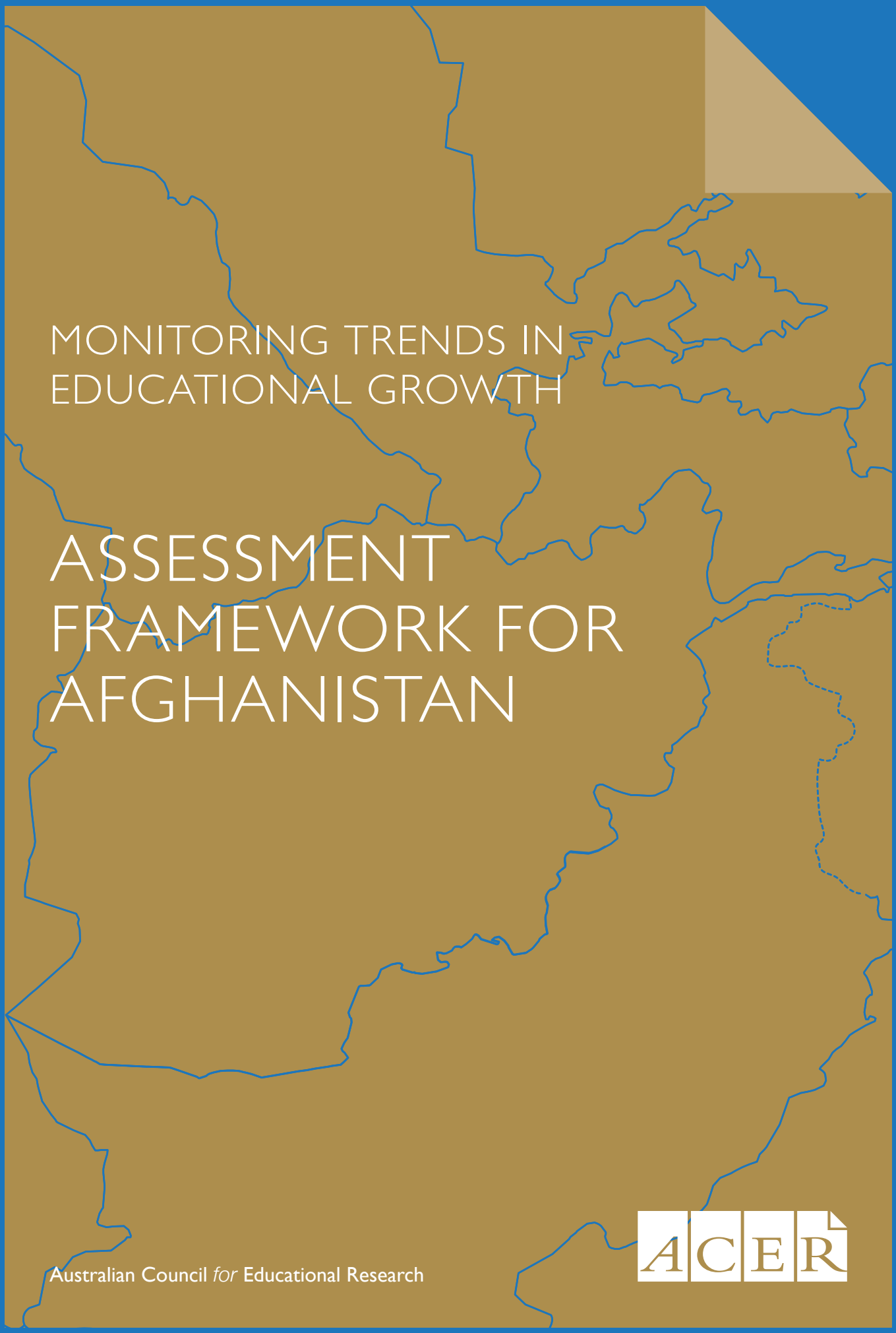
MONITORING TRENDS IN EDUCATIONAL GROWTH

ASSESSMENT FRAMEWORK FOR AFGHANISTAN



Australian Council *for* Educational Research

ACER

This report was prepared by the ACER Centre for Global Education Monitoring (ACER-GEM).

First published by
Australian Council *for* Educational Research Ltd
19 Prospect Hill Road,
Camberwell, Victoria, 3124
Phone: +61 3 9277 5555
www.acer.edu.au

ISBN 978-1-74286-412-9

Design: ACER Creative Services
ABN 19 004 398 145

Contents

Acknowledgements

Key authors and editors of this framework document were: Tim Friedman, Petra Lietz, Tom Lumley, Juliette Mendelovits, Sally Robertson, Rachel Stanyon, Ross Turner and Maurice Walker.

The MTEG 2013 assessment of Class 6 students in Afghanistan was a collaboration between the Centre for Global Education Monitoring at the Australian Council for Educational Research (ACER-GEM) and the Ministry of Education, Afghanistan. The key staff involved are acknowledged here.

Dr. Seddiq Weera, Strategic Advisor to the Minister of Education
Prof. Ustad Naeem Baheen, Director General, General Education, Ministry of Education
Dr. Amir Mansory, Professor of Education
Dr. Amir Mohammad, Director of EQUIP

Learning Assessment team
Prof. Aminullah Amin (Coordinator)
Mujihid Karimi (Research Officer)
Khan Wali Khan (Research Officer)
Nosheen Yasheen (Research Officer)
Yama Yayha (Administrative/Finance Officer)

Technical Support to the Learning Assessment team
Farahnaz Farah (National Technical Advisor)
Hafeez Samo (ITA) – until September 2013

Test Administrators
Mursal Ahrari
Ahmad Zubair Badri
Abdul Wahid Erfan
Gul Hakim
Noorullah Haqmal
Mohammad Jawid Norzoy
Emal Safir
Jamshid Sherzad
Mohammad Salim Sunny
Mohammad Rahim Taroon
Sulaiman Zadran
Shafiqa Nikzad Zahid
Mohammad Khalid Zaland

Australian Council for Educational Research
Ray Adams (Technical advisor)
Mark Butler (Test development)
Alexander Daraganov (Data analysis)
Alan Egbert (Operations management)
Tim Friedman (Data analysis/Questionnaire development)
Musab Hayatli, cApStAn (Linguistic quality control)
Nora Kovarcikova (Survey operations)
Winson Lam (IT services)
Petra Lietz (Questionnaire development)
John Lindsey (Test development)
Tom Lumley (Test development and reporting)
Greg Macaskill (Data analysis)
Juliette Mendelovits (Test development and reporting)
Martin Murphy (Sampling)
Sally Robertson (Project coordination and reporting)
Alla Routitsky (Data analysis)
Rachel Stanyon (Reporting)
Stephanie Templeton (Project administration)
Ross Turner (Test development and reporting)
Sam Venables (IT services)
Maurice Walker (ACER project director)

Thanks are also extended to the many other individuals involved, including school staff, those who contributed to the many aspects of materials development and delivery, and the students who participated.

The critical support of sponsors is also acknowledged, in particular World Bank staff and associates Samantha de Silva, Aman Farahi, Abdul Hai Sofizada and Poul Erik Rasmussen.

Chapter 1 Introduction

Overview

Large-scale international education surveys have been conducted since the early 1980s in many parts of the world. More recently, regional and national sample-based assessments have attempted to redress some of the shortcomings of these broad international surveys by focusing on more localised concerns (Wagner, 2011). The Monitoring Trends in Educational Growth (MTEG)[1] program aims to achieve the breadth and rigour of the large-scale international surveys while also addressing the unique needs and context of Afghanistan.

The MTEG program includes features designed to cater particularly for the development contexts of Afghanistan, where universal primary education has not yet been achieved and where educational infrastructure is undergoing expansion. The program may sit alongside any regular national assessment regime that provides comprehensive measures of attainment against specific curriculum goals.

The primary focus in reporting the results of MTEG is to inform Afghanistan's policymakers of the progress of educational development for which they are responsibile. In addition, the MTEG program is designed to assist other stakeholders such as teachers, parents and students in improving learning at the local level. A key factor of the rationale that underpins this strategy is that the assessment materials and the subsequent reports provide information about the strengths and weaknesses of students in the formative years of schooling. Assessments of this nature are structured so that improvements can be implemented to enhance learning programs, and resulting changes in student achievement can be measured in subsequent cycles of assessment.

The MTEG program will provide an ongoing measure of students' educational progress at key stages of learning: middle primary school (Class 3), towards the end of primary school (Class 6), and towards the end of compulsory secondary schooling (Class 9).

Aims of MTEG

The MTEG initiative has three core goals:

- To provide policymakers with relevant, sound and comparable data on contextual and learning outcomes that can directly inform local education policy development.
- To develop indicators of educational outcomes that enable meaningful comparisons of quality.
- To enhance the existing capacities of local teams to design data collection activities that will assist all aspects of the policy cycle: to develop and implement a reliable, valid and rigorous survey-based assessment and reporting program; and to appropriately analyse, interpret and disseminate assessment data with a view to informing education policy through relevant evidence.

Locally relevant policy-related outcomes

While the primary goal of all assessment programs is the collection of assessment data to contribute to the development of educational policies, the success of current programs in achieving this goal varies. In a review of the impact of national and international assessment programs on educational policy in developing countries, Best et al. (2012) pointed out that prioritising local policy concerns was key to the uptake of resulting information for educational policy development.

1 Previously MTEG was known as Monitoring Educational Development (MED).

Local policy concerns refer to those areas most salient to the national context. For example, comparisons between religious and secular schooling may be important; one region may have challenges in providing adequate school facilities, while another may have challenges in improving attendance levels. To be most effective, assessment data must address such local policy concerns.

Specific adaptations of, or additions to the MTEG instruments, are discussed by the ACER team and the Afghanistan Ministry of Education. This enables MTEG to address specific interests and concerns of Afghanistan that may be missing from or inadequately covered in the core material.

Comparisons

The MTEG program has four broad comparative aspects: growth between class levels, within-country sub-population comparisons, international comparisons and change over time.

A key feature of MTEG is the monitoring of growth as students move through primary and secondary class levels. This is essential information for the development of an education system as it allows policymakers to identify how much value is being added to students' educational outcomes by different stages of their schooling. Research questions here include: How much improvement in mathematical literacy (for example) is being achieved between Class 3 and Class 6? How much improvement is there between Class 6 and Class 9? Is the improvement the same for all groups of students or are some groups not progressing as well as others? What factors are associated with greater or less improvement? The answers to such questions help policymakers and practitioners better identify the appropriate stages and target groups for educational intervention or reform.

As with all assessment surveys, sub-population comparisons are essential. Education policymakers and practitioners need information on areas of strength and weakness for sub-populations variously defined by such characteristics as gender, socioeconomic status, geographic region, degree of urbanisation, language of instruction, and ethnicity. In addition, policymakers often want to compare educational outcomes across administratively distinct school types: public or private, religious or secular, vocational or academic. MTEG identifies relevant policy issues and has built the capacity for comparisons into the sample design specifically for Afghanistan.

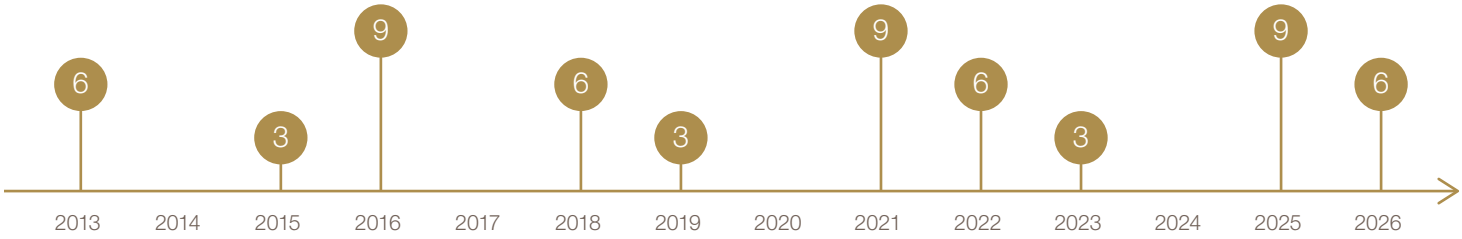
An external frame of reference is essential to an informed perspective on evaluating progress, and it provides a source of new ideas and possibilities for approaches to policy development and implementation. For example, international population comparison surveys such as the Progress in International Reading Literacy Study (PIRLS), the Trends in International Mathematics and Science Study (TIMSS) and the Programme for International Student Assessment (PISA) allow policymakers to monitor development of their education systems by providing outcome data on other countries as benchmarks. This is particularly useful when countries have commonalities such as a shared educational heritage (for example, colonial), similar cultural milieu (for example, language, ethnicity or religion), or a similar level of economic development.

Finally, trends – changes over time – are the most powerful approach to monitoring movement towards goals. Three categories of trends are particularly useful:

1. the change in achievement at a class level over time.
2. the change in growth between class levels over time.
3. the change in differences between sub-populations over time.

The proposed MTEG assessment schedule is shown in Exhibit 1. As can be seen in Exhibit 1,

Exhibit 1 MTEG assessment schedule in Afghanistan



the first assessment of Class 6 students took place in 2013. Another assessment of Class 6 students is planned for 2018, which will allow for measuring trends in Class 6 achievement over time. In addition to measuring class achievement, growth in achievement is measured by testing the same cohort as it progresses through class levels: the cohort tested in Class 3 in 2015 will be the same cohort tested in Class 6 in 2018, and again in Class 9 in 2021. Thus, the design of the program allows for conclusions to be drawn about changes in the achievement of Class 6 students at regular intervals, as well as about changes in achievement as students progress from one class level to another.

MTEG is facilitating the establishment of each of these forms of trend measurement according to the interests and needs of Afghanistan.

Capacity development

Education systems vary in their technical capacity to gather, process, analyse and interpret data in support of the development and review of educational policy. A central goal of the MTEG program is to build upon and enhance those existing capacities. Capacity development will occur at three levels:

- enhancing capacity to design data collection activities that assist all aspects of the policy cycle (Sutcliffe & Court, 2005) in terms of agenda setting, policy formulation and implementation, and the monitoring and evaluation of policy implementation
- enhancing capacity to develop and implement reliable, valid and rigorous survey-based assessment and reporting programs
- enhancing capacity to appropriately analyse, interpret and disseminate assessment data with a view to informing education policy through relevant evidence.

The purpose of an assessment framework

An assessment framework is an explicit statement and discussion about what an assessment intends to measure. It lays out the principles upon which an assessment is built.

The assessment framework serves a number of purposes and audiences. First, it gives a common language to stakeholders for discussion of the domain area.

Secondly, it guides test development, ensuring that the instrument serves the intended purposes and covers the domain in the way agreed upon at the outset.

Thirdly, it ensures that, where continuity from one year or one class level to another is of concern, there is an articulated plan for the assessment. This provides stability – or, where change is desired, it can be made explicit and implemented deliberately.

Finally, it communicates the purpose and features of the assessment program beyond the immediate stakeholders, and consequently helps in public interpretation of the results.

The development process for the MTEG assessment framework

The MTEG assessment framework outlines an approach to assessing mathematical literacy (Chapter 2), reading literacy (Chapter 3) and writing literacy (Chapter 4). It also puts forward a conceptual framework for the context questionnaires (Chapter 5).

Ideally, framework development occurs alongside test development. The development of an assessment framework often occurs post hoc: the test has been partially developed, fully developed or even administered before the construction of a framework that articulates the structure of the test. Post-hoc development is not ideal; nor is having the framework completely finished and finalised before test development begins. While the framework may be initiated before test development begins, in order to provide some structure and guidance, it is important that the process includes the capacity to review and revise the framework in light of its application, as the instruments are being developed. The development of the MTEG framework follows this ideal model.

During the period from 2012 to mid-2014, a group of researchers from ACER, comprising members of the international surveys team, psychometricians, test developers and questionnaire experts, collaborated in developing the MTEG assessment framework. The process began with a series of meetings at ACER in Melbourne, Australia, laying out the shape and principles of the framework. The first draft was shared with the Afghanistan Ministry of Education in December 2012. Subsequently, domain experts from the group drafted chapters for mathematics, reading and writing, which were reviewed and refined internally. At the same time, the first set of sample tasks for each domain was being drafted, with close reference to the first iterations of the framework. Two more drafts were circulated for comment and review by the ministry before this final version of the assessment framework was published.

General considerations in the design of the MTEG instruments

The literacy concept

MTEG aims to measure both curricular and cross-curricular knowledge, skills and understanding that are likely to allow school-aged students to progress successfully through school, and ultimately to play a constructive and fulfilling role as citizens in society. MTEG does not aim to comprehensively measure the Afghanistan curriculum. Rather, it adopts broad definitions for the domains of mathematics, reading and writing that are termed 'literacies'.[2] To convey this breadth and the parallel ways in which these three domains are conceived, the domains are referred to as mathematical literacy, reading literacy and writing literacy. The assessment includes the fundamental precursor skills that a student needs in order to be considered literate in each domain. Very importantly, together with the skills, knowledge and understandings that are inherent in each of the domains, the notion of a literacy includes the ability to acquire and *apply* such knowledge, skills and understanding in mathematics, reading and writing across a range of *contexts*, both within school and in extracurricular settings. The assessment of literacy in mathematics, reading and writing embraces the essential knowledge, skills and understanding of these curricular areas. It also investigates the extent to which such knowledge, skills and understanding can be used.

Literacy involves acquiring and applying skills, knowledge and understanding ...

The notion of mathematical literacy, for example, focuses on mathematical ways of thinking, the understanding of concepts and principles, and the ability to apply mathematical knowledge to solve problems in everyday contexts. Similarly, the

2 This discussion of the literacy concept is indebted to a discussion paper prepared by Professor Geoff Masters for the inaugural meeting of the PISA expert functional groups, Melbourne 1998.

concepts of reading literacy and writing literacy in MTEG are ultimately focused on reading and writing as means of expressing, communicating and understanding the world of ideas and information. For students at the beginning of their formal education, the development of reading and writing literacy will inevitably involve mastery of precursors of reading and writing literacy, such as decoding, phonemic awareness and basic vocabulary development, and in the development of mathematical literacy, concepts such as number and spatial development. These elements, although not ends in themselves, are essential stepping stones on the path to development of literacy within these domains, and may therefore be included in the assessment of literacy in MTEG.

... in a range of contexts

In their everyday lives, in their relations with family and friends, at school, at work and in the community, people use mathematics, reading and writing in countless ways. MTEG's aim of measuring students' ability to deal with the demands of life both at and beyond school therefore situates the sets of assessment tasks across a wide range of contexts.

MTEG has identified three broadly defined contexts in which the knowledge, skills and understanding related to the cognitive domains are likely to be enacted: personal, local and the wider world. An additional area included in the MTEG assessment, labelled 'intra-domain', deals with tasks provided without a context. *Personal* tasks relate to those matters that affect the individual, involving an inward focus. *Local* tasks pertain to contexts that require engagement with other individuals or with elements of the immediately surrounding environment. Tasks that have a *wider-world* context focus on issues relevant to whole communities or countries, and may even take a global perspective. Each of the domains will elaborate *personal*, *local* and *wider-world* contexts in somewhat different ways, but all will include tasks that assess students' proficiency across these three contexts to ensure that the instruments cover the range of areas in which mathematics, reading and writing are applied.

While the intention is generally to contextualise tasks in real-life contexts, a number of *intra-domain* tasks (tasks without context) are also included. For example, in the early stages of conceptual development, these comprise tasks that permit students to show their understanding of precursor skills within each domain, such as number sentences in mathematical literacy, recognition of letters and single words in reading literacy, and production of letters or single words in writing literacy. In addition, allowing for some context-free items permits the inclusion of items that reflect a wider range of current classroom practice.

Structure of the MTEG instruments

The core cognitive domains assessed in MTEG are mathematical literacy and reading literacy for all classes, and writing literacy for Class 6 and Class 9. Each student sampled for MTEG is administered assessment material in both reading and mathematical literacy for Class 3, and in all three domains for Class 6 and Class 9.

A substantial amount of test material is developed for the literacy domains in order to allow good coverage of the knowledge, skills and understanding involved in each. However, it is not necessary for every student to complete all of the tasks; indeed, to do so would make the assessment unreasonably long. Just as MTEG assesses a sample of students to gain an overall picture of the whole population's proficiency, so each sampled student completes only a subset of tasks from each domain. This design allows robust reporting of population and subgroup performance, and for comparisons to be made of performance in the different domains. Appendix A shows the assessment booklet designs for the assessment administered in 2013.

Background questionnaires are also included as part of the program. A student background questionnaire is administered to every participating student in Class 6 and Class 9 (where necessary, this may involve the assistance of teachers), allowing investigation of the relationship between performance on the cognitive domains and the background characteristics of students, such as gender, family type, home language and socioeconomic status. Due to the young age of Class 3 students, only a small number of background questions are administered to these students.

School principals are requested to complete a separate questionnaire that yields school-level data such as school type, number of teachers and available physical resources. Again, this information can be used to better understand factors associated with the performance of students in the mathematical, reading and writing literacy assessments.

Response formats

'Response format' refers to the kind of response that students are invited to give to an assessment task. In large-scale studies, typically two main response formats are employed: selected response, in which test-takers choose among options provided, and constructed response, in which test-takers generate their own response. The choice of response format for a task must be appropriate to the mode of delivery (for example, oral, paper-based or computer-based), to essential characteristics of the domain, and to the specific aspect of the domain being measured in a given task. The choice must also take into account practical considerations, such as the amount of testing time available, the feasibility of collecting reliable data from students, and the resources demanded for coding (scoring) the data.

The current MTEG Class 6 mathematical and reading literacy assessments are paper-based, and use both selected-response and constructed-response task formats. Typically, the selected-response format used in paper-based MTEG is the multiple-choice question, in which test-takers select one option from four or more alternatives. The constructed-response format is a short written response (a number or a solution showing working in mathematical literacy; a word or one or two sentences in reading literacy).

Research has shown that the format in which mathematics and reading tasks are administered has a significant impact on student performance. For example, Routitsky and Turner (2003) showed that in an international assessment of mathematics a mixture of task formats should be used, because students at different ability levels from different countries performed differently according to the format of the tasks. Monseur and Lafontaine (2009) found that there was a significant gender effect related to the two main task formats in reading assessments. In addition to these issues of fairness, construct considerations suggest that both multiple-choice and constructed-response formats be used. Including constructed-response tasks is important in ensuring that some elements of the domain can be adequately measured: for example, constructed response tasks are particularly useful when the focus of a task is to assess the quality or process of students' thinking, rather than to elicit a correct/incorrect response. For these reasons – to ensure proper coverage of the ability ranges in different cultural contexts, to ensure fairness between boys and girls, and to reflect the range of skills relevant to the domains – tasks of both multiple-choice and constructed-response formats are used in the mathematical and reading literacy assessments. Taking account of the additional resources required for coding constructed-response tasks, this format is used sparingly, with no more than 30 per cent of the mathematical and reading literacy tasks in constructed-response format.

MTEG's writing literacy assessment, because of the intrinsic nature of writing, consists entirely of tasks in which students are asked to generate a written response: constructed-response tasks. A variety of response formats is employed, ranging from asking students to provide a single word or

phrase (for example, to label an image), to short responses (for example, filling in a form) and more extended pieces of writing (composing a narrative or a letter) through which a student's capacity to develop ideas and sustain coherence in a piece of writing can be assessed.

Computer-delivered assessment is gaining increasing currency worldwide, and a computer-based assessment has been developed for the Class 3 assessment of MTEG Afghanistan. This assessment (mathematical literacy and reading literacy) is entirely based on a variety of selected-response formats, since requiring students at this stage of education to write answers would be likely to interfere with the measurement of their mathematical and reading proficiency.

Analysis and reporting

Using item response theory methodology, the tasks for each domain are arranged along a scale that indicates progressively the level of difficulty for students and the level of skill required to answer each task correctly. The scale summarises both the proficiency of a person in terms of his or her ability and the complexity of a task in terms of its difficulty. The assessment instruments are designed using common tasks to link between class levels (vertical linking) so that student proficiencies from lower primary to middle secondary are calibrated on the same scale, thus allowing reporting on the value added as students progress through school. Common tasks will also be used over time at the same class level (longitudinal linking) to link assessments from one cycle to the next, so that a system can monitor whether proficiency is improving (or declining) at a given class level.

The results for mathematical literacy, reading literacy and writing literacy are each reported on a described proficiency scale, which gives both quantitative results about the proportion of students performing at different levels of proficiency, and qualitative descriptions of the kinds of skills, knowledge and understanding that are associated with each level.

MTEG reporting will initially be designed for use by a wide range of policymakers, including those responsible for resource distribution, curriculum development and teacher training. Other versions of the results, with different emphases, will also be published, such as to help teachers use the data to inform their practice, or to communicate the outcomes to interested members of the public, including parents. For example, for the Class 6 results, a series of thematic reports and summary pamphlets were developed on Class 6 proficiency, Class 6 girls and boys, and school factors.[3]

Reporting will draw upon information from the student and school background questionnaires. These data will be analysed in relation to the domain-related outcomes to describe the characteristics of schools, families and students associated with stronger and weaker performance in the cognitive domains.

The analyses will provide evidence to guide effective and purposeful improvements in a rapidly developing education system, and will allow nuanced interpretation of the impact of educational reforms.

3 These documents are available from https://www.acer.edu.au/gem/key-areas/system-strengthening/mteg

Chapter 2 Mathematical literacy

The importance of mathematical literacy

An understanding of mathematics is central to a young person's future educational success and their preparedness for life. Mathematics assessment at a particular class level typically focuses on the mathematics knowledge and skills taught in that year or perhaps previous years. The primary focus of MTEG in mathematics, however, is on a broader set of mathematical skills and in particular on the extent to which students are able to make use of their mathematical knowledge and skills to solve problems and to deal with the kinds of challenges they meet in a variety of contexts, where mathematics may be relevant to those problems and challenges.

A set of underlying skills or competencies is a primary driver of a student's ability to effectively use their mathematical knowledge in a variety of contexts. Students need *communication* skills, both to recognise and process information and to express their reasoning and conclusions. Mathematical literacy often requires students to *devise strategies for solving problems*. This involves a set of critical control processes that guide an individual to recognise, formulate and solve problems, and to monitor and direct their progress through the solution process. When dealing with problems presented in various contexts, students need to be able to transform the information as presented into a mathematical form ready for the application of relevant procedural knowledge. When mathematical results and conclusions are found, these often need to be interpreted in relation to the original context. These steps of transformation and interpretation are often referred to as steps in the *mathematisation* process.

Students need to be able to work with different *representations* of mathematical objects and information, such as graphs, tables, charts, diagrams and equations. They need to develop *reasoning and argumentation* skills, in order to explore and link problem elements, to make inferences, and to justify conclusions. It is essential for students to have a repertoire of specific procedural knowledge and skills, and to recognise when a particular piece of knowledge might be relevant to the problem at hand. They therefore need to be able to use *symbolic, formal and technical language and operations* in order to interpret, manipulate and make use of symbolic expressions within a mathematical context that are governed by various conventions and rules. This may also involve using *mathematical tools* that might be relevant to a particular problem situation, such as measuring instruments, calculation devices and computer-based tools, knowing when a particular tool would be appropriate and also the limitations of such a tool.

These competencies are fundamental to mathematical literacy and are called on to varying degrees by the MTEG assessment tasks. They are based on work originally done by Mogens Niss and his colleagues in Denmark (Niss, 2003; Niss & Højgaard, 2011). The PISA 2012 framework uses a modified formulation of this set of capabilities, condensing the number from eight to seven based on investigation of the operation of the competencies through previously administered PISA items by the PISA mathematics expert group (Turner, Dossey, Blum, & Niss, 2013).

The MTEG assessment program for mathematical literacy is modelled on the concepts and structure of the OECD PISA 2012 mathematical literacy framework for 15-year-olds (OECD, 2013), but adapted for a broader range of target age groups. As such, it also includes precursor skills such as fundamental mathematical concepts (for example, magnitude), the use of positional and relational

language, numeration, arithmetic operations, classification of objects, shape recognition, elementary algebraic thinking (for example, simple number sentences), measurement, and the use and interpretation of data.

Defining the domain

MTEG is designed for students in middle primary school, upper primary school and middle secondary school. While knowledge and skills are acknowledged as important and necessary, most of the assessment questions focus on the student's capacity to take actions that will lead to a solution for a problem arising in any of the contexts they may encounter. Some questions will focus on mathematical knowledge and skills isolated from potential applications.

The working definition of mathematical literacy for MTEG is as follows:

> *MTEG mathematical literacy is a person's capacity, given a problem in a context that is of interest or importance to them to translate the problem into a suitable mathematical formulation, to apply mathematical knowledge and skills to find a solution, and to interpret the mathematical results in relation to the context and to review the merits or limitations of those results.*

The following remarks are intended to clarify the MTEG definition of mathematical literacy.

Mathematical literacy …

The term 'mathematical literacy' is used to emphasise the focus on using mathematical knowledge and skills (including those learned in the mathematics classroom) to solve problems that arise in contexts beyond the classroom.

… is a person's capacity, given a problem …

Action is required by a person to solve a problem. Success in solving the problem depends on the person's capacity to focus their mathematical competencies – their skills in communication, devising strategies, mathematisation, representation, reasoning and argumentation, using symbolic, formal and technical language and operations, using mathematical tools – on the problem.

… in a context that is of interest or importance to them …

This focus on problems in context helps the person to recognise and appreciate the role of mathematics in the world and the actions they need to practise to make sense of their world. That the problem is of interest or importance to the person provides a reason for students to engage with the problem and encourages their enthusiasm and persistence in finding a solution.

… to translate the problem into a suitable mathematical formulation …

Part of the action that needs to be taken to solve the given problem involves reformulating it in mathematical language in a form that can lead to a mathematical solution.

… to apply mathematical knowledge and skills to find a solution …

This action gives results in mathematical language.

… and to interpret the mathematical results in relation to the context and to review the merits or limitations of those results.

The suitability of the mathematical results is tested in the problem context to see whether they constitute a solution to the problem.

Organisation of the mathematical literacy domain framework

There are three components contributing to the MTEG definition of mathematical literacy:

- context: the situation in which the problem to be solved has arisen
- process: the actions required to solve the problem

- content: the mathematical knowledge and skills required to find a mathematical solution.

An assessment of a person's mathematical literacy therefore needs to have questions that:

- are set in a context of interest or importance to the person, involving one or more of the actions required to solve a problem in the context
- use broad mathematical competencies as well as a particular set of mathematical knowledge or skills appropriate to the stage of development or level of mathematical knowledge of the person.

Context

Test items and tasks used in the MTEG survey instruments are each associated with a *context* type. A context is the situation within which the details of a test item or task are located, or the situation that generated the stimulus material for the task. Contexts help to define the focus of thought or action with which people responding to problems or challenges must engage.

The main purpose of the defined contexts is to ensure that the set of items or tasks covers a range of situations in which students meet problems and challenges, and a range of different purposes for which the problems and challenges have been devised, to encourage engagement with the broadest possible range of individual interests and with a range of situations in which individuals typically operate in the 21st century.

The MTEG program uses four context types: *personal* contexts, *local* contexts, *wider-world* contexts and *intra-mathematical* contexts.

Personal contexts have an individual focus. The problem or challenge primarily affects the individual, and engagement with the task involves an inward focus. Problems fitting this context type include more abstract challenges that may have limited external purpose, and challenges focusing on personal concerns that are likely to be of interest and relevance only to the individual involved, such as games and puzzles, personal health, personal transport or travel, and personal finance.

Local contexts have an interactive focus requiring engagement with other individuals or with elements of the immediate surrounding environment. Problems fitting this context type involve day-to-day situations and activities at home or school, in the local community or at work, where the focus of thought and action lies in connections and interactions with nearby people or objects.

Wider-world contexts have an external focus on broader situations that may affect whole communities or countries, or have a wider relevance at a global level. Problems fitting this context type involve broad social issues such as public policy, transport systems, advertising, and broad scientific issues such as weather, climate, ecology or medicine.

Intra-mathematical contexts refer to problems where all the elements involved belong in the world of mathematics, without reference to any external or real-world contextual elements.

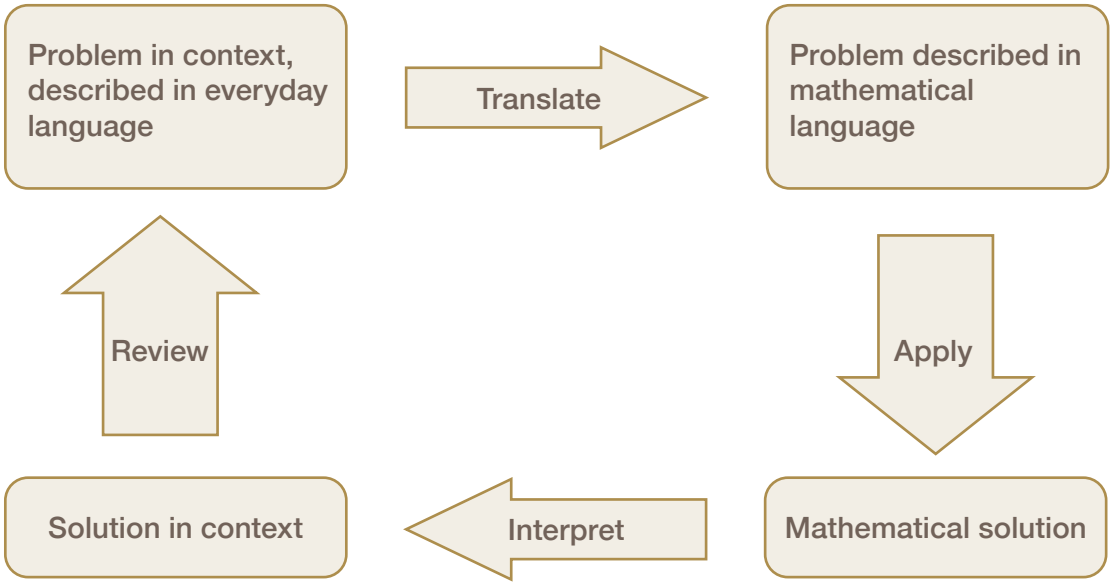
Process

Three processes have been defined for MTEG mathematical literacy assessment (see Exhibit 2). These are based on the conceptual model of mathematical literacy in the OECD PISA framework.

The processes shown in Exhibit 2 are as follows.

- *Translate* refers to the process of expressing the problem in mathematical language, thus taking it from the context to a mathematical formulation suitable for finding a solution.
- *Apply* refers to the process of using mathematical knowledge and skills to find a mathematical solution or to generate mathematical results: this process deals mainly with mathematical ideas, objects and techniques.

Exhibit 2 MTEG mathematical literacy processes



- *Interpret* refers to the process of retranslating the mathematical solution to the context of the problem. This may include a *review* of the solution to see whether it is reasonable and makes sense in context, and identifying any limitations for the solution.

In the assessment, a particular question may involve only one step in the solution cycle. For example, in a question focusing on the *apply* step, the *translate* step is included as part of the question and the required answer is the mathematical solution.

An example of a problem which does not require translation is shown in Exhibit 3. The problem has been formulated in clear *mathematical terms*, without any context other than the mathematical elements included (hence it is in the *intra-mathematical* category). The solution process involves reading and understanding the numbers and symbols, applying arithmetic skills to carry out the multiplication shown, then choosing the correct answer from the options provided. It is an example of an intra-mathematical item, presented in a clear mathematical formulation, with no *translate* step required.

Some questions require two or more processes. The step that is of greater significance to the solution cycle will determine the question category. For example, if the *interpret* step of a question is more significant than the *apply* step, the question would be categorised as 'interpret'.

The problem *Mass of Apple* (see Exhibit 4) is an example of this. Here students must read a small amount of information, presented in a shopping context (this problem is in the *local* context category), then translate that information into a mathematical problem – in this case, to divide 850 by 7. Students must then carry out the division and select from the given options the one that matches the result of their calculation. While both the *apply* and *translate* process categories are involved in this problem, it has been assigned to the *apply* category because carrying out the division is likely to be the more challenging aspect of this problem for most children. This is an example of an item that is set in the *local* context category, that illustrates the *apply* process category; and for which the *translate* process is also involved.

Exhibit 3 Multiplication: an intra-mathematical item

$13 \times 6 = ?$

- ○ 68
- ○ 78
- ○ 603
- ○ 618

Exhibit 4 *Mass of Apple*: an *apply* item

Najia buys 7 apples.

They have a mass of 850 grams altogether.

What is the approximate mass of one apple?

- ○ about 12 grams
- ○ about 80 grams
- ○ about 120 grams
- ○ about 600 grams

Content

Content is the specific mathematical knowledge and skills needed to find the solution to a problem. This framework uses three general content categories usually found in mathematics curricula: *number and algebra*, *measurement and geometry* and *chance and data*.

Pomegranates, a question in the *number and algebra* category, is shown in Exhibit 5. This problem requires students to recognise which one of four possible mathematical formulations is appropriate, in order to translate the problem depicted in the graphic stimulus into mathematical terms. The solution options provided show how algebraic thinking can be used to formulate the required calculation, but in a very practical and concrete *personal* context. The problem does not focus on carrying out the calculation shown, but on recognising how the calculation should be written mathematically, hence it is in the *translate* process category. The example item *Pomegranates*, illustrates the *number and algebra* content category, the *personal* context category and the *translate* process.

Problems arising in real life do not necessarily fall neatly into one content category. It is part of the student's role as problem-solver to choose knowledge and skills appropriate to the problem from their repertoire of mathematical knowledge and skills, combining aspects of different content areas as required, and employing their general mathematical competencies to do this.

The main purpose of this categorisation is to ensure that a wide set of mathematics knowledge and skills is represented in the problem-solution cycle. MTEG mathematical literacy includes the use of basic number skills and other fundamental mathematical conceptual understanding and skills, but encompasses much more than these with its focus on the use of those skills in a variety of contexts. It is also designed to be of interest to, and to provide a challenge for, students across a wide range of proficiency at a given level of schooling.

The problem *Buying Walnuts* (see Exhibit 6) illustrates the *measurement and geometry* content category. This problem is set in a *local* context and involves carefully interpreting a graphic stimulus to understand the measure of mass that is displayed on the face of each of the sets of scales shown. This item is in the *translate* process category because it asks students to interpret real-world contextual elements (the sets of measuring scales and the quantities of walnuts) and decide which image displays the specified mathematical quantity (400 grams on the scale displayed in kilograms, so working with different units of measurement is also involved). *Buying Walnuts* is set in the *local context and illustrates the measurement and geometry* content category.

Exhibit 5 *Pomegranates*: a *translate* item

Which of these shows how to work out how many pomegranates there are?

4 + 3

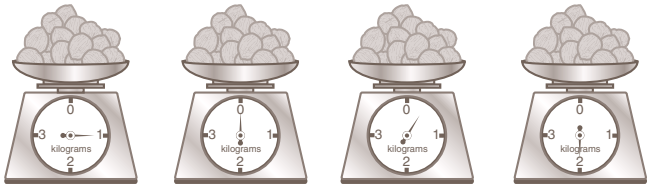
3 + 33

4 ÷ 3

4 × 3

Exhibit 6 *Buying Walnuts*: an *interpret* item

Which state of scales shows 400 grams of walnuts?



Assessing mathematical literacy

The three components that contribute to the definition of *MTEG* mathematical literacy – content, process and context – also provide the structure for the assessment of *MTEG* mathematical literacy.

Targets are established for each of these components to ensure a sensible coverage and overall balance for the assessment instrument, taking into account the level of schooling being assessed. This in turn ensures that a broad selection of problems or problem components is included to provide a fair, engaging and challenging assessment of mathematical literacy. The questions in each instrument cover a wide range of difficulty appropriate to the level of schooling.

It is assumed for tests at all levels that a calculator is available. At middle primary and upper primary levels, the test questions are structured to be as 'calculator-neutral' as possible – they can be done without a calculator, and using a calculator is not a significant advantage. At middle secondary school, because of the nature of the content now being assessed, some questions may require a calculator.

Establishing context is important for MTEG mathematical literacy, so language is an important component of mathematical literacy questions. The amount of language used and its level of difficulty are carefully monitored and reviewed to minimise the reading load while ensuring the questions are accurate, clear and unambiguous.

Target distribution of score points by content, process and context

Exhibit 7 shows the target percentages of content categories for each of the three class levels. The targets are given as ranges to emphasise that there is flexibility in the compilation of the assessment, with the overall aim being to achieve a sensible and appropriate balance of problems from each of the content categories.

Exhibit 8 shows the target percentages for process categories for each of the three school levels. Again, the targets are given as ranges to indicate flexibility while achieving overall coverage and balance in the assessment. The balance in this case is an approximately equal weighting between the two processes that link to the context (*translate* and *interpret/review*) and the process that provides a mathematical solution (*apply*).

The three main context types (*personal*, *local* and *wider-world*) should be represented approximately equally, with a smaller proportion of intra-mathematical problems also included. These target ranges are the same for all class levels. Exhibit 9 shows the target percentages of tasks in each content category in the MTEG 2013 mathematical literacy assessment.

Exhibit 7 Target percentages for mathematical literacy content categories, by class level

	Content categories Target percentage of tasks		
Class level	Number and algebra	Measurement and geometry	Chance and data
Middle primary	45–55	25–35	15–25
Upper primary	35–45	35–45	15–25
Middle secondary	35–45	30–40	20–30

Exhibit 8 Target percentages for mathematical literacy process categories, by class level

	Process categories Target percentage of tasks		
Class level	Translate	Apply	Interpret and review
Middle primary	15–25	50–70	15–25
Upper primary	20–30	40–60	20–30
Middle secondary	20–30	40–60	20–30

Exhibit 9 Target percentages of mathematical literacy context categories for all classes, MTEG 2013

	Context categories			
	Personal	Local	Wider-world	Intra-mathematical
Target (all class levels)	25–35	25–35	25–35	5–15

Exhibit 10 Target percentages for mathematical literacy response format categories, by class level

	Response format categories		
Class level	Selected response (MC and CMC)	Closed constructed response (CCR)	Open constructed response (OCR)
Middle primary	60–80	20–30	0–10
Upper primary	60–80	15–25	5–15
Middle secondary	50–70	15–25	15–25

Unit structure, response formats and scoring

An MTEG mathematical literacy assessment consists of a series of units, each of which has a stimulus to establish a context and one or more questions that require one or more of the processes (*translate*, *apply*, *interpret/review*) to be used to find an answer.

Four categories of response format are included in assessments of MTEG mathematical literacy.

Two of the categories are of the selected-response type, where the student selects one or more correct answers from a set of options.

- Multiple-choice (MC) tasks have four or five options, only one of which is the correct answer and the other three or four are plausible but incorrect answers.
- Complex multiple-choice (CMC) tasks present statements or propositions, and require students to select one or more correct response options to each statement from a set of possible options, such as ‘true or false’ or ‘always, sometimes, never’.

Two of the categories are varieties of constructed-response format, requiring students to write an answer, complete a drawing or mark a position.

- Closed constructed-response (CCR) tasks provide a structured format for the student response, which might be a single number, a word or a mark on a diagram.
- Open constructed-response (OCR) items typically need a more extended process to reach the required answer.

Some task formats provide opportunities to award partial credit for some items, where a student shows some progress towards a solution but does not give a response deserving full credit.

Exhibit 10 shows the target percentages for response formats for each of the three class levels. (The two categories of selected-response format are combined.)

Chapter 3 Reading literacy

The importance of reading literacy

Reading literacy is a foundational skill. It underlies success not only in school subjects but also in many areas of adult life (Smith, Mikulecky, Kibby, & Dreher, 2000). Acquiring skill in reading literacy benefits the individual not only by assisting participation in education and literate society, but also by shaping their thinking processes (Olson, 1994). Reading literacy is of fundamental importance to individuals in meeting their personal goals. At a broader level, a literate population is central to a nation's pursuit of its economic and social goals.

In the early stages of reading development, a number of *precursor skills* need to be acquired to support the central activity of reading for meaning. Precursor skills include letter and word recognition, fluency and speed in oral decoding of sentences and passages, and listening comprehension. While these precursor skills remain subsidiary to reading literacy, it is useful to track and measure progress in their acquisition, so that systems, schools, teachers and parents can understand what aspects of children's reading development may need attention as their reading progresses.

Defining the domain

The working definition of reading literacy for MTEG is as follows:

> *Reading literacy is understanding, using and responding to written texts, in order to meet personal, social, economic and civic needs.*

Reading literacy …

The term 'reading literacy' is used in preference to the word 'reading' alone to emphasise that what is being assessed goes beyond the simple decoding of words, though it also includes that. Reading literacy includes a range of cognitive skills such as locating and interpreting information, as well as knowledge of words and knowledge of linguistic structures and features. The term 'reading literacy' also encompasses the idea that reading is done in a context and for a purpose. Thus reading literacy includes the notion of relating one's knowledge about the world to texts, and using texts to develop and reappraise one's knowledge of the world.

… is understanding, using and responding to …

These verbs are intended to give a sense of the broad range of purposes for which texts might be read. 'Understanding' involves comprehension, while 'using' and 'responding to' acknowledge both that the reader is actively involved in the construction of meaning, and that reading is functional.

… written texts …

The term 'written texts' indicates that the focus is on the written word. It comprises handwritten, printed and digital texts, but *excludes* spoken texts. Visual artefacts such as diagrams, pictures, maps and tables may be regarded as components of written texts if they contain words, or where they support the meaning of the written text.

… in order to meet personal, social, economic and civic needs.

People read for a variety of purposes, from meeting their individual learning needs or other aspects of personal development, to communicating with others, meeting the demands of their job, or informing themselves about local and global issues.

Organisation of the reading literacy domain framework

The MTEG reading literacy framework is primarily described in terms of content (the text variables: text format and text type), context (the situation

Exhibit 11 *The Hole*: a *locate* item

'I can see something shiny at the bottom,' said Samsur. 'Maybe it's a gold coin.'

'Don't be silly,' said Nazneen, peering into the hole. Her younger brother was always seeing things, creating objects out of nothing.

'Maybe it's a sword,' continued Samsur. 'Maybe a king buried a gold sword in the ground many years ago, and then forgot about it.'

'Maybe it's dirt, covered in dirt, covered in more dirt,' said Nazneen. 'It's just a hole, probably made by a wild animal.'

'You are wrong!' exclaimed Samsur. 'No animal could make a hole as big as this!'

'Well, if you are so sure this is not an animal's hole, perhaps you should climb down into it.'

Samsur began to turn pale. 'Erm … No. I cannot go in the hole … because … I have a sore foot!'

Nanzeen smiled; it had nothing to do with Samsur's foot. A big hole could mean a big animal.

'I have have an idea,' she said, picking up a stone that lay beside her. 'I will drop this into the hole. If we hear a clink, there is treasure. If we hear a thud, there is dirt. If we hear a yelp, there is an animal.'

Nanzeen dropped the stone and they hear nothing for a moment.

Then they heard a splash.

Nanzeen says 'I have an idea'.

What is her idea?

A. to push her brother into the hole

B. to go into the hole to explore

C. to throw a coin into the hole

D. to drop a stone into the hole

to which texts are relevant) and process (the cognitive processes readers use). As an adjunct, the inclusion of precursor skills contributes to elaborating the constituents of the domain at the early stages of reading development. The precursors are described in terms of constituent skills such as word recognition.

Content: text variables

Content in the reading framework is represented by the text variables of text format and text type.

Text format

Text format refers to the way texts are organised or laid out on the page, in very broad terms. MTEG uses three categories of text format: *continuous*, *non-continuous* and *composite*.

Many texts are in the form of *continuous text*, or prose. Continuous texts are composed of sentences and paragraphs (see Exhibit 11 for an example).

Other texts that readers are required to engage with in daily life are constructed in *non-continuous* formats (see Exhibit 12 for an example). These include diagrams, table, maps and lists (Kirsch & Mosenthal, 1990).

This broad distinction between continuous and non-continuous texts is a common one in reading frameworks, such as PISA (OECD, 2010), PIRLS (Mullis, Martin, Kennedy, Trong, & Sainsbury, 2009), Southern and Eastern Africa Consortium for Monitoring Educational Quality (SACMEQ) (Ercikan, Arim, Oliveri, & Sandilands, 2008) and Programme for the International Assessment of Adult Competencies (PIACC) (OECD, 2009), though there are some variations in terminology.

A *composite* text involves more than one part. It could be a text containing both continuous and non-continuous parts (such as a page from a newspaper that comprises prose text and graphs), or it could be several texts on a single theme but in one format (for example, several opinion pieces by different authors related to a single issue).

Text type

Text type refers to the genre, orientation or broad purpose of a text. MTEG uses six categories of text type: *narrative*, *descriptive*, *persuasive*, *instructional*, *transactional* and *label.*

Narrative texts present and develop characters, events and themes, and deal with questions relating to *when*, or *in what sequence*. Examples of narration are short stories, recounts of recent activities, diary entries and stories of a person's life. *The Hole* (see Exhibit 11) is a *narrative* text, telling the story of two children's adventure.

Descriptive texts present information about people, objects and abstract concepts or constructs; these kinds of texts address *what* and some *how* questions. Description includes forms of writing sometimes referred to as exposition. Examples of description include describing a person or a place, a plant or a problem, a feeling or a phenomenon, or, at the level of precursor skills, a label for an image. *A Country Fact File* (see Exhibit 12), which provides information describing features of various countries, is an example of a descriptive text.

Persuasive texts deal with opinions and points of view, and are used to persuade the reader. They address some *which* and *why* questions. Examples of persuasive texts are a letter to the editor, a book review, an advertisement, a job application letter and a discussion of the benefits or disadvantages of a public policy.

Instructional texts explain what to do in order to complete a specified task, and thus address some *how* and *when* questions. Examples of instructional texts are giving directions for finding a location, listing materials and steps required to make an object, and explaining what to do in an emergency.

Transactional texts aim to achieve a specific purpose involving an exchange of information between two or more parties, such as arranging for something to be done. Transaction is represented by reading tasks such as a message from a friend or correspondence related to the delivery of goods. Transaction as a text type follows the definition as used in the PISA 2009 reading literacy framework (OECD, 2010).

A *label* is a text consisting of a single word or a small set of words, used to identify something. This text type is used to categorise images or words that are presented in isolation, as a stimulus to assess some of the precursor skills of reading. Exhibit 14 is an example of a text with a *label*. This text consists of an image accompanied by four words, one of which is a suitable label for the image.

Processes

Locate

A common purpose for reading is to *locate* information. The information required might be very specific, such as which character performed a particular action in a narrative, or it might be more general, such as finding evidence that supports an argument. Sometimes, the information to be located is found in a single sentence, and sometimes it must be gleaned from several paragraphs. This kind of reading has been called 'reading the lines' (Gray, 1960), because no inference, or only minimal inference, is required to complete this kind of task. An example of an item requiring students to *locate* information is given in Exhibit 11. *The Hole* example item is a *locate* item, presented in *continuous* format, of *narrative* type, set in a *personal* context.

In order to identify what Nazneen's idea is, students need to find the part of the text that contains the quotation 'I have an idea', towards the end of the text. They then need to continue reading the words that follow, which state first that she picks up a stone and secondly that she 'will drop this into the hole'. There is some minor inference required in order to recognise that 'this' refers to the stone that she has picked up, and to relate both of these to her immediately preceding statement, 'I have an idea'. However,

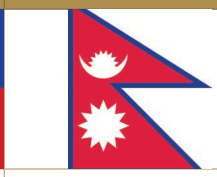
since all the information is explicitly stated, with students able to rely on direct word matches between the question and the text ('I have an idea', 'stone', 'drop ... into the hole'), this item is classified as relying essentially upon ability to locate information.

Interpret

Interpretation is the process of making meaning from a text. Gray (1960) refers to this kind of task as 'reading between the lines': it involves understanding ideas that are present in a text but not directly stated. Interpretation might involve parts of a text or the whole text. A wide variety of cognitive tasks may be included in this process, such as recognising relationships between ideas, understanding assumptions made, synthesising different pieces of information, or identifying a main idea. An example of an item requiring students to *interpret* information in a text is given in Exhibit 12. *Country Fact File* question 4 (shown in Exhibit 12) is an example of an *interpret* item, presented in *non-continuous* format, providing a *description* of features of various countries, set in a *wider-world* context.

The question shown in Exhibit 12 asks students to use information in the *Country Fact File* text to identify a country that exports the same goods as Afghanistan. In order to answer this question, students need to identify the row 'Typical exports' and read across that row to determine which goods Afghanistan exports ('fruit and nuts, carpets, saffron'). They then need to continue reading across that row, comparing the information about the other three countries, represented by the columns in the table, in order to identify a similarity. The relevant information ('carpets') is found in the cell describing Nepal. Nowhere does the table state explicitly that Afghanistan and Nepal export one category of similar goods (carpets), nor does the question indicate which category of 'Typical exports' students should focus on. Although the information that leads students to the answer takes the form of a word match between two cells, the task requires students to interpret the expression 'the same goods' and to compare multiple pieces of information in multiple cells of the table in order to identify one single

Exhibit 12 *Country Fact File Q4*: an *interpret* item

	Afghanistan	Vietnam	Philippines	Nepal
				
Climate	arid to semi-arid; freezing winters and hot summers	tropical in south; monsoonal in north	usually hot and humid	subtropical in south; cool summers and severe winters in north
Geography	landlocked and mountainous	the fertile Mekong river delta covers a large part of south western Vietnam	made up of 7,107 islands	landlocked; contains eight of the world's 10 highest peaks
Main crops	wheat, fruits, nuts; wool, sheepskins	paddy rice, coffee, rubber, cotton, fish	sugarcane, coconuts, rice	rice, corn, wheat, sugarcane, milk
Typical exports (goods sold to other countries)	fruits and nuts, carpet, saffron	crude oil, marine products, rice, coffee, rubber, garments	electronic equipment, transport equipment, garments	carpets, clothing, leather goods
Wildlife	the Marco Polo sheep: it has the longest horns of any sheep	the saola (a kind of antelope): one of the world's rarest mammals	the Philippine Eagle: the largest eagle in the world	the one-horned rhinoceros: the world's fourth largest land mammal

According to the text, which country exports the same goods as Afghanistan?

Exhibit 13 *Country Fact File* Q8: a *reflect* item

How is information shown in this text?

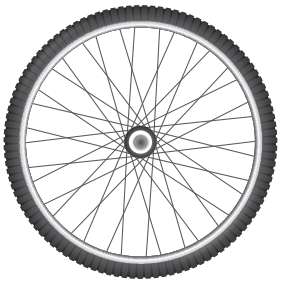
A. in sentences

B. in paragraphs

C. in a table

D. in a map

Exhibit 14 *Wheel*: a *recognise words* item



A. Car

B. Shoe

C. Wheel

D. Goat

similarity between two countries. They then need to write the word 'Nepal' for this *constructed-response* item. The need for a series of actions involving identifying relevant information followed by multiple comparisons means that this item is classified as *interpret*.

Reflect

Active readers constantly relate what they are reading to what they already know, and adjust what they know to accommodate what they have read. The *reflect* process refers to this aspect of reading, in which information within the text is related to knowledge outside the text; in other words, the reader situates the text within the wider context of his or her experience. Because this skill goes beyond the text itself, it has been called 'reading beyond the lines' (Gray, 1960). The broad range of tasks categorised under this process include those focusing on the intended audience of a text or the attitude of the writer; making an evaluation of an argument or a judgement about a character; explaining the effect of a text feature such as its layout; and comparing the behaviour of a character in a story with that of acquaintances.

An example of a *reflect* item is given in Exhibit 13. This item also comes from the unit *Country Fact File*.

The question asks students to identify the form in which information is shown in the *Country Fact File* text. In order to answer the question, students need to draw on information beyond the text. In this case, they need to use real-world knowledge to understand the differences between sentences, paragraphs, a table and a map, and relate this information to the text in order to recognise that the information is presented in a table. Items that focus on the layout of a text are classified as *reflect* questions.

Recognise words

A basic element of reading literacy is knowledge of words. Knowledge comprises both *recognising* the written form of the language and *conceptual recognition* of the meaning of a word – its correlate in the non-linguistic world. Recognising words means relating the written form of a word with its meaning (for example, as represented in picture form). An example of an item requiring students to recognise words is given in Exhibit 14.

Here students are presented with an image of a familiar object and a set of four words, from which they need to select the one that describes the picture of a wheel.

Contexts

Test items and tasks in MTEG instruments are generally associated with a *context* type – though for assessment of some of the precursor skills of reading, a context is not provided. Other than tasks of this type, however, the reading *context* is the situation within which the text is likely to be read or for which it is likely to be used. The main purpose of defining the *contexts* is to ensure that the set of items or tasks covers a range of situations in which students are likely to read.

The MTEG program uses three context types: *personal* contexts, *local* contexts and *wider-world* contexts.

Personal contexts have an individual focus such as personal health, personal transport or travel. Reading tasks fitting a personal context include those that are primarily for personal enjoyment or development, such as reading a story or a TV guide. The story *The Hole* (Exhibit 11) is an example of a text set in a *personal* context.

Local contexts have an interactive focus requiring engagement with other individuals or with elements of the immediate surrounding environment. Reading in this type of context involves day-to-day situations and activities at home, at school, in the local community, or at work, where the focus of thought and action lies in connections and interactions with nearby people or objects. Reading texts reflecting a local context include a letter from a friend, a school timetable or a description of one's hometown. The item *Wheel* (Exhibit 14), dealing with a familiar everyday object, is an example of an item set in a local context.

Wider-world contexts have an external focus on broader situations that may affect whole communities or countries, or have an even wider, global relevance. Texts fitting this context type include those dealing with broad social issues such as public policy, transport systems and advertising. Reading texts that reflect a wider-world context include a newspaper report or a historical description. The *Country Fact File* text (Exhibit 12), describing features of various countries, is an example of a text set in a *wider-world* context.

Assessing reading literacy

Target distribution of score points by content, process and context

The distributions presented in this section show the targets for Class 6 (see Exhibits 15, 16 and 17). The percentages may be adjusted for other classes.

Response formats

The reading literacy assessment includes both selected-response and constructed-response tasks. The majority of selected-response tasks are simple multiple-choice format, in which the test-taker selects one of four options. A small number of tasks may involve complex multiple-choice, in which test-takers are required to make several decisions, for example, by responding to a series of yes/no questions. Constructed-response tasks comprise approximately 30 per cent of the entire set, and the exact proportion is determined depending on the class level of assessment. The percentage of constructed-response tasks in the MTEG Class 6 reading literacy assessment for 2013 was 32 per cent.

Exhibit 15 Target percentages for Class 6 reading literacy text format categories, MTEG 2013

Text format	Target percentage of tasks
Continuous	50–60
Non-continuous	30–40
Composite	5–15

Exhibit 16: Target percentages for Class 6 reading literacy text type categories, MTEG 2013

Text type	Target percentage of tasks
Narrative	25–35
Descriptive	25–35
Persuasive	10–20
Instructional	5–15
Transactional	0–10
Label	10–20

Exhibit 17 Target percentages for Class 6 reading process categories, MTEG 2013

Process	Target percentage of tasks
Locate	35–45
Interpret	30–40
Reflect	10–20
Recognise word	10–20

Chapter 4 Writing literacy

The importance of writing literacy

Like reading and mathematics, writing is a foundational skill for future learning and for full participation in the economic, political and social life of adults. In school contexts, writing is a basic tool for learning. In later life, writing is essential for participation in many aspects of everyday life, such as communicating with friends and family, or with government departments. In the workplace even routine jobs increasingly rely on high-level cognitive skills – including written communication – rather than on manual skills. In the digital age, personal and social communication is increasingly conducted in written text, through social media. In the 21st century, written language is as at least as important for the individual as it has ever been.

As John Wirt puts it:

> *Effective writing skills are important in all stages of life from early education to future employment. In the business world, as well as in school, students must convey complex ideas and information in a clear, succinct manner. Inadequate writing skills, therefore, could inhibit achievement across the curriculum and in future careers, while proficient writing skills help students convey ideas, deliver instructions, analyse information, and motivate others. (Wirt et al., 1998, p. 70).*

While this statement is not new, and is addressed primarily to an American audience, its message remains relevant and applies to developing education systems as well as to more developed ones.

The MTEG assessment of writing is restricted to Class 6 and Class 9, since a large proportion of students in Class 3 in developing countries are still at an early stage of writing development, focused mostly on precursor skills of writing rather than on *writing literacy* as described in the following section.

Defining the domain

The working definition of writing literacy for MTEG is as follows:

> *Writing literacy is constructing meaning by generating written texts to express oneself and communicate with others, in order to meet personal, social, economic and civic needs.*

Writing literacy …

The term 'writing literacy' is used in preference to the word 'writing' to emphasise that what is being assessed goes beyond simply copying or forming words, although the ability to write words in legible handwriting and to use correct spelling or character formation are essential components of writing. The term 'writing literacy' is meant to convey the idea that writing is done in a context, for an audience and with a purpose. Writing literacy includes a range of cognitive skills such as generating and organising ideas, applying vocabulary and drawing on knowledge of linguistic structures and textual features.

… is constructing meaning by generating written texts …

The term 'construct' is used here to emphasise that meaning comes from the writer. Written texts contain ideas developed by the writer, using knowledge of language and text, rather than being simply a written copy of others' ideas.

… to express oneself and communicate with others, …

While most typically people write in order to convey ideas and information to a specific audience, writing can also be for oneself, an act of personal expression.

… in order to meet personal, social, economic and civic needs.

Writing may be done for a variety of purposes, from keeping personal records to showing one's knowledge in the classroom; from sharing one's experiences with others to getting things done; and from meeting the demands of one's job to participating in public life.

Organisation of the writing literacy domain framework

Like mathematical literacy and reading literacy, writing literacy is described in terms of content, context and process. Content in writing literacy refers to types of written text. Context refers to the situations that give rise to the writing. Process refers to the skills applied by writers in constructing texts.

Content: text types

Content in writing literacy refers to the text types included as assessment tasks. These are *narration*, *description*, *persuasion* (or argument), *instruction* and *transaction*. These categories are widely used in literacy frameworks, such as the PISA 2009 reading literacy framework (OECD, 2010), although there are minor differences in the categorisation of text types from one framework to another. MTEG adds the category *label* to include tasks directed at early-stage writers.

Narrative texts present and develop characters and sequences of events. Narration is a fundamental and universal form of writing. Writing a narrative allows students to exercise their imagination and give shape to ideas and feelings. Examples of narrative texts are short stories, recounts of recent activities, diary entries and stories of a person's life. *Brothers' Race* (see Exhibit 18) is an example of a narrative text type in the writing literacy assessment. The task presents an image together with instructions to write a story. An introduction, including the name of one of the brothers, is given to assist students who may be unsure how to begin writing.

Descriptive texts present information about concrete objects – people, places, items or events – or abstract concepts or ideas; these kinds of texts explain how things are. Description includes forms of writing sometimes referred to as exposition. Students need to be able to write descriptions for many school tasks, as well as for broader everyday contexts. Examples of this text type include describing a person or a place, a plan or a problem, a feeling or a phenomenon. *Celebration* (see Exhibit 19) is an example of a descriptive text type. Students are asked to describe the Eid celebration in a way that is interesting for a reader who is unfamiliar with how this festival is celebrated in Afghanistan.

Persuasive texts communicate opinions and argue a point of view. In writing persuasive texts, students express their own thoughts, values and beliefs, and attempt to influence others. Examples of persuasive texts are a letter to the editor, a book review, an advertisement for a product, a job application letter and a discussion of the benefits or disadvantages of a public policy.

Instructional texts explain how to complete a task. Examples of instructional texts are giving directions for finding a location, listing the materials and steps required to make something, and explaining what to do in an emergency.

Transactional texts aim to achieve a specific purpose, such as asking for information about a state of affairs, or arranging for something to be done. Transaction is represented by tasks such as writing a message to a friend or ordering goods. Transaction as a text type follows the definition as used in the PISA reading literacy framework:

> *Transaction represents the kind of text that aims to achieve a specific purpose outlined in the text, such as requesting that something is done, organising a meeting or making a social engagement with a friend. (OECD, 2013, p. 66)*

Label is a text consisting of a single word or a small set of words to identify something. This text type is used to categorise images or words that

Exhibit 18 *Brothers' Race*: a *narrative* text

Use the picture to help you write a story. Write as much as you can.

One day, Kamyar challenged his older brother to a race.

.....

.....

.....

.....

Exhibit 19 *Celebration*: a descriptive text

Write a letter to a friend in another country to describe what happens in Afghanistan during the Eid celebration.

Tell you friend about:

- Places and times
- People
- Food
- Dress

Your description should be interesting. Write your letter on the lines below.

Dear Friend,

.....

.....

.....

.....

.....

.....

.....

.....

From your friend,

are presented in isolation, as a stimulus to assess some of the precursor skills of writing.

Context

Test items and tasks are each associated with a *context* type. A context is the situation within which the writing task is likely to take place. The main purpose of the defined *contexts* is to ensure that the set of tasks covers a broad range of the situations in which students need to write, and a broad range of the purposes and audiences for writing.

The MTEG program uses three contexts: *personal*, *local* and *wider-world* contexts.

Personal contexts have an individual focus. The primary audience of writing tasks in personal contexts is the writer him- or herself. Writing tasks fitting a personal context include those that are primarily for individual needs, enjoyment or development (such as writing a story or a personal shopping list), or for personal expression (such as keeping a diary).

Local contexts have an interactive focus, requiring engagement with other individuals or with elements of the immediate surrounding environment. Tasks fitting this context type involve day-to-day situations and activities at home, at school, in the local community, or at work, where the focus of thought and action

lies in connections and interactions with nearby people or objects. Writing tasks reflecting a local context might include a letter to a family member, a friend or a teacher; a household shopping list; or a description of one's hometown.

Wider-world contexts have an external focus on broader situations that may affect whole communities or countries, or have an even wider, global relevance. Writing tasks fitting this context type might focus on broad social issues such as public policy, transport systems, ecology, medicine or advertising. Writing texts that reflect a wider-world context might include a letter to the editor or a description of a famous person.

Processes

Writing entails drawing on knowledge of language (both written and oral) and a range of skills. In the writing literacy domain, this set of knowledge and skills comprises the process dimension of the framework. Five processes have been identified as intrinsic to writing literacy: *generating ideas*, *controlling text structure and organisation*, *managing coherence*, *using vocabulary*, and *controlling syntax and grammar*. A sixth variable, *other, language-specific features*, is included here to accommodate other important features that are not assessable across all languages.

Generate ideas

Writing tasks typically require the creation, selection and crafting of ideas. The quantity and quality of the ideas and their appropriateness for the task are constituents of this skill. The nature of the ideas will vary from one text type to another. For example, in story writing (narrative), strong characterisation and storyline are important. In persuasive writing, the logic, relevance and persuasiveness of the argument are important, as is the ability to maintain critical distance. In descriptive writing, the completeness of the description, the salience of the details included, and the precision and richness of the picture created for the reader are all important.

The extract of the marking guide for the writing task *Brothers' Race* (see Exhibit 20) illustrates the *generate ideas* process. The criterion assessed for this task is *development of narrative* (elaboration of ideas). A score from 0 to 4 is awarded to each piece of student writing, depending on how well the writing shows evidence of ability to elaborate ideas relevant to the picture in order to develop a narrative in accordance with the task. As the quantity of relevant ideas increases, together with the level of detail provided, so does the score given.

Control text structure and organisation

Different text types have different structures. Effective writers have knowledge of the structural features of texts and select a suitable organisational form for the writing task. For example, if writing a recipe, the writer will start with a set of ingredients, and then describe or list a sequence of steps. If writing a narrative, they know that, conventionally, they will start with an orientation, follow this with a complication, and end with a resolution. They also know what to include in each of these sections. For example, the orientation will introduce the main characters and establish the setting.

An example of how ability to *control text structure and organisation* is assessed is given in the marking guide used for the criterion *story elements* for the *Brothers' Race* task (see Exhibit 21). The focus in this criterion is not on the quantity of ideas, but rather on whether students demonstrate ability to link their ideas into a narrative. Students who do no more than describe the elements of the picture provided, for example, would be likely to receive a score of 1.

Manage coherence

Good writers are able to structure texts in such a way that the links between ideas are clear to the reader. Coherence is achieved through a logical progression of ideas that express meaning consistent with the reader's general world

Exhibit 20 *Brothers' Race* marking guide for the criterion *development of narrative*

Development of narrative (elaboration of ideas)	0	Evidence of a response but no relevant information is included
	1	Fragments: few ideas or no complete ideas
	2	Limited writing related to the picture
	3	Simple writing related to the picture; limited detail
	4	Detailed writing with many relevant ideas

Exhibit 21 *Brothers' Race* marking guide for the criterion *story elements*

Story elements	0	Evidence of a response but no relevant information is included
	1	Ideas are present but not a narrative
	2	Ideas are linked into a narrative

knowledge, as well as through syntactic features such as reference, and lexical features such as discourse markers and connectives. Good writers make use of paragraphing to group ideas around a central topic, or use other graphical means, such as headings, to indicate the relationship between ideas.

Control of coherence is a mark of relatively sophisticated writing, and may not be taught to students in Class 6. Coherence can most easily be observed in texts of several paragraphs; students in Class 6 are typically expected to produce rather short texts, where this aspect of writing cannot easily be assessed.

Use vocabulary

Writing involves not just knowledge of words but also an understanding of how they can be used in specific contexts. Good writers are able to draw on a wide vocabulary to present ideas precisely and concisely. They choose words that are appropriate to the purpose, audience and context. A wide vocabulary allows writers to present arguments effectively, and to give life to images in descriptive or narrative writing.

An example of how vocabulary is assessed is shown in the marking guide for the *vocabulary* criterion for the *Celebration* task (see Exhibit 22), where students can be awarded a score of 0, 1 or 2, depending on their ability to use vocabulary to convey their message. Providing detail in a written text requires a relatively broad vocabulary.

Control syntax and grammar

Writers need to understand implicitly how the rules of grammar govern the way words are put together to form phrases, clauses and sentences. Good writers produce grammatically correct, meaningful sentences and make use of a range of syntactic structures. They link ideas with a variety of cohesive devices and use sentence structures appropriate to the writing task.

An example of how ability to control syntax and grammar is assessed is shown in Exhibit 24, for the task *Scenes We See: Bird over mountain* (see Exhibit 23). The marking guide recognises that students may still be at the stage of gaining control of simple sentences, while also perhaps attempting to write more complex ones. In this task students are asked to write two sentences, but the marking guide gives some credit (score 1) to students who demonstrate the ability to write a single sentence correctly. If they attempt more complex sentences, they are more likely to make errors; they receive a full credit (score 2) if they demonstrate the ability to write a correctly formed complex or compound sentence, as well as if they write two correctly formed simple sentences.

Other, language-specific features

Other, language-specific features are not defined in the framework. This category allows description of writing skills judged intrinsic to writing literacy in individual languages or language groups, which would be irrelevant in others. Character formation for some Asian languages is one example in this category. Spelling, in languages such as Dari, Pashto and English, is another. (Spelling is considered by many to be an important feature of writing literacy in English, but less so in Hindi or Spanish, in which the relationship between sound and written form is very regular – for discussion, see Share, 2008.)

An example of one way in which *other, language-specific features* are assessed in the Class 6 assessment for Afghanistan is shown in Exhibit 25, the marking guide for *punctuation* for the task *Brothers' Race*. The same marking guide is applied to both languages, but scores on this criterion are analysed separately.

Assessing writing literacy

The writing literacy assessment includes tasks of varying lengths. At the simplest level, in order to take into account the assessment needs of writers at an early stage of development, students are required to write single words as labels for images. Some tasks require students to develop one or two sentences, and for other tasks they need to write a more extended piece of prose. The longest tasks take 15 minutes, so none of the writing tasks requires more than a page or so of composition.

The distributions presented in this section show the targets for Class 6. The percentages may be adjusted for other classes.

Text types

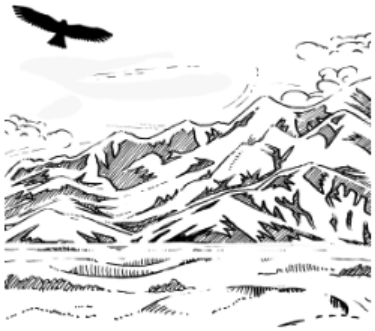
Tasks of varying difficulty and length are each categorised according to one of five text types: narrative, descriptive, persuasive, instructional, transactional and label. Exhibit 26 shows the target distribution of score points across the tasks by text type for Class 6.

Exhibit 22 *Celebration* marking guide for the criterion *vocabulary*

Vocabulary	0	Little control of relevant vocabulary
	1	Vocabulary used shows limited ability to convey a message
	2	Vocabulary is adequate to convey detail of message

Exhibit 23 *Scenes We See*: a ***descriptive*** text

Write two sentences to describe this picture.



1. .....

2. .....

Exhibit 24 *Scenes We See* marking guide for the criterion *syntax and sentence structure*

Syntax/ sentence structure	0	Isolated words or sentence fragments only
	1	Some errors but comprehensible, or one simple sentence correctly formed
	2	Two simple sentences correctly formed, or one complex/compound sentence correctly formed

Exhibit 25 *Brothers' Race* marking guide for the criterion *punctuation*

Criterion	Score	Description
Punctuation	0	No evidence of ability to use punctuation (no commas or full stops correctly used)
	1	Some correct use but some problems with punctuation
	2	Correct use of punctuation

Exhibit 26 Target percentages for Class 6 writing text type categories, MTEG 2013

Text type	Target percentage of tasks
Narrative	15–25
Descriptive	25–35
Persuasive	15–25
Instructional	5–15
Transactional	5–15
Label	5–15

Writing processes, coding and scoring

The scoring of writing literacy tasks is based on criteria that reflect the writing processes. Some of these criteria are specific to a particular text type and others are more generic. For example, *developing characters* belongs to the *generate ideas* process and is applicable only to narratives, but a criterion such as *precision*, developed for the process *use vocabulary*, is applicable to all text types.

The criteria are operationalised in the form of rating scales with a number of described categories (codes). The rating scales vary in length: some are dichotomous (with only two codes, code 0 and code 1), and some have up to five coding categories (code 0, code 1, code 2, code 3 and code 4). The number of codes for a criterion depends on the number of defined and distinguishable categories into which students' responses can be divided.

Some writing tasks, especially those designed to measure the proficiency of emerging writers, will be very constrained. Because of their brevity, they will be more likely to be assessed dichotomously, as right or wrong. Examples of such tasks include writing a single word to label an image, spelling a word correctly or manipulating sentence structures.

A major challenge in measuring writing literacy in a bilingual or multilingual assessment is achieving equivalence across languages. In order to meet this challenge, the MTEG writing literacy assessment model treats some aspects of writing as common across languages, while others may be treated as applicable only to one language, or a group of languages. This approach will yield some comparisons between writing performance in different languages, while recognising the particular characteristics of individual languages.

Assessment of the processes *generate ideas*, *control structure and organisation*, *manage coherence* and *use vocabulary* may be applied across all languages, using common coding criteria. The process *control syntax and grammar* may be assessed using criteria that are customised in accordance with the features of the individual languages. The sixth process, *other, language-specific features*, may also be assessed using language-specific criteria. Most if not all tasks are assessed on multiple criteria, including some that are comparable across languages (for example, criteria focusing on the vocabulary required to express particular concepts) and some that are language-specific (for example, criteria focusing on linguistic rules associated with spelling or syntax). For Pashto and Dari, for example, common coding criteria have been used for *control syntax and grammar*, whereas the *spelling* criteria are treated as separate for the two languages.

Some of the criteria used to code writing literacy are used for both Class 6 and Class 9, while others are used for only one of the two classes. This is technically equivalent to the practice in a reading or mathematics assessment administered to multiple classes and calibrated on a single scale: some items are administered to two different classes, and others to only one class or the other. Such an approach allows the measurement of different features of writing literacy at different class levels, while at the same time allowing a comparison of ability across classes.

Exhibit 27 shows a model for how the assessment is designed to ensure coverage of all writing

Exhibit 27 Model for writing assessment across languages, task types and class levels

	Process					
	Generate ideas	**Control structure**	**Manage coherence**	**Use vocabulary**	**Control syntax**	**Other language-specific features**
Criteria typically:	*apply across languages; vary by task type*	*apply across languages; vary by task type*	*apply across languages; apply across task types*	*apply across languages; apply across task types*	*may vary by language; apply across task types*	*may vary by language; apply across task types*
Class 9	C	F	I	L	O	R
Class 6	A B	D E	G H	J K	M N	P Q

Exhibit 28 Target percentages for Class 6 writing process categories, MTEG 2013

Process	Target percentage of tasks	Criteria are typically:
Generating ideas	20–30	Comparable across languages
Controlling text structure and organisation	10–20	Comparable across languages
Managing coherence	10–20	Comparable across languages
Using vocabulary	10–20	Comparable across languages
Controlling syntax and grammar	15–25	Language-specific – comparable only within a language, or across closely related languages
Other, language-specific features (e.g. spelling, character formation, punctuation)	5–15	Language-specific – comparable only within a language or across closely related languages

processes. It also shows how the design will be extended to Class 9. Each column in the exhibit represents one of the six writing process variables. Each rectangle (A to R) represents one criterion against which the writing is assessed. Each writing process is assessed through a number of criteria across a range of tasks; the categories within each criterion can be thought of as 'items' that will ultimately be calibrated on a writing literacy scale.

Rectangles labelled A, D, G, J and M represent the criteria measuring content (ideas), organisation (structure), coherence, vocabulary and syntax respectively that would be scored only at Class 6. These criteria would be associated with tasks including those that allow the assessment of basic levels of writing.

Rectangles B, E, H, K and N are the criteria for each writing process that would be scored at both Class 6 and Class 9.

Rectangles C, F, I, L and O are the criteria for each writing process that would be scored at Class 9

only. These criteria would be associated with tasks including those that allow assessment of more sophisticated aspects of writing.

Items A, B, C, D, E and F would typically be common across languages but vary for different task types.

Items G, H, I, J, K and L would be common across languages and possibly also task types.

Items M, N and O would vary for different languages but would tend to be common across task types.

Items P, Q and R would be associated with other, language-specific features of writing.

Chapter 5 Contextual questionnaires

The importance of contextual questionnaires

As mentioned in Chapter 1, the first goal of the MTEG program in Afghanistan is to provide policymakers with relevant, sound and comparable data on contextual and learning outcomes that can directly inform local education policy development.

While it is essential to develop high-quality measures of student learning outcomes in mathematical, reading and writing literacy, it is equally important to produce measures of the context in which student learning occurs. Contextual information collected through student and school questionnaires can provide valuable data on factors associated with the performance of students in the mathematical, reading and writing literacy assessments.

The conceptual framework

Categorising contextual factors

A range of contextual factors can contribute to students' learning outcomes. For example, an individual's beliefs and attitudes towards school and learning, as well as activities and resources at home, in the classroom, at school and in the wider community can all influence learning outcomes. Contextual factors of student learning are frequently categorised as inputs or antecedents, processes and outcomes (for example, OECD, 2016; Schulz, Fraillon, Ainley, Losito, & Kerr, 2008). These contextual factors are observed at various levels – country, community, school, classroom/teacher, home and individual student – with the higher levels (such as community) frequently influencing the lower levels (such as student).

Inputs/antecedents

Inputs or antecedents are factors that affect how student learning takes place. These factors are often less easily influenced by other contextual factors and include resources, demographics or structural information. At the student level, inputs include demographic information such as gender, age or language spoken, and structural information about the home, such as parental education. At the school level, inputs include resources such as the number of toilets, internet access, and the presence of a classroom or school library. At the school level there are also structural factors, for example, student enrolments, diversity of the student body, the number of students per teacher, teacher qualifications, the number of male and female teachers, and whether the school is located in a rural or urban community.

The distinction between structure/demographics and resources within the input category has been made in earlier frameworks for variables influencing student learning in schools (Keeves, 1972; Peaker, 1967). This distinction is considered particularly relevant to the situation of emerging education systems such as the one in Afghanistan, which are frequently confronted with structural circumstances such as population growth or increased educational participation of age cohorts which, in turn, shape resourcing decisions. For example, a large and rapid increase in the number of students who participate in schooling puts heavy demands on the resources of an education system in terms of the number of schools, or classes within schools, their equipment, and appropriately qualified teachers that are needed.

Processes

Processes are factors related to student learning, including values, practices and behaviours, and these processes are constrained by antecedents. Practices or behaviours refer to the activities undertaken by the various actors in an education system, namely students, parents, teaching and

support staff, school leaders, school communities, education departments and ministries, as well as funding bodies such as governments, private businesses, and donors. At the student level, practices and behaviours include reading habits, homework effort and time spent on tasks. Teacher-level processes refer to factors such as instructional strategies, time allocation to different tasks and the amount and type of homework set. Practices and behaviours at the school level include decision-making processes, frequency and content of staff meetings, and the evaluation of teaching staff.

The distinction within the processes category between practices and behaviours and values has been made in earlier frameworks (Keeves, 1972; Peaker, 1967). Values are largely socially based, such as the education of girls or the aim of having a literate workforce, and more influenced by education policymakers.

Outcomes

While increasing participation in education systems is a key goal in systems that are in a state of rapid development, increasing educational quality is also a critical outcome. MTEG measures student learning by way of cognitive tests in reading, mathematics and writing. Other important outcomes to measure are sometimes called non-cognitive (OECD, 2016) or affective-behavioural outcomes (Schulz et al., 2008). They are frequently measured in terms of students' attitudes towards school and learning, as well as their interests in various subject matters.

The framework

Exhibit 29 illustrates the conceptual framework for variables which affect student learning at school that has guided the development of the MTEG student and school questionnaires. In Exhibit 29, the far left-hand column of the grid lists the different levels to which information collected corresponds, while the different dimensions of each level are specified across the grid.

The grid allows the categorisation of any variable depending on the level and the dimension with which it is associated. Thus, for example, the education budget is a country-level variable that is associated with the resource dimension, while the observation of teachers is a process variable at the school level. In general, factors further down in the table are considered to be frequently dependent on factors further up in the table. Similarly, factors further to the right in the table are considered to often be dependent on factors further to the left in the table. As can be seen, outcomes in this framework are considered at the individual student level, although such measures – if aggregated correctly – may be used as outcome measures at higher levels, for example, at the school or country levels.

Content of the MTEG school and student questionnaires

Instruments may be designed to collect data at any level specified in Exhibit 29. The student performance measures used within this framework for MTEG are designed as literacy measures, with a focus on what students can do at the assessed class levels. MTEG also collects contextual data from students and principals. These instruments allow information to be collected about other levels, such as the community and home background. The instruments used in MTEG are:

- the school questionnaire – completed by principals at schools participating in the 2013 assessment of Class 6 students
- the student questionnaire – completed by students participating in the 2013 Class 6 assessment
- the student assessment – reading, mathematics and writing assessments completed by Class 6 students.

Exhibit 30 maps the variables collected through the different MTEG instruments. Information at the country level was not collected through MTEG, however, information from other sources may be used to gather information on the antecedents and

Exhibit 29 Conceptual framework for MTEG Afghanistan

Dimensions					
	Antecedents/Inputs		Processes		Outcomes
Level	*Structure/ demographics*	*Resources*	*Values*	*Practices/ behaviours*	
Country	Size of population	Education budget	Literate workforce	Secondary school entry admission	
Community	Rural/urban	Public library	Literate workforce	Support for girls attending school	
School	Proportion of female teachers	School library	Academic excellence	Teacher observation	
Classroom/ teacher	Teacher gender	Subject textbooks	Expectation of student performance	Monitoring of attendance	
Home	Language spoken at home	Educational resources at home	Value academic achievement	Cooperation with schools	
Student	Student gender	Pen and paper	Motivation towards academic achievement	Engagement with reading	Learning Performance Attitudes Interests

Adapted from Lietz, 1996.

processes at the country level, such as information about the number of schools in different provinces, and the curriculum priorities. While some information was collected at the community level, the focus was on the school, classroom/teacher, home and individual student levels. MTEG did not survey teachers, therefore, data on the classroom/teacher level is from school principals and students.

The data gathered from the contextual questionnaires provides information to education policymakers, donors and education practitioners on variables that may be associated with student achievement. Antecedents and processes at different levels can contribute to learning outcomes for students. For example, a student's home background, their beliefs and their attitudes towards school and learning may be linked to performance.

Contextual information and policymaking

In evidence-based policymaking (Sackett, Rosenberg, Gray, Haynes, & Richardson, 1996; Solesbury, 2001; Sutcliffe & Court, 2005), analyses of relationships between contextual information and student learning outcome data enable policymakers to understand what works and in which contexts. There are limits to how much MTEG data can reveal about the *causes* of specific learning outcomes, however, information can be gathered on the effects of factors related to growth of the same cohort over time, e.g. from Class 3 to Class 6.

Defining policy needs

While policy analysis can be informed by the data collected in MTEG, sound policy analysis can also assist with decisions about what contextual data is collected through the questionnaires.

Education policies may be concerned with content, instruction, resources and assessment. At the system level, these policies target educational issues such as curriculum development, the allocation of resources in education, the use of learning assessments, and the development of achievement standards, as well as standards of teacher qualifications and teaching and learning practices, among others.

Exhibit 30 Mapping of variables to the contextual framework

Dimensions			
Level	**Inputs/antecedents**	**Processes**	**Outcomes**
Country			StA and StQ: Assessment results, student attitudes and interests
Community	ScQ: Rural/urban, weather and security issues affecting school, distance from school to community facilities	ScQ: Support for girls attending school StQ: Support for attending school	
School	ScQ: School principal background, school characteristics (e.g., type of school, language of instruction, enrolments by gender, school shifts, length of lessons), school facilities (e.g., toilets, food provided for students)	ScQ: School inspections, school improvement plans	
Classroom/ teacher	ScQ: Classroom characteristics and resources (e.g., textbooks), teacher background (e.g., gender, education level)	ScQ: Monitoring of teacher attendance StQ: Frequency of receiving homework, teacher behaviours and pedagogy, support for attending school	
Home	StQ: Language spoken at home, home environment, time and method of travel to school, home resources and materials, parental level of education	StQ: Assistance with homework, support for attending school	
Student	StQ: Student gender, age, possessions, meals per day and food received from school, prior education, age started school	ScQ: Monitoring student attendance StQ: Amount of time spent studying, grade repetition, highest expected level of education, attitudes towards school, attitudes towards reading and maths, materials read	

Note: ScQ = school questionnaire, StQ = student questionnaire, StA = student assessment

A model commonly used in policymaking is that of a policy cycle with separate stages. A number of variations of the policy cycle model have been proposed, generally involving six to eight stages (Bridgman & Davis, 2004; Haddad, 1995; Young & Quinn, 2002). Exhibit 31 outlines a simplified policymaking cycle, developed by Sutcliffe and Court (2005).

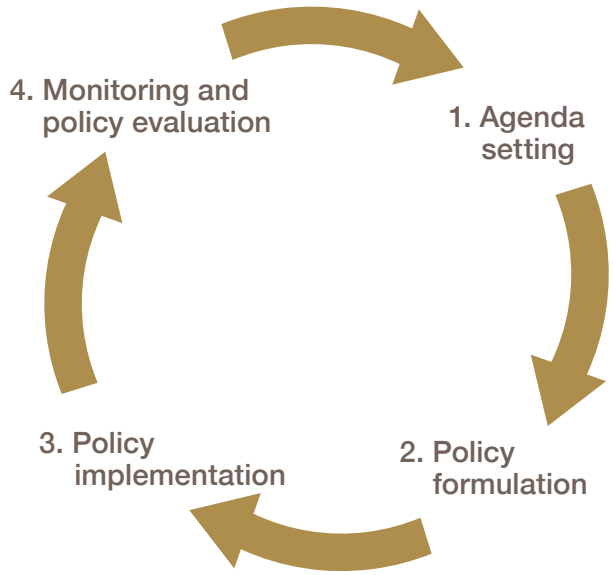
The four stages in the policy cycle as shown in Exhibit 31 are:

- agenda setting: the awareness of and priority given to an issue or problem
- policy formulation: the ways (analytical and political) in which options and strategies are constructed
- policy implementation: the forms and nature of policy administration and activities on the ground
- monitoring and policy evaluation: the nature of monitoring and evaluation of policy need, design, implementation and impact.

It is likely that the MTEG data can be used at different stages of the policy cycle, and this will be part of an ongoing dialogue with various stakeholders including the Afghanistan government and donors. Stakeholder dialogue for effective policy analysis will occur through interviews, face-to-face workshops, webinars, telephone conferences and the exchange and analysis of policy documents. These stakeholder discussions could include investigating:

- current and upcoming priorities, for instance through an analysis of Afghanistan's National Education Strategic Plan (Afghanistan Ministry of Education, 2015) and discussion about upcoming strategy documents.

Exhibit 31 Simplified model of the policy cycle



- policy process, by clarifying the policy processes in Afghanistan, identifying specific upcoming opportunities to inform policy debate, and identifying who informs, influences and approves policy decisions, including actors within and outside the ministry
- communication, dissemination and expected outputs: How do decision-makers currently access and engage with research? How can the MTEG assessment outcomes be most effectively communicated and to whom?

Policy priority example: girls' education

An example of a policy priority that informed the development of the contextual questionnaires is the focus on girls' education. Increasing girls' participation in education has been a priority in the National Education Strategic Plan for Afghanistan (2010–2014) (Afghanistan Ministry of Education, 2015), and a communication strategy to raise general awareness and positively influence community attitudes on this topic was developed as part of the Afghanistan Girls' Education Initiative (AGEI). Previous analyses indicate that unfavourable school-level factors have a greater negative effect on girls enrolling at school than they do on boys. That is, girls' enrolment suffers more when it is more difficult to travel to school, when there are fewer toilets and when there is a lack of free meals (Mingat, Tan & Sosale, 2003).

Data collected from the school and student questionnaires can provide descriptive information to address questions such as:

- Are more girls enrolled in schools with a lower number of students per toilet?
- Are fewer girls enrolled in schools where students have longer travel times to get to school?
- Are more girls enrolled in Class 6 in schools with a higher proportion of female teachers?

When data obtained from the cognitive skills test and the questionnaires are combined, questions such as the following can be addressed:

- Is there a relationship between attitudes towards schooling and performance, and is this relationship the same for girls and boys?
- Are girls performing at a higher level in schools where they receive greater support for attending schools?
- If gender differences in performance can be observed, are these greater in schools with a smaller proportion of female teachers?

Conclusion

In summary, this chapter has put forward a conceptual framework that allows factors relating to student outcomes to be categorised as input, process and outcome factors. An approach to systematic policy analysis to accompany the assessment program has also been outlined. Through this two-pronged approach of collecting high-quality cognitive data on student learning outcomes and relating these to the information obtained from the context questionnaires and policy analysis, MTEG will provide rich and relevant information for evidence-based policymaking.

Appendix A

Cluster and booklet design

The cognitive instruments for each administration of the survey comprise a total of 90 minutes of mathematical literacy material, 90 minutes of reading literacy material and 90 minutes of writing literacy material. This amount of assessment material allows good coverage of each domain. While there is a total of 270 minutes of cognitive instrumentation, each student only completes 30 minutes of assessment for each domain, as well as the student questionnaire (totalling 120 minutes).

The material is arranged in six clusters of tasks per domain, with each cluster representing 15 minutes of testing time. The item clusters are placed in test booklets according to a rotated test design, in which each booklet contains either two clusters of mathematical literacy material and two clusters of reading literacy material, or two clusters of writing literacy and a student questionnaire. Each student is administered one booklet of mathematics and reading, and one booklet of writing and questionnaire: a total of two one-hour sessions for each student.

Exhibit 32 shows a possible rotated booklet design for the first one-hour session of the assessment, comprising mathematical literacy and reading literacy. M1 to M6 represent the six 15-minute mathematical literacy clusters and R1 to R6 represent the six 15-minute reading clusters.

Exhibit 32 Rotated booklet design for mathematical literacy and reading literacy

Book 1	Book 2	Book 3	Book 4	Book 5	Book 6	Book 7	Book 8	Book 9	Book 10	Book 11	Book 12
M1	M2	M3	M4	M5	M6	R1	R2	R3	R4	R5	R6
M2	M3	M4	M5	M6	M1	R5	R6	R1	R2	R3	R4
R1	R2	R3	R4	R5	R6	M4	M5	M6	M1	M2	M3
R2	R3	R4	R5	R6	R1	M3	M4	M5	M6	M1	M2

Similarly, Exhibit 33 shows a possible rotated booklet design for the second one-hour session, with W1 to W6 representing the six writing clusters, and StQ representing the 30-minute student background questionnaire.

Exhibit 33 Rotated booklet design for writing literacy and the student background questionnaire

Book 13	Book 14	Book 15	Book 16	Book 17	Book 18
W1	W2	W3	W4	W5	W6
W2	W3	W4	W5	W6	W1
StQ	StQ	StQ	StQ	StQ	StQ

References

Afghanistan Ministry of Education. (2015). *National Education Strategic Plan for Afghanistan 2015–2020. Draft (NESP III)*. Kabul: Ministry of Education.

Best, M., Knight, P., Lietz, P., Lockwood, C., Nugroho, D., & Tobin, M. (2012). *The impact of national and international assessment programmes on educational policy, particularly policies regarding resource allocation and teaching and learning practices in developing countries. Draft report*. London: Evidence for Policy and Practice Information and Co-ordinating Centre (EPPI-Centre), Social Science Research Unit, Institute of Education, University of London.

Bridgman, P., & Davis, G. (2004). *The Australian policy handbook* (3rd ed.). Sydney: Allen & Unwin.

Ercikan, K., Arim, R., Oliveri, M., & Sandilands, D. (2008). *Evaluation of dimensions of the work of the Southern and Eastern Africa Consortium for Monitoring Educational Quality (SACMEQ) and of its programme of cooperation with the International Institute for Educational Planning (IIEP)*. Paris: UNESCO.

Gray, W. S. (1960). The major aspects of reading. In H. Robinson (Ed.), *Sequential development of reading abilities, Supplementary Educational Monographs No. 90*, (pp. 8–24). Chicago: University of Chicago Press.

Haddad, W. D. (1995). *Education policy-planning process: An applied framework.* Paris: UNESCO – International Institute for Educational Planning.

Keeves, J. P. (1972). *Educational environment and student achievement.* Stockholm: Almqvist and Wiksell.

Kirsch, I., & Mosenthal, P. B. (1990). Exploring document literacy: Variables underlying the performance of young adults. *Reading Research Quarterly, 25*(1), 5–30.

Lietz, P. (1996). *Changes in reading comprehension across cultures and over time*. Münster/New York: Waxmann.

Mingat, A., Tan, J. & Sosale, S. (2003). *Tools for education policy analysis*. Washington, D.C.: World Bank.

Monseur, C., & Lafontaine, D. (2009). Gender gap in comparative studies of reading comprehension: To what extent do the test characteristics make a difference? *European Educational Research Journal*, *8*(1), 69–79.

Mullis, I. V. S., Martin, M. O., Kennedy, A. M., Trong, K. L., & Sainsbury, M. (2009). *PIRLS 2011 Assessment Framework*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College.

Niss, M. (2003). Mathematical competencies and the learning of *m*athematics: The Danish KOM Project. In A. Gagatsis & S. Papastavridis (Eds.), *3rd Mediterranean Conference on Mathematics Education* (pp. 115–124). Athens: The Hellenic Mathematical Society and Cyprus Mathematical Society.

Niss, M., & Højgaard, T. (Eds.). (2011). *Competencies and mathematical learning: Ideas and inspiration for the development of mathematics teaching and learning in Denmark*. (English ed.). Roskilde, Denmark: Roskilde University, Dept of Science, Systems and Models, IMFUFA.

OECD. (2009). PIAAC literacy: A conceptual framework. *OECD Education Working Paper No. 34.* Paris: OECD.

OECD. (2010). *PISA 2009 assessment framework: Key competencies in reading, mathematics and science*. Paris: OECD.

OECD. (2013). *PISA 2012 assessment and analytical framework – mathematics, reading, science, problem solving and financial literacy.* Paris: OECD.

OECD. (2016). PISA 2015 context questionnaires framework. *PISA 2015 assessment and analytical framework* (pp. 101–127). Paris: OECD.

Olson, D. R. (1994). *The world on paper.* Cambridge: Cambridge University Press.

Peaker, G. F. (1967). The regression analyses of the National Survey. In Central Advisory Council for Education (England) (Ed.), *Children and their primary schools* (Vol. 2, pp. 179–221). London: Her Majesty;s Staionery Office.

Routitsky, A., & Turner, R. (2003). *Item format types and their influences on cross-national comparisons of student performance.* Paper presented at the annual meeting of the American Educational Research Association (AERA), Chicago.

Sackett, D. L., Rosenberg, W. M., Gray, J. A., Haynes, R. B., & Richardson, W. S. (1996). Evidence based medicine: What it is and what it isn't [Editorial]. *British Medical Journal*, 312.

Schulz, W., Fraillon, J., Ainley, J., Losito, B., & Kerr, D. (2008). *International Civic and Citizenship Education Study. Assessment framework.* Amsterdam: International Association for the Evaluation of Educational Achievement (IEA).

Share, D. (2008). On the Anglocentricities of current reading research and practice: The perils of overreliance on an "outlier" orthography. *Psychological Bulletin, 134*(4), 584–615.

Smith, M. C., Mikulecky, L., Kibby, M. W., & Dreher, M. J. (2000). What will be the demands of literacy in the workplace in the next millennium? *Reading Research Quarterly, 35*(3), 378–383.

Solesbury, W. (2001). *Evidence based policy: Whence it came and where it's going. (Working Paper 1.)*. Swindon: ESRC UK Centre for Evidence Based Policy and Practice.

Sutcliffe, S., & Court, J. (2005). *Evidence-based policymaking: What is it? How does it work? What relevance for developing countries?* London: Overseas Development Institute.

Turner, R., Dossey, J., Blum, W., & Niss, M. (2013). Using mathematical competencies to predict item difficulty in PISA. In M. Prenzel, M. Kobarg, K. Schöps, & S. Rönnebeck (Eds.), *Research on PISA: Research outcomes of the PISA Research Conference 2009* (pp. 23–27). New York: Springer.

Wagner, D. A. (2011). Smaller, quicker, cheaper. *Improving Learning Assessments for Developing Countries. Education for All Fast Track Initiative Secretariat.* Paris & Washington, D.C.: UNESCO & Education For All – Fast Track Initiative.

Wirt, J., Snyder, T., Sable, J., Choy, S. P., Bae, Y., Stennett, J., … Perie, M. (1998). *The condition of education 1998* (pp. 70). Washington, D.C.: U.S. Department of Education, National Center for Education Statistics.

Young, E., & Quinn, L. (2002). *Writing effective public policy papers: A guide for policy advisers in Central and Eastern Europe.* Budapest: Local Government and Public Service Reform Initiative, Open Society Institute.

Australian Council *for* Educational Research