary significantly, and although PIAAC specifies strict guidelines about changing the magnitude of monetary amounts in order to try to keep them at the same time realistic and mathematically comparable, at this lowest level of mathematical complexity this may be difficult to achieve. PISA copes with this by having all its monetary-related items set in a fictional country, Zedland, with a fictional currency of zeds and zedcents; this is again not likely to be able to be applied to PIAAC.

The Review Team suggests that the development of a numeracy components assessment in PIAAC should research the best way to incorporate aspects of everyday use and understanding of local currency and monetary transactions in order to make them comparable across countries and currencies.

One challenge with context-based items is that where the mathematics is embedded within texts and stimuli, some of the targeted cohort will not be easily able to read, interpret and hence engage with and understand the mathematics required to be used due to their potential low level of literacy skills. This is also discussed further in the next section below. The reading accessibility of items embedded in real-life texts and situations could be supported if, for example, it were possible to have the questions read out to the respondent to minimise the reading demands. Therefore, it may be the case that items need to be made available both within accessible and familiar settings, while some are, as much as possible, text-free, but with oral support available where appropriate. As argued earlier in previous sections, the use of supporting images, simulations or videos can potentially assist with minimising dependence on reading skills. The challenge will be to avoid the requirement to use formal mathematical language and terminology in the stimulus or the questions to allow respondents to use what we have described above as the informal ways of doing mathematics.

Representations, reading demands, and authenticity-related issues

As mentioned above, it will be essential to make the reading demands minimal for this assessment of numeracy components, hence the suggestion to offer oral/spoken support in some form or other, either from the administrator, or if conducted on a laptop or tablet, through audio or video support. Consideration could also be given to the administrator recording oral answers for the respondent.

Another recommendation is that the stimuli should be based on photos or videos of realistic representations of real life objects, which would help to make them accessible, more familiar and more realistic and authentic, while potentially helping to reduce reading load.

Another option, as with the existing numeracy item where an actual hard copy photo and a ruler are provided, is to use real items or objects for some test items. These could be used for tasks such as comparing, sorting, or classifying. This would make the numeracy components assessment more accessible, practical and hands-on.

Additionally, or alternatively, technology could be used so that similar actions could be done on screen, such as using drag-and-drop items on a laptop or tablet using touch screen capabilities. For example, respondents could be asked to order objects representing quantities by dragging and dropping rather than writing down an ordered list. Technology allows the use of photos and realistic representations of real life objects, which can help make them accessible and more familiar. The existing Pre-Level 1 item on counting the number of bottles is a good example of how this can be done using a photo and little text; in fact, even without being able to read the question text it is highly likely that respondents would be able to assume what the question was asking. If audio support were also made available, then this would make such questions even more accessible. This is a good model to follow across the numeracy components assessment.

Responses/actions

In section 3.1.2 Facet 2: Responses (PIAAC Numeracy Expert Group, 2009, pp. 24-26), the current framework describes and elaborates how in different real-life situations, adults may have to react to a numeracy problem with different types of responses. The current framework groups those under three broad headings: *Identify, locate, or access*; *Act upon or use*; and *Interpret, evaluate/analyse, communicate*. It is highly likely that the same range of responses would be expected of respondents being assessed on a components of numeracy assessment, with the possibility that the focus might be on the first two categories.

Item formats

Examples of the most appropriate item formats and delivery that could be expected of respondents being assessed on numeracy components items have been discussed above. This will be dependent on the method of delivery, but some of the recommended options, which would support a numeracy components assessment, would include:

- use of photos and realistic representations of real-life objects
- use of real items or objects for tasks such as comparing, sorting or classifying
- drag and drop items or click on items on laptop or a tablet using touch screen capabilities
- video or audio supported items
- having the administrator record oral answers for the respondent.

Timing/fluency

Given that in the delivery of the reading components each component was timed in order to be able to get an estimate of fluency and automaticity, this should also be made available for the numeracy components. This would support and enable estimates and comparisons to be made in relation to the fluency and automaticity of estimation skills and mental arithmetic calculations.

SECTION VI. IMPLICATIONS AND RECOMMENDATIONS FOR PIAAC

In this section, the Review Team synthesises the previous recommendations regarding the definition and descriptions of the construct of numeracy and the priorities for development of the numeracy assessment framework and content for the second cycle of PIAAC.

The PIAAC numeracy assessment framework has defined the construct of numeracy, has guided the development of the items and the interpretation of results. It has been robust and has served valuable purposes and stood the test of time since its initial development as the numeracy assessment framework of the Adult Literacy and Life Skills Survey (ALL). However, it is opportune to reflect on its status and to review how the existing framework and assessment can be enhanced and refined as we move further into the 21st century.

The OECD should address a number of overlapping areas stemming from the Review Team's work in the next iteration of the PIAAC numeracy assessment. These areas include:

- framework definition and description
- assessment content and delivery
- numeracy components assessment
- other issues.

Framework definition and description

There are a number of specific issues that should be addressed in relation to the PIAAC numeracy framework definition and descriptions. These have arisen from the following aspects:

- Theoretical developments
- 21st century representations and tools
- Lessons from PISA and other numeracy models.

Theoretical developments

The Review Team recommends consideration of four related issues arising from research into theoretical developments of adult numeracy in updating and refining the existing PIAAC framework definition and description:

- disposition to use mathematics
- the ability to see mathematics in a numeracy situation
- critical reflection
- degree of accuracy.

These four issues should be addressed explicitly in the background information and related discussions in the framework and, where appropriate, by reworking the definition of numeracy itself and/or in expanding on the framework's elaboration and description of numerate behaviour through its set of facets.

21st century digital and technological representations and tools

Being numerate in the 21st century means being able to cope with the aspects of the world as we encounter it, which includes the digital and technological aspects of information and society—society generally already has all kinds of techno-mathematical aspects. Research shows that much 21st century workplace mathematics and numeracy practice interacts with technology. The review found that 21st century digital technologies provide tools and processes that mediate thinking as well as action and are not just devices that can be used to complete manual, hands-on tasks more efficiently. These tools and processes often change the numeracy task itself and so transform practices within adults' lives and within the workplace. The use and application of a range of techno-mathematical literacies underpins much of this. This aspect of 21st century representations and tools was missing from much of the existing PIAAC numeracy framework discussions, and not reflected in the definition and elaborations. This needs to be addressed across many aspects of the numeracy framework and construct.

However, it is important to acknowledge that in addressing this issue PIAAC is a survey of adult competencies across **all** aspects of life, not just about workplace and employment, and not just about engaging with numeracy and mathematics actions within technologically rich environments. It is recommended that a balance be kept between numeracy activities in digital and technological environments and those embedded in other, non-digital media; between numeracy demands and situations met as an individual and those encountered as part of society; and between work and employment settings and home and family activities.

Lessons from PISA and other numeracy models

It is recommended that both the PISA problem solving model (see Figure 4, p. 34) and its set of processes and the model for numeracy in the twenty-first century (see Figure 5, p. 35) be utilised in refining PIAAC's definition of numeracy, particularly in how PIAAC elaborates and describes numerate behaviour through its facets. These considerations would also help address the above issues resulting from recent theoretical developments related to the disposition to use mathematics, the ability to see mathematics in a numeracy situation, critical reflection and degree of accuracy.

Another benefit from looking at the PISA model, and other models, is to refine and update the different parameters of the PIAAC assessment construct, including for its descriptions of:

- Content
- Context
- Responses and actions.

Content

The existing descriptions of the four Content facets in the PIAAC numeracy framework (quantity and number; dimension and shape; pattern, relationships, change; and data and chance) should be reviewed in light of the research about 21st digital information, representations, devices and applications including of techno-mathematical literacies.

Although consistent in their descriptions, the labels for the four Content areas across PIAAC and PISA are different. Therefore, the recommendation is that it would be appropriate to bring consistency and coherence to the labelling of what are essentially the same Content areas in PISA and PIAAC, and for PIAAC to consider using the same four Content labels.

Context

There is little discussion of the connections between 21st digital and technological representations and tools across the different, named Contexts within the current framework document. This needs to be reflected in the revised PIAAC numeracy framework.

There are a number of different labels used for naming the Contexts within different existing PIAAC documents, which should be addressed with the labels being made consistent.

The four existing labels and descriptions for the Contexts in numeracy need refining and updating, taking the two above issues into account. The following are recommendations regarding three of the existing four Contexts:

- The PISA Context label *Personal* is preferable to the PIAAC label of *Everyday*.
- There is a need to refresh the description of the *Work-related* Context in PIAAC. This includes making a stronger connection to techno-mathematical literacies in workplaces, and updating the existing use of "quantitative situations"—work-related environments and settings cover a broad range of mathematical skills and knowledge, not just quantitative. It is also recommended that the phrase in the PISA description "Occupational contexts may relate to any level of the workforce, from unskilled work to the highest levels of professional work" is an aspect that could be better reflected in the PIAAC description. Finally, the PISA label *Occupational* is preferable to the PIAAC label of *Work-related*.

- The fourth Context, *Further learning* in PIAAC and *Scientific* in PISA, is the Context that needs significant thought and consideration. The Review Team suggests that the PISA classification and description of *Scientific* can be used to reflect on the *Further learning* Context of PIAAC. Another possibility is the name *Technical*, which would possibly work as an alternative. Therefore, the review of the PIAAC Framework needs to address what this fourth Context of *Further learning* covers and includes, and the question of what the most appropriate name for it is.

Responses and actions

Unlike a number of the other facets of numerate behaviour in PIAAC and their related descriptions, the facet of responses or actions has the least consistency between PISA and PIAAC. This facet is also requiring revision and enhancement in relation to other identified issues. It is recommended that this aspect of the PIAAC framework needs major review, and could potentially learn from the processes described in PISA.

Assessment content and delivery

The Review Team recommends improvements and enhancements to the assessment content and delivery mechanisms be made for the next iteration of PIAAC .

It is important to make a distinction and achieve a balance between the drive for change that stems from development of technologies that can be used in assessments (technology-driven) and the drive to design an assessment that is closely related to the concepts conceived from the construct of numeracy (concept-driven). The Review Team recognises that in the last PIAAC assessment, there was a gap between the sophistication of the concept of numeracy used and the design of the assessment. As acknowledged earlier, the assessments that exist in the current item pool are simple and quite one-dimensional.

More sophisticated assessments utilising some of the technological possibilities outlined earlier, especially in Sections II and III, are not necessarily aimed at more complex or higher order skills, but focus more on the multifaceted and multimodal nature of numeracy problem situations encountered in real life. To assess a sophisticated concept of numeracy there is the need for multimodal options to better represent reality, in which the respondents can show their competence (or not).

There are three related aspects recommended for consideration:

- Representations, reading demands, and authenticity-related issues
- Item formats and responses
- A dimension for reviewing assessment possibilities.

Representations, reading demands, and authenticity-related issues

As developments in the 21st century impact on how mathematical and numerical information is represented, the facet, *Representations of mathematical information*, in the PIAAC framework should be updated to reflect these changes. The Review Team recommends that these could be reorganised under new categories such as physical, representational, and digital.

For an assessment of numeracy in PIAAC, the assessment stimuli and items should be, as far as possible, authentic and representative of what adults might meet in their lives. As described earlier, digital and technological aspects need to be included in the next iteration of PIAAC, alongside the capabilities of 21st century assessments to facilitate more enhanced representations of stimuli.

A strong recommendation is that the next iteration of PIAAC harness the potential of technology to support a more effective and representative 21st century assessment, for example, through greater use of non-text based media, such as the use of videos or animations, in assessment items. Use of different technology, media and associated supports has the potential to transcend not only some of the challenges with literacy and/or language that may impede some adults' ability to demonstrate their numeracy capability, but also to make the assessment more relevant to the 21st Century.

As mentioned earlier, the existing PIAAC numeracy item pool is based mainly around static images and the items are more like paper-based assessments transferred onto a computer, which does not seem to reflect how many numeracy tasks and actions are now situated and practised in the 21st century. The Review Team recommends that the next delivery of the PIAAC numeracy assessment builds on assessment developments as described in the earlier sections.

It is also recommended, however, that a balance be kept between numeracy and mathematics tasks and actions embedded in 21st century digital and technological environments with those embedded in modes that are more traditional.

Item formats and responses

The Review Team recommends that the next delivery of the PIAAC numeracy assessment should build on assessment developments and lessons such as those learnt from the successful delivery of CBAM in PISA 2012 and introduce new stimulus and item formats that better reflect 21st century related numeracy tasks and assessment responses.

A key challenge in PIAAC is how to assess some of these 21st century digital and technological aspects of numeracy use in a generic way, so that it is accessible to most of the population across a spread of countries.

Another challenge facing PIAAC is that open response format items have not been used, as in the last cycle of PIAAC where the assessment is adaptive and all items needed to be autoscored. An advantage of considering the use of open response type questions that are human-marked, is that they are more realistic, taking more of an adult approach and closer to problem solving situations in real life. Although human marking is more expensive and time consuming, it could be worth reconsidering in order to do justice to the multifaceted concepts being assessed in PIAAC.

A dimension for reviewing assessment possibilities

In order to facilitate the development of newer assessment content and delivery mechanisms, the Review Team developed a dimension of assessment possibilities (see Figure 3, p. 31) that could form a starting point for monitoring and balancing the range of PIAAC numeracy item formats and types. Having a spread of stimuli and items developed and selected from across this spectrum would enable PIAAC numeracy to be better representative of the framework and construct, add to the issue of authenticity, and hence better assess adults' capabilities and competencies around numeracy practices in the 21st Century.

It is recommended that this dimension of possibilities, or a similar spectrum or description of delivery and assessment options and modalities, be used as a set of parameters to classify items and to create a spread across these categories to show the multimodality of numeracy problem situations and the multifaceted nature of numeracy practices and the PIAAC numeracy construct.

A proportion of the existing PIAAC numeracy items will continue to be used as link items, and they tend to be at the left hand end of the dimension of assessment possibilities as decribed. It is recommended that any new test items developed to complement the existing pool should take on board potential enhancements and innovations due to 21st century requirements and assessment capabilities. This would then provide the potential for a balance across the spectrum of assessment possibilities.

Challenges and issues

As mentioned in the discussions about 21st Century Digital Implications, the results from the PSTRE assessment in PIAAC need to be considered in relation to the implementation of the PIAAC numeracy assessment, and that presumably there will remain in place similar filtering processes to direct respondents to the most appropriate delivery mode. Due to the ICT-related capabilities of adults across the PIAAC age-range of 16-65 years, a variety of item types and representations from across the proposed dimension of possibilities will need to be balanced against the risks of using too much sophisticated technology. There needs to be a balance between an individuals' ability to utilise digital tools to answer questions and their ability to solve the presented numeracy problem.

On a pragmatic level too, some of the innovations and developments arising from technologically-driven advancements in educational measurement and assessment in the 21st century will need to be carefully reviewed and considered as to the feasibility of their use and practical implementation for PIAAC in 2021-22. The issues the Review team identified included:

- the costs—some technologies, media and tools would potentially be expensive to use and implement
- the time available for testing
- the translation issues of producing such materials in potentially 30 different languages
- the suitability and performance of the use of any innovations.

It would make sense that as part of the revision of the numeracy framework and the associated development of new assessment items a staged, incremental process documenting what is feasible in the immediate future for the next delivery of PIAAC and what could be considered as part of longer term, horizon planning be developed. A timeframe for implementation and possible delivery in 2021-22 would need to be included so that sufficient time was allowed for implementation.

No matter what representations or item types are used, the development of the numeracy assessment items must cover principles of good assessment, and the items developed should be as accessible, valid, reliable, fair, and authentic as possible. Cultural fairness and applicability is a significant issue in international assessments and brings its own set of challenges. The standard quality assurance practices should be implemented and used: a set of items developed against the construct and its parameters, which are trialled and psychometrically checked for validity, reliability, and fairness, including across countries. From that set of items a representative set of successfully performing final items will be used for the implementation of PIAAC.

Numeracy Components

The Review Team strongly recommends the development of a PIAAC numeracy components assessment. The purpose of the numeracy components assessment would have parallel aims to the existing reading components: to develop a set of "fine-grained tasks" so that "at least some of these adults would demonstrate some level of **numeracy** knowledge and skills" and would provide insights into the skills and knowledge of adults with low levels of numeracy.

The percentages of adults performing at the lowest level in numeracy in PIAAC are significantly higher than in literacy, therefore there is a very strong argument for developing an equivalent to the reading components assessment in the PIAAC numeracy assessment.

Given that such a numeracy component assessment has not been researched, theorised or examined quantitatively for the adult world, considerable work would need to be undertaken to establish what such components are and how they relate to the existing PIAAC below level 1 items and their descriptions.

The Review Team considers the following as an initial list of tasks across different content areas to research and consider in the development of a numeracy components assessment:

- sorting tasks (sort according to attributes – shape, colour, size, …)
- ordering and comparing objects or collections (e.g., smallest to largest; counters; area; volume)
- identifying and recognising different representations for numbers, including orally
- counting, basic calculating and estimating with numbers from 1 to 20, 10-fold and 100-fold numbers
- identifying and recognising dates and times
- interpreting data in a simple pictograph or table of values
- interpreting relational words (above, next to, …)

- identifying a location on a diagram, or a point on a simple graph or chart.

Issues and challenges

There are a number of issues and challenges surrounding the development of a numeracy components assessment as part of PIAAC. These are detailed in Section V, but below is a summary of them and where known, how they might be addressed.

Implementation practicalities

One important issue to address in the development of a numeracy components assessment is that of the constraints imposed by the practicalities of delivering such an assessment. Factors to consider include: the time available; the number and range of items in terms of content areas, difficulty levels, and amount of text; the delivery and item types (oral instructions and support by administrator; pen-and-paper versus online delivery; use of real, physical objects; use of video or animations; interactive or not) and more.

The role of contextualised tasks and the intersection with literacy skills

Mathematical knowledge is acquired through both formal and informal learning, and the latter is as valuable as the former. A numeracy components assessment has to be able to assess adults' skills and abilities from a non-school world perspective since those individuals who would be assessed at this basic level would likely have limited schooling.

Different people will have very different settings and situations in which they may comfortably and more confidently use their mathematical knowledge. Finding the right situation to embed a problem for an individual so that they can demonstrate their understanding of mathematics concepts will be a challenge. At this more basic level of mathematical knowledge, the familiarity of the situation could be critical. A potential solution could be to use a form of adaptive delivery to allow respondents to be able to select from a range of settings and situations where the same content and level of mathematics content is embedded.

Given the likelihood of an interaction between low numeracy and low literacy skill levels, delivery of a numeracy components assessment needs to take special account of the reading demands of the assessment, and find means to reduce this influence and connection. It is highly recommended that accessibility of stimuli and questions should be supported, for example, by having the questions read out to the respondent to minimise the reading demands, or through providing video or audio support.

Another recommendation is that the stimuli are based on physical objects, photos, videos or animations of realistic representations of real life objects, which would help to make them accessible, more familiar, more realistic and authentic.

Consideration could also be given to the administrator recording oral answers for the respondent.

A further challenge will be to avoid the requirement to use formal mathematical language and terminology in the stimulus or the questions to allow respondents to use their informal ways of doing mathematics.

Research time required

Given that a numeracy components assessment for adults will almost be starting from scratch and has a number of challenges, considerable work would need to be undertaken to establish what such components are and how they relate to the existing PIAAC below level 1 items and their descriptions. The Review Team recommends that a team of adult numeracy researchers and assessment experts be created to research and develop models for both the content and delivery of such an assessment and oversee substantial testing and trialling prior to their possible implementation as part of a numeracy components assessment in the next iteration of PIAAC.

Timing/fluency

In order to get an estimate of fluency and automaticity it is recommended that each section of the numeracy components assessment be timed. This would support and enable estimates and comparisons to be made in relation to the fluency and automaticity of, for example, any estimation skills and mental arithmetic calculations.

Other issues

Linking PISA and PIAAC scales

As recommended above, there is much to gain from making significant and more explicit connections between the PIAAC numeracy framework and the PISA mathematical literacy framework and assessment. To facilitate this it is recommended that a method be implemented so that the two scales can be aligned psychometrically and anchored against each other. This would make analysis and comparison across the two assessments stronger and more useful for research purposes, both within countries and internationally. This could be done either by inserting suitable PISA items into a delivery of PIAAC as a one-off linking study, or alternatively, without affecting the delivery of PIAAC, the relationship between the scales could be established through an appropriately designed separate linking study.

Drivers/indicators of mathematical proficiency

One of the important features of both the PISA and PIAAC frameworks is the way that each describes a schema of aspects of test items that drive item difficulty, and which indicate the mathematical proficiency of tested individuals and populations. It is recommended that further research could be usefully undertaken to establish the degree to which the PISA capabilities could enrich the PIAAC complexity scheme, and vice versa.

REFERENCES

Askew, M., (2015) Numeracy for the 21st century: a commentary. *ZDM Mathematics Education* (2015) 47:707–712

Atkinson, R. K. (2005). Multimedia learning of mathematics. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 31-48). Cambridge, U.K.; New York: Cambridge University Press.

Australian Association of Mathematics Teachers (AAMT) & Australian Industry Group (AiGroup), *Tackling the School–Industry Mathematics Divide*, Commonwealth of Australia 2014

Bakker, A. (2014). Characterising and developing vocational mathematical knowledge. *Educational Studies in Mathematics, 86*(2), 151–156. doi:10.1007/s10649-014-9560-4

Bennett, R. E. (1998). *Reinventing assessment: Speculations on the future of large-scale educational testing*. Princeton, NJ: Policy Information Center, Educational Testing Service. Retrieved from https://www.ets.org/research/policy_research_reports/publications/report/1998/cgln Accessed 31st October 2016.

Bennett, R. E. (2010). Technology for large-scale assessment. In P. Peterson, E. Baker, & B. McGaw (Eds.), *International encyclopedia of education* (3rd ed., Vol. 8, pp. 48–55). Oxford, England: Elsevier.

Bennett, R. E. (2015). The Changing Nature of Educational Assessment, in *Review of Research in Education, March 2015*, Vol. 39, pp. 370–407

Bower, M., Howe, C., McCredie, N., Robinson, A., & Grover, D. (2014). Augmented reality in education – cases, places and potentials. *Educational Media International, 51*(1), 1-15. doi:10.1080/09523987.2014.889400

Brown, M., Küchemann, D., & Hodgen, J. (2010). The struggle to achieve multiplicative reasoning 11-14. In M. Joubert & P. Andrews (Eds.), *Proceedings of the Seventh British Congress of Mathematics Education (BCME7)* (Vol. 30, pp. 49-56). Manchester: BSRLM.

Buckley, Sarah (2013). Deconstructing maths anxiety: Helping students to develop a positive attitude towards learning maths. (*ACER Occasional Essays*) Melbourne Vic: ACER. Available from: http://research.acer.edu.au/cgi/viewcontent.cgi?article=1016&context=learning_processes Accessed 3rd January 2017.

Carpentieri, J. D., Litster, J., & Frumkin, L. (2010). *Adult numeracy: A review of research*. London, England, National Research and Development Centre for Adult Literacy and Numeracy.

Carraher, T. N., Carraher, D. W., & Schliemann, A. D. (1985). Mathematics in the streets and in schools. *British Journal of Developmental Psychology, 3*, 21-29.

Chisman, F. (2011). *Facing the challenge of numeracy in adult education*. New York, Council for Advancement of Adult Literacy

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.* Washington D.C.: American Institutes for Research.

D'Ambrosio, U. (1985). Ethnomathematics. *For the Learning of Mathematics, 5*(1), 44-48.

Department for Education (2014). *Statutory framework for the early years foundation stage: Setting the standards for learning, development and care for children from birth to five*, Department for Education, London

Department of Education and Training (2013). *Framework of Mathematical Learning*. Department of Education and Training, East Melbourne. Available from: http://www.education.vic.gov.au/school/teachers/teachingresources/discipline/maths/Pages/enrpframe.aspx Accessed 11th October 2016.

FitzSimons, G. (2014). Commentary on vocational mathematics education: where mathematics education confronts the realities of people's work. *Educational Studies in Mathematics, 86*(2), 291–305.

FitzSimons, G. E., & Coben, D. (2009). Adult numeracy for work and life: Curriculum and teaching implications of recent research. In R. Maclean & D. Wilson (Eds.), *UNESCO-UNEVOC International Handbook of Technical and Vocational Education and Training. Bridging Academic and Vocational Education* (Vol. 6, pp. 2731-2745). Dordrecht: Springer.

Gal, I. 2016, Assessment of adult numeracy skills: Background paper commissioned for the UNESCO Global Education Monitoring Report 2016. Report number: ED/GEMR/MRT/2016/P1/8 REV, Affiliation: UNESCO, Global Education Monitoring Report

Gal, I. and Tout, D (2014). *Comparison of PIAAC and PISA Frameworks for Numeracy and Mathematical Literacy*. OECD Education Working Papers, No. 102. Paris: OECD Publishing.

Geiger, V., Goos, M. & Forgasz, H. (2015) A rich interpretation of numeracy for the 21st century: a survey of the state of the field. *ZDM Mathematics Education* (2015) 47:531–548.

Geiger, V., Goos, M., & Dole, S. (2014). Students' perspectives on their numeracy development across the learning areas. In Y. Li & G. Lappan (Eds.), *Mathematics curriculum in school education* (pp. 473–492). New York: Springer.

Geisinger Kurt F. (2016) 21st Century Skills: What Are They and How Do We Assess Them?, *Applied Measurement in Education, 29:4*, 245-249, DOI: 10.1080/08957347.2016.1209207

Ginsburg, L., Manly, M., & Schmitt, M.J. (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.

Goos, M., Geiger, V., & Dole, S. (2014). Transforming professional practice in numeracy teaching. In Y. Li, E. Silver, & S. Li (Eds.), *Transforming mathematics instruction: multiple approaches and practices* (pp. 81–102). New York: Springer.

Griffiths, Graham and Stone, Rachel. *Teaching Adult Numeracy: Principles and Practice*. Maidenhead: Open University Press, 2013.

Grotlüschen, A. et al. (2016), "Adults with Low Proficiency in Literacy or Numeracy", *OECD Education Working Papers, No. 131*, OECD Publishing, Paris.

Harris, M. 1991, *Schools, Mathematics and Work*, Falmer Press, London.

Hoogland, K., Pepin, B., Bakker, A., de Koning, J., & Gravemeijer, K. (2016). Representing contextual mathematical problems in descriptive or depictive form: Design of an instrument and validation of its uses. *Studies in Educational Evaluation, 50*, 22-32. doi:10.1016/j.stueduc.2016.06.005

Hoyles, C., Noss, R., Kent, P., & Bakker, A. (2010). *Improving mathematics at work: The need for technomathematical literacies*. London: Routledge.

Hoyles, C., Wolf, A., Molyneux-Hodgson, S., & Kent, P. (2002). *Mathematical skills in the workplace*. London: The Science Technology and Mathematics Council.

Hung, H.-T., Chiu, Y.-C. J. and Yeh, H.-C. (2013) Multimodal assessment of and for learning: A theory-driven design rubric. In *British Journal of Educational Technology 44(3)*, May 2013, 400–409 doi:10.1111/j.1467-8535.2012.01337.x

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

Jorgensen, R. (2011). Young workers and their dispositions towards mathematics: Tensions of a mathematical habitus in the retail industry. *Educational Studies in Mathematics, 76*, 87-100

Kent, P., Bakker, A., Hoyles, C. & Noss, R. (2011). Measurement in the workplace: the case of process improvement in manufacturing industry. *ZDM Mathematics Education* (2011) 43:747–758.

Kindler, J., Kenrick, R., Marr, B., Tout, D., Wignall, L. (1996). *Certificates in General Education for Adults*. Victoria, Melbourne: Adult, Community and Further Education Board.

Lackey, S., & Shumaker, R. (Eds.). (2016). Virtual, Augmented and Mixed Reality: 8th International Conference, VAMR 2016, Held as Part of HCI International 2016, Toronto, Canada, July 17-22, 2016. Proceedings. Cham: Springer International Publishing.

Lowrie, T., & Diezmann, C. M. (2009). National numeracy tests: a graphic tells a thousand words. *Australian Journal of Education*, 53(2), 141–158.

Ma, X. (1999). A Meta-Analysis of the Relationship between Anxiety toward Mathematics and Achievement in Mathematics. *Journal for Research in Mathematics Education, 30*(5), 520-540. doi:10.2307/749772

Maguire, T. and A. Smith (2016) Maths Eyes- A Concept with Potential. Invited paper presented at TSG 6, ICME 13, (13th International Congress on Mathematical Education, Hamburg, 24-31 July 2016).

Maguire, T. and J. O'Donoghue, J. (2002). A grounded approach to practitioner training in Ireland: Some findings from a national survey of practitioners in adult basic education. In *Numeracy for empowerment and democracy? Proceedings of the 8th International Conference of Adult Learning Mathematics*, ed. L. O. Johansen and T. Wedege, 120-132. Roskilde, Denmark: Roskilde University, Centre for Research in Learning Mathematics. Stevenage, UK: Avanti Books.

Marr, B., & Hagston, J. (2007). *Thinking Beyond Numbers: Learning numeracy for the future workplace*. Adelaide: NCVER http://www.ncver.edu.au/publications/1795.html Accessed 31st October 2016.

Matthijsse W. (2000) "Adult Numeracy at the Elementary Level: Addition and Subtraction Up to 100". In *Adult Numeracy Development: Theory, Research, Practice. Series on Literacy: Research, Policy, and Practice*. I. Gal (ed.) Hampton Press, Inc., 23 Broadway, Cresskill, NJ 07626

Mayer, R. E. (2005). *The Cambridge handbook of multimedia learning*. New York: Cambridge University Press.

Mayer, R. E. (2009). *Multimedia learning (2nd edition)*. New York, NY: Cambridge University Press.

McIntosh, A. (2005). *Mental computation: a strategies approach (2nd Edition)*. Hobart: Department of Education, Tasmania

McIntosh, A., Reys, B. J., & Reys, R. E. (1992). A proposed framework for examining basic number sense. *For the Learning of Mathematics, 12*(3), 2-8.

McLean, P., Perkins, K., Tout, D., Wyse, L. & Brewer, K. (2012). *Australian core skills framework: 5 core skills, 5 levels of performance, 3 domains of communication*. Canberra, Australia: Australian Government.

Milgram, P., & Colquhoun, H. (1999). A taxonomy of real and virtual world display integration. In Y. Ohta & H. Tamura (Eds.), *Mixed Reality* (pp. 1-26). Berlin, Germany: Springer.

Murray, S., Y. Clermont and M. Binkley (eds) (2005). *Measuring Adult Literacy and Life Skills: New frameworks for Assessment*, Statistics Canada, Ottawa, Catalogue No. 89-552-MIE, No. 13.

NIACE (National Institute of Adult Continuing Education) (2011), *Numeracy Counts, NIACE Committee of Inquiry on Adult Numeracy Learning*, NIACE, Leicester

Nuñes, T., Schliemann, A. D., & Carraher, D. W. (1993). *Street Mathematics and School Mathematics*. Cambridge: Cambridge University Press.

Nys, J., Ventura, P., Fernandes, T., Querido, L., Leybaert, J., & Content, A. (2013). Does math education modify the approximate number system? A comparison of schooled and unschooled adults. *Trends in Neuroscience and Education*. 2, 13-22.

OECD (2005). *The Definition and Selection of Key Competencies. Executive Summary*. Retrieved from Paris: https://www.oecd.org/pisa/35070367.pdf Accessed 31st October 2016.

OECD (2012). *Literacy, Numeracy and Problem Solving in Technology-Rich Environments: Framework for the OECD Survey of Adult Skills*, OECD Publishing, Paris. http://dx.doi.org/10.1787/9789264128859-en

OECD (2013a). Mathematics Framework. In *PISA 2012 Assessment and Analytical Framework: Mathematics, Reading, Science, Problem Solving and Financial Literacy*, OECD Publishing, Paris.

OECD (2013b). *PISA 2012 Assessment and Analytical Framework: Mathematics, Reading, Science, Problem Solving and Financial Literacy*, OECD Publishing. http://dx.doi.org/10.1787/9789264190511-en

OECD (2013c). *OECD Skills Outlook 2013: First Results from the Survey of Adult Skills*, OECD Publishing. http://dx.doi.org/10.1787/9789264204256-en

OECD (2013d). *The Survey of Adult Skills: Reader's Companion*, OECD Publishing. http://dx.doi.org/10.1787/9789264204027-en

OECD (2016). *Skills Matter: Further Results from the Survey of Adult Skills*, OECD Skills Studies, OECD Publishing, Paris.

OECD (2016). PISA 2015 Mathematics Framework. In *PISA 2015 Assessment and Analytical Framework: Science, Reading, Mathematics and Financial Literacy*. Paris: OECD.

Palm, T. (2006). Word problems as simulations of real-world situations: a proposed framework. *For the Learning of Mathematics, 26*(1), 42–47.

Palm, T. (2008a). Impact of authenticity on sense making in word problem solving. *Educational Studies in Mathematics 67*(1): 37-58.

Palm, T. (2008b). Performance assessment and authentic assessment: a conceptual analysis of the literature. *Practical Assessment Research & Evaluation*, *13*(4).

Palm, T. (2009). Theory of authentic task situations. Words and worlds - modelling verbal descriptions of situations. L. Verschaffel, B. Greer, W. V. Dooren and S. Mukhopadhyay. Rotterdam, the Netherlands, Sense: 3-20.

Parshall, C. G., Spray, J. A., Kalohn, J. C., & Davey, T. (2002). *Practical considerations in computer-based testing*. New York: Springer-Verlag.

PIAAC (2014). *Technical Report of the Survey of Adult Skills (PIAAC)*, OECD Publishing. http://www.oecd.org/site/piaac/_Technical%20Report_17OCT13.pdf

PIAAC Numeracy Expert Group (2009). *PIAAC Numeracy: A Conceptual Framework, OECD Education Working Papers*, No. 35, OECD Publishing. http://dx.doi.org/10.1787/220337421165

PISA Mathematics Expert Group (2009). *PISA CBAM Item Types*. Australian Council for Educational research, Melbourne. (Unpublished manuscript)

Powell, Arthur B. and Frankenstein, Marilyn (Ed's) (1997). *Ethnomathematics: challenging Eurocentrism in mathematics education*, State University of New York Press, Albany, New York.

Quality Qualifications Ireland QQI (2016). *General Learning P1GL0*. http://qsearch.qqi.ie/WebPart/AwardDetails?awardCode=P1GL0 Accessed August 2016.

Ricker, K. L. (2006). Setting cut-scores: A critical review of the Angoff and Modified Angoff Methods. *Alberta Journal of Educational Research, 52*(1), 53-64.

Roth, W.-m. (2012). The Work of Seeing Mathematically. Alternative Forms of Knowing. In *Mathematics: Celebrations of Diversity of Mathematical Practices*. S. Mukhopadhyay and W.-M. Roth. Rotterdam, SensePublishers**:** 227-245.

Rychen, D. S. (2004). An overarching conceptual framework for assessing key competences in an international context. Lessons from an interdisciplinary and policy-oriented approach. In P. Descy & M. Tessaring (Eds.), *The Foundations of Evaluation and Impact Research. Third report on vocational training research in Europe: Background report*. Luxembourg: Office for Official Publications of the European Communities. http://www.cedefop.europa.eu/files/BgR1_Rychen.pdf Accessed 31st October 2016.

Sabatini, J. P. and K. M. Bruce (2009), "PIAAC Reading Component: A Conceptual Framework", *OECD Education Working Papers*, No. 33, OECD Publishing. doi: 10.1787/220367414132

Schnotz, W. (2002). Commentary: Towards an integrated view of learning from text and visual displays. *Educational Psychology Review, 14*(1), 101-120. doi:10.1023/A:1013136727916

Schnotz, W. (2005). An integrated model of text and picture comprehension. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 31-48). Cambridge, UK: Cambridge University Press.

Schnotz, W., & Bannert, M. (2003). Construction and interference in learning from multiple representation. *Learning and Instruction, 13*(2), 141-156. doi:10.1016/s0959-4752(02)00017-8

Schnotz, W., & Kürschner, C. (2008). External and internal representations in the acquisition and use of knowledge: Visualization effects on mental model construction. *Instructional Science, 36*(3), 175-190. doi:10.1007/s11251-007-9029-2

Schnotz, W., Baadte, C., Müller, A., & Rasch, R. (2010). Creative thinking and problem solving with depictive and descriptive representations. In L. Verschaffel, E. De Corte, T. De Jong, & J. Elen (Eds.), *Use of representations in reasoning and problem solving - analysis and improvement* (pp. 11-35). London, UK: Routledge.

Schüler, A., et al. (2015). "Processing multimedia material: Does integration of text and pictures result in a single or two interconnected mental representations?" Learning and Instruction 35(0): 62-72.

Shute, Valerie J., Leighton, Jacqueline P., Jang Eunice E. & Chu, Man-Wai (2016) Advances in the Science of Assessment, *Educational Assessment, 21:1*, 34-59

Siemon, Dianne, Breed Margarita & Virgona Jo, (2010) *From Additive to Multiplicative Thinking – The Big Challenge of the Middle Years*. RMIT University, School of Education, Melbourne.

Smith, A. (2004). *Making Mathematics Count: The Report of Professor Adrian Smith's Inquiry into Post-14 Mathematics Education* (2/04 937764). London: The Stationery Office http://www.mathsinquiry.org.uk/ Accessed 31st October 2016.

Sommerauer, P., & Müller, O. (2014). Augmented reality in informal learning environments: A field experiment in a mathematics exhibition. *Computers & Education, 79*, 59-68. doi:http://dx.doi.org/10.1016/j.compedu.2014.07.013

Stacey, K. (2015). The Real World and the Mathematical World . In K. Stacey & R. Turner (Eds.), *Assessing mathematical literacy: the PISA experience* (pp. 57–84). New York: Springer.

Steen, L. A. (2004). Data, Shapes, Symbols: Achieving balance in School Mathematics. In B. Madison & L. A. Steen (Eds.), *Quantitative Literacy: Why numeracy matters for schools and colleges* (pp. 53-74). Washington, DC: Mathematical Association of America.

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

Sweller, J. (2005). Implications of cognitive load theory for multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 31-48). Cambridge, U.K.; New York: Cambridge University Press.

Sweller, J. (2010). Element interactivity and intrinsic, extraneous, and germane cognitive load. *Educational Psychology Review, 22*(2), 123-138. doi:10.1007/s10648-010-9128-5

Tertiary Education Commission (2008) *Learning Progressions for Adult Numeracy*, Wellington, New Zealand

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

Tout, D. and Gal, I (2015). Perspectives on numeracy: reflections from international assessments. *ZDM Mathematics Education* (2015) 47: 691–706.

Tout, D., & Spithill, J. (2015). The challenges and complexities of writing items to test mathematical literacy. In K. Stacey & R. Turner (Eds.), *Assessing mathematical literacy: the PISA experience* (pp. 145–171). New York: Springer.

Tsuchida, S., & Matsuura, S. (2016). A Role of Augmented Reality in Educational Contents: Intermediating Between Reality and Virtual Reality. In S. Lackey & R. Shumaker (Eds.), *Virtual, Augmented and Mixed Reality: 8th International Conference, VAMR 2016, Held as Part of HCI International 2016, Toronto, Canada, July 17-22, 2016. Proceedings* (pp. 736-745). Cham: Springer International Publishing.

Turner, R., Blum, W., & Niss, M. (2015). Using competencies to explain mathematical item demand: A work in progress. In K. Stacey & R. Turner (Eds.), *Assessing Mathematical Literacy: The PISA Experience* (pp. 85-115). New York: Springer.

U.S. Department of Education (DoE), Office of Vocational and Adult Education. (2013). *College and Career Readiness Standar ds for Adult Education. DoE:* Washington, D.C.

UNESCO, (2016). Chapter 15: Literacy and Numeracy in *Global Education Monitoring Report 2016. Education for people and planet: Creating sustainable futures for all*. United Nations Educational, Scientific and Cultural Organization, Paris, France http://unesdoc.unesco.org/images/0024/002457/245752e.pdf Accessed 31st October 2016.

Van Gog, T., Paas, F., & Sweller, J. (2010). Cognitive load theory: Advances in research on worked examples, animations, and cognitive load measurement. *Educational Psychology Review, 22*(4), 375-378. doi:10.1007/s10648-010-9145-4

Victorian Curriculum and Assessment Authority (VCAA) (2008). *Curriculum Planning Guide: Literacy and Numeracy Skills Strand Numeracy Skills Units*. 1st ed. Victoria: Victorian Curriculum and Assessment Authority.

Wake, G. (2015). Preparing for workplace numeracy: a modelling perspective. *ZDM Mathematics Education, 47(4*). Pp. 675–689.

Weeks, K. W., Higginson, R., Clochesy, J. M., & Coben, D. (2013). Safety in Numbers 7: Veni, vidi, duci: A grounded theory evaluation of nursing students' medication dosage calculation problem-solving schemata construction. *Nurse Education in Practice, 13*(2), e78-87. doi:http://dx.doi.org/10.1016/j.nepr.2012.10.014

Windisch, H. C. (2015). Adults with low literacy and numeracy skills: A literature review on policy intervention, OECD Education Working Papers, No. 123. Paris: OECD Publishing.

Zevenbergen, R. (2004). Technologizing numeracy: Intergenerational differences in working mathematically in new times. *Educational Studies in Mathematics*, *56*(1), 97-117