PISA

PISA 2018 Assessment and Analytical Framework

Programme for International Student Assessment

OECD

PISA 2018 Assessment and Analytical Framework

BETTER POLICIES FOR BETTER LIVES

This work is published under the responsibility of the Secretary-General of the OECD. The opinions expressed and arguments employed herein do not necessarily reflect the official views of OECD member countries.

This document, as well as any data and any map included herein, are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.

Please cite this publication as:
OECD (2019), *PISA 2018 Assessment and Analytical Framework*, PISA, OECD Publishing, Paris, *https://doi.org/10.1787/b25efab8-en*.

ISBN 978-92-64-94031-4 (print)
ISBN 978-92-64-47759-9 (pdf)

PISA
ISSN 1990-8539 (print)
ISSN 1996-3777 (online)

The statistical data for Israel are supplied by and under the responsibility of the relevant Israeli authorities. The use of such data by the OECD is without prejudice to the status of the Golan Heights, East Jerusalem and Israeli settlements in the West Bank under the terms of international law.

Photo credits: Cover © LuminaStock/iStock; © Dean Mitchell/iStock; © bo1982/iStock; © karandaev/iStock; © IA98/Shutterstock; © Tupungato/Shutterstock.

Corrigenda to OECD publications may be found on line at: *www.oecd.org/about/publishing/corrigenda.htm*.

Foreword

The OECD Programme for International Student Assessment (PISA) assesses the extent to which 15 year old students near the end of their compulsory education have acquired the knowledge and skills that are essential for full participation in modern societies. The assessment does not just ascertain whether students can reproduce knowledge; it also examines how well students can extrapolate from what they have learned and can apply that knowledge in unfamiliar settings, both in and outside of school. This approach reflects the fact that modern economies reward individuals not for what they know, but for what they can do with what they know.

The triennial assessment, launched in 1997, focuses on the core school subjects of reading, mathematics and science. Students' proficiency in an innovative domain is also assessed; in 2018, this domain was global competence. This publication presents the theory underlying the PISA 2018 assessment – the seventh since the programme's inception. It includes frameworks for assessing the three core subjects of reading, mathematics and science, the framework for the third assessment of students' financial literacy, and the framework for assessing the innovative domain, global competence. These chapters outline the content knowledge that students need to acquire in each domain, the processes that students need to be able to perform, and the contexts in which this knowledge and these skills are applied. The publication also discusses how each domain is assessed. It concludes with the frameworks for the various questionnaires distributed to students, school principals, parents and teachers, and the framework for the new well-being questionnaire distributed to students.

In PISA 2018, reading was the major domain of assessment, as it was in 2000 and 2009. The three assessment subscales used in 2000 and 2009 were renamed "locating information", "understanding" and "evaluating and reflecting" for 2018. Two new subscales were also used to describe students' literacy with single-source and multiple-source texts. The reading scale was also extended by adding Level 1c, which better describes the proficiency of the lowest-performing students.

PISA is the product of a collaborative effort between OECD and the governments of both OECD countries and its partner countries/economies. The assessments are developed co-operatively, agreed by participating countries/economies, and implemented by national organisations. The co-operation of students, teachers and principals in participating schools has been crucial to the success of PISA during all stages of development and implementation.

The reading framework was developed by the reading expert group with the guidance of John de Jong and Peter Foltz from Pearson. The reading expert group was chaired by Jean-François Rouet (University of Poitiers, France). Other experts who contributed to the reading framework are Paul van den Broek (Universiteit Leiden, the Netherlands), Kevin Chung (University of Hong Kong, China), Sascha Schroeder (Max Planck Institute for Human Development, Berlin, Germany), Sari Sulkunen (University of Jyväskylä, Finland; also served as the liaison to the PISA global competence expert group), and Dominique

Lafontaine (Université de Liège, Belgium; also served as the liaison to the PISA questionnaire expert group).

The global competence framework was developed by Mario Piacentini of the OECD Secretariat with Martyn Barrett (University of Surrey, Guildford, UK), Veronica Boix Mansilla (Harvard University and Project Zero, Cambridge, USA), Darla Deardorff (Duke University, Durham, USA) and Hye Won Lee (Korea Institute for Curriculum and Evaluation, Jincheon, Korea), with additional help from Rose Bolognini and Natalie Foster (OECD Secretariat), Natasha Robinson (University of Oxford, UK) and Mattia Baiutti (Fondazione Intercultura, Colle di Val d'Elsa, Italy and the University of Udine, Italy). This framework built on earlier work from experts who led the first part of the development of the global competence assessment: Darla Deardorff (Duke University, Durham, USA), David Kerr (University of Reading, UK and YoungCitizens, London, UK), Peter Franklin (HTWG Konstanz University of Applied Sciences, Germany), Sarah Howie (University of Pretoria, South Africa), Wing On Lee (Open University of Hong Kong, China), Jasmine B Y Sim (National Institute of Education, Singapore), and Sari Sulkunen (University of Jyväskylä, Finland).

The framework for the PISA 2018 questionnaires was developed by the questionnaire expert group with the guidance of John de Jong and Christine Rozunick from Pearson. The questionnaire expert group was chaired by Fons van de Vijver (Tilburg University, the Netherlands; the North-West University, Potchefstroom, South Africa; and the University of Queensland, Brisbane, Australia). Other experts who contributed to the development of the questionnaire framework are Dominique Lafontaine (Université de Liège, Belgium), Sarah Howie (University of Pretoria, South Africa), Andrew Elliot (University of Rochester, USA), Therese Hopfenbeck (University of Oxford, UK) and David Kaplan (University of Wisconsin-Madison, USA).

The framework for the well-being questionnaire was developed by Jonas Bertling (ETS). All other frameworks were based on versions developed for previous PISA cycles.

Pearson facilitated the development of the reading and questionnaire frameworks. The Educational Testing Service (ETS) was responsible for managing and overseeing this survey, developing the instruments, scaling, analysis, and developing the electronic platform. Other partners or subcontractors involved with ETS include the Department of Experimental and Theoretical Pedagogy at the Université de Liège (aSPe) in Belgium and the Educational Measurement and Research Centre (EMACS) of the University of Luxembourg in Luxembourg. Westat assumed responsibility for survey operations and sampling with the subcontractor, the Australian Council for Educational Research (ACER). cApStAn Linguistic Quality Control assumed responsibility for ensuring the linguistic equivalence of all language versions.

The frameworks were reviewed by expert panels in each of the participating countries. The chapters were drafted by the respective expert groups under the direction of their chairs. The members of the expert groups are listed in Annex B.

The publication was prepared by the OECD Secretariat. Jeffrey Mo coordinated the preparation of the framework, with contributions from Marilyn Achiron, Hélène Guillou and Miyako Ikeda. Rebecca Tessier oversaw the production of this revised edition, and Hanna Varkki provided editorial support.

The report is published under the responsibility of the Secretary General of the OECD.

Table of contents

Foreword **3**

Table of contents **5**

1. What is PISA? **11**

What makes PISA unique 13
The PISA 2018 test 13
An overview of what is assessed in each domain 14
The evolution of reporting student performance in PISA 16
The context questionnaires 17
A collaborative project 18

2. PISA 2018 Reading Framework **21**

Introduction 22
Defining reading literacy 27
Organising the domain 30
Assessing reading literacy 41
Reporting proficiency in reading 53
Notes 56
References 57
Annex 2.A. Main changes in the reading framework, 2000-2015 67
Annex 2.B. Sample tasks 68
Task 1: Sample reading ease and efficiency task 68
Tasks 2-4: Sample scenario with three embedded tasks 68

3. PISA 2018 Mathematics Framework **73**

Introduction 74
Defining mathematical literacy 75
Organising the domain of mathematics 76
Assessing mathematical literacy 89
Notes 93
References 94

4. PISA 2018 Science Framework **97**

Introduction: Scientific literacy and why it matters 98
Scientific literacy: Towards a definition 98
Organisation of the domain 102
Assessment of the Domain 108
References 116

5. PISA 2018 Financial Literacy Framework **119**

Introduction 120
Defining financial literacy 128
Organising the domain 130
Assessing financial literacy 152
The impact of knowledge and skills in other domains on financial literacy 155
Reporting financial literacy 157
Notes 159
References 160

6. PISA 2018 Global Competence Framework 165

Introduction: The importance of an international global competence assessment 166
The assessment of global competence in PISA 180
Conclusions 198
Notes 200
References 201
Annex 6.A. Illustrative examples of scenarios for the cognitive assessment of global understanding 209
Are global temperatures rising? 209
A talented player 210
A song in Quechua 211
Annex 6.B. Description of possible topics for the scenarios of the cognitive test 212
Culture and intercultural relations 212
Socio-economic development and interdependence 213
Environmental sustainability 214
Institutions, conflicts and human rights 215

7. PISA 2018 Questionnaire Framework 217

Introduction 218
Defining the questionnaire core in PISA 2018 218
Coverage of policy issues in PISA 2018 226
Notes 243
References 244

8. PISA 2018 Well-being Framework 257

Executive summary 258
Introduction 259
Well-being as a multi-dimensional construct 262
Addressing measurement challenges 264
Suggested quality of life indicators 267
Notes 280
References 281

Annex A. PISA 2018 Background questionnaires 299

PISA 2018 Background questionnaires 300

Annex B. PISA 2018 Expert Groups 301

Reading expert group (REG) 302
Global Competence expert group (GEG) 302
Questionnaire expert group (QEG) 304

Tables

Table 2.1. Mapping of the 2018 process typology to 2018 reporting scales and to 2009-2015 cognitive aspects ... 37
Table 2.2. Approximate distribution of tasks by targeted process and text source ... 42
Table 2.3. Item difficulty as a function of task and source dimensions ... 44
Table 2.4. An overview of reading proficiency levels adapted from the descriptions in PISA 2015 ... 55
Table 3.1. Desired distribution of mathematics items, by process category ... 80
Table 3.2. Desired distribution of mathematics items, by content category ... 86
Table 3.3. Desired distribution of mathematics items, by context ... 89
Table 4.1. Aspects of the scientific literacy assessment framework for PISA 2015/2018 ... 102
Table 4.2. Contexts for the PISA 2018 scientific literacy assessment ... 103
Table 4.3. PISA 2018 scientific competencies: Explaining phenomena scientifically ... 104
Table 4.4. PISA 2018 scientific competencies: Evaluating and designing scientific enquiry ... 104
Table 4.5. PISA 2018 scientific competencies: Interpreting data and evidence scientifically ... 105
Table 4.6. Content knowledge in the PISA 2018 science assessment ... 106
Table 4.7. Procedural knowledge in the PISA 2018 science assessment ... 107
Table 4.8. Epistemic knowledge in the PISA 2018 science assessment ... 108
Table 4.9. Target distribution of score points by knowledge type and content knowledge category .. 114
Table 4.10. Target distribution of score points for scientific competencies ... 114
Table 4.11. Summary description of the seven levels of proficiency in science in PISA 2015 ... 115
Table 5.1. Approximate distribution of score points in financial literacy, by content ... 155
Table 5.2. Approximate distribution of score points in financial literacy, by process ... 155
Table 5.3. Approximate distribution of score points in financial literacy, by contexts ... 155
Table 6.1. Typologies of cognitive processes by level in the PISA 2018 test of global competence . 186
Table 6.2. Content domains and subdomains of the scenarios ... 190
Table 6.3. Dimensions and levels of complexity of the scenarios ... 191
Table 8.1. Dimensions of well-being in other frameworks ... 263

Annex Table 2.A.1. Main changes in the reading framework, 2000-2015 ... 67

Figures

Figure 2.1. Factors that contribute to reading literacy ... 31
Figure 2.2. PISA 2018 Reading framework processes ... 33
Figure 3.1. A model of mathematical literacy in practice ... 77
Figure 3.2. Relationship between mathematical processes (top row) and fundamental mathematical capabilities (left-most column) ... 82
Figure 3.3. Summary description of the six levels of mathematics proficiency in PISA 2018 ... 92
Figure 4.1. Inter-relations between the three aspects ... 103
Figure 4.2. PISA 2015/2018 framework for cognitive demand ... 110
Figure 4.3. A tool for constructing and analysing assessment units and items ... 112
Figure 5.1. Illustrative PISA Financial literacy item 1 – BANK STATEMENT ... 133
Figure 5.2. Illustrative PISA Financial literacy item 2 – MUSIC SYSTEM ... 134
Figure 5.3. Illustrative PISA Financial literacy item 3 – ZCYCLE ... 136
Figure 5.4. Illustrative PISA Financial literacy item 3 – ZCYCLE – Question ... 137
Figure 5.5. Illustrative PISA Financial literacy item 4 – MOTORBIKE INSURANCE ... 139
Figure 5.6. Illustrative PISA Financial literacy item 5 – MOBILE PHONE CONTRACT ... 141
Figure 5.7. Illustrative PISA Financial literacy item 6 – PAY SLIP ... 142

Figure 5.8. Illustrative PISA Financial literacy item 7 – PHONE PLANS 144
Figure 5.9. Illustrative PISA Financial literacy item 8 – ONLINE SHOPPING 146
Figure 5.10. Illustrative PISA Financial literacy item 9 – RINGTONES 147
Figure 5.11. Relationship between the content of financial literacy and mathematical literacy in PISA 156
Figure 6.1. The dimensions of global competence 169
Figure 6.2. The PISA approach to assessing global competence 180
Figure 6.3. Elements of a typical PISA 2018 global competence test unit 181
Figure 6.4. The relationship between the cognitive test of global understanding and the dimensions of global competence.. 185
Figure 7.1. PISA 2018 Questionnaire modules 220
Figure 8.1. Framework overview 258

Annex Figure 2.B.1. Task 1. Sample reading ease and efficiency task.. 68
Annex Figure 2.B.2. Task 2. Scanning and locating (single text).. 69
Annex Figure 2.B.3. Task 3. Multiple text inference 70
Annex Figure 2.B.4. Task 4. Evaluating and reflecting 71
Annex Figure 6.A.1. Global temperature 209
Annex Figure 6.A.2. Global temperature 210

Boxes

Box 1.1. Key features of PISA 2018 12
Box 1.2. Definitions of the domains... 14
Box 2.1. The definition of reading literacy in earlier PISA cycles... 27
Box 2.2. The 2018 definition of reading literacy 28
Box 2.3. Characteristics used to classify texts in the PISA 2009, 2012 and 2015 reading frameworks 38
Box 2.4. Text availability and its impact on comprehension 44
Box 2.5. Categorisation of situations 45
Box 2.6. Classification of text types... 46
Box 2.7. The status of writing skills in the PISA 2018 reading literacy assessment............................ 49
Box 3.1. The 2012 definition of mathematical literacy.. 75
Box 4.1. Scientific knowledge: PISA 2015/2018 terminology 100
Box 4.2. The 2015/2018 definition of scientific literacy... 100
Box 5.1. OECD activities in relation to financial education 123
Box 6.1. Integrating global and intercultural issues in the curriculum.. 171
Box 6.2. Pedagogies for promoting global competence... 174
Box 6.3. Perspectives on global competence from different cultures 177
Box 6.4. Teaching attitudes and values related to global competence 178

Follow OECD Publications on:

http://twitter.com/OECD_Pubs

http://www.facebook.com/OECDPublications

http://www.linkedin.com/groups/OECD-Publications-4645871

http://www.youtube.com/oecdilibrary

OECD Alerts

http://www.oecd.org/oecddirect/

1. What is PISA?

The OECD Programme for International Student Assessment (PISA), now in its seventh cycle, seeks to determine what is important for citizens to know and be able to do. PISA assesses the extent to which 15-year-old students near the end of their compulsory education have acquired the knowledge and skills that are essential for full participation in modern societies.

The triennial assessment focuses on the core school subjects of reading, mathematics and science. Students' proficiency in an innovative domain is also assessed; in 2018, this domain was global competence. The assessment does not just ascertain whether students can reproduce knowledge; it also examines how well students can extrapolate from what they have learned and can apply that knowledge in unfamiliar settings, both in and outside of school. This approach reflects the fact that modern economies reward individuals not for what they know, but for what they can do with what they know.

PISA is an ongoing programme that monitors trends in the knowledge and skills that students around the world, and in demographic subgroups within each country, have acquired. In each round of PISA, one of the core domains is tested in detail, taking up roughly one-half of the total testing time. The major domain in 2018 was reading, as it was in 2000 and 2009. Mathematics was the major domain in 2003 and 2012, and science was the major domain in 2006 and 2015.

Through questionnaires distributed to students and school principals, and optional questionnaires distributed to parents and teachers, PISA also gathers information about students' home background, their approaches to learning and their learning environments.

With this alternating schedule of major domains, a thorough analysis of achievement in each of the three core areas is presented every nine years; an analysis of trends is offered every three years. Combined with the information gathered through the various questionnaires, the PISA assessment provides three main types of outcomes:

- Basic indicators that provide a profile of the knowledge and skills of students
- Indicators derived from the questionnaires that show how such skills relate to various demographic, social, economic and educational variables
- Indicators on trends that show changes in outcomes and their distributions, and in relationships between student-, school- and system-level background variables and outcomes.

Policy makers around the world use PISA findings to gauge the knowledge and skills of the students in their own country/economy compared with those in other participating countries/economies, establish benchmarks for improvements in the education provided and/or in learning outcomes, and understand the relative strengths and weaknesses of their own education systems.

This publication presents the theory underlying the PISA 2018 assessment – the seventh since the programme's inception. It includes frameworks for assessing the three core subjects of reading, mathematics and science (Chapters 2, 3 and 4, respectively), the framework for the third assessment of students' financial literacy (Chapter 5), and the framework for assessing the innovative domain, global competence (Chapter 6). These chapters outline the knowledge content that students need to acquire in each domain, the processes that students need to be able to perform, and the contexts in which this knowledge and these skills are applied. They also discuss how each domain is assessed. The publication concludes with the frameworks for the various questionnaires distributed to students, school principals, parents and teachers (Chapter 7), and the framework for the new well-being questionnaire distributed to students (Chapter 8).

Box 1.1. Key features of PISA 2018

The content

PISA not only assesses whether students can reproduce knowledge, but also whether they can extrapolate from what they have learned and apply their knowledge in new situations. It emphasises the mastery of processes, the understanding of concepts, and the ability to function in various types of situations.

The PISA 2018 survey focused on reading, with mathematics and science as minor domains of assessment. For the first time, global competence was assessed as an innovative domain. PISA 2018 also included an assessment of young people's financial literacy, which was optional for countries and economies.

The students

Approximately 710 000 students completed the PISA 2018 assessment, representing over 31 million 15-year-olds in the schools of the 79 participating countries and economies.

The assessment

Computer-based tests were used, with assessments lasting a total of two hours for each student.

Test items were a mixture of multiple-choice questions and questions requiring students to construct their own responses. The items were organised in groups based on a passage setting out a real-life situation. About 930 minutes of test items were used, with different students taking different combinations of test items.

Students also answered a background questionnaire that took 35 minutes to complete. The questionnaire sought information about the students themselves, their homes, and their school and learning experiences. School principals completed a questionnaire that covered the school system and the learning environment.

To obtain additional information, some countries/economies decided to distribute a questionnaire to teachers to learn about their training and professional development, their teaching practices and their job satisfaction. In some countries/economies, optional questionnaires were distributed to parents, who were asked to provide information on their perceptions of and involvement in their child's school, their support for learning in the home, and their own engagement with reading and with other cultures.

Countries/economies could also choose three other optional questionnaires for students: one asked students about their familiarity with and use of information and communications technologies; one sought information about students' education to date, including any interruptions in their schooling, and whether and how they are preparing for a future career; and one, distributed for the first time in PISA 2018, examined students' well-being and life satisfaction.

Countries/economies that conducted the optional financial literacy assessment also distributed a financial literacy questionnaire.

What makes PISA unique

PISA is the most comprehensive and rigorous international programme to assess student performance and to collect data on the student, family and institutional factors that can help explain differences in performance. Decisions about the scope and nature of the assessments and the background information to be collected are made by leading experts in participating countries, and are steered jointly by governments on the basis of shared, policy-driven interests. Substantial efforts and resources are devoted to achieving cultural and linguistic breadth and balance in the assessment materials. Stringent quality-assurance mechanisms are applied in translation, sampling and data collection. As a consequence, results from PISA have a high degree of validity and reliability.

PISA's unique features include its:

- **policy orientation**, which links data on student learning outcomes with data on students' backgrounds and attitudes towards learning, and on key factors that shape their learning in and outside of school; this exposes differences in performance and identifies the characteristics of students, schools and education systems that perform well
- **innovative concept of "literacy"**, which refers to students' capacity to apply knowledge and skills, and to analyse, reason and communicate effectively as they identify, interpret and solve problems in a variety of situations
- **relevance to lifelong learning**, as PISA asks students to report on their motivation to learn, their beliefs about themselves and their learning strategies
- **regularity**, which enables countries to monitor their progress in meeting key learning objectives
- **breadth of coverage**, which, in PISA 2018, encompasses all 37 OECD countries and 42 partner countries and economies.

The PISA 2018 test

The PISA 2018 assessment was conducted principally via computer, as was the case, for the first time, in 2015. Paper-based assessment instruments were provided for countries that chose not to test their students by computer; but the paper-based assessment was limited to reading, mathematics and science trend items only (i.e. those items that had already been used in prior paper-based assessments). New items were developed only for the computer-based assessment.

The 2018 computer-based assessment was designed to be a two-hour test. Each test form distributed to students comprised four 30-minute clusters of test material. This test design included six clusters from both of the domains of mathematics and science to measure trends. For the major domain of reading, material equivalent to 15 30-minute clusters was developed. This material was organised into units instead of clusters, as the PISA 2018 reading assessment adopted an adaptive approach, whereby students were assigned units based on their performance in earlier units. In addition, four clusters of global competence items were developed for the countries that chose to participate in that assessment.

There were different test forms for countries that participated in the global competence assessment. Students spent one hour on the reading assessment (composed of a core stage followed by two stages of either greater or lesser difficulty) plus one hour on one or two other subjects – mathematics, science or global competence. For the countries/economies that chose not to participate in the global competence assessment, 36 test forms were prepared.

Countries that chose paper-based delivery for the main survey measured student performance with 30 paper-and-pencil forms containing trend items from the three core PISA domains. The reading items in these paper-based forms were based on the 2009 reading literacy framework and did not include any items based on the new 2018 reading literacy framework.

Each test form was completed by a sufficient number of students to allow for estimations of proficiency on all items by students in each country/economy and in relevant subgroups within a country/economy, such as boys and girls, or students from different social and economic backgrounds.

The assessment of financial literacy was offered as an option in PISA 2018 based on the same framework as that developed for PISA 2012, which was also used in 2015. Within PISA-participating schools, a sample of students different from the main sample sat the financial literacy test. In addition to the one-hour financial literacy test, these students also sat either a one-hour reading or one-hour mathematics assessment.

An overview of what is assessed in each domain

Box 1.2 presents definitions of the three domains assessed in PISA 2018. The definitions all emphasise the functional knowledge and skills that allow one to participate fully in society. Such participation requires more than just the ability to carry out tasks imposed externally by, for example, an employer; it also involves the capacity to participate in decision making. The more complex tasks in PISA require students to reflect on and evaluate material, not just answer questions that have one correct answer.

Box 1.2. Definitions of the domains

Reading literacy: An individual's capacity to understand, use, evaluate, reflect on and engage with texts in order to achieve one's goals, develop one's knowledge and potential, and participate in society.

Mathematical literacy: An individual's capacity to formulate, employ and interpret mathematics in a variety of contexts. It includes reasoning mathematically and using mathematical concepts, procedures, facts and tools to describe, explain and predict phenomena.

Scientific literacy: The ability to engage with science-related issues, and with the ideas of science, as a reflective citizen. A scientifically literate person is willing to engage in reasoned discourse about science and technology, which requires the competencies to explain phenomena scientifically, evaluate and design scientific enquiry, and interpret data and evidence scientifically.

Reading literacy (Chapter 2) is defined as students' ability to understand, use, evaluate, reflect on and engage with text to achieve their purposes.

PISA assesses students' performance in reading through questions that involve a variety of:

- **Processes (aspects)**: Students are not assessed on the most basic reading skills, as it is assumed that most 15-year-old students will have acquired these. Rather, students are expected to demonstrate their proficiency in locating information, including both accessing and retrieving information within a piece of text, and searching for and selecting relevant text; understanding text, including both acquiring a representation of the literal meaning of text and constructing an integrated representation of text; and evaluating and reflecting on text, including both assessing its quality and credibility, and reflecting on content and form.
- **Text formats**: PISA uses both single-source and multiple-source texts; static and dynamic texts; continuous texts (organised in sentences and paragraphs); non-continuous texts (e.g. lists, forms, graphs or diagrams); and mixed texts.
- **Situations**: These are defined by the use for which the text was constructed. For example, a novel, personal letter or biography is written for people's personal use; official documents or announcements are for public use; a manual or report is for occupational use; and a textbook or worksheet is for educational use. Since some students may perform better in one type of reading situation than another, a range of reading situations is included in the test.

New forms of reading that have emerged since the framework was last updated in 2009, especially digital reading and the growing diversity of material available in both print and digital forms, have been incorporated into the revised PISA 2018 reading framework.

Mathematical literacy (Chapter 3) is defined as students' ability to analyse, reason and communicate ideas effectively as they pose, formulate, solve and interpret solutions to mathematical problems in a variety of situations.

PISA assesses students' performance in mathematics through questions related to:

- **Processes**: PISA defines three categories of processes: formulating situations mathematically; employing mathematical concepts, facts, procedures and reasoning; and interpreting, applying and evaluating mathematical outcomes. They describe what students do to connect the context of a problem with the mathematics involved and thus solve the problem. These three processes each draw on seven fundamental mathematical capabilities: communicating; mathematising; representing; reasoning and arguing; devising strategies for solving problems; using symbolic, formal and technical language and operations; and using mathematical tools. All of these capabilities draw on the problem solver's detailed mathematical knowledge about individual topics.

- **Content**: These are four ideas (quantity; space and shape; change and relationships; and uncertainty and data) that are related to familiar curricular subjects, such as numbers, algebra and geometry, in overlapping and complex ways.
- **Contexts**: These are the settings in a student's world in which the problems are placed. The framework identifies four contexts: personal, educational, societal and scientific.

Scientific literacy (Chapter 4) is defined as the ability to engage with science-related issues, and with the ideas of science, as a reflective citizen. A scientifically literate person is willing to engage in reasoned discourse about science and technology, which requires the competencies to explain phenomena scientifically, evaluate and design scientific enquiry, and interpret data and evidence scientifically.

PISA assesses students' performance in science through questions related to:

- **Contexts**: These include personal, local/national and global issues, both current and historical, that demand some understanding of science and technology.
- **Knowledge**: This is the understanding of the major facts, concepts and explanatory theories that form the basis of scientific knowledge. Such knowledge includes knowledge of both the natural world and technological artefacts (content knowledge), knowledge of how such ideas are produced (procedural knowledge), and an understanding of the underlying rationale for these procedures and the justification for their use (epistemic knowledge).
- **Competencies**: These are the ability to explain phenomena scientifically, evaluate and design scientific enquiry, and interpret data and evidence scientifically.

The evolution of reporting student performance in PISA

Results from PISA are reported using scales. Initially, the average score across OECD countries for all three subjects was 500 with a standard deviation of 100, which meant that two-thirds of students across OECD countries scored between 400 and 600 points. These scores represent degrees of proficiency in a particular domain. Scores in subsequent cycles of PISA are calibrated so as to be directly comparable to those in previous cycles; hence the average score across OECD countries in subsequent cycles has fluctuated slightly around the original 500.

Reading literacy was the major domain in 2000, and the reading scale was divided into five proficiency levels of knowledge and skills. The main advantage of this approach is that it is useful for describing what substantial numbers of students can do with tasks at different levels of difficulty. Results were also presented through three "aspect" subscales of reading: accessing and retrieving information; integrating and interpreting texts; and reflecting and evaluating texts.

PISA 2009 marked the first time that reading literacy was re-assessed as a major domain. Trend results were reported for all three domains – reading, mathematics and science. PISA 2009 added a Level 6 to the reading scale to describe very high levels of reading proficiency. The bottom level of proficiency, Level 1, was renamed Level 1a. Another level, Level 1b, was introduced to describe the performance of students who would previously have been rated as "below Level 1", but who show proficiency in relation to new items that were easier than those included in previous PISA assessments. These changes allowed countries to know more about what kinds of tasks students with very high and very low reading proficiency were capable of completing.

Reading was once again the major domain of assessment in PISA 2018. The three subscales described above were renamed “locating information”, “understanding”, and “evaluating and reflecting”. Two new subscales that describe students’ literacy with single-source and multiple-source texts were also developed. In addition, the reading scale was extended by adding Level 1c, which better describes the proficiency of the lowest-achieving students. These students show minimal reading literacy; what they could do in reading was not described in the previous PISA reading literacy scales.

The context questionnaires

To gather contextual information, PISA asks students and the principals of their schools to respond to questionnaires. These take about 35 and 45 minutes, respectively, to complete. The responses to the questionnaires are analysed with the assessment results to provide at once a broader and more nuanced picture of student, school and system performance. Chapter 7 presents the questionnaire framework in detail. Some countries/economies asked students to complete an additional well-being questionnaire, new to PISA 2018; the framework for this questionnaire is presented in Chapter 8. The questionnaires from all assessments since PISA’s inception are available on the PISA website: www.oecd.org/pisa/.

The questionnaires seek information about:

- Students and their family background, including their economic, social and cultural capital
- Aspects of students’ lives, such as their attitudes towards learning, their habits and life in and outside of school, and their family environment
- Aspects of schools, such as the quality of the schools’ human and material resources, public and private management and funding, decision-making processes, staffing practices, and the school’s curricular emphasis and extracurricular activities offered
- Context of instruction, including institutional structures and types, class size, classroom and school climate, and reading activities in class
- Aspects of learning, including students’ interest, motivation and engagement.

In PISA 2018, five additional questionnaires were offered as options:

- **Computer familiarity questionnaire**, focusing on the availability and use of information and communications technology (ICT) and on students’ ability to carry out computer tasks and their attitudes towards computer use
- **Well-being questionnaire**, new to PISA 2018, on students’ perceptions of their health, life satisfaction, social connections, and in- and outside-of-school activities
- **Educational career questionnaire**, which collects additional information on interruptions in schooling, preparation for students’ future career, and support with language learning

- **Parent questionnaire**, focusing on parents' perceptions of and involvement in their child's school, their support for learning at home, school choice, their child's career expectations, and their background (immigrant/non-immigrant)
- **Teacher questionnaire**, which asks about teachers' initial training and professional development, their beliefs and attitudes, and their teaching practices; separate questionnaires were developed for teachers of the test language and for other teachers in the school.

The contextual information collected through the student, school and optional questionnaires comprises only a part of the information available to PISA. Indicators describing the general structure of education systems (their demographic and economic contexts, such as their costs, enrolments, school and teacher characteristics, and some classroom processes) and their effect on labour market outcomes are routinely developed and applied by the OECD (e.g. in the annual OECD publication, *Education at a Glance*).

A collaborative project

PISA is the result of a collaborative effort between OECD and partner governments. The assessments are developed co-operatively, agreed by participating countries/economies, and implemented by national organisations. The co-operation of students, teachers and principals in participating schools has been crucial to the success of PISA during all stages of development and implementation.

The PISA Governing Board (PGB), composed of representatives at the senor policy level from all participating countries/economies, determines the policy priorities for PISA in the context of OECD objectives. It also oversees adherence to these priorities during the implementation of the programme. The PGB sets priorities for developing indicators, establishing assessment instruments and reporting results. Experts from participating countries/economies also serve on working groups tasked with linking PISA policy objectives with the best available technical expertise in the different assessment domains. By participating in these expert groups, countries/economies ensure that the instruments are internationally valid and take into account differences in cultures and education systems.

Participating countries/economies implement PISA at the national level through National Centres managed by National Project Managers, subject to the agreed administration procedures. National Project Managers play a vital role in ensuring that the implementation is of high quality. They also verify and evaluate survey results, analyses, reports and publications.

The reading framework was developed by the reading expert group with the guidance of John de Jong and Peter Foltz from Pearson. The reading expert group was chaired by Jean-François Rouet (University of Poitiers, France). Other experts who contributed to the reading framework are Paul van den Broek (Universiteit Leiden, the Netherlands), Kevin Chung (University of Hong Kong), Sascha Schroeder (Max Planck Institute for Human Development, Berlin, Germany), Sari Sulkunen (University of Jyväskylä, Finland; also served as the liaison to the PISA global competence expert group), and Dominique Lafontaine (Université de Liège, Belgium; also served as the liaison to the PISA questionnaire expert group).

The global competence framework was developed by Mario Piacentini of the OECD Secretariat with Martyn Barrett (University of Surrey, Guildford, UK), Veronica Boix

Mansilla (Harvard University and Project Zero, Cambridge, USA), Darla Deardorff (Duke University, Durham, USA) and Hye-Won Lee (Korea Institute for Curriculum and Evaluation, Jincheon, Korea), with additional help from Rose Bolognini and Natalie Foster (OECD Secretariat), Natasha Robinson (University of Oxford, UK) and Mattia Baiutti (Fondazione Intercultura, Colle di Val d'Elsa, Italy and the University of Udine, Italy). This framework built on earlier work from experts who led the first part of the development of the global competence assessment: Darla Deardorff (Duke University, Durham, USA), David Kerr (University of Reading, UK and YoungCitizens, London, UK), Peter Franklin (HTWG Konstanz University of Applied Sciences, Germany), Sarah Howie (University of Pretoria, South Africa), Wing On Lee (Open University of Hong Kong, China), Jasmine B-Y Sim (National Institute of Education, Singapore), and Sari Sulkunen (University of Jyväskylä, Finland).

The framework for the PISA 2018 questionnaires was developed by the questionnaire expert group with the guidance of John de Jong and Christine Rozunick from Pearson. The questionnaire expert group was chaired by Fons van de Vijver (Tilburg University, the Netherlands; the North-West University, Potchefstroom, South Africa; and the University of Queensland, Brisbane, Australia). Other experts who contributed to the development of the questionnaire framework are Dominique Lafontaine (Université de Liège, Belgium), Sarah Howie (University of Pretoria, South Africa), Andrew Elliot (University of Rochester, USA), Therese Hopfenbeck (University of Oxford, UK) and David Kaplan (University of Wisconsin-Madison, USA).

The framework for the well-being questionnaire was developed by Jonas Bertling (ETS). The frameworks for the mathematics and science assessments received their last major updates when they were the major domain of assessment (2012 for mathematics, 2015 for science).

Pearson facilitated the development of the reading and questionnaire frameworks. The Educational Testing Service (ETS) was responsible for managing and overseeing this survey; developing the instruments, scaling and analysis; and creating the electronic platform. Other partners or subcontractors involved with ETS include the Department of Experimental and Theoretical Pedagogy at the Université de Liège (aSPe) in Belgium and the Educational Measurement and Research Centre (EMACS) of the University of Luxembourg in Luxembourg. Westat assumed responsibility for survey operations and sampling with the subcontractor, the Australian Council for Educational Research (ACER). cApStAn Linguistic Quality Control assumed responsibility for ensuring the linguistic equivalence of all language versions.

The OECD Secretariat has overall managerial responsibility for the programme, monitors its implementation on a day-to-day basis, acts as the secretariat for the PGB, builds consensus among countries, and serves as the interlocutor between the PGB and the contractors charged with implementation. The OECD Secretariat is also responsible for producing the indicators, and for the analysis and preparation of the international reports and publications, in co-operation with the contractors and in close consultation with participating countries/economies at both the policy (PGB) and implementation (National Project Managers) levels.

2. PISA 2018 Reading Framework

Reading is the major domain of assessment of the 2018 cycle of the Programme for International Student Assessment (PISA). This chapter defines reading literacy as it is assessed in PISA 2018. It describes the types of processes and scenarios exhibited in the tasks that PISA uses to assess reading literacy. Moreover, it describes how the nature of reading literacy has changed over the past two decades, notably through the growing presence of digital texts. The chapter also explains how PISA assesses the ease and efficiency with which a student reads, and how it measures various metacognitive aspects of students' reading practices. It then discusses how student performance in reading is measured and reported. Various sample items from the reading assessment are included at the end of this chapter.

Introduction

Reading as the major domain

PISA 2018 marks the third time that reading is a major domain and the second time that the reading literacy framework receives a major revision. Such a revision must reflect the changing definition of reading literacy as well as the changing contexts in which reading is used in citizens' lives. Thus, the present revision of the framework builds on contemporary and comprehensive theories of reading literacy and considers how students acquire and use information in a variety of contexts.

We live in a rapidly changing world in which both the quantity and variety of written materials are increasing and where people are expected to use these materials in new and increasingly complex ways. It is now generally accepted that our understanding of reading literacy evolves as society and culture themselves change. The reading literacy skills needed for individual growth, educational success, economic participation and citizenship 20 years ago are different from those required today, and it is likely that in 20 years' time they will change further still.

The goal of education has continued to shift its emphasis from the collection and memorisation of information to a broader concept of knowledge: "whether a technician or a professional person, success lies in being able to communicate, share, and use information to solve complex problems, in being able to adapt and innovate in response to new demands and changing circumstances, in being able to marshal and expand the power of technology to create new knowledge and expand human capacity and productivity" (Binkley et al., 2011[1]). The ability to locate, access, understand and reflect on all kinds of information is essential if individuals are to be able to participate fully in our knowledge-based society. Reading literacy is not only a foundation for achievement in other subject areas within the educational system but also a prerequisite for successful participation in most areas of adult life (Cunningham and Stanovich, 1997[2]; OECD, 2013[3]; Smith et al., 2000[4]). The PISA framework for assessing the reading literacy of students towards the end of compulsory education, therefore, must focus on reading literacy skills that include finding, selecting, interpreting, integrating and evaluating information from the full range of texts associated with situations that extend beyond the classroom.

Changes in the nature of reading literacy

Evolving technologies have rapidly changed the ways in which people read and exchange information, both at home and in the workplace. The automation of routine jobs has created a demand for people who can adapt to quickly changing contexts and who can find and learn from diverse sources of information. In 1997, when the first PISA framework for reading began to be discussed, just 1.7% of the world's population used the Internet. By 2014, the number had grown to a global penetration rate of 40.4%, representing almost three billion people (International Telecommunications Union, 2014[5]). Between 2007 and 2013, the number of mobile phone subscriptions doubled: in 2013, there were almost as many active subscriptions as people on earth (95.5 subscriptions per 100 people) and the number of mobile broadband subscriptions had increased to almost two billion worldwide (International Telecommunications Union, 2014[6]). The Internet increasingly pervades the life of all citizens, from learning in and out of school, to working in real or virtual workplaces, to dealing with personal matters such as taxes, health care or holiday planning. Personal and professional development is a lifelong endeavour and the students of tomorrow will need to be skilled with digital tools in order to successfully manage the increased complexity and quantity of information available.

In the past, the primary interest when evaluating student reading literacy proficiency was the ability to understand, interpret and reflect upon single texts. While these skills remain important, greater emphasis on the integration of information technologies into citizens' social and work lives requires that the definition of reading literacy be updated and extended. It must reflect the broad range of newer skills associated with literacy tasks required in the 21st century (Ananiadou and Claro, 2009[7]; Kirsch et al., 2002[8]; Rouet, 2006[9]; Spiro et al., 2015[10]). This necessitates an expanded definition of reading literacy encompassing both basic reading processes and higher-level digital reading skills while recognising that what constitutes literacy will continue to change due to the influence of new technologies and changing social contexts (Leu et al., 2013[11]; 2015[12]).

As the medium through which we access textual information moves from print to computer screens to smartphones, the structure and formats of texts have changed. This in turn requires readers to develop new cognitive strategies and clearer goals in purposeful reading. Therefore, success in reading literacy should no longer be defined by just being able to read and comprehend a single text. Although the ability to comprehend and interpret extended pieces of continuous texts – including literary texts – remains valuable, success will also require the deployment of complex information-processing strategies, including the analysis, synthesis, integration and interpretation of relevant information from multiple text (or information) sources. In addition, successful and productive citizens will need to use information from across domains, such as science and mathematics, and employ technologies to effectively search, organise and filter a wealth of information. These will be the key skills necessary for full participation in the labour market, in further education as well as in social and civic life in the 21st century (OECD, 2013[13]).

Continuity and change in the framework from 2000 to 2015

With the changes in the nature of reading literacy, the framework also has changed. Reading literacy was the major domain assessed during the first PISA cycle (PISA 2000). For the fourth PISA cycle (PISA 2009), it was the first to be revisited as a major domain, requiring a full review of its framework and the development of new instruments that represent it. For the seventh PISA cycle (2018), the framework is once again being revised.

The original reading literacy framework for PISA was developed for the PISA 2000 cycle (from 1998 to 2001) through a consensus-building process involving experts in reading selected by the participating countries to form the PISA 2000 reading expert group (REG). The definition of reading literacy evolved in part from the IEA Reading Literacy Study (1992) and the International Adult Literacy Survey (IALS, 1994, 1997 and 1998). In particular, it reflected the IALS emphasis on the importance of reading skills for active participation in society. It was also influenced by contemporary – and still current – theories of reading, which emphasise the multiple linguistic-cognitive processes involved in reading and their interactive nature (Britt, Goldman and Rouet, 2013[14]; Kamil et al., 2000[15]; Perfetti, 1985[16]; 2007[17]; Snow and the RAND Corporation, 2002[18]; Rayner and Reichle, 2010[19]), models of discourse comprehension (Kintsch, 1998[20]; Zwaan and Singer, 2003[21]) and theories of performance in solving information problems (Kirsch, 2001[22]; Kirsch and Mosenthal, 1990[23]; Rouet, 2006[9]).

Much of the substance of the PISA 2000 framework was retained in the PISA 2009 framework, respecting one of the central purposes of PISA: to collect and report trend information about performance in reading, mathematics and science. However, the PISA domain frameworks are designed to be evolving documents that adapt to and integrate new developments in theory and practice, reflecting both an expansion in our understanding of the nature of reading and changes in the world. This evolution is shown in greater detail in Appendix A, which provides an overview of the primary changes in the reading framework from 2000 to 2015.

Changes in our concept of reading since 2000 have led to an expanded definition of reading literacy, which recognises the motivational and behavioural characteristics of reading alongside the cognitive characteristics. Both reading engagement and metacognition – an awareness and understanding of how one develops an understanding of text and uses reading strategies – were referred to briefly at the end of the first PISA framework for reading under “Other issues” (OECD, 2000[24]). In the light of recent research, reading engagement and metacognition were featured more prominently in the PISA 2009 and 2015 reading frameworks as elements that can be developed and fostered as components of reading literacy.

A second major modification of the framework from PISA 2000 to PISA 2009 was the inclusion of digital texts, in recognition of the increasing role of such texts in both individual growth and active participation in society (OECD, 2011[25]). This modification was concomitant with the new computer-based format of the assessment and thus involved the presentation of texts on a computer screen. PISA 2009 was the first large-scale international study to assess the reading of digital texts.

During PISA 2015, reading was a minor domain and the description and illustration of reading literacy developed for PISA 2009 were kept. However, PISA 2015 involved important changes in the test administration procedures, some of which required adjustments in the wording of the reading framework. For example, the reading assessment in the 2015 cycle was administered primarily on computer. As a result, the “environment” and “medium” dimensions were revisited and further elaborated with the inclusion of the terms “fixed” and “dynamic”.

Revising the framework for PISA 2018

The PISA 2018 reading literacy framework retains aspects of the 2009/2015 frameworks that are still relevant to PISA 2018. However, the framework has been enhanced and revised in the following ways:

- The framework fully integrates reading in a traditional sense together with the new forms of reading that have emerged over the past decades and that continue to emerge due to the spread of digital devices and digital texts.
- The framework incorporates constructs involved in basic reading processes. These constructs, such as fluent reading, literal interpretation, inter-sentence integration, extraction of the central themes and drawing inferences, are critical skills for processing complex or multiple texts for specific purposes. If students fail at performing higher-level text processing functions, it is critical to know whether the failure was due to difficulties in these basic skills in order to provide appropriate support to these students.
- The framework revisits the way in which the domain is organised to incorporate reading processes such as evaluating the veracity of texts, seeking information, reading from multiple sources and integrating/synthesising information across sources. The revision rebalances the prominence of different reading processes to

reflect the global importance of the different constructs, while ensuring there is a link to the prior frameworks in order to be able to measure trends in achievement.

- The revision considers how new technology options and the use of scenarios involving print and digital text can be harnessed to achieve a more authentic assessment of reading, consistent with the current use of texts around the world.

The importance of digital reading literacy

Reading in today's world is very different from what it was just 20 years ago. Up to the mid-1990s, reading was mostly performed on paper. Printed matter existed and continue to exist in many different forms, shapes and textures, from children's books to lengthy novels, from leaflets to encyclopaedias, from newspapers and magazines to scholarly journals, from administrative forms to notes on billboards.

In the early 1990s, only a small percentage of people owned computers and most such computers were mainframes or desktop PCs. Very few people owned laptops for their personal use, whereas digital tablets and smartphones were still yet to become popular. Computer-based reading was limited to specific users and uses, typically a specialised worker dealing with technical or scientific information. In addition, due to mediocre display quality, reading on the computer was slower, more error-prone and more tiring than reading on paper (Dillon, 1994[26]). Initially acclaimed as a means to "free" the reader from the printed text "straightjacket", emerging hypertext technology, such as the linking of digital information pages allowing each reader to dynamically construct their own route through chunks of information (Conklin, 1987[27]), also led to disorientation and cognitive overload, as the Web was still in its infancy (Foltz, 1996[28]; Rouet and Levonen, 1996[29]). But at that time, only a very small fraction of the world population had access to the newly-born World Wide Web.

In less than 20 years, the number of computers in use worldwide grew to an estimated 2 billion in 2015 (International Telecommunications Union, 2014[6]). In 2013, 40% of the world's population had access to the Internet at home, with a sharp contrast between developed countries, where access reached 80% of the population, and some less developed countries, where access lagged below 20% (International Telecommunications Union, 2014[6]). The last decade has witnessed a dramatic expansion of portable digital devices, with wireless Internet access overtaking fixed broadband subscriptions in 2009 (OECD, 2012[30]). By 2015, computer sales were slowing, while sales of digital pads, readers and cell phones were still growing at two-digit rates (Gartner, 2014[31]).

As a notable consequence of the spread of information and communication technology (ICT) among the general public, reading is massively shifting from print to digital texts. For example, computers have become the second most-used source of news for American citizens, after TV and before radio and printed newspapers and magazines (American Press Institute, 2014[32]). Similarly, British children and teenagers prefer to read digital rather than printed texts (Clark, 2014[33]), and a recent UNESCO report showed that two thirds of users of a phone-based reader across five developing countries indicated that their interest in reading and time spent reading increased once it was possible to read on their phones (UNESCO, 2014[34]). This shift has important consequences for the definition of reading as a skill. Firstly, the texts that people read on line are different from traditional printed texts. In order to enjoy the wealth of information, communication and other services offered through digital devices, online readers have to cope with smaller displays, cluttered screens and challenging networks of pages. In addition, new genres of digital-based communication have appeared, such as e-mail, short messaging, forums and social networking applications.

It is important to stress that the rise of digital technology means that people need to be selective in what they read while they must also read more, more often and for a broader range of purposes. Reading and writing are even replacing speech in some everyday communication acts, such as using chat systems rather than telephoning help desks. A consequence is that readers have to understand these new text-based genres and socio-cultural practices.

Readers in the digital age also have to master several new skills. They have to be minimally ICT literate in order to understand and operate devices and applications. They also have to search for and access the texts they need through the use of search engines, menus, links, tabs and other paging and scrolling functions. Due to the uncontrolled profusion of information on the Internet, readers also have to be discerning in their choice of information sources and must assess of information quality and credibility. Finally, readers have to read across texts to corroborate information, to detect potential discrepancies and conflicts and to resolve them. The importance of these new skills was clearly illustrated in the OECD's PISA 2009 digital reading study, whose report noted the following:

Navigation is a key component of digital reading, as readers "construct" their text through navigation. Thus, navigational choices directly influence what kind of text is eventually processed. Stronger readers tend to choose strategies that are suited to the demands of the individual tasks. Better readers tend to minimise their visits to irrelevant pages and locate necessary pages efficiently. (OECD, 2011, p. 20[25])

In addition, a 2015 study of student use of computers in the classroom (OECD, 2015, p. 119[35]) shows, for instance, that "students' average navigation behaviour explains a significant part of the differences in digital reading performance between countries/economies that is not accounted for by differences in print-reading performance"; see also Naumann (2015[36]).

Thus, in many parts of the world, skilful digital reading literacy is now key to one's ability to achieve one's goals and participate in society. The 2018 PISA reading framework has been revised and expanded so as to encompass those skills that are essential for reading and interacting with digital texts.

Reading motivation, practices and metacognition

Individuals' reading practices, motivation and attitudes towards reading, as well as an awareness of how effective reading strategies are, play a prominent role in reading. Students who read more frequently, be it with print or on-screen, who are interested in reading, who feel confident in their reading abilities and who know which strategies to use, to, for instance, summarise a text or search for information on the Internet, tend to be more proficient in reading.

Moreover, practices, motivation, and metacognition deserve close attention not only because they are potential predictors of reading achievement and growth but also because they can be considered important goals or outcomes of education, potentially driving life-long learning (Snow and the RAND Corporation, 2002[18]). Furthermore, they are malleable variables, amenable to change. For instance, there is strong evidence that reading engagement and metacognition (awareness of strategies) can be enhanced through teaching and supportive classroom practices (Brozo and Simpson, 2007[37]; Guthrie, Wigfield and You, 2012[38]; Guthrie, Klauda and Ho, 2013[39]; Reeve, 2012[40]). Reading motivation, practices and metacognition are briefly discussed in the reading literacy framework since they are critical factors of reading. However, they are assessed in the questionnaire and are thus covered in more detail in the questionnaire framework.

The structure of the reading literacy framework

Having addressed what is meant by the term "reading literacy" in PISA and introduced the importance of reading literacy in today's society in this introduction, the remainder of the framework is organised as follows. The second section defines reading literacy and elaborates on various phrases that are used in the reading framework, along with the assumptions underlying the use of these words. The third section focuses on the organisation of the domain of reading literacy and discusses the characteristics that will be represented in the tasks included in the PISA 2018 assessment. The fourth section discusses some of the operational aspects of the assessment and how reading literacy will be measured, and presents sample items. Finally, the last section describes how the reading literacy data will be summarised and outlines plans for reporting the results.

Defining reading literacy

Definitions of reading and reading literacy have changed over time to reflect changes in society, economy, culture and technology. Reading is no longer considered an ability acquired only in childhood during the early years of schooling. Instead it is viewed as an expanding set of knowledge, skills and strategies that individuals build on throughout life in various contexts, through interaction with their peers and the wider community. Thus, reading must be considered across the various ways in which citizens interact with text-based artefacts and its role in life-long learning.

Cognitively-based theories of reading emphasise the constructive nature of comprehension, the diversity of cognitive processes involved in reading and their interactive nature (Binkley, Rust and Williams, 1997[41]; Kintsch, 1998[20]; McNamara and Magliano, 2009[42]; Oakhill, Cain and Bryant, 2003[43]; Snow and the RAND Corporation, 2002[18]; Zwaan and Singer, 2003[21]). The reader generates meaning in response to text by using previous knowledge and a range of text and situational cues that are often socially and culturally derived. When constructing meaning, competent readers use various processes, skills and strategies to locate information, to monitor and maintain understanding (van den Broek, Risden and Husbye-Hartmann, 1995[44]) and to critically assess the relevance and validity of the information (Richter and Rapp, 2014[45]). These processes and strategies are expected to vary with context and purpose as readers interact with multiple continuous and non-continuous texts both in print and when using digital technologies (Britt and Rouet, 2012[46]; Coiro et al., 2008[47]).

Box 2.1. The definition of reading literacy in earlier PISA cycles

The PISA 2000 definition of reading literacy was as follows:

Reading literacy is understanding, using and reflecting on written texts, in order to achieve one's goals, to develop one's knowledge and potential, and to participate in society.

The PISA 2009 definition of reading literacy, also used in 2012 and 2015, added engagement in reading as part of reading literacy:

Reading literacy is understanding, using, reflecting on and engaging with written texts, in order to achieve one's goals, to develop one's knowledge and potential, and to participate in society.

For 2018 the definition of reading literacy includes the evaluation of texts as an integral part of reading literacy and removes the word “written”.

Box 2.2. The 2018 definition of reading literacy

Reading literacy is understanding, using, evaluating, reflecting on and engaging with texts in order to achieve one’s goals, to develop one’s knowledge and potential and to participate in society.

Each part of the definition is considered in turn below, taking into account the original elaboration and some important developments in the definition of the domain that use evidence from PISA and other empirical studies and that take into account theoretical advances and the changing nature of the world.

Reading literacy...

The term “reading literacy” is used instead of the term “reading” because it is likely to convey to a non-expert audience more precisely what the survey is measuring. “Reading” is often understood as simply decoding (e.g., converting written text into sounds), or even reading aloud, whereas the intention of this assessment is to measure much broader and more encompassing constructs. Reading literacy includes a wide range of cognitive and linguistic competencies, from basic decoding to knowledge of words, grammar and the larger linguistic and textual structures needed for comprehension, as well as integration of meaning with one’s knowledge about the world. It also includes metacognitive competencies: the awareness of and ability to use a variety of appropriate strategies when processing texts. Metacognitive competencies are activated when readers think about, monitor and adjust their reading activity for a particular goal.

The term “literacy” typically refers to an individual’s knowledge of a subject or field, although it has been most closely associated with an individual’s ability to learn, use and communicate written and printed information. This definition seems close to the notion that the term “reading literacy” is intended to express in this framework: the active, purposeful and functional application of reading in a range of situations and for various purposes. PISA assesses a wide range of students. Some of these students will go on to university, possibly to pursue an academic or professional career; some will pursue further studies in preparation for joining the labour force; and some will enter the workforce directly upon completion of secondary schooling. Regardless of their academic or labour-force aspirations, reading literacy will be important to students’ active participation in their community and in their economic and personal lives.

... is understanding, using, evaluating, reflecting on...

The word “understanding” is readily connected with the widely accepted concept of “reading comprehension”, which states that all reading involves some level of integrating information from the text with the reader's pre-existing knowledge. Even at the earliest stages of reading, readers must draw on their knowledge of symbols (e.g., letters) to decode texts and must use their knowledge of vocabulary to generate meaning. However, this process of integration can also be much broader, including, for instance, the development of mental models of how texts relate to the world. The word “using” refers to the notions of application and function – doing something with what we read. The term “evaluating” was added for PISA 2018 to incorporate the notion that reading is often goal-directed, and consequently the reader must weigh such factors as the veracity of the arguments in the text, the point of view of the author and the relevance of a text to the reader's goals. "Reflecting on" is added to "understanding", "using" and "evaluating" to emphasise the notion that reading is interactive: readers draw on their own thoughts and experiences when engaging with text. Every act of reading requires some reflection, where readers review and relate information within the text with information from outside the text. As readers develop their stores of information, experience and beliefs, they constantly test what they read against outside knowledge, thereby continually reviewing and revising their sense of the text. Reflecting on texts can include weighing the author's claim(s), their use of rhetorical and other means of discourse, as well as inferring the author's perspective. At the same time, incrementally and perhaps imperceptibly, readers' reflections on texts may alter their sense of the world. Reflection might also require readers to consider the content of the text, apply their previous knowledge or understanding or think about the structure or form of the text. Each of these skills in the definition – "understanding", "using", "evaluating" and "reflecting on" – are necessary, but none is sufficient for successful reading literacy.

...and engaging with...

A person who is literate in reading not only has the skills and knowledge to read well, but also values and uses reading for a variety of purposes. It is therefore a goal of education to cultivate not only proficiency but also engagement with reading. Engagement in this context implies the motivation to read and comprises a cluster of affective and behavioural characteristics that include an interest in and enjoyment of reading, a sense of control over what one reads, involvement in the social dimension of reading and diverse and frequent reading practices.

...texts...

The phrase "texts" is meant to include all language as used in its graphic form: handwritten, printed or screen-based. In this definition, we exclude as texts purely aural language artefacts such as voice recordings, film, TV, animated visuals and pictures without words. Texts do include visual displays such as diagrams, pictures, maps, tables, graphs and comic strips, which include some written language (for example, captions). These visual texts can exist either independently or they can be embedded within larger texts.

Dynamic texts, which give the reader some level of decision-making power as to how to read them, differ from fixed texts in a number of respects, including the lack of physical clues allowing readers to estimate the length and quantity of text (e.g. the dimensions of paper-based documents are hidden in virtual space); the way different parts of a piece of text and different texts are connected with one another through hypertext links; whether multiple summarised texts are shown as the result of a search. As a result of these differences, readers also typically engage differently with dynamic texts. To a much greater extent than with text that is printed, readers need to construct their own pathways to complete any reading activity associated with dynamic texts.

The term "texts" was chosen instead of the term "information" because of its association with written language and because it more readily connotes literary as well as information-focused reading.

...in order to achieve one's goals, to develop one's knowledge and potential and to participate in society.

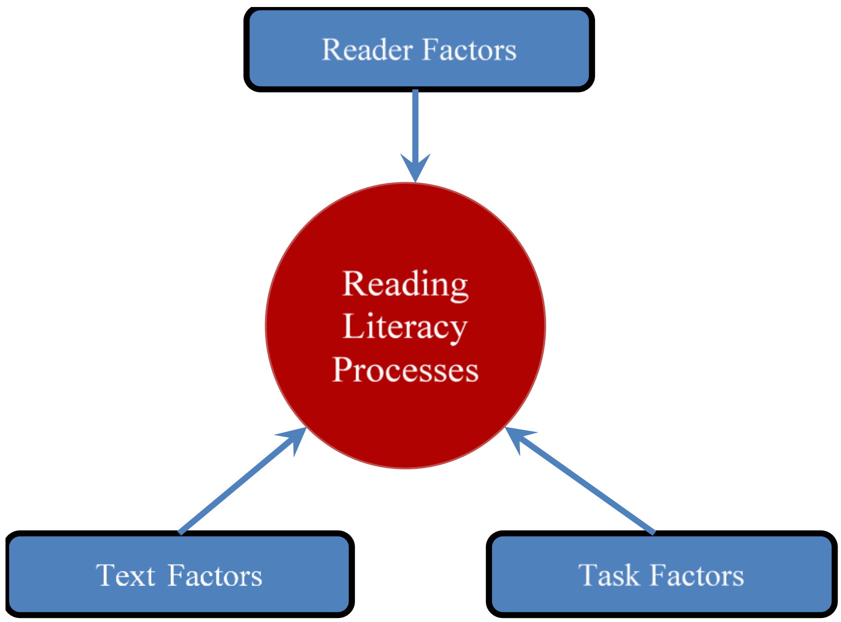
This phrase is meant to capture the full scope of situations in which reading literacy plays a role, from private to public, from school to work, from formal education to lifelong learning and active citizenship. "To achieve one's goals" and "to develop one's knowledge and potential" both spell out the long-held idea that reading literacy enables the fulfilment of individual aspirations – both defined ones such as graduating or getting a job, and those less defined and less immediate that enrich and extend one's personal life and that contribute to lifelong education (Gray and Rogers, 1956[48]). The PISA definition of reading literacy also embraces the new types of reading in the 21st century. It conceives of reading literacy as the foundation for full participation in the economic, political, communal and cultural life of contemporary society. The word "participate" is used because it implies that reading literacy allows people to contribute to society as well as to meet their own needs: "participating" includes social, cultural and political engagement (Hofstetter, Sticht and Hofstetter, 1999[49]). For instance, literate people have greater access to employment and more positive attitudes toward institutions (OECD, 2013[3]). Higher levels of reading literacy have been found to be related to better health and reduced crime (Morrisroe, 2014[50]). Participation may also include taking a critical stance, a step toward personal liberation, emancipation and empowerment (Lundberg, 1991[51]).

Organising the domain

Reading as it occurs in everyday life is a pervasive and highly diverse activity. In order to design an assessment that adequately represents the many facets of reading literacy, the domain is organized according to a set of dimensions. The dimensions will in turn determine the test design and, ultimately, the evidence about student proficiencies that can be collected and reported.

Snow and the RAND Reading Group's (2002[18]) influential framework defined reading comprehension as the joint outcome of three combined sources of influence: the *reader*, the *text* and the activity, *task* or purpose for reading. Reader, text and task dimensions interact within a broad sociocultural context, which can be thought of as the diverse range of situations in which reading occurs. PISA adopts a similar view of the dimensions of reading literacy, as illustrated in Figure 2.1. A reader brings a number of *reader factors* to reading, which can include motivation, prior knowledge, and other cognitive abilities. The reading activity is a function of *text factors* (i.e. the text or texts that are available to the reader at a given place and time). These factors include the format of the text, the complexity of the language used, and the number of pieces of text a reader encounters. The reading activity is also a function of *task factors* (i.e. the requirements or reasons that motivate the reader's engagement with text). Task factors include the potential time and other practical constraints, the goals of the task (e.g. whether reading for pleasure, reading for deep understanding or skimming for information) and the complexity or number of tasks to be completed. Based on their individual characteristics and their perception of text and task factors, readers apply a set of *reading literacy processes* in order to locate and extract information and construct meaning from texts to achieve tasks.

Figure 2.1. Factors that contribute to reading literacy



The PISA cognitive assessment measures reading literacy by manipulating task and text factors. An additional questionnaire assays some of the reader factors, such as motivation, disposition and experience.

In designing the PISA reading literacy assessment, the two most important considerations are, first, to *ensure broad coverage* of what students read and for what purposes they read, both in and outside of school, and, second, to *represent a natural range of difficulty in texts and tasks*. The PISA reading literacy assessment is built on three major characteristics: *text* – the range of material that is read; *processes* – the cognitive approach that determines how readers engage with a text; and *scenarios* – the range of broad contexts or purposes for which reading takes place. Within scenarios are *tasks* – the assigned goals that readers must achieve in order to succeed. All three contribute to ensuring broad coverage of the domain. In PISA, task *difficulty* can be varied by manipulating text features and task goals, which then require deployment of different cognitive processes. Thus, the PISA reading literacy assessment aims to measure students' mastery of reading *processes* (the possible cognitive approaches of readers to a text) by varying the dimensions of *text* (the range of material that is read) and *scenarios* (the range of broad contexts or purposes for which reading takes place) with one or more thematically related texts. While there may be individual differences in *reader* factors based on the skills and background of each reader, these are not manipulated in the cognitive instrument but are captured through the assessment in the questionnaire.

These three characteristics must be operationalised in order to use them to design the assessment. That is, the various values that each of these characteristics can take on must be specified. This allows test developers to categorise the materials they work with and the tasks they construct so that they can then be used to organise the reporting of the data and to interpret results.

Processes

The PISA typology of the *cognitive aspects* involved in reading literacy was designed at the turn of the 21st century (OECD, 2000[24]). A revision of these aspects in the 2018 PISA reading literacy framework is needed for at least three reasons:

a) A definition of reading literacy must reflect contemporary developments in school and societal literacy demands, namely, the increasing amount of text information available in print and digital forms and the increasing diversity and complexity of situations involving text and reading. These developments are partly driven by the spread of digital information technology and in particular by increased access to the Internet worldwide.

b) The PISA 2018 framework should also reflect recent developments in the scientific conceptualisation of reading and be as consistent as possible with the terminology used in current theories. There is a need to update the vocabulary that was used to designate the cognitive processes involved in reading, taking into account progress in the research literature.

c) Finally, a revision is needed to reassess the necessary trade-off between the desire to stay faithful to the precise definition of the aspects as described in the framework and the limited possibility to account for each of these individual aspects in a large-scale international assessment. Such a reassessment of the reading framework is particularly relevant in the context of PISA 2018, in which reading literacy is the main domain.

The 2018 framework replaces the phrase “cognitive aspects”, used in previous versions of the framework, with the phrase “cognitive processes” (not to be confused with the reading literacy processes described above). The phrase “cognitive processes” aligns with the terminology used in reading psychology research and is more consistent with a description of reader skills and proficiencies. The term “aspects” tended to confound the reader's actual cognitive processes with the requirements of various types of tasks (i.e. the demands of specific types of questions). A description of reading processes permits the 2018 framework to map these processes to a typology of tasks.

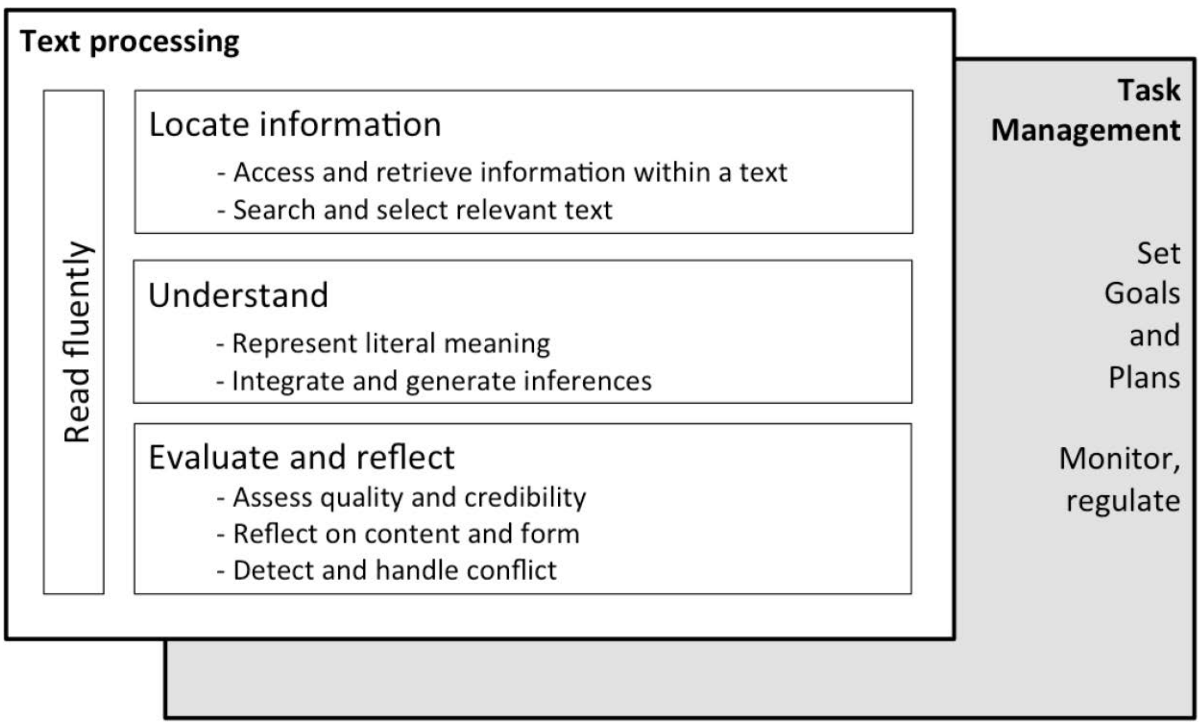
Recent theories of reading literacy emphasise the fact that "reading does not take place in a vacuum" (Snow and the RAND Corporation, 2002[18]; McCrudden and Schraw, 2007[52]; Rouet and Britt, 2011[53]). Indeed, most reading activities in people's daily lives are motivated by specific purposes and goals (White, Chen and Forsyth, 2010[54]). Reading as a cognitive skill involves a set of specific reading processes that competent readers use when engaging with texts in order to achieve their goals. Goal setting and goal achievement drive not only readers' decisions to engage with texts, their selection of texts and passages of text, but also their decisions to disengage from a particular text, to re-engage with a different text, to compare, and to integrate information across multiple texts (Britt and Rouet, 2012[46]; Goldman, 2004[55]; Perfetti, Rouet and Britt, 1999[56]).

To achieve reading literacy as defined in this framework, an individual needs to be able to execute a wide range of processes. Effective execution of these processes, in turn, requires that the reader have the cognitive skills, strategies and motivation that support the processes.

The PISA 2018 reading framework acknowledges the goal-driven, critical and intertextual nature of reading literacy (McCrudden and Schraw, 2007[52]; Vidal-Abarca, Mañá and Gil, 2010[57]). Consequently, the former typology of reading aspects (OECD, 2000[24]) has been revised and extended so as to explicitly represent the fuller range of processes from which skilled readers selectively draw depending on the particular task context and information environment.

More specifically, two broad categories of reading processes are defined for PISA 2018: text processing and task management (Figure 2.2). This distinction is consistent with current views of reading as a situated and purposeful activity, see e.g. (Snow and the RAND Corporation, 2002[18]). The focus of the cognitive assessment is on processes identified in the text processing box.

Figure 2.2. PISA 2018 Reading framework processes



Text processing

The 2018 typology of reading processes specifically identifies the process of reading fluently as distinct from other processes associated with text comprehension.

Reading fluently

Reading fluency can be defined as an individual's ability to read words and text accurately and automatically and to phrase and process these words and texts in order to comprehend the overall meaning of the text (Kuhn and Stahl, 2003[58]). In other words, fluency is *the ease and efficiency of reading texts for understanding*. There is considerable empirical evidence demonstrating a link between reading ease/efficiency/fluency and reading comprehension (Chard, Pikulski and McDonagh, 2006[59]; Kuhn and Stahl, 2003[58]; Wagner et al., 2010[60]; Wayman et al., 2007[61]; Woodcock, McGrew and Mather, 2001[62]; Jenkins et al., 2003[63]). The chief psychological mechanism proposed to explain this relationship is that the ease and efficiency of reading text is indicative of expertise in the foundational reading skills of decoding, word recognition and syntactic parsing of texts.

Fluent reading frees up attention and memory resources, which can be allocated to higher-level comprehension processes. Conversely, weaknesses in reading fluency divert resources from comprehension towards the lower-level processes necessary to process printed text, resulting in weaker performance in reading comprehension (Cain and Oakhill, 2008[64]; Perfetti, Marron and Foltz, 1996[65]). Acknowledging this strong link between fluency and comprehension, the National Reading Panel (2000) in the United States recommended fostering fluency in reading to enhance students' comprehension skills.

Locating information

Competent readers can carefully read an entire piece of text in order to comprehend the main ideas and reflect on the text as a whole. On a daily basis, however, readers most often use texts for purposes that require the location of specific information, with little or no consideration for the rest of the text (White, Chen and Forsyth, 2010[54]). Furthermore, locating information is an obligatory component of reading when using complex digital information such as search engines and websites (Brand-Gruwel, Wopereis and Vermetten, 2005[66]; Leu et al., 2013[11]). The 2018 framework defines two processes whereby readers find information within and across texts:

Accessing and retrieving information within a piece of text. Locating information from tables, text chapters or whole books is a skill in and by itself (Dreher and Guthrie, 1990[67]; Moore, 1995[68]; Rouet and Coutelet, 2008[69]). Locating information draws on readers' understanding of the demands of the task, their knowledge of text organisers (e.g., headers, paragraphs) and their ability to assess the relevance of a piece of text. The ability to locate information depends on readers' strategic awareness of their information needs and their capacity to quickly disengage from irrelevant passages (McCrudden and Schraw, 2007[52]). In addition, readers sometimes have to skim through a series of paragraphs in order to retrieve specific pieces of information. This requires an ability to modulate one's reading speed and depth of processing and to know when to keep in consideration or dismiss the information in the text (Duggan and Payne, 2009[70]). Access and retrieval tasks in PISA 2018 require the reader to scan a single piece of text in order to retrieve target information composed of a few words, phrases or numerical values. There is little or no need to comprehend the text beyond the phrase level. The identification of target information is achieved through literal or close to literal matching of elements in the question and in the text, although some tasks may require inferences at the word or phrase level.

Searching for and selecting relevant text. Proficient readers are able to select information when faced with not just one, but also when faced with several pieces of text. In electronic environments, the amount of available information often largely exceeds the amount readers are able to actually process. In these multiple-text reading situations, readers have to make decisions as to which of the available pieces of text is the most important, relevant, accurate or truthful (Rouet and Britt, 2011[53]). These decisions are based on readers' assessment of the qualities of the pieces of text, which are made from partial and sometimes opaque indicators, such as the information contained in a web link (Gerjets, Kammerer and Werner, 2011[71]; Mason, Boldrin and Ariasi, 2010[72]; Naumann, 2015[36]; Rieh, 2002[73]). Thus, one's ability to search for and select a piece of text from among a set of texts is an integral component of reading literacy. In PISA 2018, text search and selection tasks involve the use of text descriptors such as headers, source information (e.g. author, medium, date), and embedded or explicit links such as search engine result pages.

Understanding

A large number of reading activities involve the parsing and integration of extended passages of text in order to form an understanding of the meaning conveyed in the passage. Text understanding (also called comprehension) may be seen as the construction by the reader of a mental representation of what the text is about, which Kintsch (1998[20]) defines as a “situation model”. A situation model is based on two core processes: the construction of a memory **representation of the literal meaning** of the text; and the **integration** of the contents of the text with one's prior knowledge through mapping and **inference processes** (McNamara and Magliano, 2009[42]; Zwaan and Singer, 2003[21]).

Acquiring a representation of the literal meaning of a text requires readers to comprehend sentences or short passages. Literal comprehension tasks involve a direct or paraphrased match between the question and target information within a passage. The reader may need to rank, prioritise or condense information at a local level. (Note that tasks requiring integration at the level of an entire passage, such as identifying the main idea, summarizing the passage, or giving a title to the passage, are considered to be integration tasks; see below.)

Constructing an integrated text representation requires working from the level of individual sentences to the entire passage. The reader needs to **generate various types of inferences**, ranging from simple connecting inferences (such as the resolution of anaphora) to more complex coherence relationships (e.g. spatial, temporal, causal or claim-argument links) (van den Broek, Risden and Husbye-Hartmann, 1995[44]). Inferences might link different portions of the text together, or they may link the text to the question statement. Finally, the production of inferences is also needed in tasks where the reader must identify the implicit main idea of a given passage, possibly in order to produce a summary or a title for the passage.

When readers are faced with more than one text, integration and inference generation may need to be performed based on pieces of information located in different pieces of texts (Perfetti, Rouet and Britt, 1999[56]). One specific problem that may arise when integrating information across multiple pieces of text is that they might provide inconsistent or conflicting information. In those cases, readers must engage in evaluation processes in order to acknowledge and handle the conflict (Bråten, Strømsø and Britt, 2009[74]; Stadtler and Bromme, 2014[75]) (see below).

Evaluating and reflecting

Competent readers can reason beyond the literal or inferred meaning of the text. They can reflect on the content and form of the text and critically assess the quality and validity of the information therein.

Assessing quality and credibility. Competent readers can evaluate the quality and credibility of the information in a piece of text: whether the information is valid, up-to-date, accurate and/or unbiased. Proficient evaluation sometimes requires the reader to identify and assess the source of the information: whether the author is competent, well-informed and benevolent.

Reflecting on content and form. Competent readers must also be able to reflect on the quality and style of the writing. This reflection involves being able to evaluate the form of the writing and how the content and form together relate to and express the author’s purposes and point of view. Reflecting also involves drawing upon one's knowledge, opinions or attitudes beyond the text in order to relate the information provided within the text to one's own conceptual and experiential frames of reference. Reflection items may be thought of as those that require readers to consult their own experience or knowledge to compare, contrast or hypothesise different perspectives or viewpoints. Evaluation and reflection were arguably always part of reading literacy, but their importance has increased with the increased amount and heterogeneity of information readers are faced with today.

Detecting and handling conflict. When facing multiple pieces of text that contradict each other, readers need to be aware of the conflict and to find ways to deal with it (Britt and Rouet, 2012[46]; Stadtler and Bromme, 2013[76]; 2014[75]). Handling conflict typically requires readers to assign discrepant claims to their respective sources and to assess the soundness of the claims and/or the credibility of the sources. As these skills underlie much of contemporary reading, it is an issue of critical importance to measure the extent to which 15-year-olds can meet the new challenges of comprehending, comparing and integrating multiple pieces of texts (Bråten et al., 2011[77]; Coiro et al., 2008[47]; Goldman, 2004[55]; Leu et al., 2015[12]; Mason, Boldrin and Ariasi, 2010[72]; Rouet and Britt, 2014[78]).

Task management processes

In the context of any assessment, but also in many everyday reading situations (White, Chen and Forsyth, 2010[54]), readers engage with texts because they receive some kind of assignment or external prompt to do so. Reading literacy involves one's ability to accurately represent the reading demands of a situation, to set up task-relevant reading goals, to monitor progress toward these goals, and to self-regulate their goals and strategies throughout the activity (see, e.g., Hacker (1998[79]) and Winne and Hadwin, (1998[80]), for discussions of self-regulated reading).

Task-oriented goals fuel the reader's search for task-relevant texts and/or passages within a text (McCrudden and Schraw, 2007[52]; Rouet and Britt, 2011[53]; Vidal-Abarca, Mañá and Gil, 2010[57]). Finally, monitoring (metacognitive) processes enable the dynamic updating of goals throughout the reading activity. Task management is represented in the background of text processing to emphasise the fact that it constitutes a different, metacognitive level of processing.

While readers' own interpretation of a task's requirements is an important component of the task management processes, the construction of reading goals extends beyond the explicit task instructions as goals may be self-generated based on one's own interests and initiative. However, the PISA reading literacy assessment only considers those goals that readers form upon receiving external prompts to accomplish a given task. In addition, due to implementation constraints, task management processes are represented but not directly and independently assessed as part of PISA 2018. However, portions of the background questionnaire will estimate readers' awareness of reading strategies. Future cycles may consider the use of computer-generated process indicators (such as how often and at what time intervals a student visits a particular page of text or the number of looks back at a question a student makes) as part of the assessment of task management skills.

Summary of reading processes

To summarise, the 2018 framework features a comprehensive and detailed typology of the cognitive processes involved in purposeful reading activities as they unfold in single or multiple text environments. Due to design constraints, it is not possible to distinguish each of these processes in a separate proficiency scale. Instead, the framework defines a smaller list of processes that will form the basis for scaling and reporting (Table 2.1).

It is worth noting that the 2018 process typology also permits an analysis of changes over time in students' proficiency at the level of broad reading processes, as the former "cognitive aspects" featured in previous frameworks can be mapped onto specific processes in the new typology. Table 2.1 shows the correspondence between the 2018 typology and the 2009 typology (which was also used in 2012 and 2015). The distinction between single and multiple text processes is discussed in greater detail below.

Table 2.1. Mapping of the 2018 process typology to 2018 reporting scales and to 2009-2015 cognitive aspects

2018 Cognitive processes	Superordinate Category Used for Scaling in 2018	2009-2015 Aspects
Reading fluently	Reported on PISA scale[1]	Not assessed
Accessing and retrieving information within a text	Locating information	Accessing and retrieving
Searching for and selecting relevant text		
Representing literal meaning	Understanding	Integrating and interpreting
Integrating and generating inferences		
Assessing quality and credibility	Evaluating and reflecting	Reflecting and evaluating
Reflecting on content and form		
Detecting and handling conflict		Complex

Note 1. Reading fluency items were scaled in three steps. First, only the (other) reading items were scaled. Second, these reading items were finalised and item fits were evaluated in a way that was not affected by reading fluency items. Third, reading fluency items were added to the scaling procedure and item fits were evaluated. As reading fluency items reflect the orthography of the test language, it was expected that such items had stronger item-to-country/language associations than other items in the assessment.

Texts

Reading necessarily requires material for the reader to read. In an assessment, that material – a piece of text or a set of texts related to a particular task – must include sufficient information for a proficient reader to engage in meaningful comprehension and resolve the problem posed by the task. Although it is obvious that there are many different kinds of text and that any assessment should include a broad range of texts, there was never a single agreed-upon categorisation of the many different kinds of text that readers encounter. With the advent of digital media and the profusion of new text genres and text-based communication services – some of which may not survive the next decade, some of which may be newly created in the same time span – this issue becomes even more complex. Box 2.3 outlines a categorisation that was used between PISA 2009 and PISA 2015.

Box 2.3. Characteristics used to classify texts in the PISA 2009, 2012 and 2015 reading frameworks

The previous reference framework (2009) included four major dimensions to characterise texts:

1) Medium: print or electronic
2) Environment: authored or message-based
3) Text format: continuous, non-continuous, mixed or multiple
4) Text type: description, narration, exposition, argumentation, instruction or transaction

A Digital Reading Assessment was offered as an optional component in 2009 and 2012.

For the 2015 reading literacy assessment, only texts that had their origin as paper-based print documents were used, albeit presented on computer. For clarity, these were referred to as fixed and dynamic texts under the heading "text display space" instead of medium (in an attempt to clarify that while their origin was paper-based print, students were in fact reading them on a computer screen, hence on an electronic medium). Because reading literacy was a minor domain in 2015, no new tasks were designed and implemented. Consequently, dynamic texts, i.e. texts such as websites designed to take advantage of hyperlinks, menus, and other navigational features of an electronic medium, were not part of PISA 2015.[1]

Reading is the major domain in 2018 and with a revised framework, a broader range of texts can now be represented in the assessment. These include texts that are typical of the print medium but also the ever-expanding category of texts typical of the digital medium. Just like printed texts, some digital texts are "static" in that they come with a minimal set of tools for interaction (scrolling, paging and a find function). This describes, for instance, documents intended to be printed but displayed on a computer screen (e.g. word processing documents or PDF files). However, many digital texts come with innovative features that increase the possibilities for the reader to interact with the material, hence their characterisation as "dynamic texts". Features of dynamic text include embedded hyperlinks that take the reader to other sections, pages or websites; advanced search functions that provide ad hoc indexes of the searched keywords and/or highlight these words in the text; and social interaction as in interactive text-based communication media such as e-mail, forums and instant messaging services.

The 2018 framework defines four dimensions of texts: source (single, multiple); organisational and navigational structure (static, dynamic); format (continuous, non-continuous, mixed); and type (description, narration, exposition, argumentation, instruction, interaction, transaction). The design of test materials that vary along these four dimensions will ensure a broad coverage of the domain and a representation of traditional as well as emerging reading practices.

Source

In the PISA 2018 framework, a source is a unit of text. *Single-source texts* may be defined by having a definite author (or group of authors), time of writing or publication date, and reference title or number. Authors may be defined precisely, like in most traditional printed books, or more vaguely like the pseudonyms in a blog post or the sponsors of a website. A single-source text may also be construed as such because it is presented to the reader in isolation from other texts, even if it does not explicitly bear any source indication. *Multiple-source texts* are defined by having different authors, or by being published at different times, or by bearing different titles or reference numbers. Note that in the PISA framework, "title" is meant in the sense of a bibliographical catalogue unit. Lengthy texts that feature several sections with titles and subtitles are still single texts, to the extent that they were written by a definite author (or group of authors) at a given date. Likewise, multi-page websites are single-source texts as long as there is no explicit mention of a different author or date. Multiple-source texts may be represented on a single page. This is the case in printed newspapers and in many textbooks, but also in forums, customer reviews and question-and-answer websites. Finally, a single text may contain embedded sources, that is, references to other authors or texts (Rouet and Britt, 2014[78]; Strømsø et al., 2013[81]).

In sum, the multiple texts considered in previous versions of the framework correspond to multiple-source texts in the PISA 2018 framework as long as they involve several sources. All the other texts are subsumed under the category of single-source texts.

Organisational and navigational structure

Screen sizes vary dramatically in digital environments, from cell phone displays, which are smaller than a traditional index card, to large, multiple screen displays for simultaneously showing multiple screen windows of information. At the time of the drafting of this framework, however, the typical computer screen (such as the 15" or 17" screen that comes with ordinary desktop and laptop computers) features a display resolution of 1024x768 pixels. Assuming a typical font size, this is enough to display about a half-page of A4 or US-Letter page, that is, a very short piece of text. Given the wide variation in the "landscape" available on screens to display text, digital texts come with a number of tools meant to let the user access and display specific passages. These tools range from generic tools, such as the scroll bar and tabs (also found in a number of other software applications like spreadsheets and word processors) and tools to resize or position the text on the screen, to more specific tools such as menus, tables of contents and embedded hyperlinks to move between text segments. There is growing evidence that navigation in digital text requires specific skills (OECD, 2011[25]; Rouet, Vörös and Pléh, 2012[82]). Therefore, it is important to assess readers' ability to handle texts featuring a high density of navigational tools. For reasons of simplicity, the PISA 2018 framework distinguishes "static" texts, with a simple organisation and low density of navigational tools (typically, one or several screen pages arranged linearly), from "dynamic" texts, which feature a more complex, non-linear organisation and a higher density of navigational devices. Note that the term "density" is preferred to "number" to mark the fact that dynamic texts do not have to be longer than static texts.

In order to ensure a broad coverage of the domain and to maintain consistency with past frameworks, the 2018 framework also retains two former dimensions of the classification of texts, "format" and "type", that remain for the most part unchanged from the previous framework.

Text format

An important way to classify texts, and one at the heart of the organisation of the PISA 2000 framework and assessment, is to distinguish between continuous and non-continuous texts. Continuous texts are typically composed of sentences that are, in turn, organised into paragraphs. These may fit into even larger structures such as sections, chapters and books. Non-continuous texts are most frequently organised in matrix format, based on combinations of lists.

Texts in continuous and non-continuous formats can be either fixed or dynamic texts. Mixed and multiple format texts can also be fixed texts but are particularly often dynamic texts. Each of these four formats is elaborated below.

Other non-text-formatted objects are also commonly used in conjunction with fixed texts and particularly with dynamic texts. Pictures and graphic images occur frequently in fixed texts and can legitimately be regarded as integral to such texts. Static images as well as videos, animations and audio files regularly accompany dynamic texts and can, also, be regarded as integral to those texts. As a reading literacy assessment, PISA does not include non-text formatted objects in their own right, but any such objects may, in principle, appear in PISA as part of a (verbal) text. However, in practice, the use of video and animation is very limited in the current assessment. Audio is not used at all because of practical limitations such as the need for headphones and audio translation.

Continuous texts

Continuous texts are formed by sentences organised into paragraphs. Examples of continuous texts include newspaper reports, essays, novels, short stories, reviews and letters.

Graphically or visually, text is organised by its separation into sentences and paragraphs with spacing (e.g. indentation) and punctuation conventions. Texts also follow a hierarchical structure signalled by headings and content that help readers to recognise its organisation. These markers also provide clues to text boundaries (showing section completion, for example). The location of information is often facilitated by the use of different font sizes, font types such as italic and boldface, and borders and patterns. The use of typographical and format clues is an essential subskill of effective reading.

Discourse markers also provide organisational information. For example, sequence markers (“first”, “second”, “third”, etc.) signal the relation of each of the units introduced to each other and indicate how the units relate to the larger surrounding text. Causal connectors (“therefore”, “for this reason”, “since”, etc.) signify cause-and-effect relationships between parts of a text.

Non-continuous texts

Non-continuous texts are organised differently to continuous texts and therefore require a different kind of reading approach. Most non-continuous texts are composed of a number of lists (Kirsch and Mosenthal, 1990[23]). Some are single, simple lists, but most consist of several simple lists possibly crossed with one another.

Examples of non-continuous text objects are lists, tables, graphs, diagrams, advertisements, schedules, catalogues, indices and forms. These text objects may be either fixed or dynamic.

Mixed texts

Many fixed and dynamic texts are single, coherent objects consisting of a set of elements in both continuous and non-continuous formats and are therefore known as *mixed* texts. Examples of mixed texts include a paragraph together with a picture, or a graph with an explanatory legend. If such mixed texts are well-constructed, the components (for example, a graph or table with an associated prose explanation) support one another through coherent and cohesive links both at local (e.g., locating a city on a map) and global (e.g., discussing the trend represented in a graph) levels.

Mixed text is a common format in fixed-text magazines, reference books and reports, where authors employ a variety of representations to communicate information. Among dynamic texts, authored web pages are typically mixed texts, with combinations of lists, paragraphs of prose and often graphics. Message-based texts, such as online forms, e-mail messages and forums, also combine texts that are continuous and non-continuous in format.

The "multiple" format defined in the previous versions of the framework is now represented as one modality of the new "source" dimension defined above.

Assessing reading literacy

The previous section outlined the conceptual framework for reading literacy. The concepts in the framework must in turn be represented in tasks and questions in order to measure students' proficiencies in reading literacy.

In this section, we consider the use of scenarios, factors affecting item difficulty, dimensions ensuring coverage of the domain and some of the other major issues in constructing and operationalising the assessment.

Scenarios

Reading is a purposeful act that occurs within the context of particular goals. In many traditional reading assessments, test takers are presented with a series of unrelated passages on a range of general topics. Students answer a set of discrete items on each passage and then move on to the next unrelated passage. In this traditional design, students are effectively expected to "forget" what they have read previously when answering questions on later passages. Consequently, there is no overarching purpose for reading other than to answer discrete questions (Rupp, Ferne and Choi, 2006[83]). In contrast, a scenario-based assessment approach can enhance students' engagement with the tasks and thus enable a more accurate assessment of what they can do (Sabatini et al., 2014[84]; 2015[85]).

The PISA 2018 assessment will include scenarios in which students are provided an overarching purpose for reading a collection of thematically related texts in order to complete a higher-level task (e.g responding to some larger integrative question or writing a recommendation based on a set of texts), along with traditional standalone PISA reading units. The reading purpose sets up a collection of goals, or criteria, that students use to search for information, evaluate sources, read for comprehension and/or integrate across texts. The collection of sources can be diverse and may include a selection from literature, textbooks, e-mails, blogs, websites, policy documents, primary historical documents and so forth. Although the prompts and tasks that will evolve from this framework may not grant student test takers the freedom to choose their own purposes for reading and the texts related to those individual purposes, the goal of this assessment is to offer test takers some freedom in choosing the textual sources and paths they will use to respond to initial prompts. In this way, goal-driven reading can be assessed within the constraints of a large-scale assessment.

Tasks

Each scenario is made up of one or more tasks. In each task, students may be asked questions about the texts contained therein ranging from traditional comprehension items (locating information, generating an inference) to more complex tasks such as the synthesis and integration of multiple texts, evaluating web search results or corroborating information across multiple texts. Each task is designed to assess one or more of the processes identified in the framework. Tasks in a scenario can be ordered from least difficult to most difficult to measure student abilities. For instance, a student might encounter an initial task in which he or she must locate a particular document based on a search result. In the second task, the student might have to answer a question about information that is specifically stated in the text. Finally, in the third task, the student might need to determine if the author’s point of view in the first text is the same as in a second text. In each case, these tasks can be scaffolded so that if a student fails to find the correct document in the first task, he or she is then provided with the correct document in order to complete the second task. In this way, complex multipart scenarios do not become an “all or none activity”, but are rather a way to triangulate the level of different student skills through a realistic set of tasks. Thus, scenarios and tasks in the PISA 2018 reading literacy assessment correspond to units and items in previous assessments.

A scenario-based assessment mimics the way an individual interacts with and uses literacy source material in a more authentic way than a traditional, decontextualised assessment would. It presents students with realistic problems and issues to solve, and it involves the use of both basic and higher-level reading and reasoning skills (O’Reilly and Sabatini, 2013[86]).

Scenarios represent a natural extension of the traditional, unit-based approach in PISA. A scenario-based approach was used in the PISA 2012 assessment of problem solving and the PISA 2015 assessment of collaborative problem solving. Tasks 2-4 in Appendix B illustrate a sample scenario with multiple items.

Distribution of tasks

Each task will primarily assess one of the three main categories of cognitive process defined earlier. As such, they can be thought of as individual assessment items. The approximate distribution of tasks for the 2018 reading literacy assessment are shown below in Table 2.2 and are contrasted with the distribution of tasks for the 2015 assessment.

Table 2.2. Approximate distribution of tasks by targeted process and text source

2015 FRAMEWORK	2018 FRAMEWORK		
	SINGLE Text		**MULTIPLE Text**
Accessing and retrieving 25%	Scanning and locating 15%		Searching for and selecting relevant text 10%
Integrating and interpreting 50%	Literal Comprehension 15% Inferential Comprehension 15%		Multiple-text Inferential Comprehension 15%
Reflecting and evaluating 25%	Assessing quality and credibility Reflecting on content and form	20 %	Corroborating/handling conflict 10%

Items will be reused from previous PISA reading literacy assessments in order to allow for the measurement of trends. In order to achieve the desired proportion of tasks involving multiple pieces of text, and because prior PISA assessments focused on tasks involving only single texts, the development of new items will mostly require the creation of tasks involving multiple texts (e.g. searching for and selecting relevant text, multiple-text inferential comprehension and corroborating/handling conflict). At the same time, a sufficient number of single-text items need to be present to ensure that future trend items cover the entire framework.

Factors affecting item difficulty

The PISA reading literacy assessment is designed to monitor and report on the reading proficiency of 15-year-olds as they approach the end of compulsory education. Each task in the assessment is designed to gather a specific piece of evidence about that proficiency by simulating a reading activity that a reader might carry out either inside or outside school, as an adolescent or as an adult.

The PISA reading literacy tasks range from straightforward locating and comprehension activities to more sophisticated activities requiring the integration of information across multiple pieces of text. Drawing on Kirsch and Mosenthal's work (Kirsch, 2001[22]; Kirsch and Mosenthal, 1990[23]), task difficulty can be manipulated through the process and text format variables. In Table 2.3 below, we outline the factors on which the difficulty of different types of tasks depend. Box 2.4 discusses how the availability of text – whether the student can see the text when answering questions about it – is related to their performance on comprehension questions.

Table 2.3. Item difficulty as a function of task and source dimensions

Single	Multiple
In **scanning and locating** tasks, the difficulty depends on the number of pieces of information that the reader needs to locate, the number of inferences the reader must make, the amount and prominence of competing information and the length and complexity of the piece of text.	The difficulty of **searching** through multiple pieces of text depends on the number of pieces of text, the complexity of the document hierarchy (depth and breadth), the reader's familiarity with the hierarchy, the amount of non-hierarchical linking, the salience of target information, the relevance of the headers and the degree of similarity between different source texts.
In **literal and inferential comprehension** tasks, the difficulty depends on the type of interpretation required (for example, making a comparison is easier than finding a contrast); the number of pieces of information to be considered and the distance among them; the degree and prominence of competing information in the text; and the nature of the text (the longer, less familiar and the more abstract the content and organisation of ideas, the more difficult the task is likely to be).	In tasks involving multiple documents, the difficulty of making **inferences** depends on the number of pieces of text, the relevance of the headers, the similarity of the content between the pieces of text (e.g. between the arguments and points of view), and the similarity of the physical presentation/structure of the sources.
In **reflecting on content and form** tasks, the difficulty depends on the nature of the knowledge that the reader needs to bring to the piece of text (a task is more difficult if the reader needs to draw on narrow, specialised knowledge rather than broad and common knowledge); on the abstraction and length of the piece of text; and on the depth of understanding of the piece of text required to complete the task. For **assessing quality and credibility** tasks, the difficulty depends on whether the credentials and intention of the author are explicit or left for the reader to guess, and whether the text genre (e.g., a commercial message vs. a public health statement) is clearly marked.	In tasks involving multiple documents, the difficulty of tasks requiring readers to **corroborate or handle conflict** is likely to increase with the number of pieces of text, the dissimilarity of the content or arguments across texts, differences in the amount of information available about the sources, its physical presentation, and organisation.
In **scanning and locating** tasks, the difficulty depends on the number of pieces of information that the reader needs to locate, the number of inferences the reader must make, the amount and prominence of competing information and the length and complexity of the piece of text.	The difficulty of **searching** through multiple pieces of text depends on the number of pieces of text, the complexity of the document hierarchy (depth and breadth), the reader's familiarity with the hierarchy, the amount of non-hierarchical linking, the salience of target information, the relevance of the headers and the degree of similarity between different source texts.
In **literal and inferential comprehension** tasks, the difficulty depends on the type of interpretation required (for example, making a comparison is easier than finding a contrast); the number of pieces of information to be considered and the distance among them; the degree and prominence of competing information in the text; and the nature of the text (the longer, less familiar and the more abstract the content and organisation of ideas, the more difficult the task is likely to be).	In tasks involving multiple documents, the difficulty of making **inferences** depends on the number of pieces of text, the relevance of the headers, the similarity of the content between the pieces of text (e.g. between the arguments and points of view), and the similarity of the physical presentation/structure of the sources.

Box 2.4. Text availability and its impact on comprehension

In the last decade, there has been some debate as to whether memory-based measures of reading comprehension, i.e. answering comprehension question while the text is not available to students after initial reading, might be a better indicator of students' reading comprehension skills than answering questions with the text available. Answering comprehension questions with the text by one's side might be more ecologically valid because many reading settings (especially in the digital age) allow the reader to refer back to the text. In addition, if the text is not available to students, their performance on the comprehension questions might be confounded with their ability to remember the content of the text. On the other hand, answering comprehension questions when the text is no longer available is also a common situation (e.g. answering questions during a class session about a textbook chapter that was read the evening before). Empirical studies (Ozuru et al., 2007[87]; Schroeder, 2011[88]) provide some evidence that comprehension questions without text availability might be more sensitive to the quality of the processes that are executed while students are reading a text and the strength of the resulting memory representation. At the same time, however, both measures are highly correlated and are thus difficult to dissociate empirically. At present, therefore, there is not enough evidence to justify any major changes in the way the PISA reading assessment is administered. However, to further explore this issue, future cycles of PISA could consider measuring the time spent during the initial reading of a piece of text, the time spent re-reading the text when answering questions, and the total time spent on a task.

Factors improving the coverage of the domain

Situations

Scenarios can be developed to simulate a wide range of potential reading situations. The word "situation" is primarily used to define the contexts and uses for which the reader engages with the text. Most often, contexts of use match specific text genres and author purposes. For instance, textbooks are typically written for students and used by students in educational contexts. Therefore, the situation generally refers to both the context of use and the supposed audience and purpose of the text. Some situations, however, involve the use of texts that belong to various genres, such as when a history student works from both a first-hand account of an event (e.g., a personal diary, a court testimony) and a scholarly essay written long after the event (Wineburg, 1991[89]).

The framework categorises situations using a typology adapted from the Common European Framework of Reference (CEFR) developed for the Council of Europe. The situations may be personal, public, occupational or educational; these terms are defined in Box 2.5.

Box 2.5. Categorisation of situations

A personal situation is intended to satisfy an individual's personal interests, both practical and intellectual. This category also includes leisure or recreational activities that are intended to maintain or develop personal connections with other people through a range of text genres such as personal letters, fiction, biography and informational texts (e.g., a gardening guide). In the electronic medium, they include reading personal e-mails, instant messages and diary-style blogs.

A public situation is one that relates to the activities and concerns of the larger society. This category makes use of official documents as well as information about public events. In general, the texts associated with this category involve more or less anonymous contact with others; therefore, they also include message boards, news websites and public notices that are encountered both on line and in print.

Educational situations make use of texts designed specifically for the purpose of instruction. Printed textbooks, electronic textbooks and interactive learning software are typical examples of material generated for this kind of reading. Educational reading normally involves acquiring information as part of a larger learning task. The materials are often not chosen by the reader but are instead assigned by an instructor.

A typical occupational reading situation is one that involves the accomplishment of some immediate task. The task could be to find a job, either in a print newspaper's classified advertisement section or online; or it could be following workplace directions. Texts written for these purposes, and the tasks based on them, are classified as occupational in PISA. While only some of the 15-year-olds who are assessed are currently working, it is important to include tasks based on work-related texts since the assessment of young people's readiness for life beyond compulsory schooling and their ability to use their knowledge and skills to meet real-life challenges is a fundamental goal of PISA.

Many texts used in classrooms are not specifically designed for classroom use. For example, a piece of literary text may typically be read by a 15-year-old in a mother-tongue language or literature class, yet the text was written (presumably) for readers' personal enjoyment and appreciation. Given its original purpose, such a text is classified as being of a personal situation in PISA. As Hubbard (1989[90]) has shown, some kinds of reading usually associated with out-of-school settings for children, such as rules for clubs and records of games, often take place informally at school as well. These are classified as public situations in PISA. Conversely, textbooks are read both in schools and at home, and the process and purpose probably differ little from one setting to another. These are classified as educational situations in PISA.

It should be further emphasised that many texts can be cross-classified as pertaining to different situations. In practice, for example, a piece of text may be intended both to delight and to instruct (personal and educational); or to provide professional advice, which is also general information (occupational and public). The intent of sampling texts of a variety of situations is to maximise the diversity of content that will be included in the PISA reading literacy test.

Text types

The construct of text type refers both to the intent and the internal organisation of a text. Major text types include: description, narration, exposition, argumentation, instruction and transaction (Meyer and Rice, 1984[91]).[2] Real-world texts tend to cut across text type categories are typically difficult to categorise. For example, a chapter in a textbook might include some definitions (exposition), some directions on how to solve particular problems (instruction), a brief historical account of the discovery of the solution (narration) and descriptions of some typical objects involved in the solution (description). Nevertheless, in an assessment like PISA, it is useful to categorise texts according to text type, based on the predominant characteristics of the text, in order to ensure that a range of types of reading is represented.

The classification of text types used in PISA 2018 is adapted from the work of Werlich (1976[92]) and is shown in Box 2.6. Again, many texts can be cross-classified as belonging to multiple text types.

Box 2.6. Classification of text types

Description texts are texts where the information refers to properties of objects in space. Such texts typically provide an answer to *what* questions. Descriptions can take several forms. Impressionistic descriptions present subjective impressions of relations, qualities and directions in space. Technical descriptions, on the other hand, are objective observations in space. Technical descriptions frequently use non-continuous text formats such as diagrams and illustrations. Examples of description-type text objects are a depiction of a particular place in a travelogue or diary, a catalogue, a geographical map, an online flight schedule and a description of a feature, function or process in a technical manual.

The information in *narration texts* refer to properties of objects in time. Narration texts typically answer questions relating to *when*, *in what sequence*, and *why* characters in stories behave as they do. Narration can take different forms. *Narratives* record actions and events from a subjective point of view. *Reports* record actions and events from an objective point of view, one which can be verified by others. *News stories* intend to enable readers to form their own independent opinion of facts and events without being influenced by the reporter's own views. Examples of *narration*-type text objects are novels, short stories, plays, biographies, comic strips and newspaper reports of events.

Exposition texts present information as composite concepts or mental constructs, or those elements through which such concepts or mental constructs can be analysed. The text provides an explanation of how the different elements interrelate and form a meaningful whole and often answers questions about *how*. Expositions can take various forms. *Expository essays* provide a simple explanation of concepts, mental constructs or experiences from a subjective point of view. *Definitions* relate terms or names to mental concepts, thereby explaining their meaning. *Explications* explain how a mental concept can be linked with words or terms. The concept is treated as a composite whole that can be understood by breaking it down into its constituent elements and then listing the relationships between those elements. *Summaries* explain and communicate texts in a shorter form than the original text. *Minutes* are a record of the results of meetings or presentations. *Text interpretations* explain the abstract concepts that are discussed in a particular (fictional or non-fictional) piece of text or group of texts. Examples of *exposition*-type text objects are scholarly essays, diagrams showing a model of how a biological system (e.g. the heart) functions, graphs of population trends, concept maps and entries in an online encyclopaedia.

Argumentation texts present the relationship among concepts or propositions. Argumentative texts often answer why questions. An important subclassification of argumentation texts is persuasive and opinionative texts, referring to opinions and points of view. *Comments* relate events, objects and ideas to a private system of thoughts, values and beliefs. *Scientific argumentation* relates events, objects and ideas to systems of thought and knowledge so that the resulting propositions can be verified as valid or non-valid. Examples of argumentation-type text objects might be letters to the editor, poster advertisements, posts in an online forum and web-based reviews of books or films.

Instruction texts, also sometimes called *injunction* texts, provide directions on what to do. *Instructions* are the directions to complete a task. *Rules*, *regulations* and *statutes* specify certain behaviours. Examples of instruction-type text objects are recipes, a series of diagrams showing a first-aid procedure and guidelines for operating digital software.

Transaction texts aim to achieve a specific purpose, such as requesting that something be done, organising a meeting or making a social engagement with a friend. Before the spread of electronic communication, the act of transaction was a significant component of some kinds of letters and the principal purpose of many phone calls. Werlich's categorisation (1976[92]), used until now in the PISA framework, did not include transaction texts.

The term "transaction" is used in PISA not to describe the general process of extracting meaning from texts (as in reader-response theory), but the type of text written for the kinds of purposes described here. Transactional texts are often personal in nature, rather than public, and this may help to explain why they do not appear to be represented in many text typologies. For example, this kind of text is not commonly found on websites (Santini, 2006[93]). With the ubiquity of e-mails, text messages, blogs and social networking websites today as means of personal communication, transactional text has become much more significant as a reading text type in recent years. Transactional texts often build on the possibly private knowledge and understanding common to those involved in the transaction – though clearly, such prior relationships are difficult to replicate in a large-scale assessment. Examples of transaction-type text objects are everyday e-mail and text message exchanges between colleagues or friends that request and confirm arrangements.

Narration-type texts occupy a prominent position in many national and international assessments. Some such texts are presented as accounts of the world as it is (or was) and therefore claim to be factual or non-fictional. Fictional accounts bear a more metaphorical relation to the world, presenting it as how it might be or of how it seems to be. In other large-scale reading studies, particularly those for school students (the National Assessment of Educational Progress [NAEP]; the IEA Reading Literacy Study [IEARLS] and the IEA Programme in International Reading Literacy Study [PIRLS]), the major classification of texts is between fictional or literary texts and non-fictional texts (*reading for literary experience and reading for information or to perform a task in NAEP; for literary experience and to acquire and use information in PIRLS*). This distinction is becoming increasingly blurred as authors use formats and structures typical of factual texts when writing fiction. The PISA reading literacy assessment includes both factual and fictional texts, and texts that may not be clearly one or the other. PISA, however, does not attempt to measure differences in reading proficiency between factual and fictional texts. In PISA, fictional texts are classified as narration-type texts.

Response formats

The form in which evidence of student ability is collected – the *response format* – varies depending on the kinds of evidence that is being collected, and also according to the pragmatic constraints of a large-scale assessment. As in any large-scale assessment, the range of feasible item formats is limited. However, computer-based assessment makes possible response formats that involve interactions with text, such as highlighting and dragging-and-dropping. Computer-based assessments can also include multiple choice and short constructed-response items (to which students write their own answer), just as paper-based assessments do.

Students with different characteristics might perform differently with different response formats. For example, closed and some multiple-choice items are typically more dependent on decoding skills than open constructed-response items, because readers have to decode distractors or items (Cain and Oakhill, 2006[94]). Several studies based on PISA data suggest that the response format has a significant effect on the performance of different groups: for example, students at different levels of proficiency (Routitsky and Turner, 2003[95]), students in different countries (Grisay and Monseur, 2007[96]), students with different levels of intrinsic reading motivation (Schwabe, McElvany and Trendtel, 2015[97]), and boys and girls (Lafontaine and Monseur, 2006[98]; 2006[99]; Schwabe, McElvany and Trendtel, 2015[97]). Given this variation, it is important to maintain a similar proportion of response formats in the items used in each PISA cycle so as to measure trends over time.

A further consideration in the reading literacy assessment is that open constructed-response items are particularly important to assess the reflecting and evaluating process, where the intent is often to assess the quality of a student's thinking rather than the student's final response itself. Nevertheless, because the focus of the assessment is on reading and not on *writing*, constructed response items should not be designed to put great emphasis on assessing writing skills such as spelling and grammar (see Box 2.7 for more on the place of writing skills in the reading literacy assessment). Finally, various response formats are not equally familiar to students in different countries. Including items in a variety of formats is therefore likely to provide all students, regardless of nationality, the opportunity to see both familiar and less familiar formats.

In summary, to ensure proper coverage of the ability ranges, to ensure fairness given the inter-country and gender differences observed and to ensure a valid assessment of the *reflecting and evaluating* process, both multiple choice and open constructed-response items continue to be used in PISA reading literacy assessments regardless of the change in delivery mode. Any major change in the distribution of item types from that used in the paper-based reading assessment would also impact the measurement of trends.

Box 2.7. The status of writing skills in the PISA 2018 reading literacy assessment

Readers are often required to write comments, explanations or essays in response to questions, and they might choose to make notes, outlines and summaries, or simply write down their thoughts and reflections about texts, while achieving their reading goals. They also routinely engage in written communication with others (e.g. teachers, fellow students or acquaintances) for educational reasons (e.g. to e-mail an assignment to a teacher) or for social reasons (e.g. to chat with peers about text or in other school literacy contexts). The PISA 2018 reading framework considers writing to be an important correlate of reading literacy. However, test design and administration constraints prohibit the inclusion of an assessment of writing skills, where writing is in part defined as the quality and organization of the production. However, a significant proportion of test items require readers to articulate their thinking into written answers. Thus, the assessment of reading skills also draws on readers' ability to communicate their understanding in writing, although such aspects as spelling, quality of writing and organization are not measured in PISA.

Assessing ease and efficiency

The PISA 2018 reading literacy assessment will include an assessment of reading fluency, defined as the ease and efficiency with which students can read simple texts for understanding. This will provide a valuable indicator for describing and understanding differences between students, especially those in the lower reading proficiency levels. Students with low levels of foundational reading skills may be exerting so much attention and cognitive effort on lower-level skills of decoding, word recognition and sentence parsing that they have fewer resources left to perform higher-level comprehension tasks with single or multiple texts. This finding applies to young as well as to teenage readers (Rasinski et al., 2005[100]; Scammacca et al., 2007[101]).

The computerised administration and scoring in PISA 2018 allows for the measurement of ease and efficiency with which 15-year-olds can read texts accurately with understanding. While not all slow reading is poor reading, as noted above, a large body of evidence documents how and why a lack of automaticity in one's basic reading processes can be a bottleneck to higher-level reading proficiency and is associated with poor comprehension (Rayner et al., 2001[102]). Thus, it is valuable to have an indicator of ease and efficiency to better describe and interpret very low-level performance on PISA reading comprehension tasks. A basic indicator of reading rate under low-demand conditions can also be used for other purposes, such as investigating how much students regulate their reading rate or strategic processes in the face of more complex tasks or larger volumes of text.

It is further worth noting that with the exponential expansion of text content available on the Internet, there is an ever greater need for 21st century students to be not only proficient readers, but also efficient readers (OECD, 2011[103]). While there are many ways to define, operationalise and measure reading ease, efficiency or fluency, the most common method when using silent reading tasks, those where the reader does not read aloud, are indicators of accuracy and rate. Oral reading fluency measures, where the reader does read aloud, can also be used to estimate prosody and expressiveness of the reader, but unfortunately, it is currently infeasible to implement and score oral reading fluency in all the languages in which PISA is administered. Furthermore, it is not yet established whether such oral reading fluency measures add value to the silent reading indicators of accuracy and rate (Eason et al., 2013[104]; Kuhn, Schwanenflugel and Meisinger, 2010[105]). Thus, a silent reading task design is most feasible for a PISA administration.

In order to better understand the challenges facing 15-year-olds scoring at lower levels on the PISA reading literacy assessment, a specific task can be administered near the start of the assessment to measure reading ease and efficiency. Performance on this task can be scaled and reported independently from the main proficiency scale. As noted above, inefficient reading can be a symptom of low foundational skills. However, some individuals, such as non-native speakers of the assessment language, might be relatively slow readers and yet possess compensatory or strategic processes that permit them to be higher-level readers when given sufficient time to complete complex tasks. Thus, it seems most prudent to use the ease of reading indicator as a descriptive variable to help differentiate students who may have foundational skill deficits from those who are slow, but nonetheless proficient readers.

In addition, a measure of reading ease and efficiency could be, as one of several indicators, used to place students in a level for adaptive testing (see section below on "Considerations for adaptive testing"). However, for the reasons cited in the previous paragraph, the measure would not be suitable as the sole indicator of reading level.

One task that has been used effectively as an indicator of reading ease and efficiency in other surveys requires students to read a sentence and make a judgment of the plausibility of the sentence based on general knowledge or the internal logical consistency of the sentence. The measure takes into account both the accuracy of the student's understanding of the text and the time it takes to read and respond. This task has been used in the Woodcock Johnson Subtest of Reading Fluency (Woodcock, McGrew and Mather, 2001[62]) and the Test of Silent Reading Efficiency and Comprehension (TOSREC) (Wagner et al., 2010[60]). It is also used in the PIAAC Reading Components task set

(OECD, 2013[13]; Sabatini and Bruce, 2009[106]). A similar task was used in the Austrian PISA 2000 assessment and was highly correlated (r = .64) with the final test score (Landerl and Reiter, 2002[107]). Task 1 in Appendix B shows a sample reading ease and efficiency item taken from the PIAAC Reading Components task set.

In PISA 2018, data from complex reading literacy tasks will not be used to measure reading fluency. The design and instructions accompanying reading fluency tasks should specifically target the reading fluency construct. The texts therefore need to be simple and short so that students do not make use of strategic or compensatory processes when responding to questions. In addition, the task demands should require minimal reasoning so as to not confound decision time with reading time. The more complex the task, the less likely it is that it evaluates solely reading fluency.

However, it is recommended that the log files from this cycle be analysed to evaluate whether there are indicators within the new PISA Reading Literacy task set that are strongly correlated with the sentence-level efficiency task.

Assessing students' reading motivation, reading practices and awareness of reading strategies

Since PISA 2000, the reading literacy framework has highlighted the importance of readers' motivational attributes (such as their attitude toward reading) and reading practices (e.g. the reader factors in Figure 2.1); accordingly, items and scales have been developed to measure these constructs in the student questionnaire. It is important to note that reading motivation and reading strategies may vary with the context and type of text. Therefore, questionnaire items assessing motivation and reading strategies should refer to a range of situations that students may find themselves in. In addition to being more relevant, items referring to more specific and concrete situations are known to decrease the risk of response bias that comes with ratings and self-reports.

Intrinsic motivation and interest in reading

"While *motivation* refers to goals, values, and beliefs in a given area, such as reading, *engagement* refers to behavioural displays of effort, time, and persistence in attaining desired outcomes" (Klauda and Guthrie, 2015, p. 240[108]). A number of studies have shown that reading engagement, motivation and practices are strongly linked with reading proficiency (Becker, McElvany and Kortenbruck, 2010[109]; Guthrie et al., 1999[110]; Klauda and Guthrie, 2015[108]; Mol and Bus, 2011[111]; Morgan and Fuchs, 2007[112]; Pfost, Dörfler and Artelt, 2013[113]; Schaffner, Philipp and Schiefele, 2016[114]; Schiefele et al., 2012[115]). In PISA 2000, engagement in reading (comprising interest, intrinsic motivation, avoidance and practices) was strongly correlated with reading proficiency, even more so than socio-economic status was (OECD, 2002[116]; 2010[117]). In other studies, reading engagement has been shown to explain reading achievement more than any other variable besides previous reading achievement (Guthrie and Wigfield, 2000[118]). Furthermore, perseverance as a characteristic of engagement has also been linked to successful learning and achievement outside of school (Heckman and Kautz, 2012[119]). Thus, motivation and engagement are powerful variables and levers on which one can act in order to enhance reading proficiency and reduce gaps between groups of students.

During the previous PISA cycles in which reading literacy was the major domain (PISA 2000 and PISA 2009), the main motivational construct investigated was *interest in reading and intrinsic motivation.* The scale measuring interest and intrinsic motivation also captured *reading avoidance*, which is a lack of interest or motivation and is strongly and negatively associated with achievement, especially among struggling readers (Klauda and Guthrie, 2015[108]; Legault, Green-Demers and Pelletier, 2006[120]). For PISA 2018, in accordance with what was done in mathematics and science, two other motivational constructs will be investigated as part of the PISA questionnaire: *self-efficacy,* an individual's perceived capacity of performing specific tasks, and *self-concept,* an individual's own perceived abilities in a domain.

Reading practices

Reading practices were previously measured as *the self-reported frequencies of reading different types of texts in various media, including online reading*. In PISA 2018, the list of online reading practices will be updated and extended in order to take into account emerging practices (such as e-books, online search, short messaging and social networking).

Awareness of reading strategies

Metacognition is an individual's ability to think about and control his or her reading and comprehension strategies. A number of studies have found an association between reading proficiency and metacognitive strategies (Artelt, Schiefele and Schneider, 2001[121]; Brown, Palincsar and Armbruster, 1984[122]). Explicit or formal instruction of reading strategies leads to an improvement in text understanding and information use (Cantrell et al., 2010[123]). More specifically, it is assumed that the reader becomes independent of the teacher after these strategies have been acquired and can be applied without much effort. By using these strategies, the reader can effectively interact with the text by conceiving of reading as a problem-solving task that requires the use of strategic thinking to accomplish reading comprehension.

In previous PISA cycles, engagement and metacognition proved to be robust predictors of reading achievement, mediators of gender or socioeconomic status (OECD, 2010[124]) and also potential levers to reduce achievement gaps. The measures of motivational, metacognition and reader practices have been updated and extended in the questionnaire in order to take into account recent and emerging practices in reading as well as to better measure the teaching practices and the classroom support that support reading growth.

Skilled reading requires students to know and employ strategies in order to optimise the knowledge they gain from a piece of text given their purposes and goals. For instance, students must know when it is appropriate to scan a passage or when the task requires the sustained and complete reading of the passage. PISA 2009 collected information about reading strategies through two reading scenarios. In the first scenario, students were asked to evaluate the effectiveness of different reading and text comprehension strategies to fulfil the goal of *summarising information*; in the second, students had to evaluate the effectiveness of other strategies for *understanding and remembering a text*. In accordance with the new characterisation of reading processes (Figure 2.2), PISA 2018 will also collect information about knowledge of reading strategies specifically linked to the goal of *assessing the quality and credibility of sources*, which is particularly prominent in digital reading and when reading multiple pieces of text.

Teaching practices and classroom support for reading growth and engagement

There is strong research evidence showing that classroom practices, such as the direct teaching of reading strategies, contribute to growth in reading skills (Pressley, 2000[125]; Rosenshine and Meister, 1997[126]; Waters and Schneider, 2010[127]). In addition, teachers' scaffolding and support for autonomy, competence and ownership of their tasks improve students' reading proficiency, awareness of strategies, and engagement in reading (Guthrie, Klauda and Ho, 2013[39]; Guthrie, Wigfield and You, 2012[38]). While in most education systems, reading is no longer taught as a subject matter to 15-year-old students in the same way that mathematics and science are, some reading instruction may be explicitly or incidentally given in language lessons and in other disciplines (such as social science, science, foreign languages, civic education or ICT). The dispersed nature of reading instruction represents a challenge for articulating questionnaire items that measure the classroom practices and opportunities to learn reading strategies that best support the development of students' reading skills, practices and motivation.

Considerations for adaptive testing

The deployment of computer-based assessment in PISA creates the opportunity to implement adaptive testing. Adaptive testing enables higher levels of measurement precision using fewer items per individual student. This is accomplished by presenting students with items that are aligned to their ability level.

Adaptive testing has the potential to increase the resolution and sensitivity of the assessment, most particularly at the lower end of the distribution of student performance. For example, students who perform poorly on items that assess their reading fluency will likely struggle on highly complex multiple text items. Future cycles of PISA could provide additional lower-level texts to those students to better assess specific aspects of their comprehension.

Reporting proficiency in reading

Reporting scales

PISA reports students' results through proficiency scales that can be interpreted in educational policy terms. In PISA 2000, when reading was the major domain, the results of the reading literacy assessment were first summarised on a single composite reading literacy scale with a mean of 500 and a standard deviation of 100. In addition to the composite scale, student performance was also represented on five subscales: three process (aspect) subscales (retrieving information, interpreting texts, and reflection and evaluation) and two text format subscales (continuous and non-continuous) (Kirsch et al., 2002[8]). These five subscales made it possible to compare mean scores and distributions among subgroups and countries in each of the components of the reading literacy construct. Although these subscales were highly correlated, there were interesting differences across subscales. Such differences could be examined and linked to the curriculum and teaching methodology in the countries concerned. Reading was again the major domain in PISA 2009, which again reported a composite scale as well as subscales.

In PISA 2003, 2006, 2012 and 2015, when reading was a minor domain and fewer reading items were administered to participating students, a single reading literacy scale was reported based on the overall composite scale (OECD, 2004[128]; OECD, 2007[129]; OECD, 2014[130]). In 2018, reading is again the major domain, and reporting on subscales is again possible.

For PISA 2018, the reporting subscales will be (see also Table 2.1):

1) Locating information, which is composed of tasks that require students to search for and select relevant texts, and access relevant information within texts;

2) Understanding, which is composed of tasks that require students to represent the explicit meaning of texts as well as integrate information and generate inferences; and

3) Evaluating and reflecting, which is composed of tasks that require the student to assess the quality and credibility of information, reflect on the content and form of a text and detect and handle conflict within and across texts.

Interpreting and using the scales

Just as students can be ordered from least to most proficient on a single scale, reading literacy tasks are arranged along a scale that indicates their level of difficulty and the level of skill required to answer the task correctly.

Reading literacy tasks used in PISA vary widely in situation, text format and task requirements, and they also vary in difficulty. This range is captured through what is known as an item map. The task map provides a visual representation of the reading literacy skills demonstrated by students at different points along the scale.

Difficulty is in part determined by the length, structure and complexity of the text itself. However, what the reader has to do with that text, as defined by the question or instruction, also affects the overall difficulty. A number of variables that can influence the difficulty of any reading literacy task have been identified, including the complexity and sophistication of the mental processes integral to the task process (retrieving, interpreting or reflecting), the amount of information the reader needs to assimilate and the familiarity or specificity of the knowledge that the reader must draw on both from within and from outside the text.

Defining levels of reading literacy proficiency

In an attempt to capture this progression of complexity and difficulty in PISA 2000, the composite reading literacy scale and each of the subscales were divided into six levels (Below Level 1, 1, 2, 3, 4 and 5). These levels as they were defined for PISA 2000 were kept for the composite scale used to measure trends in PISA 2009 and 2015. However, newly constructed items helped to improve descriptions of the existing levels of performance and to furnish descriptions of levels of performance above and below those established in PISA 2000. Thus, the scales were extended to Level 6, and Level 1 was divided into Levels 1a and 1b (OECD, 2012[30]).

The levels provide a useful way to explore the progression of reading literacy demands within the composite scale and each subscale. The scale summarises both the proficiency of a person in terms of his or her ability and the complexity of an item in terms of its difficulty. The mapping of students and items on one scale represents the idea that students are more likely to be able to successfully complete tasks mapped at the same level on the scale (or lower), and less likely to be able to successfully complete tasks mapped at a higher level on the scale.

As an example, the reading proficiency scale for the PISA 2015 study is presented in Table 2.4. The left-hand column shows, for each proficiency level, the lower score boundary (i.e. all students who score at or above this boundary perform at this proficiency level or higher), and the percentage of students who are able to perform tasks at the level (on average across OECD countries). The right-hand column, as adapted from OECD (2013[13]), describes what students can do at each level.

Table 2.4. An overview of reading proficiency levels adapted from the descriptions in PISA 2015

Level	What Students Can Do
6 698 1.1%	Readers at Level 6 typically can make multiple inferences, comparisons and contrasts that are both detailed and precise. They demonstrate a full and detailed understanding of one or more texts and may integrate information from more than one text. Tasks may require the reader to deal with unfamiliar ideas in the presence of prominent competing information, and to generate abstract categories for interpretations. Students can hypothesise about or critically evaluate a complex text on an unfamiliar topic, taking into account multiple criteria or perspectives and applying sophisticated understandings from beyond the text. A salient condition for accessing and retrieving tasks at this level is the precision of analysis and fine attention to detail that is inconspicuous in the texts.
5 626 8.4%	At Level 5, readers can locate and organise several pieces of deeply embedded information, inferring which information in the text is relevant. Reflective tasks require critical evaluation or hypothesis-making, drawing on specialised knowledge. Both interpreting and reflecting tasks require a full and detailed understanding of a text whose content or form is unfamiliar. For all aspects of reading, tasks at this level typically involve dealing with concepts that are contrary to expectations.
4 553 29.5%	At Level 4, readers can locate and organise several pieces of embedded information. They can also interpret the nuances of language in a section of text by taking into account the text as a whole. In other interpreting tasks, students demonstrate understanding and application of categories in an unfamiliar context. In addition, students at this level can use formal or public knowledge to hypothesise about or critically evaluate a text. Readers must demonstrate an accurate understanding of long or complex texts whose content or form may be unfamiliar.
3 480 58.6%	Readers at Level 3 can locate, and in some cases recognise the relationship between, several pieces of information that must meet multiple conditions. They can also integrate several parts of a text in order to identify a main idea, understand a relationship or construe the meaning of a word or phrase. They need to take into account many features in comparing, contrasting or categorising. Often the required information is not prominent or there is much competing information; or there are other text obstacles, such as ideas that are contrary to expectations or negatively worded. Reflecting tasks at this level may require connections, comparisons, and explanations, or they may require the reader to evaluate a feature of the text. Some reflecting tasks require readers to demonstrate a fine understanding of the text in relation to familiar, everyday knowledge. Other tasks do not require detailed text comprehension but require the reader to draw on less common knowledge.
2 407 82.0%	Readers at Level 2 can locate one or more pieces of information, which may need to be inferred and may need to meet several conditions. They can recognize the main idea in a text, understand relationships, or construe meaning within a limited part of the text when the information is not prominent and the reader must make low-level inferences. Tasks at this level may involve comparisons or contrasts based on a single feature in the text. Typical reflecting tasks at this level require readers to make a comparison or several connections between the text and outside knowledge, by drawing on personal experience and attitudes.
1a 335 94.3%	Readers at Level 1a can locate one or more independent pieces of explicitly stated information; they can recognise the main theme or author's purpose in a text about a familiar topic, or make a simple connection between information in the text and common, everyday knowledge. Typically, the required information in the text is prominent and there is little, if any, competing information. The student is explicitly directed to consider relevant factors in the task and in the text.
1b 262 98.7%	Readers at Level 1b can locate a single piece of explicitly stated information in a prominent position in a short, syntactically simple text with a familiar context and text type, such as a narrative or a simple list. Texts in Level 1b tasks typically provide support to the reader, such as repetition of information, pictures or familiar symbols. There is minimal competing information. Level 1b readers can interpret texts by making simple connections between adjacent pieces of information.

Given that the top of the reading literacy scale currently has no bounds, there is arguably some uncertainty about the upper limits of proficiency of extremely high-performing students. However, such students can still be described by the scale as performing tasks at the highest level of proficiency. There is a greater issue for students at the bottom end of the reading literacy scale. Although it is possible to measure the reading proficiency of students performing below Level 1b, at this stage their proficiency – what they can do – cannot be described. In developing new material for PISA 2018, items were designed to measure reading skill and understanding located at or below the current Level 1b.

Notes

[1] Some dynamic navigation features were incidentally included in the 2015 assessment. This was a result of the adaptation of trend materials, which were formerly presented in print, for screen presentation. Many of these so-called fixed texts were used in previous cycles. Although they were adapted to mimic the printed texts as closely as possible, they had to be reformatted to the smaller screen size typical of computer displays. Therefore, tabs and other very simple navigation tools were included to let the reader navigate from one page to another.

[2] In the PISA 2000 reading framework, these text types were subcategories of the continuous text format. In the PISA 2009 cycle, it was acknowledged that non-continuous texts (and the elements of mixed and multiple texts) also have a descriptive, narrative, expository, argumentative or instructional purpose.

References

American Press Institute (2014), *How Americans get their news*, http://www.americanpressinstitute.org/wp-content/uploads/2014/03/The_Media_Insight_Project_The_Personal_News_Cycle_Final.pdf. [32]

Ananiadou, K. and M. Claro (2009), "21st Century Skills and Competences for New Millennium Learners in OECD Countries", *OECD Education Working Papers*, No. 41, OECD Publishing, Paris, http://dx.doi.org/10.1787/218525261154. [7]

Artelt, C., U. Schiefele and W. Schneider (2001), "Predictors of reading literacy", *European Journal of Psychology of Education*, Vol. XVI/3, pp. 363-383, http://dx.doi.org/10.1007/BF03173188. [121]

Becker, M., N. McElvany and M. Kortenbruck (2010), "Intrinsic and extrinsic reading motivation as predictors of reading literacy: A longitudinal study", *Journal of Educational Psychology*, Vol. 102/4, pp. 773-785, http://dx.doi.org/10.1037/a0020084. [109]

Binkley, M. et al. (2011), "Defining Twenty-First Century Skills", in *Assessment and Teaching of 21st Century Skills*, Springer Netherlands, Dordrecht, http://dx.doi.org/10.1007/978-94-007-2324-5_2. [1]

Binkley, M., K. Rust and T. Williams (eds.) (1997), *Reading literacy in an international perspective*, U.S. Department of Education, Washington D.C. [41]

Brand-Gruwel, S., I. Wopereis and Y. Vermetten (2005), "Information problem solving by experts and novices: Analysis of a complex cognitive skill", *Computers in Human Behavior*, Vol. 21, pp. 487–508, http://dx.doi.org/10.1016/j.chb.2004.10.005. [66]

Bråten, I. et al. (2011), "The role of epistemic beliefs in the comprehension of multiple expository texts: Toward an integrated model", *Educational Psychologist*, Vol. 46/1, pp. 48-70, http://dx.doi.org/10.1080/00461520.2011.538647. [77]

Bråten, I., H. Strømsø and M. Britt (2009), "Trust Matters: Examining the Role of Source Evaluation in Students' Construction of Meaning within and across Multiple Texts", *Reading Research Quarterly*, Vol. 44/1, pp. 6-28, http://dx.doi.org/10.1598/RRQ.41.1.1. [74]

Britt, M., S. Goldman and J. Rouet (eds.) (2013), *Reading--From words to multiple texts*, Routledge, New York. [14]

Britt, M. and J. Rouet (2012), "Learning with multiple documents: Component skills and their acquisition.", in Kirby, J. and M. Lawson (eds.), *Enhancing the Quality of Learning: Dispositions, Instruction, and Learning Processes*, Cambridge University Press, New York, http://dx.doi.org/10.1017/CBO9781139048224.017. [46]

Brown, A., A. Palincsar and B. Armbruster (1984), "Instructing Comprehension-Fostering Activities in Interactive Learning Situations", in Mandl, H., N. Stein and T. Trabasso (eds.), *Learning and Comprehension of Text*, Lawrence Erlbaum Associates, Hillsdale, NJ. [122]

Brozo, W. and M. Simpson (2007), *Content Literacy for Today's Adolescents: Honoring Diversity and Building Competence*, Merrill/Prentice Hall, Upper Saddle River, NJ. [37]

Cain, K. and J. Oakhill (2008), *Children's Comprehension Problems in Oral and Written Language*, Guilford Press, New York. [64]

Cain, K. and J. Oakhill (2006), "Assessment matters: Issues in the measurement of reading comprehension", *British Journal of Educational Psychology*, Vol. 76, pp. 697–708, http://dx.doi.org/10.1348/000709905X69807. [94]

Cantrell, S. et al. (2010), "The impact of a strategy-based intervention on the comprehension and strategy use of struggling adolescent readers", *Journal of Educational Psychology*, Vol. 102/2, pp. 257-280, http://dx.doi.org/10.1037/a0018212. [123]

Chard, D., J. Pikulski and S. McDonagh (2006), "Fluency: The Link between Decoding and Comprehension for Struggling Readers", in Rasinski, T., C. Blachowicz and K. Lems (eds.), *Fluency Instruction: Research-Based Best Practices*, Guilford Press, New York. [59]

Clark, C. (2014), *Children's and Young People's Writing in 2013: Findings from the National Literacy Trust's Annual Literacy Survey*, National Literacy Trust, London, http://www.literacytrust.org.uk. [33]

Coiro, J. et al. (2008), "Central Issues in New Literacies and New Literacies Research", in Coiro, J. et al. (eds.), *Handbook of Research on New Literacies*, Lawrence Erlbaum Associates, New York. [47]

Conklin, J. (1987), "Hypertext: An Introduction and Survey", *Computer*, Vol. 20, pp. 17-41, http://dx.doi.org/10.1109/MC.1987.1663693. [27]

Cunningham, A. and K. Stanovich (1997), "Early reading acquisition and its relation to reading experience and ability 10 years later", *Developmental Psychology*, Vol. 33/6, pp. 934-945. [2]

Dillon, A. (1994), *Designing Usable Electronic Text Ergonomics Aspects of Human Information Usage*, Taylor & Francis, London. [26]

Dreher, M. and J. Guthrie (1990), "Cognitive processes in textbook chapter search tasks", *Reading Research Quarterly*, Vol. 25/4, pp. 323-339, http://dx.doi.org/10.2307/747694. [67]

Duggan, G. and S. Payne (2009), "Text skimming: the process and effectiveness of foraging through text under time pressure", *Journal of Experimental Psychology: Applied*, Vol. 15/3, pp. 228-242, http://dx.doi.org/10.1037/a0016995. [70]

Eason, S. et al. (2013), "Examining the Relationship Between Word Reading Efficiency and Oral Reading Rate in Predicting Comprehension Among Different Types of Readers", *Scientific Studies of Reading*, Vol. 17/3, pp. 199-223, http://dx.doi.org/10.1080/10888438.2011.652722. [104]

Foltz, P. (1996), "Comprehension, Coherence and Strategies in Hypertext and Linear Text", in Rouet, J. et al. (eds.), *Hypertext and Cognition*, Lawrence Erlbaum Associates, Hillsdale, NJ. [28]

Gartner (2014), *Forecast: PCs, Ultramobiles and Mobile Phones, Worldwide, 2011-2018, 4Q14 Update*, https://www.gartner.com/doc/2945917/forecast-pcs-ultramobiles-mobile-phones. [31]

Gerjets, P., Y. Kammerer and B. Werner (2011), "Measuring spontaneous and instructed evaluation processes during Web search: Integrating concurrent thinking-aloud protocols and eye-tracking data.", *Learning and Instruction*, Vol. 21/2, pp. 220-231, http://dx.doi.org/10.1016/j.learninstruc.2010.02.005. [71]

Goldman, S. (2004), "Cognitive aspects of constructing meaning through and across multiple texts", in Shuart-Faris, N. and D. Bloome (eds.), *Uses of intertextuality in classroom and educational research*, Information Age Publishing, Greenwich, CT. [55]

Gray, W. and B. Rogers (1956), *Maturity in Reading: Its nature and appraisal*, The University of Chicago Press, Chicago, IL. [48]

Grisay, A. and C. Monseur (2007), "Measuring the Equivalence of item difficulty in the various versions of an international test", *Studies in Educational Evaluation*, Vol. 33, pp. 69-86, http://dx.doi.org/10.1016/j.stueduc.2007.01.006. [96]

Guthrie, J., S. Klauda and A. Ho (2013), "Modeling the relationships among reading instruction, motivation, engagement, and achievement for adolescents", *Reading Research Quarterly*, Vol. 48/1, pp. 9-26, http://dx.doi.org/10.1002/rrq.035. [39]

Guthrie, J. and A. Wigfield (2000), "Engagement and motivation in reading.", in Kamil, M. et al. (eds.), *Handbook of reading research*, Lawrence Erlbaum Associates , Mahwah, NJ. [118]

Guthrie, J. et al. (1999), "Motivational and Cognitive Predictors of Text Comprehension and Reading Amount", *Scientific Studies of Reading*, http://dx.doi.org/10.1207/s1532799xssr0303_3. [110]

Guthrie, J., A. Wigfield and W. You (2012), "Instructional Contexts for Engagement and Achievement in Reading", in Christenson, S., A. Reschly and C. Wylie (eds.), *Handbook of Research on Student Engagement*, Springer Science, New York, http://dx.doi.org/10.1007/978-1-4614-2018-7. [38]

Hacker, D. (1998), "Self-regulated comprehension during normal reading", in Hacker, D., J. Dunlosky and A. Graesser (eds.), *The Educational Psychology Series. Metacognition in Educational Theory and Practice*, Lawrence Erlbaum Associates, Mahwah, NJ. [79]

Heckman, J. and T. Kautz (2012), "Hard Evidence on Soft Skills", *Discussion Paper Series*, No. 6580, The Institute for the Study of Labor , Bonn, Germany. [119]

Hofstetter, C., T. Sticht and C. Hofstetter (1999), "Knowledge, Literacy, and Power", *Communication Research*, Vol. 26/1, pp. 58-80, http://dx.doi.org/10.1177/009365099026001004. [49]

Hubbard, R. (1989), "Notes from the Underground: Unofficial Literacy in One Sixth Grade", *Anthropology & Education Quarterly*, Vol. 20/4, pp. 291-307, http://www.jstor.org/stable/3195740. [90]

International Telecommunications Union (2014), *Key 2005-2014 ICT data for the world, by geographic regions and by level of development [Excel file]*, http://www.itu.int/en/ITU-D/Statistics/Pages/publications/mis2014.aspx. [6]

International Telecommunications Union (2014), *Measuring the Information Society Report 2014*, International Telecommunication Union, Geneva. [5]

Jenkins, J. et al. (2003), "Sources of Individual Differences in Reading Comprehension and Reading Fluency.", *Journal of Educational Psychology*, Vol. 95/4, pp. 719-729, http://dx.doi.org/10.1037/0022-0663.95.4.719. [63]

Kamil, M. et al. (eds.) (2000), *Handbook of Reading Research. Volume III*, Lawrence Erlbaum Associates, Mahwah, NJ. [15]

Kintsch, W. (1998), *Comprehension: A paradigm for cognition*, Cambridge University Press, Cambridge, MA. [20]

Kirsch, I. (2001), *The International Adult Literacy Survey (IALS): Understanding What Was Measured*, Educational Testing Service, Princeton, NJ. [22]

Kirsch, I. et al. (2002), *Reading for Change Performance and Engagement Across Countries: Results From PISA 2000*, OECD, Paris. [8]

Kirsch, I. and P. Mosenthal (1990), "Exploring Document Literacy: Variables Underlying the Performance of Young Adults", *Reading Research Quarterly*, Vol. 25/1, pp. 5-30, http://dx.doi.org/10.2307/747985. [23]

Klauda, S. and J. Guthrie (2015), "Comparing relations of motivation, engagement, and achievement among struggling and advanced adolescent readers", *Reading and Writing*, Vol. 28, pp. 239–269, http://dx.doi.org/10.1007/s11145-014-9523-2. [108]

Kuhn, M., P. Schwanenflugel and E. Meisinger (2010), "Aligning Theory and Assessment of Reading Fluency: Automaticity, Prosody, and Definitions of Fluency", *Reading Research Quarterly*, Vol. 45/2, pp. 230–251, http://dx.doi.org/10.1598/RRQ.45.2.4. [105]

Kuhn, M. and S. Stahl (2003), "Fluency: A review of developmental and remedial practices", *Journal of Educational Psychology*, Vol. 95/1, pp. 3–21, http://dx.doi.org/10.1037/0022-0663.95.1.3. [58]

Lafontaine, D. and C. Monseur (2006), *Impact of item choice on the measurement of trends in educational achievement*, Paper presented at the 2006 Annual AERA meeting, San Francisco, CA, https://convention2.allacademic.com/one/aera/aera06/index.php?click_key=1&cmd=Multi+Search+Search+Load+Publication&publication_id=49435&PHPSESSID=h5edlqu9juanjmlaaa6ece3e74. [98]

Lafontaine, D. and C. Monseur (2006), *Impact of test characteristics on gender equity indicators in the Assessment of Reading Comprehension*, University of Liège, Liège. [99]

Landerl, K. and C. Reiter (2002), "Lesegeschwindigkeit als Indikator für basale Lesefertigkeiten", in Wallner-Paschon, C. and G. Haider (eds.), *PISA Plus 2000. Thematische Analysen nationaler Projekte*, Studien Verlag, Innsbruck. [107]

Legault, L., I. Green-Demers and L. Pelletier (2006), "Why do high school students lack motivation in the classroom? Toward an understanding of academic a motivation and the role of social support", *Journal of Educational Psychology*, Vol. 98/3, pp. 567-582, http://dx.doi.org/dx.doi.org/10.1037/0022-0663.98.3.567. [120]

Leu, D. et al. (2015), "The new literacies of online research and comprehension: Rethinking the reading achievement gap", *Reading Research Quarterly*, Vol. 50/1, pp. 37–59, http://dx.doi.org/10.1002/rrq.85. [12]

Leu, D. et al. (2013), "New Literacies: A dual-level theory of the changing nature of literacy instruction and assessment", in Alvermann, D., Norman J. Unrau and R. Ruddell (eds.), *Theoretical Models and Processes of Reading*, International Reading Association, Newark. [11]

Lundberg, I. (1991), "Reading as an individual and social skill", in Lundberg, I. and T. Höien (eds.), *Literacy in a world of change: perspective on reading and reading disability; proceedings*, Literacy Conference at Stavanger Forum, Center for Reading Research/UNESCO. [51]

Mason, L., A. Boldrin and N. Ariasi (2010), "Searching the Web to Learn about a Controversial Topic: Are Students Epistemically Active?", *Instructional Science: An International Journal of the Learning Sciences*, Vol. 38/6, pp. 607-633. [72]

McCrudden, M. and G. Schraw (2007), "Relevance and goal-focusing in text processing", *Educational Psychology Review*, Vol. 19/2, pp. 113–139, http://dx.doi.org/10.1007/s10648-006-9010-7. [52]

McNamara, D. and J. Magliano (2009), "Toward a Comprehensive Model of Comprehension", *The Psychology of Learning and Motivation*, Vol. 51, pp. 297-384, http://dx.doi.org/10.1016/S0079-7421(09)51009-2. [42]

Meyer, B. and G. Rice (1984), "The Structure of Text", in Pearson, P. et al. (eds.), *Handbook of Reading Research*, Longman, New York. [91]

Mol, S. and A. Bus (2011), "To read or not to read: a meta-analysis of print exposure from infancy to early adulthood", *Psychological Bulletin*, Vol. 137/2, pp. 267-96, http://dx.doi.org/10.1037/a0021890. [111]

Moore, P. (1995), "Information Problem Solving: A Wider View of Library Skills", *Contemporary Educational Psychology*, Vol. 20/1, pp. 1-31, http://dx.doi.org/10.1006/ceps.1995.1001. [68]

Morgan, P. and D. Fuchs (2007), "Is there a bidirectional relationship between children's reading skills and reading motivation?", *Exceptional Children*, Vol. 73/2, pp. 165-183, http://dx.doi.org/10.1177%2F001440290707300203. [112]

Morrisroe, J. (2014), *Literacy Changes Lives 2014: A new perspective on health, employment and crime*, National Literacy Trust, London, http://www.literacytrust.org.uk. [50]

Naumann, J. (2015), "A model of online reading engagement: Linking engagement, navigation, and performance in digital reading", *Computers in Human Behavior*, Vol. 53, pp. 263–277, http://dx.doi.org/10.1016/j.chb.2015.06.051. [36]

Oakhill, J., K. Cain and P. Bryant (2003), "The dissociation of word reading and text comprehension: Evidence from component skills", *Language and Cognitive Processes*, Vol. 18/4, pp. 443-468, http://dx.doi.org/10.1080/01690960344000008. [43]

OECD (2015), *Students, Computers and Learning: Making the Connection*, OECD, Paris, http://dx.doi.org/10.1787/19963777. [35]

OECD (2014), *PISA 2012 Results: What Students Know and Can Do (Volume I, Revised edition, February 2014): Student Performance in Mathematics, Reading and Science*, PISA, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264208780-en. [130]

OECD (2013), *OECD Skills Outlook 2013: First Results from the Survey of Adult Skills*, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264204256-en. [3]

OECD (2013), *PISA 2015 Draft Frameworks*, OECD, Paris, http://www.oecd.org/pisa/pisaproducts/pisa2015draftframeworks.htm. [13]

OECD (2012), *OECD Internet Economy Outlook 2012*, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264086463-en. [30]

OECD (2011), "Do Students Today Read for Pleasure?", *PISA in Focus*, No. 8, OECD Publishing, Paris, http://dx.doi.org/10.1787/5k9h362lhw32-en. [103]

OECD (2011), *PISA 2009 Results: Students On Line. Digital Technologies and Performance (Volume VI)*, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264112995-en. [25]

OECD (2010), *PISA 2009 Assessment Framework*, OECD, Paris, http://dx.doi.org/10.1787/19963777. [117]

OECD (2010), *PISA 2009 Results: Learning to Learn: Student Engagement, Strategies and Practices (Volume III)*, PISA, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264083943-en. [124]

OECD (2007), *PISA 2006: Science Competencies for Tomorrow's World: Volume 1: Analysis*, PISA, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264040014-en. [129]

OECD (2004), *Learning for Tomorrow's World: First Results from PISA 2003*, PISA, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264006416-en. [128]

OECD (2002), *Reading for Change: Performance and Engagement across Countries: Results from PISA 2000*, PISA, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264099289-en. [116]

OECD (2000), *Measuring Student Knowledge and Skills. The PISA 2000 Assessment of Reading, Mathematical and Scientific Literacy*, OECD, Paris. [24]

O'Reilly, T. and J. Sabatini (2013), *Reading for Understanding: How Performance Moderators and Scenarios Impact Assessment Design*, ETS Research Report RR-13-31, http://www.ets.org/Media/Research/pdf/RR-13-31.pdf. [86]

Ozuru, Y. et al. (2007), "Influence of question format and text availability on the assessment of expository text comprehension", *Cognition and Instruction*, Vol. 25/4, pp. 399-438, http://dx.doi.org/10.1080/07370000701632371. [87]

Perfetti, C. (2007), "Reading ability: Lexical quality to comprehension", *Scientific Studies of Reading*, Vol. 11/4, pp. 357-383, http://dx.doi.org/10.1080/10888430701530730. [17]

Perfetti, C. (1985), *Reading Ability*, Oxford University Press, New York. [16]

Perfetti, C., M. Marron and P. Foltz (1996), "Sources of Comprehension Failure: Theoretical Perspectives and Case Studies", in Cornoldi, C. and J. Oakhill (eds.), *Reading Comprehension Difficulties: Processes and Intervention*, Lawrence Erlbaum Associates, Mahwah, NJ. [65]

Perfetti, C., J. Rouet and M. Britt (1999), "Toward Theory of Documents Representation", in van Oostendorp, H. and S. Goldman (eds.), *The Construction of Mental Representations During Reading*, Lawrence Erlbaum Associates, Mahwah, NJ. [56]

Pfost, M., T. Dörfler and C. Artelt (2013), "Students' extracurricular reading behavior and the development of vocabulary and reading comprehension", *Learning and Individual Differences*, Vol. 26, pp. 89-102, http://dx.doi.org/10.1016/j.lindif.2013.04.008. [113]

Pressley, M. (2000), "What Should Comprehension Instruction be the Instruction Of?", in Kamil, M. et al. (eds.), *Handbook of reading research, Volume III*, Lawrence Erlbaum Associates, Mahwah, NJ. [125]

Rasinski, T. et al. (2005), "Is Reading Fluency a Key for Successful High School Reading?", *Journal of Adolescent & Adult Literacy*, Vol. 49/1, pp. 22-27, http://dx.doi.org/10.1598/JAAL.49.1.3. [100]

Rayner, K. et al. (2001), "How Psychological Science Informs the Teaching of Reading", *Psychological Science in the Public Interest*, Vol. 2/2, pp. 31-74, http://dx.doi.org/10.1111/1529-1006.00004. [102]

Rayner, K. and E. Reichle (2010), "Models of the Reading Process", *Wiley Interdisciplinary Reviews. Cognitive Science*, Vol. 1/6, pp. 787-799, http://dx.doi.org/10.1002/wcs.68. [19]

Reeve, J. (2012), "A Self-determination Theory Perspective on Student Engagement", in *Handbook of research on student engagement*, Springer, Boston, MA, http://dx.doi.org/10.1007/978-1-4614-2018-7_7. [40]

Richter, T. and D. Rapp (2014), “Comprehension and Validation of Text Information: Introduction to the Special Issue”, *Discourse Processes*, Vol. 51/1-2, pp. 1-6, http://dx.doi.org/10.1080/0163853X.2013.855533. [45]

Rieh, S. (2002), “Judgment of Information Quality and Cognitive Quthority in the Web”, *Journal of the Association for Information Science and Technology*, Vol. 53/2, pp. 145-161, http://dx.doi.org/10.1002/asi.10017. [73]

Rosenshine, B. and C. Meister (1997), “Cognitive strategy instruction in reading”, in Stahl, S. and D. Hayes (eds.), *Instructional Models in Reading*, Lawrence Erlbaum Associates, Mahwah, NJ. [126]

Rouet, J. (2006), *The Skills of Document Use: From Text Comprehension to Web-based Learning*, Lawrence Erlbaum Associates, Mahwah, NJ. [9]

Rouet, J. and M. Britt (2014), “Multimedia Learning from Multiple Documents”, in Mayer, R. (ed.), *Cambridge handbook of multimedia learning*, Cambridge University Press, Cambridge, MA. [78]

Rouet, J. and M. Britt (2011), “Relevance Processes in Multiple Document Comprehension”, in McCrudden, M., J. Magliano and G. Schraw (eds.), *Text Relevance and Learning from Text*, Information Age, Greenwich, CT. [53]

Rouet, J. and B. Coutelet (2008), “The Acquisition of Document Search Strategies in Grade School Students”, *Applied Cognitive Psychology*, Vol. 22, pp. 389–406, http://dx.doi.org/10.1002/acp.1415. [69]

Rouet, J. and J. Levonen (1996), “Studying and Learning with Hypertext: Empirical Studies and Their Implications”, in Rouet, J. et al. (eds.), *Hypertext and Cognition*, Lawrence Erlbaum Associates, Mahwah, NJ. [29]

Rouet, J., Z. Vörös and C. Pléh (2012), “Incidental Learning of Links during Navigation: The Role of Visuo-Spatial Capacity”, *Behaviour & Information Technology*, Vol. 31/1, pp. 71-81. [82]

Routitsky, A. and R. Turner (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), April 2003. [95]

Rupp, A., T. Ferne and H. Choi (2006), “How Assessing Reading Comprehension With Multiple-choice Questions Shapes the Construct: A Cognitive Processing Perspective”, *Language Testing*, Vol. 23/4, pp. 441–474, http://dx.doi.org/10.1191/0265532206lt337oa. [83]

Sabatini, J. and K. Bruce (2009), “PIAAC Reading Components: A Conceptual Framework”, *OECD Education Working Papers*, No. 33, OECD, Paris, http://www.oecd.org/edu/workingpapers. [106]

Sabatini, J. et al. (2014), “Broadening the Scope of Reading Comprehension Using Scenario-Based Assessments: Preliminary findings and challenges”, *L’Année Psychologique*, Vol. 114/4, pp. 693-723, http://dx.doi.org/10.4074/S0003503314004059. [84]

Sabatini, J. et al. (2015), "Improving Comprehension Assessment for Middle and High School Students: Challenges and Opportunities", in Santi, K. and D. Reed (eds.), *Improving Reading Comprehension of Middle and High School Students (Literacy Studies)*, Springer, New York. [85]

Santini, M. (2006), "Web pages, Text types, and Linguistic Features: Some Issues", *International Computer Archive of Modern and Medieval English (CAME)*, Vol. 30, pp. 67-86. [93]

Scammacca, N. et al. (2007), *Interventions for Adolescent Struggling Readers: A Meta-Analysis With Implications for Practice*, Center on Instruction at RMC Research Corporation, Portsmouth, NH, http://www.centeroninstruction.org. [101]

Schaffner, E., M. Philipp and U. Schiefele (2016), "Reciprocal Effects Between Intrinsic Reading Motivation and Reading Competence? A Cross-lagged Panel Model for Academic Track and Nonacademic Track Students", *Journal of Research in Reading*, Vol. 39/1, pp. 19–36, http://dx.doi.org/10.1111/1467-9817.12027. [114]

Schiefele, U. et al. (2012), "Dimensions of Reading Motivation and Their Relation to Reading Behavior and Competence", *Reading Research Quarterly*, Vol. 47/4, pp. 427–463, http://dx.doi.org/10.1002/RRQ.030. [115]

Schroeder, S. (2011), "What readers have and do: Effects of students' verbal ability and reading time components on comprehension with and without text availability", *Journal of Educational Psychology*, Vol. 103/4, pp. 877-896, http://dx.doi.org/10.1037/a0023731. [88]

Schwabe, F., N. McElvany and M. Trendtel (2015), "The School Age Gender Gap in Reading Aachievement: Examining the Influences of Item Format and Intrinsic Reading Motivation", *Reading Research Quarterly*, Vol. 50/2, pp. 219–232, http://dx.doi.org/10.1002/rrq.92. [97]

Smith, M. et al. (2000), "What Will Be the Demands of Literacy in the Workplace in the Next Millennium?", *Reading Research Quarterly*, Vol. 35/3, pp. 378-383, http://dx.doi.org/10.1598/RRQ.35.3.3. [4]

Snow, C. and the RAND Corporation (2002), *Reading for Understanding: Toward an R and D Program in Reading Comprehension*, RAND Reading Study Group, Santa Monica, CA, http://www.rand.org/. [18]

Spiro, R. et al. (eds.) (2015), *Reading at a Crossroads? Disjunctures and Continuities in Current Conceptions and Practices*, Routledge, New York. [10]

Stadtler, M. and R. Bromme (2014), "The Content–Source Integration Model: A Taxonomic Description of How Readers Comprehend Conflicting Scientific Information", in Rapp, D. and J. Braasch (eds.), *Processing Inaccurate Information: Theoretical and Applied Perspectives from Cognitive Science and the Educational Sciences*, The MIT Press, Cambridge, MA. [75]

Stadtler, M. and R. Bromme (2013), "Multiple Document Comprehension: An Approach to Public Understanding of Science", *Cognition and Instruction*, Vol. 31/2, pp. 122-129, http://dx.doi.org/10.1080/07370008.2013.771106. [76]

Strømsø, H. et al. (2013), "Spontaneous Sourcing Among Students Reading Multiple Documents", *Cognition and Instruction*, Vol. 31/2, pp. 176-203, http://dx.doi.org/10.1080/07370008.2013.769994. [81]

UNESCO (2014), *Reading in the Mobile Era: A Study of Mobile Reading in Developing Countries*, UNESCO, Paris, http://www.unesco.org/open-access/terms-use-ccbysa-en. [34]

van den Broek, P., K. Risden and E. Husbye-Hartmann (1995), "The Role of Readers' Standards of Coherence in the Generation of Inferences During Reading", in Lorch, Jr., R. and E. O'Brien (eds.), *Sources of Coherence in Text Comprehension*, Lawrence Erlbaum Associates, Hilsdale, NJ. [44]

Vidal-Abarca, E., A. Mañá and L. Gil (2010), "Individual Differences for Self-regulating Task-oriented Reading Activities", *Journal of Educational Psychology*, Vol. 102/4, pp. 817-826, http://dx.doi.org/10.1037/a0020062. [57]

Wagner, R. et al. (2010), *Test of Silent Reading Efficiency and Comprehension*, Pro-Ed, Austin, TX. [60]

Waters, H. and W. Schneider (eds.) (2010), *Metacognition, Strategy Use, and Instruction*, Guilford Press, New York. [127]

Wayman, M. et al. (2007), "Literature Synthesis on Curriculum-Based Measurement in Reading", *The Journal of Special Education*, Vol. 41/2, pp. 85-120. [61]

Werlich, E. (1976), *A Text Grammar of English*, Quelle and Meyer, Heidelberg. [92]

White, S., J. Chen and B. Forsyth (2010), "Reading-Related Literacy Activities of American Adults: Time Spent, Task Types, and Cognitive Skills Used", *Journal of Literacy Research*, Vol. 42, pp. 276–307, http://dx.doi.org/10.1080/1086296X.2010.503552. [54]

Wineburg, S. (1991), "On the Reading of Historical Texts: Notes on the Breach Between School and Academy", *American Educational Research Joumal*, Vol. 28/3, pp. 495-519. [89]

Winne, P. and A. Hadwin (1998), "Studying as Self-regulated Learning", in Hacker, D., J. Dunlosky and A. Graesser (eds.), *Metacognition in Educational Theory and Practice*, Lawrence Erlbaum Associates, Mahwah, NJ. [80]

Woodcock, R., K. McGrew and N. Mather (2001), *Woodcock-Johnson III Tests of Achievement*, Riverside Publishing, Itasca, IL. [62]

Zwaan, R. and M. Singer (2003), "Text comprehension", in Graesser, A., M. Gernsbacher and S. Goldman (eds.), *Handbook of Discourse Processes*, Lawrence Erlbaum Associates, Mahwah, NJ. [21]

Annex 2.A. Main changes in the reading framework, 2000-2015

Annex Table 2.A.1. Main changes in the reading framework, 2000-2015

	2000	2009	2015
TEXT			
Format	Continuous, non-continuous, mixed	Same as 2000, plus multiple	Same as 2009
Type	Argumentation, description, exposition, narration, instruction	Same as 2000, plus “transactional”	Same as 2009
Environment	N/A	Authored, message-based	N/A
Medium	N/A	Print, electronic	N/A
Space	N/A	N/A	Fixed, dynamic
SITUATIONS	Educational, personal, professional, public	Same as 2000	Same as 2000
ASPECT	Accessing and retrieving, integrating and interpreting, reflecting and evaluating	Same as 2000, plus “complex”	Same as 2000

Annex 2.B. Sample tasks

Task 1: Sample reading ease and efficiency task.

The sentence-processing items are timed tasks that require the respondent to assess whether a sentence makes sense in terms of the properties of the real world or the internal logic of the sentence. The respondent reads the sentence and circles YES if the sentence makes sense or NO if the sentence does not make sense. This task is adapted from PISA 2012 and PIAAC's Reading Components sentence processing items.

Annex Figure 2.B.1. Task 1. Sample reading ease and efficiency task

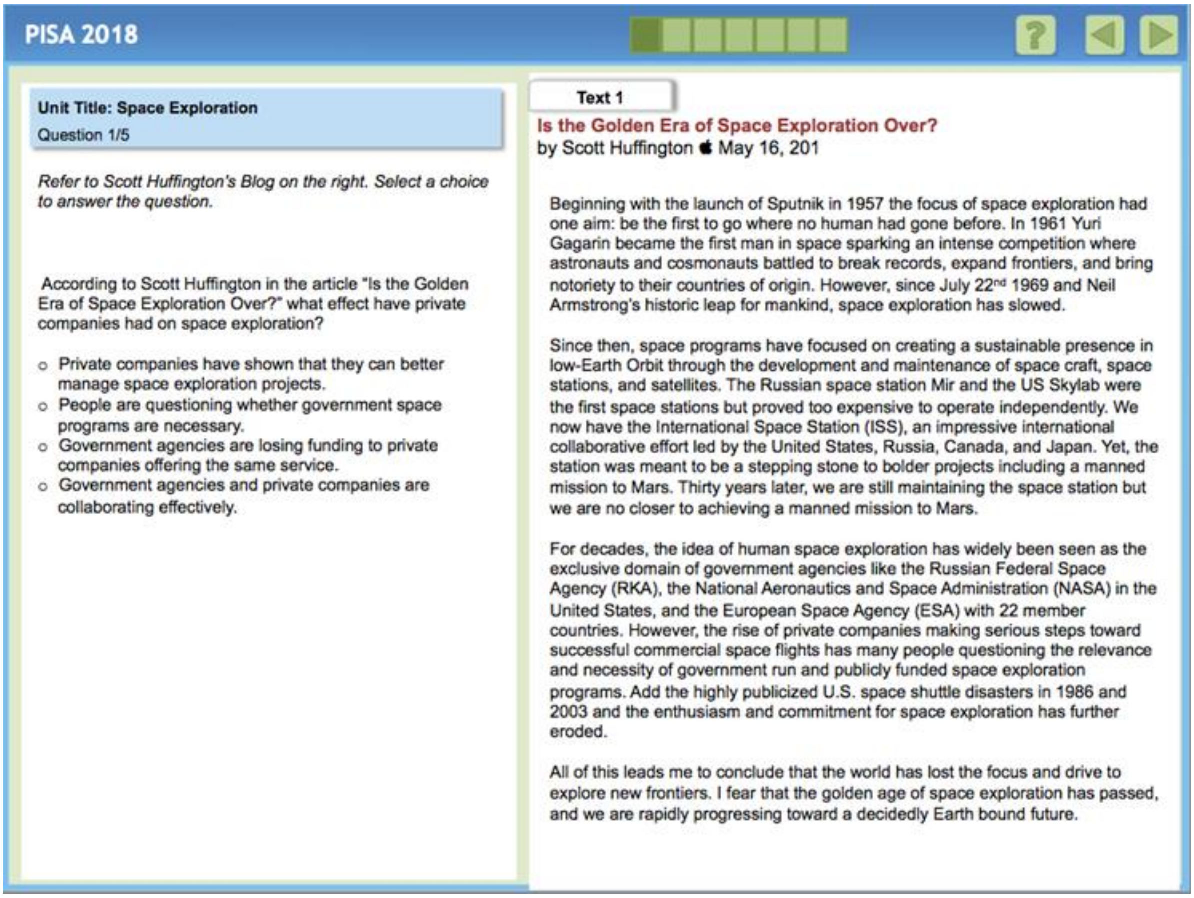
Directions: Circle **YES** if the sentence makes sense. Circle **NO** if the sentence does not make sense.

The red car had a flat tire	**YES**	**NO**
Airplanes are made of dogs	**YES**	**NO**
The happy student read the book last night	**YES**	**NO**
If the cat had stayed out all night, it would not have been in the house at 2 a.m.	**YES**	**NO**
The man who is taller than the woman and the boy is shorter than both of them	**YES**	**NO**

Tasks 2-4: Sample scenario with three embedded tasks

In this scenario, students are asked to read three sources: a blog post, the comments section that follows and an article that is referenced by one of the commenters. The articles and comments all discuss space exploration now and in the future. Students are asked to answer several questions that assess different reading processes.

Annex Figure 2.B.2. Task 2. Scanning and locating (single text)



PISA 2018

Unit Title: Space Exploration
Question 1/5

Refer to Scott Huffington's Blog on the right. Select a choice to answer the question.

According to Scott Huffington in the article "Is the Golden Era of Space Exploration Over?" what effect have private companies had on space exploration?

- Private companies have shown that they can better manage space exploration projects.
- People are questioning whether government space programs are necessary.
- Government agencies are losing funding to private companies offering the same service.
- Government agencies and private companies are collaborating effectively.

Text 1

Is the Golden Era of Space Exploration Over?
by Scott Huffington May 16, 201

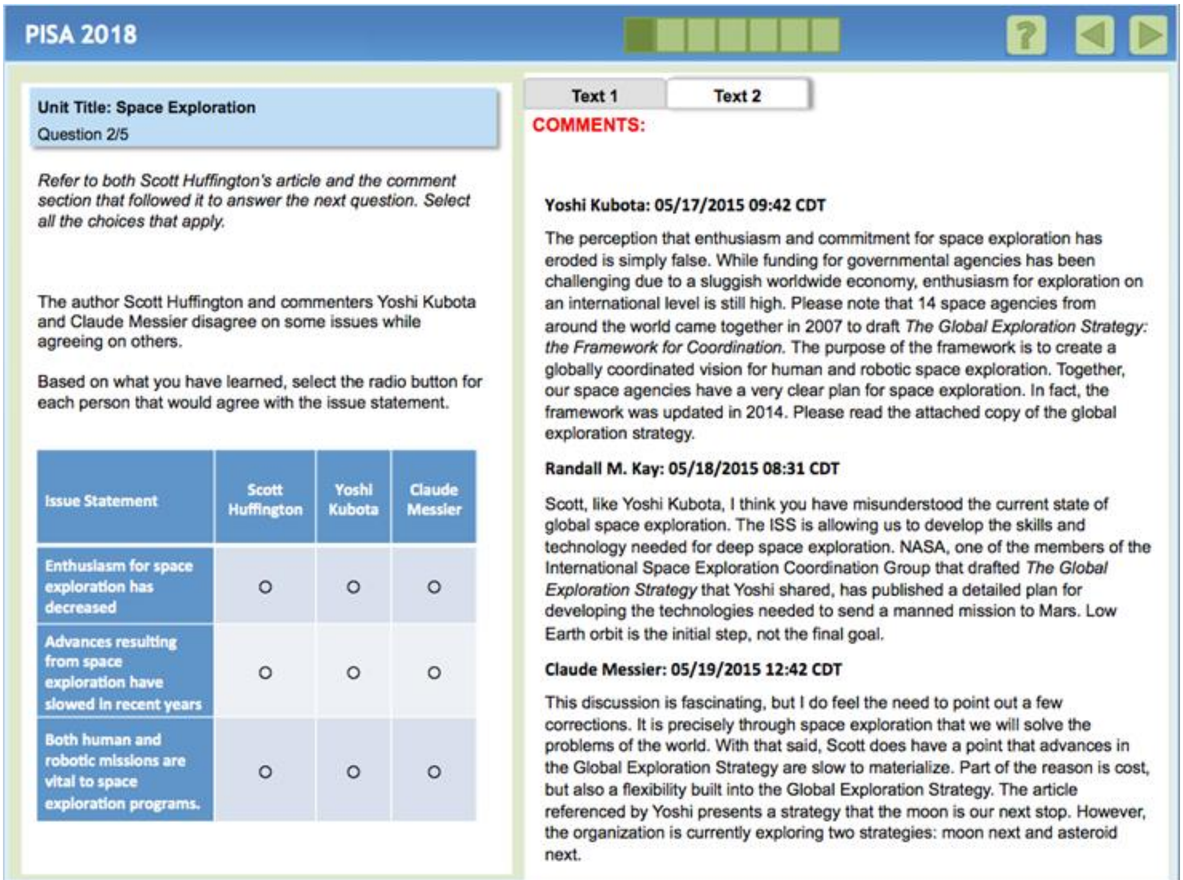
Beginning with the launch of Sputnik in 1957 the focus of space exploration had one aim: be the first to go where no human had gone before. In 1961 Yuri Gagarin became the first man in space sparking an intense competition where astronauts and cosmonauts battled to break records, expand frontiers, and bring notoriety to their countries of origin. However, since July 22nd 1969 and Neil Armstrong's historic leap for mankind, space exploration has slowed.

Since then, space programs have focused on creating a sustainable presence in low-Earth Orbit through the development and maintenance of space craft, space stations, and satellites. The Russian space station Mir and the US Skylab were the first space stations but proved too expensive to operate independently. We now have the International Space Station (ISS), an impressive international collaborative effort led by the United States, Russia, Canada, and Japan. Yet, the station was meant to be a stepping stone to bolder projects including a manned mission to Mars. Thirty years later, we are still maintaining the space station but we are no closer to achieving a manned mission to Mars.

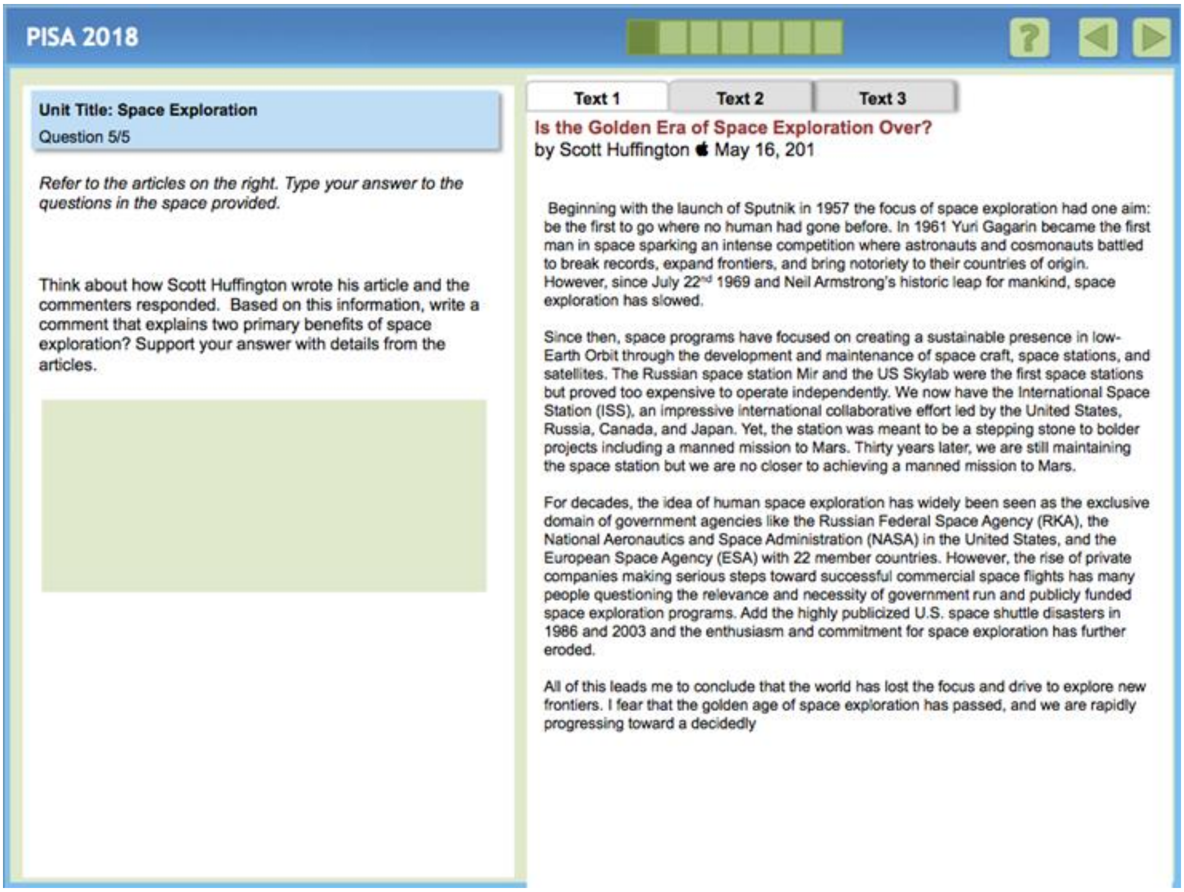
For decades, the idea of human space exploration has widely been seen as the exclusive domain of government agencies like the Russian Federal Space Agency (RKA), the National Aeronautics and Space Administration (NASA) in the United States, and the European Space Agency (ESA) with 22 member countries. However, the rise of private companies making serious steps toward successful commercial space flights has many people questioning the relevance and necessity of government run and publicly funded space exploration programs. Add the highly publicized U.S. space shuttle disasters in 1986 and 2003 and the enthusiasm and commitment for space exploration has further eroded.

All of this leads me to conclude that the world has lost the focus and drive to explore new frontiers. I fear that the golden age of space exploration has passed, and we are rapidly progressing toward a decidedly Earth bound future.

Annex Figure 2.B.3. Task 3. Multiple text inference



Annex Figure 2.B.4. Task 4. Evaluating and reflecting



Is the Golden Era of Space Exploration Over?
by Scott Huffington ● May 16, 201

Beginning with the launch of Sputnik in 1957 the focus of space exploration had one aim: be the first to go where no human had gone before. In 1961 Yuri Gagarin became the first man in space sparking an intense competition where astronauts and cosmonauts battled to break records, expand frontiers, and bring notoriety to their countries of origin. However, since July 22nd 1969 and Neil Armstrong's historic leap for mankind, space exploration has slowed.

Since then, space programs have focused on creating a sustainable presence in low-Earth Orbit through the development and maintenance of space craft, space stations, and satellites. The Russian space station Mir and the US Skylab were the first space stations but proved too expensive to operate independently. We now have the International Space Station (ISS), an impressive international collaborative effort led by the United States, Russia, Canada, and Japan. Yet, the station was meant to be a stepping stone to bolder projects including a manned mission to Mars. Thirty years later, we are still maintaining the space station but we are no closer to achieving a manned mission to Mars.

For decades, the idea of human space exploration has widely been seen as the exclusive domain of government agencies like the Russian Federal Space Agency (RKA), the National Aeronautics and Space Administration (NASA) in the United States, and the European Space Agency (ESA) with 22 member countries. However, the rise of private companies making serious steps toward successful commercial space flights has many people questioning the relevance and necessity of government run and publicly funded space exploration programs. Add the highly publicized U.S. space shuttle disasters in 1986 and 2003 and the enthusiasm and commitment for space exploration has further eroded.

All of this leads me to conclude that the world has lost the focus and drive to explore new frontiers. I fear that the golden age of space exploration has passed, and we are rapidly progressing toward a decidedly

3. PISA 2018 Mathematics Framework

This chapter defines "mathematical literacy" as assessed in the Programme for International Student Assessment (PISA) in 2018 and the competencies required for mathematical literacy. It explains the processes, content knowledge and contexts reflected in the tasks that PISA uses to measure scientific literacy, and how student performance in mathematics is measured and reported.

Introduction

In PISA 2018, mathematics is assessed as a minor domain, providing an opportunity to make comparisons of student performance over time. This framework continues the description and illustration of the PISA mathematics assessment as set out in the 2012 framework, when mathematics was re-examined and updated for use as the major domain in that cycle.

For PISA 2018, as in PISA 2015, the computer is the primary mode of delivery for all domains, including mathematical literacy. However, paper-based assessment instruments are provided for countries that choose not to test their students by computer. The mathematical literacy component for both the computer-based and paper-based instruments are composed of the same clusters of mathematics trend items. The number of trend items in the minor domains (of which mathematics is one in 2018) are increased, when compared to PISA assessments prior to 2015, therefore increasing the construct coverage while reducing the number of students responding to each question. This design is intended to reduce potential bias while stabilising and improving the measurement of trends.

The PISA 2018 mathematics framework is organised into several major sections. The first section, "Defining Mathematical Literacy," explains the theoretical underpinnings of the PISA mathematics assessment, including the formal definition of the mathematical literacy construct. The second section, "Organising the Domain of Mathematics," describes three aspects: a) the mathematical *processes* and the *fundamental mathematical capabilities* (in previous frameworks the "*competencies*") underlying those processes; b) the way mathematical *content* knowledge is organised, and the content knowledge that is relevant to an assessment of 15-year-old students; and c) the *contexts* in which students face mathematical challenges. The third section, "Assessing Mathematical Literacy", outlines the approach taken to apply the elements of the framework previously described, including the structure of the assessment, the transfer to a computer-based assessment and reporting proficiency. The 2012 framework was written under the guidance of the 2012 Mathematics Expert Group (MEG), a body appointed by the main PISA contractors with the approval of the PISA Governing Board (PGB). The ten MEG members included mathematicians, mathematics educators, and experts in assessment, technology, and education research from a range of countries. In addition, to secure more extensive input and review, a draft of the PISA 2012 mathematics framework was circulated for feedback to over 170 mathematics experts from over 40 countries. Achieve and the Australian Council for Educational Research (ACER), the two organisations contracted by the Organisation for Economic Co-operation and Development (OECD) to manage framework development, also conducted various research efforts to inform and support development work. Framework development and the PISA programme generally have been supported and informed by the ongoing work of participating countries, as in the research described in OECD (2010[1]). The PISA 2015 framework was updated under the guidance of the mathematics expert group (MEG), a body appointed by the Core 1 contractor with the approval of the PISA Governing Board (PGB). There are no substantial changes to the mathematics framework between PISA 2015 and PISA 2018.

In PISA 2012, mathematics (the major domain) was delivered as a paper-based assessment, while the computer-based assessment of mathematics (CBAM) was an optional domain that was not taken by all countries. As a result, CBAM was not part of the mathematical literacy trend. Therefore, CBAM items developed for PISA 2012 are not included in the 2015 and 2018 assessments where mathematical literacy is a minor domain, despite the change in delivery mode to computer-based.

The framework was updated for PISA 2015 to reflect the change in delivery mode, and includes a discussion of the considerations of transposing paper items to a screen and examples of what the results look like. The definition and constructs of mathematical literacy however, remain unchanged and consistent with those used in 2012.

Defining mathematical literacy

An understanding of mathematics is central to a young person's preparedness for life in modern society. A growing proportion of problems and situations encountered in daily life, including in professional contexts, require some level of understanding of mathematics, mathematical reasoning and mathematical tools, before they can be fully understood and addressed. It is thus important to understand the degree to which young people emerging from school are adequately prepared to apply mathematics in order to understand important issues and to solve meaningful problems. An assessment at age 15 – near the end of compulsory education – provides an early indication of how individuals may respond in later life to the diverse array of situations they will encounter that involve mathematics.

The construct of mathematical literacy used in this report is intended to describe the capacities of individuals to reason mathematically and use mathematical concepts, procedures, facts and tools to describe, explain and predict phenomena. This conception of mathematical literacy supports the importance of students developing a strong understanding of concepts of pure mathematics and the benefits of being engaged in explorations in the abstract world of mathematics. The construct of mathematical literacy, as defined for PISA, strongly emphasises the need to develop students' capacity to use mathematics in context, and it is important that they have rich experiences in their mathematics classrooms to accomplish this. In PISA 2012, mathematical literacy was defined as shown in Box 3.1. This definition is also used in the PISA 2015 and 2018 assessments.

Box 3.1. The 2012 definition of mathematical literacy

Mathematical literacy is an individual's capacity to formulate, employ and interpret mathematics in a variety of contexts. It includes reasoning mathematically and using mathematical concepts, procedures, facts and tools to describe, explain and predict phenomena. It assists individuals to recognise the role that mathematics plays in the world and to make the well-founded judgements and decisions needed by constructive, engaged and reflective citizens.

The focus of the language in the definition of mathematical literacy is on active engagement in mathematics, and is intended to encompass reasoning mathematically and using mathematical concepts, procedures, facts and tools in describing, explaining and predicting phenomena. In particular, the verbs "formulate", "employ", and "interpret" point to the three processes in which students as active problem solvers will engage.

The language of the definition is also intended to integrate the notion of mathematical modelling, which has historically been a cornerstone of the PISA framework for mathematics (OECD, 2004[2]), into the PISA 2012 definition of mathematical literacy. As individuals use mathematics and mathematical tools to solve problems in contexts, their work progresses through a series of stages (individually developed later in the document).

The modelling cycle is a central aspect of the PISA conception of students as active problem solvers; however, it is often not necessary to engage in every stage of the modelling cycle, especially in the context of an assessment (Blum, Galbraith and Niss, 2007, pp. 3-32[3]). The problem solver frequently carries out some steps of the modelling cycle but not all of them (e.g. when using graphs), or goes around the cycle several times to modify earlier decisions and assumptions.

The definition also acknowledges that mathematical literacy helps individuals to recognise the role that mathematics plays in the world and in helping them make the kinds of well-founded judgements and decisions required of constructive, engaged and reflective citizens.

Mathematical tools mentioned in the definition refer to a variety of physical and digital equipment, software and calculation devices. The 2015 and 2018 computer-based assessments include an online calculator as part of the test material provided for some questions.

Organising the domain of mathematics

The PISA mathematics framework defines the domain of mathematics for the PISA survey and describes an approach to assessing the mathematical literacy of 15-year-olds. That is, PISA assesses the extent to which 15-year-old students can handle mathematics adeptly when confronted with situations and problems – the majority of which are presented in real-world contexts.

For purposes of the assessment, the PISA 2012 definition of mathematical literacy – also used for the PISA 2015 and 2018 cycles – can be analysed in terms of three interrelated aspects:

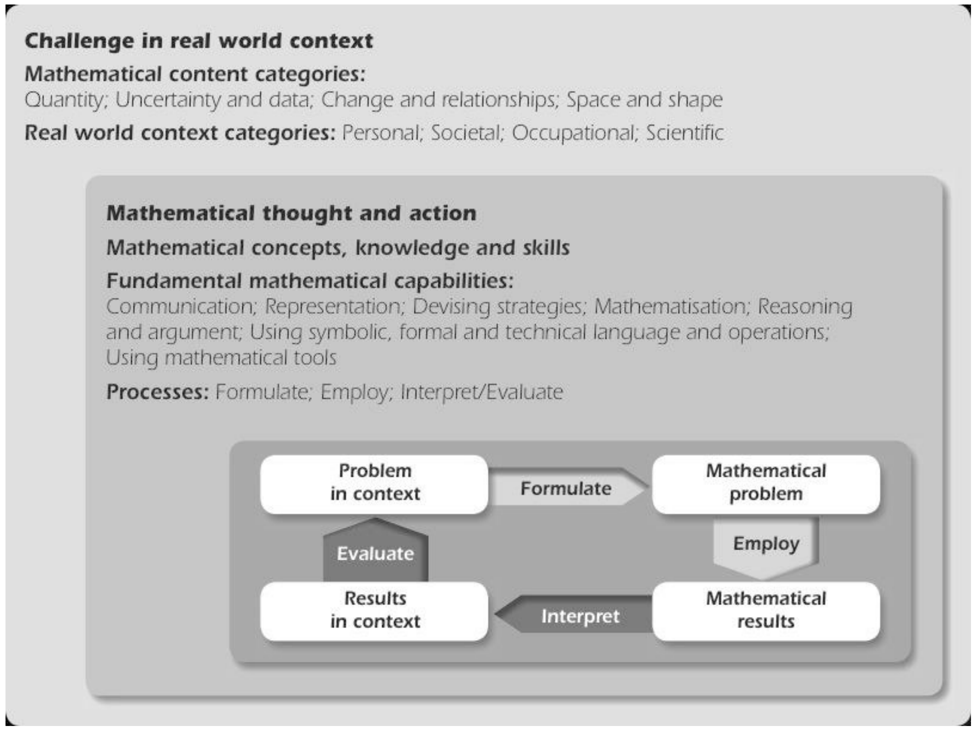
- the mathematical processes that describe what individuals do to connect the context of the problem with mathematics and thus solve the problem, and the capabilities that underlie those processes
- the mathematical content that is targeted for use in the assessment items
- the contexts in which the assessment items are located.

The following sections elaborate these aspects. In highlighting these aspects of the domain, the PISA 2012 mathematics framework, which is also used in PISA 2015 and PISA 2018, helps to ensure that assessment items developed for the survey reflect a range of processes, content and contexts, so that, considered as a whole, the set of assessment items effectively operationalises what this framework defines as mathematical literacy. To illustrate the aspects of mathematic literacy, examples are available in the PISA 2012 Assessment and Analytical Framework (OECD, 2013[4]) and on the PISA website (www.oecd.org/pisa/).

Several questions, based on the PISA 2012 definition of mathematical literacy, lie behind the organisation of this section of the framework. They are:

- What processes do individuals engage in when solving contextual mathematics problems, and what capabilities do we expect individuals to be able to demonstrate as their mathematical literacy grows?
- What mathematical content knowledge can we expect of individuals – and of 15-year-old students in particular?
- In what contexts can mathematical literacy be observed and assessed?

Figure 3.1. A model of mathematical literacy in practice



Mathematical processes and the underlying mathematical capabilities

Mathematical processes

The definition of mathematical literacy refers to an individual's capacity to formulate, employ and interpret mathematics. These three words – formulate, employ and interpret – provide a useful and meaningful structure for organising the mathematical processes that describe what individuals do to connect the context of a problem with the mathematics and thus solve the problem. Items in the PISA 2018 mathematics assessment are assigned to one of three mathematical processes:

- formulating situations mathematically
- employing mathematical concepts, facts, procedures and reasoning
- interpreting, applying and evaluating mathematical outcomes.

It is important for both policy makers and those involved more closely in the day-to-day education of students to know how effectively students are able to engage in each of these processes. The *formulating* process indicates how effectively students are able to recognise and identify opportunities to use mathematics in problem situations and then provide the necessary mathematical structure needed to formulate that contextualised problem into a mathematical form. The *employing* process indicates how well students are able to perform computations and manipulations and apply the concepts and facts that they know to arrive at a mathematical solution to a problem formulated mathematically. The *interpreting* process indicates how effectively students are able to reflect upon mathematical solutions or conclusions, interpret them in the context of a real-world problem, and determine whether the results or conclusions are reasonable. Students' facility in applying mathematics to problems and situations is dependent on skills inherent in all three of these processes, and an understanding of their effectiveness in each category can help inform both policy-level discussions and decisions being made closer to the classroom level.

Formulating situations mathematically

The word *formulate* in the definition of mathematical literacy refers to individuals being able to recognise and identify opportunities to use mathematics and then provide mathematical structure to a problem presented in some contextualised form. In the process of formulating situations mathematically, individuals determine where they can extract the essential mathematics to analyse, set up and solve the problem. They translate from a real-world setting to the domain of mathematics and provide the real-world problem with mathematical structure, representations and specificity. They reason about and make sense of constraints and assumptions in the problem. Specifically, this process of formulating situations mathematically includes activities such as the following:

- identifying the mathematical aspects of a problem situated in a real-world context and identifying the significant variables
- recognising mathematical structure (including regularities, relationships and patterns) in problems or situations
- simplifying a situation or problem in order to make it amenable to mathematical analysis
- identifying constraints and assumptions behind any mathematical modelling and simplifications gleaned from the context
- representing a situation mathematically, using appropriate variables, symbols, diagrams and standard models
- representing a problem in a different way, including organising it according to mathematical concepts and making appropriate assumptions
- understanding and explaining the relationships between the context-specific language of a problem and the symbolic and formal language needed to represent it mathematically
- translating a problem into mathematical language or a representation
- recognising aspects of a problem that correspond with known problems or mathematical concepts, facts or procedures
- using technology (such as a spreadsheet or the list facility on a graphing calculator) to portray a mathematical relationship inherent in a contextualised problem.

Employing mathematical concepts, facts, procedures and reasoning

The word *employ* in the definition of mathematical literacy refers to individuals being able to apply mathematical concepts, facts, procedures and reasoning to solve mathematically formulated problems to obtain mathematical conclusions. In the process of employing mathematical concepts, facts, procedures and reasoning to solve problems, individuals perform the mathematical procedures needed to derive results and find a mathematical solution (e.g. performing arithmetic computations, solving equations, making logical deductions from mathematical assumptions, performing symbolic manipulations, extracting mathematical information from tables and graphs, representing and manipulating shapes in space, and analysing data). They work on a model of the problem situation, establish regularities, identify connections between mathematical entities, and create mathematical arguments. Specifically, this process of employing mathematical concepts, facts, procedures and reasoning includes activities such as:

- devising and implementing strategies for finding mathematical solutions
- using mathematical tools[1], including technology, to help find exact or approximate solutions
- applying mathematical facts, rules, algorithms and structures when finding solutions
- manipulating numbers, graphical and statistical data and information, algebraic expressions and equations, and geometric representations
- making mathematical diagrams, graphs and constructions, and extracting mathematical information from them
- using and switching between different representations in the process of finding solutions
- making generalisations based on the results of applying mathematical procedures to find solutions
- reflecting on mathematical arguments and explaining and justifying mathematical results.

Interpreting, applying and evaluating mathematical outcomes

The word *interpret* used in the definition of mathematical literacy focuses on the abilities of individuals to reflect upon mathematical solutions, results, or conclusions and interpret them in the context of real-life problems. This involves translating mathematical solutions or reasoning back into the context of a problem and determining whether the results are reasonable and make sense in the context of the problem. This mathematical process category encompasses both the "interpret" and "evaluate" arrows noted in the previously defined model of *mathematical literacy* in practice (see Figure 3.1). Individuals engaged in this process may be called upon to construct and communicate explanations and arguments in the context of the problem, reflecting on both the modelling process and its results. Specifically, this process of interpreting, applying and evaluating mathematical outcomes includes activities such as:

- interpreting a mathematical result back into the real-world context
- evaluating the reasonableness of a mathematical solution in the context of a real-world problem
- understanding how the real world impacts the outcomes and calculations of a mathematical procedure or model in order to make contextual judgements about how the results should be adjusted or applied
- explaining why a mathematical result or conclusion does, or does not, make sense given the context of a problem
- understanding the extent and limits of mathematical concepts and mathematical solutions

- critiquing and identifying the limits of the model used to solve a problem.

Desired distribution of items by mathematical process

The goal in constructing the assessment is to achieve a balance that provides approximately equal weighting between the two processes that involve making a connection between the real world and the mathematical world and the process that calls for students to be able to work on a mathematically formulated problem. Table 3.1 shows the desired distribution of items by process.

Table 3.1. Desired distribution of mathematics items, by process category

Process category	Percentage of items
Formulating situations mathematically	25
Employing mathematical concepts, facts, procedures and reasoning	50
Interpreting, applying and evaluating mathematical outcomes	25
Total	**100**

Fundamental mathematical capabilities underlying the mathematical processes

A decade of experience in developing PISA items and analysing the ways in which students respond to items has revealed that there is a set of fundamental mathematical capabilities that underpins each of these reported processes and mathematical literacy in practice. The work of Mogens Niss and his Danish colleagues (Niss, 2003[5]; Niss and Jensen, 2002[6]; Niss and Højgaard, 2011[7]) identified eight capabilities – referred to as "competencies" by Niss and in the PISA 2003 framework (OECD, 2004[2]) – that are instrumental to mathematical behaviour.

The PISA 2018 framework uses a modified formulation of this set of capabilities, which condenses the number from eight to seven based on an investigation of the operation of the competencies through previously administered PISA items (Turner et al., 2013[8]). These cognitive capabilities are available to or learnable by individuals in order to understand and engage with the world in a mathematical way, or to solve problems. As the level of mathematical literacy possessed by an individual increases, that individual is able to draw to an increasing degree on the fundamental mathematical capabilities (Turner and Adams, 2012[9]). Thus, increasing activation of fundamental mathematical capabilities is associated with increasing item difficulty. This observation has been used as the basis of the descriptions of different proficiency levels of mathematical literacy reported in previous PISA surveys and discussed later in this framework.

The seven fundamental mathematical capabilities used in this framework are as follows:

- *Communication:* Mathematical literacy involves communication. The individual perceives the existence of some challenge and is stimulated to recognise and understand a problem situation. Reading, decoding and interpreting statements, questions, tasks or objects enables the individual to form a mental model of the situation, which is an important step in understanding, clarifying and formulating a problem. During the solution process, intermediate results may need to be summarised and presented. Later on, once a solution has been found, the problem solver may need to present the solution, and perhaps an explanation or justification, to others.

- *Mathematising:* Mathematical literacy can involve transforming a problem defined in the real world to a strictly mathematical form (which can include structuring, conceptualising, making assumptions, and/or formulating a model), or interpreting or evaluating a mathematical outcome or a mathematical model in relation to the original problem. The term mathematising is used to describe the fundamental mathematical activities involved.
- *Representation:* Mathematical literacy frequently involves representations of mathematical objects and situations. This can entail selecting, interpreting, translating between, and using a variety of representations to capture a situation, interact with a problem, or to present one's work. The representations referred to include graphs, tables, diagrams, pictures, equations, formulae and concrete materials.
- *Reasoning and argument:* This capability involves logically rooted thought processes that explore and link problem elements so as to make inferences from them, check a justification that is given, or provide a justification of statements or solutions to problems.
- *Devising strategies for solving problems:* Mathematical literacy frequently requires devising strategies for solving problems mathematically. This involves a set of critical control processes that guide an individual to effectively recognise, formulate and solve problems. This skill is characterised as selecting or devising a plan or strategy to use mathematics to solve problems arising from a task or context, as well as guiding its implementation. This mathematical capability can be demanded at any of the stages of the problem-solving process.
- *Using symbolic, formal and technical language and operations:* Mathematical literacy requires using symbolic, formal and technical language and operations. This involves understanding, interpreting, manipulating, and making use of symbolic expressions within a mathematical context (including arithmetic expressions and operations) governed by mathematical conventions and rules. It also involves understanding and utilising formal constructs based on definitions, rules and formal systems and also using algorithms with these entities. The symbols, rules and systems used vary according to what particular mathematical content knowledge is needed for a specific task to formulate, solve or interpret the mathematics.
- *Using mathematical tools[1]:* Mathematical tools include physical tools, such as measuring instruments, as well as calculators and computer-based tools that are becoming more widely available. In addition to knowing how to use these tools to assist them in completing mathematical tasks, students need to know about the limitations of such tools. Mathematical tools can also have an important role in communicating results.

Figure 3.2. Relationship between mathematical processes (top row) and fundamental mathematical capabilities (left-most column)

	Formulating situations mathematically	Employing mathematical concepts, facts, procedures and reasoning	Interpreting, applying and evaluating mathematical outcomes
Communicating	Read, decode, and make sense of statements, questions, tasks, objects or images, in order to form a mental model of the situation	Articulate a solution, show the work involved in reaching a solution and/or summarise and present intermediate mathematical results	Construct and communicate explanations and arguments in the context of the problem
Mathematising	Identify the underlying mathematical variables and structures in the real world problem, and make assumptions so that they can be used	Use an understanding of the context to guide or expedite the mathematical solving process, e.g. working to a context-appropriate level of accuracy	Understand the extent and limits of a mathematical solution that are a consequence of the mathematical model employed
Representation	Create a mathematical representation of real-world information	Make sense of, relate and use a variety of representations when interacting with a problem	Interpret mathematical outcomes in a variety of formats in relation to a situation or use; compare or evaluate two or more representations in relation to a situation
Reasoning and argument	Explain, defend or provide a justification for the identified or devised representation of a real-world situation	Explain, defend or provide a justification for the processes and procedures used to determine a mathematical result or solution Connect pieces of information to arrive at a mathematical solution, make generalisations or create a multi-step argument	Reflect on mathematical solutions and create explanations and arguments that support, refute or qualify a mathematical solution to a contextualised problem
Devising strategies for solving problems	Select or devise a plan or strategy to mathematically reframe contextualised problems	Activate effective and sustained control mechanisms across a multi-step procedure leading to a mathematical solution, conclusion or generalisation	Devise and implement a strategy in order to interpret, evaluate and validate a mathematical solution to a contextualised problem
Using symbolic, formal and technical language and operations	Use appropriate variables, symbols, diagrams and standard models in order to represent a real-world problem using symbolic/formal language	Understand and utilise formal constructs based on definitions, rules and formal systems as well as employing algorithms	Understand the relationship between the context of the problem and representation of the mathematical solution. Use this understanding to help interpret the solution in context and gauge the feasibility and possible limitations of the solution
Using mathematical tools	Use mathematical tools in order to recognise mathematical structures or to portray mathematical relationships	Know about and be able to make appropriate use of various tools that may assist in implementing processes and procedures for determining mathematical solutions	Use mathematical tools to ascertain the reasonableness of a mathematical solution and any limits and constraints on that solution, given the context of the problem

These capabilities are evident to varying degrees in each of the three mathematical processes. The ways in which these capabilities manifest themselves within the three processes are described in Figure 3.1.

A good guide to the empirical difficulty of items can be obtained by considering which aspects of the fundamental mathematical capabilities are required for planning and executing a solution (Turner and Adams, 2012[9]; Turner et al., 2013[8]). The easiest items will require the activation of few capabilities and in a relatively straightforward way. The hardest items require complex activation of several capabilities. Predicting difficulty requires consideration of both the number of capabilities and the complexity of activation required.

Mathematical content knowledge

An understanding of mathematical content – and the ability to apply that knowledge to the solution of meaningful contextualised problems – is important for citizens in the modern world. That is, to solve problems and interpret situations in personal, occupational, societal and scientific contexts, there is a need to draw upon certain mathematical knowledge and understandings.

Mathematical structures have been developed over time as a means to understand and interpret natural and social phenomena. In schools, the mathematics curriculum is typically organised around content strands (e.g. number, algebra and geometry) and detailed topic lists that reflect historically well-established branches of mathematics and that help in defining a structured curriculum. However, outside the mathematics classroom, a challenge or situation that arises is usually not accompanied by a set of rules and prescriptions that shows how the challenge can be met. Rather, it typically requires some creative thought in seeing the possibilities of bringing mathematics to bear on the situation and in formulating it mathematically. Often a situation can be addressed in different ways drawing on different mathematical concepts, procedures, facts or tools.

Since the goal of PISA is to assess mathematical literacy, an organisational structure for mathematical content knowledge is proposed based on the mathematical phenomena that underlie broad classes of problems and which have motivated the development of specific mathematical concepts and procedures. Because national mathematics curricula are typically designed to equip students with knowledge and skills that address these same underlying mathematical phenomena, the outcome is that the range of content arising from organising content this way is closely aligned with that typically found in national mathematics curricula. This framework lists some content topics appropriate for assessing the mathematical literacy of 15-year-old students, based on analyses of national standards from eleven countries.

To organise the domain of mathematics for purposes of assessing mathematical literacy, it is important to select a structure that grows out of historical developments in mathematics, that encompasses sufficient variety and depth to reveal the essentials of mathematics, and that also represents, or includes, the conventional mathematical strands in an acceptable way. Thus, a set of content categories that reflects the range of underlying mathematical phenomena was selected for the PISA 2018 framework, consistent with the categories used for previous PISA surveys.

The following list of content categories, therefore, is used in PISA 2018 to meet the requirements of historical development, coverage of the domain of mathematics and the underlying phenomena which motivate its development, and reflection of the major strands of school curricula. These four categories characterise the range of mathematical content that is central to the discipline and illustrate the broad areas of content used in the test items for PISA 2018:

- Change and relationships
- Space and shape
- Quantity
- Uncertainty and data

With these four categories, the mathematical domain can be organised in a way that ensures a spread of items across the domain and focuses on important mathematical phenomena, but at the same time, avoids a too fine division that would work against a focus on rich and challenging mathematical problems based on real situations. While categorisation by content category is important for item development and selection, and for reporting of assessment results, it is important to note that some specific content topics may materialise in more than one content category. Connections between aspects of content that span these four content categories contribute to the coherence of mathematics as a discipline and are apparent in some of the assessment items for the PISA 2018 assessment.

The broad mathematical content categories and the more specific content topics appropriate for 15-year-old students described later in this section reflect the level and breadth of content that is eligible for inclusion on the PISA 2018 assessment. Narrative descriptions of each content category and the relevance of each to solving meaningful problems are provided first, followed by more specific definitions of the kinds of content that are appropriate for inclusion in an assessment of mathematical literacy of 15-year-old students. These specific topics reflect commonalities found in the expectations set by a range of countries and education jurisdictions. The standards examined to identify these content topics are viewed as evidence not only of what is taught in mathematics classrooms in these countries but also as indicators of what countries view as important knowledge and skills for preparing students of this age to become constructive, engaged and reflective citizens.

Descriptions of the mathematical content knowledge that characterise each of the four categories – *change and relationships, space and shape, quantity, and uncertainty and data* – are provided below.

Change and relationships

The natural and designed worlds display a multitude of temporary and permanent relationships among objects and circumstances, where changes occur within systems of inter-related objects or in circumstances where the elements influence one another. In many cases, these changes occur over time, and in other cases changes in one object or quantity are related to changes in another. Some of these situations involve discrete change; others change continuously. Some relationships are of a permanent, or invariant, nature. Being more literate about change and relationships involves understanding fundamental types of change and recognising when they occur in order to use suitable mathematical models to describe and predict change. Mathematically this means modelling the change and the relationships with appropriate functions and equations, as well as creating, interpreting, and translating among symbolic and graphical representations of relationships.

Change and relationships is evident in such diverse settings as growth of organisms, music, and the cycle of seasons, weather patterns, employment levels and economic conditions. Aspects of the traditional mathematical content of functions and algebra, including algebraic expressions, equations and inequalities, tabular and graphical representations, are central in describing, modelling and interpreting change phenomena. Representations of data and relationships described using statistics also are often used to portray and interpret change and relationships, and a firm grounding in the basics of number and units is also essential to defining and interpreting *change and relationships*. Some interesting relationships arise from geometric measurement, such as the way that changes in perimeter of a family of shapes might relate to changes in area, or the relationships among lengths of the sides of triangles.

Space and shape

Space and shape encompasses a wide range of phenomena that are encountered everywhere in our visual and physical world: patterns, properties of objects, positions and orientations, representations of objects, decoding and encoding of visual information, navigation and dynamic interaction with real shapes as well as with representations. Geometry serves as an essential foundation for *space and shape*, but the category extends beyond traditional geometry in content, meaning and method, drawing on elements of other mathematical areas such as spatial visualisation, measurement and algebra. For instance, shapes can change, and a point can move along a locus, thus requiring function concepts. Measurement formulas are central in this area. The manipulation and interpretation of shapes in settings that call for tools ranging from dynamic geometry software to Global Positioning System (GPS) software are included in this content category.

PISA assumes that the understanding of a set of core concepts and skills is important to mathematical literacy relative to *space and shape*. Mathematical literacy in the area of *space and shape* involves a range of activities such as understanding perspective (for example in paintings), creating and reading maps, transforming shapes with and without technology, interpreting views of three-dimensional scenes from various perspectives and constructing representations of shapes.

Quantity

The notion of *Quantity* may be the most pervasive and essential mathematical aspect of engaging with, and functioning in, our world. It incorporates the quantification of attributes of objects, relationships, situations and entities in the world, understanding various representations of those quantifications, and judging interpretations and arguments based on quantity. To engage with the quantification of the world involves understanding measurements, counts, magnitudes, units, indicators, relative size, and numerical trends and patterns. Aspects of quantitative reasoning – such as number sense, multiple representations of numbers, elegance in computation, mental calculation, estimation and assessment of reasonableness of results – are the essence of mathematical literacy relative to *quantity*.

Quantification is a primary method for describing and measuring a vast set of attributes of aspects of the world. It allows for the modelling of situations, for the examination of change and relationships, for the description and manipulation of space and shape, for organising and interpreting data, and for the measurement and assessment of uncertainty. Thus mathematical literacy in the area of *quantity* applies knowledge of number and number operations in a wide variety of settings.

Uncertainty and data

In science, technology and everyday life, uncertainty is a given. Uncertainty is therefore a phenomenon at the heart of the mathematical analysis of many problem situations, and the theory of probability and statistics as well as techniques of data representation and description have been established to deal with it. The *uncertainty and data* content category includes recognising the place of variation in processes, having a sense of the quantification of that variation, acknowledging uncertainty and error in measurement, and knowing about chance. It also includes forming, interpreting and evaluating conclusions drawn in situations where uncertainty is central. The presentation and interpretation of data are key concepts in this category (Moore, 1997[10]).

There is uncertainty in scientific predictions, poll results, weather forecasts and economic models. There is variation in manufacturing processes, test scores and survey findings, and chance is fundamental to many recreational activities enjoyed by individuals. The traditional curricular areas of probability and statistics provide formal means of describing, modelling and interpreting a certain class of uncertainty phenomena, and for making inferences. In addition, knowledge of number and of aspects of algebra, such as graphs and symbolic representation, contribute to facility in engaging in problems in this content category. The focus on the interpretation and presentation of data is an important aspect of the *uncertainty and data category*.

Desired distribution of items by content category

The trend items selected for PISA 2015 and 2018 are distributed across the four content categories, as shown in Table 3.2. The goal in constructing the assessment is a balanced distribution of items with respect to content category, since all of these domains are important for constructive, engaged and reflective citizens.

Table 3.2. Desired distribution of mathematics items, by content category

Content category	Percentage of items
Change and relationships	25
Space and shape	25
Quantity	25
Uncertainty and data	25
Total	**100**

Content topics for guiding the assessment of mathematical literacy

To effectively understand and solve contextualised problems involving *change and relationships, space and shape, quantity* and *uncertainty and data* requires drawing upon a variety of mathematical concepts, procedures, facts, and tools at an appropriate level of depth and sophistication. As an assessment of mathematical literacy, PISA strives to assess the levels and types of mathematics that are appropriate for 15-year-old students on a trajectory to become constructive, engaged and reflective citizens able to make well-founded judgements and decisions. It is also the case that PISA, while not designed or intended to be a curriculum-driven assessment, strives to reflect the mathematics that students have likely had the opportunity to learn by the time they are 15 years old.

The content included in PISA 2018 is the same as that developed in PISA 2012. The four content categories of *change and relationships, space and shape, quantity* and *uncertainty and data* serve as the foundation for identifying this range of content, yet there is not a one-to-one mapping of content topics to these categories. The following content is intended to reflect the centrality of many of these concepts to all four content categories and reinforce the coherence of mathematics as a discipline. It intends to be illustrative of the content topics included in PISA 2018, rather than an exhaustive listing:

- *Functions:* the concept of function, emphasising but not limited to linear functions, their properties, and a variety of descriptions and representations of them. Commonly used representations are verbal, symbolic, tabular and graphical.
- *Algebraic expressions:* verbal interpretation of and manipulation with algebraic expressions, involving numbers, symbols, arithmetic operations, powers and simple roots.

- *Equations and inequalities:* linear and related equations and inequalities, simple second-degree equations, and analytic and non-analytic solution methods.
- *Co-ordinate systems:* representation and description of data, position and relationships.
- *Relationships within and among geometrical objects in two and three dimensions:* static relationships such as algebraic connections among elements of figures (e.g. the Pythagorean theorem as defining the relationship between the lengths of the sides of a right triangle), relative position, similarity and congruence, and dynamic relationships involving transformation and motion of objects, as well as correspondences between two- and three-dimensional objects.
- *Measurement:* quantification of features of and among shapes and objects, such as angle measures, distance, length, perimeter, circumference, area and volume.
- *Numbers and units:* concepts, representations of numbers and number systems, including properties of integer and rational numbers, relevant aspects of irrational numbers, as well as quantities and units referring to phenomena such as time, money, weight, temperature, distance, area and volume, and derived quantities and their numerical description.
- *Arithmetic operations:* the nature and properties of these operations and related notational conventions.
- *Percents, ratios and proportions:* numerical description of relative magnitude and the application of proportions and proportional reasoning to solve problems.
- *Counting principles:* simple combinations and permutations.
- *Estimation:* purpose-driven approximation of quantities and numerical expressions, including significant digits and rounding.
- *Data collection, representation and interpretation:* nature, genesis and collection of various types of data, and the different ways to represent and interpret them.
- *Data variability and its description:* concepts such as variability, distribution and central tendency of data sets, and ways to describe and interpret these in quantitative terms.
- *Samples and sampling:* concepts of sampling and sampling from data populations, including simple inferences based on properties of samples.
- *Chance and probability:* notion of random events, random variation and its representation, chance and frequency of events, and basic aspects of the concept of probability.

Contexts

The choice of appropriate mathematical strategies and representations is often dependent on the context in which a mathematics problem arises. Context is widely regarded as an aspect of problem solving that imposes additional demands on the problem solver (see (Watson and Callingham, 2003[11]) for findings about statistics). It is important that a wide variety of contexts is used in the PISA assessment. This offers the possibility of connecting with the broadest possible range of individual interests and with the range of situations in which individuals operate in the 21st century.

For purposes of the PISA 2018 mathematics framework, four context categories have been defined and are used to classify assessment items developed for the PISA survey:

- *Personal* – Problems classified in the personal context category focus on activities of one's self, one's family or one's peer group. The kinds of contexts that may be considered personal include (but are not limited to) those involving food preparation, shopping, games, personal health, personal transportation, sports, travel, personal scheduling and personal finance.
- *Occupational* – Problems classified in the occupational context category are centred on the world of work. Items categorised as occupational may involve (but are not limited to) such things as measuring, costing and ordering materials for building, payroll/accounting, quality control, scheduling/inventory, design/architecture and job-related decision making. Occupational contexts may relate to any level of the workforce, from unskilled work to the highest levels of professional work, although items in the PISA assessment must be accessible to 15-year-old students.
- *Societal* – Problems classified in the societal context category focus on one's community (whether local, national or global). They may involve (but are not limited to) such things as voting systems, public transport, government, public policies, demographics, advertising, national statistics and economics. Although individuals are involved in all of these things in a personal way, in the societal context category the focus of problems is on the community perspective.
- *Scientific* – Problems classified in the scientific category relate to the application of mathematics to the natural world and issues and topics related to science and technology. Particular contexts might include (but are not limited to) such areas as weather or climate, ecology, medicine, space science, genetics, measurement and the world of mathematics itself. Items that are intra-mathematical, where all the elements involved belong in the world of mathematics, fall within the scientific context.

PISA assessment items are arranged in units that share stimulus material. It is therefore usually the case that all items in the same unit belong to the same context category. Exceptions do arise; for example stimulus material may be examined from a personal point of view in one item and a societal point of view in another. When an item involves only mathematical constructs without reference to the contextual elements of the unit within which it is located, it is allocated to the context category of the unit. In the unusual case of a unit involving only mathematical constructs and being without reference to any context outside of mathematics, the unit is assigned to the scientific context category.

Using these context categories provides the basis for selecting a mix of item contexts and ensures that the assessment reflects a broad range of uses of mathematics, ranging from everyday personal uses to the scientific demands of global problems. Moreover, it is important that each context category be populated with assessment items having a broad range of item difficulties. Given that the major purpose of these context categories is to challenge students in a broad range of problem contexts, each category should contribute substantially to the measurement of mathematical literacy. It should not be the case that the difficulty level of assessment items representing one context category is systematically higher or lower than the difficulty level of assessment items in another category.

In identifying contexts that may be relevant, it is critical to keep in mind that a purpose of the assessment is to gauge the use of mathematical content knowledge, processes and capabilities that students have acquired by the age of 15. Contexts for assessment items, therefore, are selected in light of relevance to students' interests and lives, and the demands that will be placed upon them as they enter society as constructive, engaged and reflective citizens. National project managers from countries participating in the PISA survey are involved in judging the degree of such relevance.

Desired distribution of items by context category

The trend items selected for the PISA 2015 and 2018 mathematics assessment represent a spread across these context categories, as described in Table 3.3. With this balanced distribution, no single context type is allowed to dominate, providing students with items that span a broad range of individual interests and a range of situations that they might expect to encounter in their lives.

Table 3.3. Desired distribution of mathematics items, by context

Content category	Percentage of items
Personal	25
Occupational	25
Societal	25
Scientific	25
Total	**100**

Assessing mathematical literacy

This section outlines the approach taken to apply the elements of the framework described in previous sections to PISA 2015. This includes the structure of the mathematics component of the PISA survey, arrangements for transferring the paper-based trend items to a computer-based delivery, and reporting mathematical proficiency.

Structure of the survey instrument

In 2012, when mathematical literacy was the major domain, the paper-based instrument contained a total of 270 minutes of mathematics material. The material was arranged in nine clusters of items, with each cluster representing 30 minutes of testing time. The item clusters were placed in test booklets according to a rotated design, they also contained linked materials.

Mathematical literacy is a minor domain in 2018 and students are asked to complete fewer clusters. However the item clusters are similarly constructed and rotated. Six mathematics clusters from previous cycles, including one "easy" and one "hard", are used in one of three designs, depending on whether countries take the Collaborative Problem Solving option or not, or whether they take the test on paper. Using six clusters rather than three as was customary for the minor domains in previous cycles results in a larger number of trend items, therefore the construct coverage is increased. However, the number of students responding to each question is lower. This design is intended to reduce potential bias, thus stabilising and improving the measurement of trends. The field trial was used to perform a mode-effect study and to establish equivalence between the computer- and paper-based forms.

Response formats

Three types of response format are used to assess mathematical literacy in PISA 2018: open constructed-response, closed constructed-response and selected-response (simple and complex multiple-choice) items. Open constructed-response items require a somewhat extended written response from a student. Such items also may ask the student to show the steps taken or to explain how the answer was reached. These items require trained experts to manually code student responses.

Closed constructed-response items provide a more structured setting for presenting problem solutions, and they produce a student response that can be easily judged to be either correct or incorrect. Often student responses to questions of this type can be keyed into data-capture software, and coded automatically, but some must be manually coded by trained experts. The most frequent closed constructed-responses are single numbers.

Selected- response items require students to choose one or more responses from a number of response options. Responses to these questions can usually be automatically processed. About equal numbers of each of these response formats is used to construct the survey instruments.

Item scoring

Although most of the items are dichotomously scored (that is, responses are awarded either credit or no credit), the open constructed-response items can sometimes involve partial credit scoring, which allows responses to be assigned credit according to differing degrees of "correctness" of responses. For each such item, a detailed coding guide that allows for full credit, partial credit or no credit is provided to persons trained in the coding of student responses across the range of participating countries to ensure coding of responses is done in a consistent and reliable way. To maximise the comparability between the paper-based and computer-based assessments, careful attention is given to the scoring guides in order to ensure that the important elements are included.

Computer-based assessment of mathematics

The main mode of delivery for the PISA 2012 assessment was paper-based. In moving to computer-based delivery for 2015, care was taken to maximise comparability between the two assessments. The following section describes some of the features intrinsic to a computer-based assessment. Although these features provide the opportunities outlined below, to ensure comparability the PISA 2015 and 2018 assessments consist solely of items from the 2012 paper-based assessment. The features described here, however, will be used in future PISA assessments when their introduction can be controlled to ensure comparability with prior assessments.

Increasingly, mathematics tasks at work involve some kind of electronic technology, so that mathematical literacy and computer use are melded together (Hoyles et al., 2002[12]). For employees at all levels of the workplace, there is now an interdependency between mathematical literacy and the use of computer technology. Solving PISA items on a computer rather than on paper moves PISA into the reality and the demands of the 21st century.

There is a great deal of research evidence into paper- and computer-based test performance, but findings are mixed. Some research suggests that a computer-based testing environment can influence students' performance. Richardson et al. (2002[13]) reported that students found computer-based problem-solving tasks engaging and motivating, often despite the unfamiliarity of the problem types and the challenging nature of the items. They were sometimes distracted by attractive graphics, and sometime used poor heuristics when attempting tasks.

In one of the largest comparisons of paper-based and computer-based testing, Sandene et al. (2005[14]) found that eighth-grade students' mean score was four points higher on a computer-based mathematics test than on an equivalent paper-based test. Bennett et al. (2008[15]) concluded from his research that computer familiarity affects performance on computer-based mathematics tests, while others have found that the range of functions available through computer-based tests can affect performance. For example, Mason (Mason, Patry and Bernstein, 2001[16]) found that students' performance was negatively affected in computer-based tests compared to paper-based tests when there was no opportunity on the computer version to review and check responses. Bennett (2003[17]) found that screen size affected scores on verbal-reasoning tests, possibly because smaller computer screens require scrolling.

By contrast, Wang et al. (Wang et al., 2007[18]) conducted a meta-analysis of studies pertaining to K-12 students' mathematics achievements which indicated that administration mode has no statistically significant effect on scores. Moreover, recent mode studies that were part of the Programme for the International Assessment of Adult Competencies (PIAAC) suggested that equality can be achieved (OECD, 2014[19]). In this study, adults were randomly assigned to either a computer-based or paper-based assessment of literacy and numeracy skills. The majority of the items used in the paper delivery mode were adapted for computer delivery and used in this study. Analyses of these data revealed that almost all of the item parameters were stable across the two modes, thus demonstrating the ability to place respondents on the same literacy and numeracy scale. Given this, it is hypothesised that mathematics items used in PISA 2012 can be transposed onto a screen without affecting trend data. The PISA 2015 field trial studied the effect on student performance of the change in mode of delivery; for further details see Annex A6 of *PISA 2015 Results, Volume I* (OECD, 2016[20]).

Just as paper-based assessments rely on a set of fundamental skills for working with printed materials, computer-based assessments rely on a set of fundamental information and communications technology (ICT) skills for using computers. These include knowledge of basic hardware (e.g. keyboard and mouse) and basic conventions (e.g. arrows to move forward and specific buttons to press to execute commands). The intention is to keep such skills to a minimal, core level in the computer-based assessment.

Reporting proficiency in mathematics

The outcomes of the PISA mathematics assessment are reported in a number of ways. Estimates of overall mathematical proficiency are obtained for sampled students in each participating country, and a number of proficiency levels are defined. Descriptions of the degree of mathematical literacy typical of students in each level are also developed. For PISA 2003, scales based on the four broad content categories were developed. In Figure 3.3, descriptions of the six proficiency levels reported for the overall PISA mathematics scale in 2012 are presented. These form the basis of the PISA 2018 mathematics scale. The finalised 2012 scale is used to report the PISA 2018 outcomes. As *mathematical literacy* is a minor domain in 2018, only the overall proficiency scale is reported.

Fundamental mathematical capabilities play a central role in defining what it means to be at different levels of the scales for mathematical literacy overall and for each of the reported processes. For example, in the proficiency scale description for Level 4 (see Figure 3.3), the second sentence highlights aspects of mathematising and representation that are evident at this level. The final sentence highlights the characteristic communication, reasoning and argument of Level 4, providing a contrast with the short communications and lack of argument of Level 3 and the additional reflection of Level 5. In an earlier section of this framework and in Figure 3.2, each of the mathematical processes was described in terms of the fundamental mathematical capabilities that individuals might activate when engaging in that process.

Figure 3.3. Summary description of the six levels of mathematics proficiency in PISA 2018

Level	What students can typically do
6	At Level 6, students can conceptualise, generalise and utilise information based on their investigations and modelling of complex problem situations, and can use their knowledge in relatively non-standard contexts. They can link different information sources and representations and flexibly translate among them. Students at this level are capable of advanced mathematical thinking and reasoning. These students can apply this insight and understanding, along with a mastery of symbolic and formal mathematical operations and relationships, to develop new approaches and strategies for attacking novel situations. Students at this level can reflect on their actions, and can formulate and precisely communicate their actions and reflections regarding their findings, interpretations, arguments, and the appropriateness of these to the original situation.
5	At Level 5, students can develop and work with models for complex situations, identifying constraints and specifying assumptions. They can select, compare and evaluate appropriate problem-solving strategies for dealing with complex problems related to these models. Students at this level can work strategically using broad, well-developed thinking and reasoning skills, appropriate linked representations, symbolic and formal characterisations, and insight pertaining to these situations. They begin to reflect on their work and can formulate and communicate their interpretations and reasoning.
4	At Level 4, students can work effectively with explicit models for complex concrete situations that may involve constraints or call for making assumptions. They can select and integrate different representations, including symbolic, linking them directly to aspects of real-world situations. Students at this level can utilise their limited range of skills and can reason with some insight, in straightforward contexts. They can construct and communicate explanations and arguments based on their interpretations, arguments and actions.
3	At Level 3, students can execute clearly described procedures, including those that require sequential decisions. Their interpretations are sufficiently sound to be a base for building a simple model or for selecting and applying simple problem-solving strategies. Students at this level can interpret and use representations based on different information sources and reason directly from them. They typically show some ability to handle percentages, fractions and decimal numbers, and to work with proportional relationships. Their solutions reflect that they have engaged in basic interpretation and reasoning.
2	At Level 2, students can interpret and recognise situations in contexts that require no more than direct inference. They can extract relevant information from a single source and make use of a single representational mode. Students at this level can employ basic algorithms, formulae, procedures or conventions to solve problems involving whole numbers. They are capable of making literal interpretations of the results.
1	At Level 1, students can answer questions involving familiar contexts where all relevant information is present and the questions are clearly defined. They are able to identify information and to carry out routine procedures according to direct instructions in explicit situations. They can perform actions that are almost always obvious and follow immediately from the given stimuli.

Notes

[1] In some countries, "mathematical tools" can also refer to established mathematical procedures, such as algorithms. For the purposes of the PISA framework, "mathematical tools" refers only to the physical and digital tools described in this section.

References

Bennett, R. (2003), *Online Assessment and the Comparability of Score Meaning*, Educational Testing Service, Princeton, NJ. [17]

Bennett, R. et al. (2008), "Does it Matter if I Take My Mathematics Test on Computer? A Second Empirical Study of Mode Effects in NAEP", *The Journal of Technology, Learning, and Assessment*, Vol. 6/9, http://www.jtla.org. [15]

Blum, W., P. Galbraith and M. Niss (2007), "Introduction", in Blum, W. et al. (eds.), *Modelling and Applications in Mathematics Education*, Springer US, New York, http://dx.doi.org/10.1007/978-0-387-29822-1. [3]

Gagatsis, A. and S. Papastavridis (eds.) (2003), *Mathematical Competencies and the Learning of Mathematics: The Danish KOM Project*, Hellenic Mathematic Society, Athens. [5]

Hoyles, C. et al. (2002), "Mathematical Skills in the Workplace: Final Report to the Science Technology and Mathematics Council", *Institute of Education, University of London; Science, Technology and Mathematics Council, London. (2002)*, http://discovery.ucl.ac.uk/1515581/ (accessed on 5 April 2018). [12]

Mason, B., M. Patry and D. Bernstein (2001), "An Examination of the Equivalence Between Non-Adaptive Computer-Based And Traditional Testing", *Journal of Education Computing Research*, Vol. 24/1, pp. 29-39. [16]

Moore, D. (1997), "New Pedagogy and New Content: The Case of Statistics", *International Statistical Review*, Vol. 65/2, pp. 123-165, https://onlinelibrary.wiley.com/doi/epdf/10.1111/j.1751-5823.1997.tb00390.x. [10]

Niss, M. and T. Højgaard (2011), *Competencies and Mathematical Learning: Ideas and Inspiration for the Development of Mathematics Teaching and Learning in Denmark*, Ministry of Education, Roskilde University, https://pure.au.dk/portal/files/41669781/thj11_mn_kom_in_english.pdf. [7]

Niss, M. and T. Jensen (2002), *Kompetencer og matematiklaering Ideer og inspiration til udvikling af matematikundervisning i Danmark*, Ministry of Education, Copenhagen, http://static.uvm.dk/Publikationer/2002/kom/hel.pdf. [6]

OECD (2016), *PISA 2015 Results (Volume I): Excellence and Equity in Education*, OECD Publishing, Paris, https://doi.org/10.1787/19963777. [20]

OECD (2014), *Technical Report of the Survey of Adult Skills (PIAAC) 2013*, Pre-publication, OECD, http://www.oecd.org/skills/piaac/_Technical%20Report_17OCT13.pdf. [19]

OECD (2013), *PISA 2012 Assessment and Analytical Framework: Mathematics, Reading, Science, Problem Solving and Financial Literacy*, PISA, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264190511-en. [4]

OECD (2010), *Pathways to Success: How Knowledge and Skills at Age 15 Shape Future Lives in Canada*, PISA, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264081925-en. [1]

OECD (2004), *The PISA 2003 Assessment Framework: Mathematics, Reading, Science and Problem Solving Knowledge and Skills*, PISA, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264101739-en. [2]

Richardson, M. et al. (2002), "Challenging minds? Students' perceptions of computer-based World Class Tests of problem solving", *Computers in Human Behavior*, Vol. 18/6, pp. 633-649, http://dx.doi.org/10.1016/S0747-5632(02)00021-3. [13]

Sandene, B. et al. (2005), *Online Assessment in Mathematics and Writing: Reports from the NAEP Technology-Based Assessment Project, Research and Development Series (NCES 2005-457)*, US Department of Education, National Center for Education Statistics, US Government Printing Office. [14]

Turner, R. and R. Adams (2012), *Some Drivers of Test Item Difficulty in Mathematics: An Analysis of the Competency Rubric*, American Educational Research Association (AERA), http://research.acer.edu.au/pisa. [9]

Turner, R. et al. (2013), "Using mathematical competencies to predict item difficulty in PISA", in *Research on PISA. Research outcomes of the PISA research conference 2009, Kiel, Germany, September 14–16, 2009*, Springer, New York, https://doi.org/10.1007/978-94-007-4458-5_2. [8]

Wang, S. et al. (2007), "A meta-analysis of testing mode effects in grade K-12 mathematics tests", *Educational and Psychological Measurement*, Vol. 67/2, pp. 219-238, http://dx.doi.org/10.1177/0013164406288166. [18]

Watson, J. and R. Callingham (2003), "Statistical Literacy: A Complex Hierarchical Construct", *Statistics Education Research Journal*, Vol. 2/2, pp. 3-46, http://fehps.une.edu.au/serj. [11]

4. PISA 2018 Science Framework

This chapter defines "scientific literacy" as assessed in the Programme for International Student Assessment (PISA) in 2018. It describes the types of contexts, knowledge and competencies that are reflected in the tasks that PISA uses to measure scientific literacy. The chapter also discusses how student performance in science is measured and reported.

Introduction: Scientific literacy and why it matters

This document provides a description and rationale for the framework that forms the basis of the PISA assessment of scientific literacy – the major domain in PISA 2015 and a minor domain in PISA 2018. Previous PISA frameworks for the science assessment (OECD, 1999[1]; OECD, 2003[2]; OECD, 2006[3]) have used scientific literacy as their central construct. This framework for PISA 2015/2018 has refined and extended the previous construct, specifically the PISA 2006 framework that was used as the basis for assessment in 2006, 2009 and 2012.

Scientific literacy is developed through science education that is both broad and applied. Thus, within this framework, the concept of scientific literacy *refers both to a knowledge of science and of science-based technology*. However, science and technology differ in their purposes, processes and products. Technology seeks the optimal solution to a human problem and there may be more than one optimal solution. In contrast, science seeks the answer to a specific question about the natural material world.

Scientific literacy also requires not just knowledge of the concepts and theories of science but also a knowledge of the common procedures and practices associated with scientific enquiry and how these enable science to advance. Therefore, individuals who are scientifically literate understand the major conceptions and ideas that form the foundation of scientific and technological thought; how such knowledge has been derived; and the degree to which such knowledge is justified by evidence or theoretical explanations.

For all of these reasons, scientific literacy is perceived to be a key competency (Rychen and Salganik, 2001[4]) which is defined in terms of the ability to use knowledge and information interactively. In other words, scientific literacy includes "an understanding of how it [a knowledge of science] changes the way one can interact with the world and how it can be used to accomplish broader goals" (ibid.: 10).

The rest of this document defines *scientific literacy* and describes how PISA attempts to measure this concept.

Scientific literacy: Towards a definition

There is a widespread belief that an understanding of science is so important that it should be a feature of every young person's education (American Association for the Advancement of Science, 1989[5]; COSCE, 2011[6]; Fensham, 1985[7]; Millar and Osborne, 1998[8]; National Research Council, 2012[9]; Ständige Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland, 2005[10]; Ministry of Education, Chinese Taipei, 1999[11]). Indeed, in many countries, science is an obligatory element of the school curriculum from kindergarten until the completion of compulsory education.

Three science-specific competencies are required in order to understand and engage in critical discussion about issues that involve science and technology. The first is the ability to provide explanatory accounts of natural phenomena, technical artefacts and technologies and their implications for society. The second is the competency to use one's knowledge and understanding of scientific enquiry to identify questions that can be answered by scientific enquiry; propose ways in which such questions might possibly be addressed; and identify whether appropriate procedures have been used. The third is the competency to interpret and evaluate data and evidence scientifically and evaluate whether the conclusions are warranted.

Thus, scientific literacy in PISA 2018 is defined by the three competencies of:

- Explaining phenomena scientifically;
- Evaluating and designing scientific enquiry; and
- Interpreting data and evidence scientifically.

All of these competencies require knowledge. Explaining scientific and technological phenomena, for instance, demands a knowledge of the content of science – referred to hereafter as **content knowledge**. The second and third competencies, however, require more than just content knowledge. They also depend on an understanding of how scientific knowledge is established and the degree of confidence with which it is held. Recognising and identifying the features that characterise scientific enquiry requires a knowledge of the standard procedures that underlie the diverse methods and practices used to establish scientific knowledge – referred to here as **procedural knowledge**. Finally, these competencies require **epistemic knowledge**, defined here as an understanding of the rationale for the common practices of scientific enquiry, the status of the claims that are generated, and the meaning of foundational terms such as theory, hypothesis and data. Box 4.1 provides more examples of each of these three types of knowledge, all of which are also further developed later in this framework.

Procedural and epistemic knowledge are necessary to identify questions that are amenable to scientific enquiry, to judge whether appropriate procedures have been used to ensure that claims are justified, and to distinguish scientific issues from matters of values or economic considerations. Procedural and epistemic knowledge are also essential to deciding whether the many claims that pervade contemporary media have been derived using appropriate procedures and are warranted; after all, over their lifetimes, individuals will need to acquire knowledge, not through scientific investigations, but through the use of resources such as libraries and the Internet, and will need to evaluate such knowledge.

Box 4.1. Scientific knowledge: PISA 2015/2018 terminology

This document is based upon the view that scientific knowledge consists of three distinguishable but related elements. The first of these and the most familiar is knowledge of the facts, concepts, ideas and theories about the natural world that science has established, such as how plants synthesise complex molecules using light and carbon dioxide or the particulate nature of matter. This kind of knowledge is referred to as “content knowledge” or “knowledge of the content of science”.

Knowledge of the procedures that scientists use to establish scientific knowledge is referred to as “procedural knowledge”. This is the knowledge of the practices and concepts on which empirical enquiry is based, such as repeating measurements to minimise error and reduce uncertainty, the control of variables, and standard procedures for representing and communicating data (Millar et al., 1994[12]). More recently, these have been elaborated as a set of “concepts of evidence” (Roberts, Gott and Glaesser, 2010[13]).

Furthermore, understanding science as a practice also requires “epistemic knowledge”, which refers to an understanding of the role of specific constructs and defining features essential to the process of building scientific knowledge (Duschl, 2008[14]). Epistemic knowledge includes an understanding of the function that questions, observations, theories, hypotheses, models and arguments play in science; a recognition of the variety of forms of scientific enquiry; and understanding the role that peer review plays in establishing knowledge that can be trusted.

A more detailed discussion of these three forms of knowledge is provided in the later section on scientific knowledge in Table 4.4, Table 4.5 and Table 4.6

Scientific literacy requires all three forms of scientific knowledge. Therefore, PISA 2015 focussed on the extent to which 15-year-olds are capable of displaying these three forms of knowledge appropriately within a range of personal, local, national and global contexts. This perspective is broader than that of many school science programmes, where content knowledge often dominates.

It is such considerations that have led to the following definition of scientific literacy for PISA 2015 and 2018:

Box 4.2. The 2015/2018 definition of scientific literacy

Scientific literacy is the ability to engage with science-related issues, and with the ideas of science, as a reflective citizen.

A scientifically literate person, therefore, is willing to engage in reasoned discourse about science and technology which requires the competencies of:

- **Explaining phenomena scientifically:**

Recognising, offering and evaluating explanations for a range of natural and technological phenomena.

- **Evaluating and designing scientific enquiry:**

Describing and appraising scientific investigations and proposing ways of addressing questions scientifically.

- **Interpreting data and evidence scientifically:**

Analysing and evaluating data, claims and arguments in a variety of representations and drawing appropriate scientific conclusions.

Explanatory Notes

The following remarks are offered to clarify the meaning and use of this definition of scientific literacy for the purposes of the PISA 2018 assessment.

- The use of the term "scientific literacy" rather than "science" underscores the importance that the PISA science assessment places on the application of scientific knowledge in the context of real-world situations.
- For the purposes of the PISA assessment, these competencies will only be tested using the content, procedural and epistemic knowledge of science that 15-year-old students can reasonably be expected to have.
- Finally, throughout this document, the term "natural world" is used to refer to phenomena taking place in or associated with any object in the living or the material world.

The competencies required for scientific literacy

Competency 1: Explaining phenomena scientifically

Science has managed to develop a set of explanatory theories that have transformed our understanding of the natural world. Moreover, such knowledge has enabled the development of technologies that support human life, such as treatments for various diseases and rapid communication across the globe. The competency to explain scientific and technological phenomena thus depends on a knowledge of these major explanatory ideas of science.

Explaining some scientific phenomena, however, requires more than just the ability to recall and use theories, explanatory ideas, information, and facts (**content knowledge**). Offering scientific explanations also requires an understanding of how such knowledge has been derived and the level of confidence one can hold about any scientific claims. Hence, individuals also require a knowledge of the standard forms and procedures used in scientific enquiry to obtain such knowledge (**procedural knowledge**) and an understanding of their own role and function in justifying the knowledge produced by science (**epistemic knowledge**).

Competency 2: Evaluating and designing scientific enquiry

Scientific literacy requires students to have some understanding of the goal of scientific enquiry, which is to generate reliable knowledge about the natural world (Ziman, 1978[15]). Data obtained by observation and experiment, either in the laboratory or in the field, lead to the development of models and explanatory hypotheses that enable predictions that can then be tested experimentally. New claims and hypotheses are always provisional and may not stand up when subjected to critical peer review (Longino, 1990[16]). Hence, scientists commit to publishing or reporting their findings and the methods used in obtaining the evidence that support these findings. Measurements, however, all contain a degree of error. Much of the work of the experimental scientist is, therefore, devoted to the resolution of uncertainty by repeating measurements, collecting larger samples, building instruments that are more accurate, and using statistical techniques that assess the degree of confidence in any result.

This competency draws on **content knowledge**, a knowledge of the common procedures used in science (**procedural knowledge**) and the function of these procedures in justifying any claims advanced by science (**epistemic knowledge**). Procedural and epistemic knowledge serve two functions. First, such knowledge is required by individuals to appraise scientific investigations, thus deciding whether appropriate procedures have been followed and whether the conclusions are warranted. Second, such knowledge allows individuals to propose, at least in broad terms, how a scientific question might be investigated appropriately.

Competency 3: Interpreting data and evidence scientifically

Interpreting data is a core activity for all scientists. It typically begins by looking for patterns, perhaps through constructing simple tables or graphical visualisations. Any relationships or patterns in the data must then be read using a knowledge of standard patterns. The scientifically literate individual can also be expected to understand that uncertainty is an inherent feature of all measurement, and that one criterion for expressing our confidence in a finding is the probability that it might have occurred by chance. All of this draws on a body of **procedural knowledge**.

It is not sufficient, however, to understand the procedures that have been applied to obtain a data set. The scientifically literate individual needs to be able to judge whether these procedures are appropriate and whether the ensuing claims are justified (**epistemic knowledge**). For instance, many sets of data can be interpreted in multiple ways, and scientists must argue in support of their own interpretation while defending it from the critique of others. Resolution of which interpretation is the best requires a knowledge of science (**content knowledge**). A critical and sceptical disposition towards all empirical evidence is indeed the hallmark of the professional scientist.

Organisation of the domain

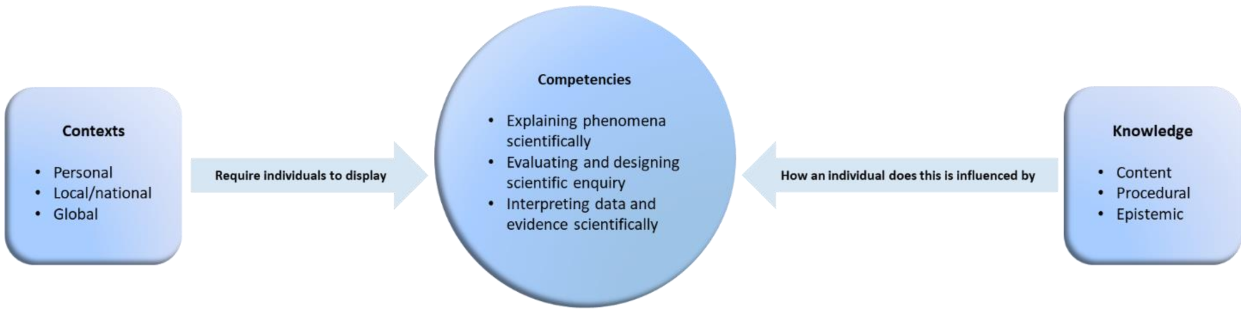
For the purposes of assessment, the PISA 2018 definition of scientific literacy can be characterised as consisting of three interrelated aspects (see Figure 4.1).

Table 4.1. Aspects of the scientific literacy assessment framework for PISA 2015/2018

Contexts	Personal, local/national and global issues, both current and historical, which demand some understanding of science and technology.
Knowledge	An understanding of the major facts, concepts and explanatory theories that form the basis of scientific knowledge. Such knowledge includes knowledge of both the natural world and technological artefacts (**content knowledge**), knowledge of how such ideas are produced (**procedural knowledge**), and an understanding of the underlying rationale for these procedures and the justification for their use (**epistemic knowledge**).
Competencies	The ability to explain phenomena scientifically, evaluate and design scientific enquiry, and interpret data and evidence scientifically.

Each of these aspects is discussed further below.

Figure 4.1. Inter-relations between the three aspects



Contexts for assessment items

PISA 2018 assesses scientific knowledge using contexts that raised pertinent issues that were often relevant to the science education curricula of participating countries. However, assessment items are not limited to school science contexts. Items in the PISA 2018 science assessment may relate to the self, family and peer groups (personal), to the community (local and national) or to life across the world (global). The context may involve technology or, in some cases, a historical element that may be used to assess students' understanding of the processes and practices involved in advancing scientific knowledge.

Contexts for items in the PISA science assessment have also been categorised into five applications of science and technology: health and disease, natural resources, environmental quality, hazards, and the frontiers of science and technology. The PISA science assessment, however, is not an assessment of contexts. Rather, it assesses competencies and knowledge in specific contexts. These contexts have been chosen in light of their relevance to students' interests and lives and because they are the areas in which scientific literacy has particular value in enhancing and sustaining quality of life and in the development of public policy.

Table 4.2 shows how these five applications interact with the personal, local/national, and global contexts described above.

Table 4.2. Contexts for the PISA 2018 scientific literacy assessment

	Personal	Local/National	Global
Health and disease	Maintenance of health, accidents, nutrition	Control of disease, food choices, community health	Epidemics, spread of infectious diseases
Natural resources	Personal consumption of materials and energy	Maintenance of human populations, quality of life, security, production and distribution of food, energy supply	Renewable and non-renewable natural systems, population growth, sustainable use of species
Environmental quality	Environmentally friendly actions, use and disposal of materials and devices	Population distribution, disposal of waste, environmental impact	Biodiversity, ecological sustainability, control of pollution, production and loss of soil/biomass
Hazards	Risk assessments of lifestyle choices	Rapid changes (e.g., earthquakes, severe weather), slow and progressive changes (e.g., coastal erosion, sedimentation), risk assessment	Climate change, impact of modern communication
Frontiers of science and technology	Scientific aspects of hobbies, personal technology, music and sporting activities	New materials, devices and processes, genetic modifications, health technology, transport	Extinction of species, exploration of space, origin and structure of the Universe

Scientific competencies

Table 4.3, Table 4.4 and Table 4.5 provide a detailed description of the tasks that make up each of the three competencies that comprise scientific literacy. This set of scientific competencies reflects a view that science is best seen as an ensemble of social and epistemic practices that are common across all of its subfields (National Research Council, 2012[9]). Hence, all of these competencies are framed as actions, conveying what the scientifically literate person both understands and is capable of doing.

Table 4.3. PISA 2018 scientific competencies: Explaining phenomena scientifically

Explaining phenomena scientifically
Recognising, offering and evaluating explanations for a range of natural and technological phenomena through: • Recalling and applying appropriate scientific knowledge; • Identifying, using and generating explanatory models and representations; • Making and justifying appropriate predictions; • Offering explanatory hypotheses; • Explaining the potential implications of scientific knowledge for society

Demonstrating the competency of *explaining phenomena scientifically* requires students to recall the appropriate content knowledge in a given situation and use it to interpret and provide an explanation for the phenomenon of interest. Such knowledge can also be used to generate tentative explanatory hypotheses for an observed phenomenon or when presented with data. A scientifically literate person is expected to be able to draw on standard scientific models to construct simple representations for everyday phenomena and then use these representations to make predictions. This competency includes the ability to describe or interpret phenomena and predict possible changes. In addition, it may involve recognising or identifying appropriate descriptions, explanations, and predictions.

Table 4.4. PISA 2018 scientific competencies: Evaluating and designing scientific enquiry

Evaluating and designing scientific enquiry
Describing and appraising scientific investigations and proposing ways of addressing questions scientifically through: • Identifying the question explored in a given scientific study; • Distinguishing questions that are possible to investigate scientifically; • Proposing a way of exploring a given question scientifically; • Evaluating ways of exploring a given question scientifically; • Describing and evaluating a range of ways that scientists use to ensure the reliability of data and the objectivity and generalisability of explanations.

The competency of *evaluating and designing scientific enquiry* is required to evaluate reports of scientific findings and investigations critically. It relies on the ability to discriminate scientific questions from other forms of enquiry, or in other words, to recognise questions that can be investigated scientifically. This competency requires a knowledge of the key features of a scientific investigation, such as what things should be measured, what variables should be changed or controlled, and what action should be taken so that accurate and precise data can be collected. It requires an ability to evaluate the quality of data, which in turn depends on recognising that data are not always completely accurate. It also requires the competency to identify if an investigation is driven by an underlying theoretical premise or, alternatively, whether it seeks to determine identifiable patterns.

A scientifically literate person should also be able to recognise the significance of previous research in judging the value of any given scientific enquiry. Moreover, students need to understand the importance of developing a sceptical disposition to all media reports in science, recognising that all research builds on previous work, that the findings of any one study are always subject to uncertainty, and that the study may be biased by its sources of funding. This competency requires students to possess both procedural and epistemic knowledge but may also draw on their content knowledge of science.

Table 4.5. PISA 2018 scientific competencies: Interpreting data and evidence scientifically

Interpreting data and evidence scientifically
Analysing and evaluating scientific data, claims and arguments in a variety of representations and drawing appropriate conclusions through: • Transforming data from one representation to another; • Analysing and interpreting data and drawing appropriate conclusions; • Identifying the assumptions, evidence and reasoning in science-related texts; • Distinguishing between arguments that are based on scientific evidence and theory and those based on other considerations; • Evaluating scientific arguments and evidence from different sources (e.g. newspaper, Internet, journals).

Students who can interpret data and evidence scientifically should be able to convey the meaning of a piece of scientific evidence and its implications to a specified audience in their own words, using diagrams or other representations as appropriate. This competency requires the use of mathematical tools to analyse or summarise data, and the ability to use standard methods to transform data to different representations.

This competency also includes accessing scientific information and producing and evaluating arguments and conclusions based on scientific evidence (Kuhn, 2010[17]; Osborne, 2010[18]). It may also involve evaluating alternative conclusions using evidence; giving reasons for or against a given conclusion; and identifying the assumptions made in reaching a conclusion. In short, the scientifically literate individual should be able to identify logical or flawed connections between evidence and conclusions.

Scientific knowledge

Content knowledge

Only a sample of the content domain of science can be assessed in the PISA 2018 science assessment. Hence, it is important that clear criteria are used to guide the selection of the knowledge that is assessed. The content knowledge that PISA assesses is selected from the major fields of physics, chemistry, biology, and earth and space sciences and:

- Is relevant to real-life situations;
- Represents an important scientific concept or major explanatory theory that has enduring utility; and
- is appropriate to the developmental level of 15-year-olds.

Table 4.6 presents the categories of content knowledge selected by applying the criteria above.

Table 4.6. Content knowledge in the PISA 2018 science assessment

Physical systems, including:
Structure of matter (e.g., particle model, bonds)
Properties of matter (e.g., changes of state, thermal and electrical conductivity)
Chemical changes of matter (e.g., chemical reactions, energy transfer, acids/bases)
Motion and forces (e.g., velocity, friction) and action at a distance (e.g., magnetic, gravitational and electrostatic forces)
Energy and its transformation (e.g., conservation, dissipation, chemical reactions)
Interactions between energy and matter (e.g., light and radio waves, sound and seismic waves)
Living systems, including:
Cells (e.g., structures and function, DNA, differences between plant and animal cells)
The concept of an organism (e.g., unicellular vs. multicellular)
Humans (e.g., health; nutrition; subsystems such as the digestive, the respiratory, the circulatory, the excretory and the reproductive and their relationship)
Populations (e.g., species, evolution, biodiversity, genetic variation)
Ecosystems (e.g., food chains, matter and energy flow)
Biosphere (e.g., ecosystem services, sustainability)
Earth and space systems, including:
Structures of the Earth (e.g., lithosphere, atmosphere, hydrosphere)
Energy in the Earth (e.g., sources, global climate)
Change in the Earth (e.g., plate tectonics, geochemical cycles, constructive and destructive forces)
Earth's history (e.g., fossils, origin and evolution)
Earth in space (e.g., gravity, solar systems, galaxies)
The history and scale of the Universe and its history (e.g., light year, Big Bang theory)

Procedural knowledge

A fundamental goal of science is to generate explanatory accounts of the material world. Tentative explanatory accounts are first developed and then tested through empirical enquiry. Empirical enquiry is reliant on certain well-established concepts and methods such as the notion of dependent and independent variables, the control of variables, various types of measurement and forms of error, methods for minimising error, a recognition of common patterns observed in data, and methods of presenting data. It is this knowledge of the standard concepts and procedures essential to scientific enquiry that underpins the collection, analysis and interpretation of scientific data. Such ideas form a body of procedural knowledge, which has also been called "concepts of evidence" (Roberts, Gott and Glaesser, 2010[13]; Millar et al., 1994[12]). Such knowledge is needed both to undertake scientific enquiry and engage in a critical review of the evidence that might be used to support particular claims. The examples listed in Table 4.7 are some examples of procedural knowledge that may be tested.

Table 4.7. Procedural knowledge in the PISA 2018 science assessment

Procedural knowledge
The concept of variables, including dependent, independent and control variables;
Concepts of measurement, e.g. quantitative measurements, qualitative observations, the use of a scale or other instruments, categorical and continuous variables;
Ways of assessing and minimising uncertainty such as repeating and averaging measurements;
Mechanisms to ensure the replicability (closeness of agreement between repeated measurements of the same quantity) and accuracy (the closeness of agreement between a measured quantity and its true value) of measurements;
Common ways of abstracting and representing data using tables, graphs and charts and their appropriate use; The control of variables and its role in experimental design;
The use of randomised controlled trials to avoid confounded findings and to identify possible causal mechanisms; The nature of an appropriate design for a given scientific question, e.g., experimental, field-based or pattern-seeking.

Epistemic Knowledge

Epistemic knowledge is a knowledge of the constructs and defining features essential to the process of knowledge building in science (e.g. hypotheses, theories and observations) and their role in justifying the knowledge produced by science (Duschl, 2008[14]). Students use epistemic knowledge to explain, with examples, the difference between a scientific theory and a hypothesis or between a scientific fact and an observation. Epistemic knowledge includes the understanding that the construction of models, be they directly representational, abstract or mathematical, is a key feature of science and that such models are akin to maps rather than accurate pictures of the material world. Students should also recognise that the word "theory" is not used the same way in science as it is in everyday language, where it is a synonym for "guess" or "hunch". Whereas procedural knowledge is required to explain what is meant by the control of variables strategy, epistemic knowledge is required to explain why the use of the control of variables strategy is central to establishing scientific knowledge.

Scientifically literate individuals will also understand that scientists draw on data to advance claims to knowledge and that argument is a commonplace feature of science. These students also understand the role and significance of peer review as the mechanism that the scientific community has established for testing new claims. Epistemic knowledge thus provides a rationale for the procedures and practices in which scientists engage and the foundation for the basis of belief in the claims that science makes about the natural world.

Table 4.8 represents what are considered to be the major components of epistemic knowledge necessary for scientific literacy.

Table 4.8. Epistemic knowledge in the PISA 2018 science assessment

Epistemic knowledge
The constructs and defining features of science, that is: • The nature of scientific observations, facts, hypotheses, models and theories; • The purpose and goals of science (to produce explanations of the natural world) as distinguished from technology (to produce an optimal solution to human need), what constitutes a scientific or technological question, and what constitutes appropriate data; • The values of science, such as a commitment to publication, objectivity and the elimination of bias; • The nature of reasoning used in science, such as deductive, inductive, inference to the best explanation (abductive), analogical and model-based; The role of these constructs and features in justifying the knowledge produced by science, that is: • How scientific claims are supported by data and reasoning in science; • The function of different forms of empirical enquiry in establishing knowledge, including both their goal (to test explanatory hypotheses or identify patterns) and their design (observation, controlled experiments, correlational studies); • How measurement error affects the degree of confidence in scientific knowledge; • The use and role of physical, system and abstract models and their limits; • The role of collaboration and critique and how peer review helps to establish confidence in scientific claims; • The role of scientific knowledge, along with other forms of knowledge, in identifying and addressing societal and technological issues.

Epistemic knowledge is most likely to be tested in a pragmatic fashion: student will typically be required to interpret and answer a question that requires some epistemic knowledge rather than being directly asked about the points in Table 4.8. For instance, students may be asked to identify whether the conclusions are justified by the data or what piece of evidence best supports the hypothesis advanced in an item and explain why.

Assessment of the Domain

Cognitive Demand

A key feature of the 2018 PISA framework is the definition of levels of cognitive demand within the assessment of scientific literacy and across all three competencies of the framework. In assessment frameworks, item difficulty, which is empirically derived, is often confused with cognitive demand. Empirical item difficulty is estimated from the proportion of the test-taking population that is successful in solving the item correctly, while cognitive demand refers to the type of mental processing required (Davis and Buckendahl, 2011[19]). An item can have a high difficulty level because it tests knowledge that is unfamiliar to most students while at the same time requiring only low cognitive demand because students only need to recall a piece of information. Conversely, an item can be cognitively demanding because it requires the individual to relate and evaluate many items of knowledge, yet still be of a low difficulty level because each of the pieces of knowledge is easily recalled (Brookhart and Nitko, 2011[20]).

Various classifications of cognitive demand schemes have been developed and evaluated since Bloom's Taxonomy was first published (Bloom, 1956[21]). These have been largely based on categorisations of knowledge types and associated cognitive processes that are used to describe educational objectives or assessment tasks.

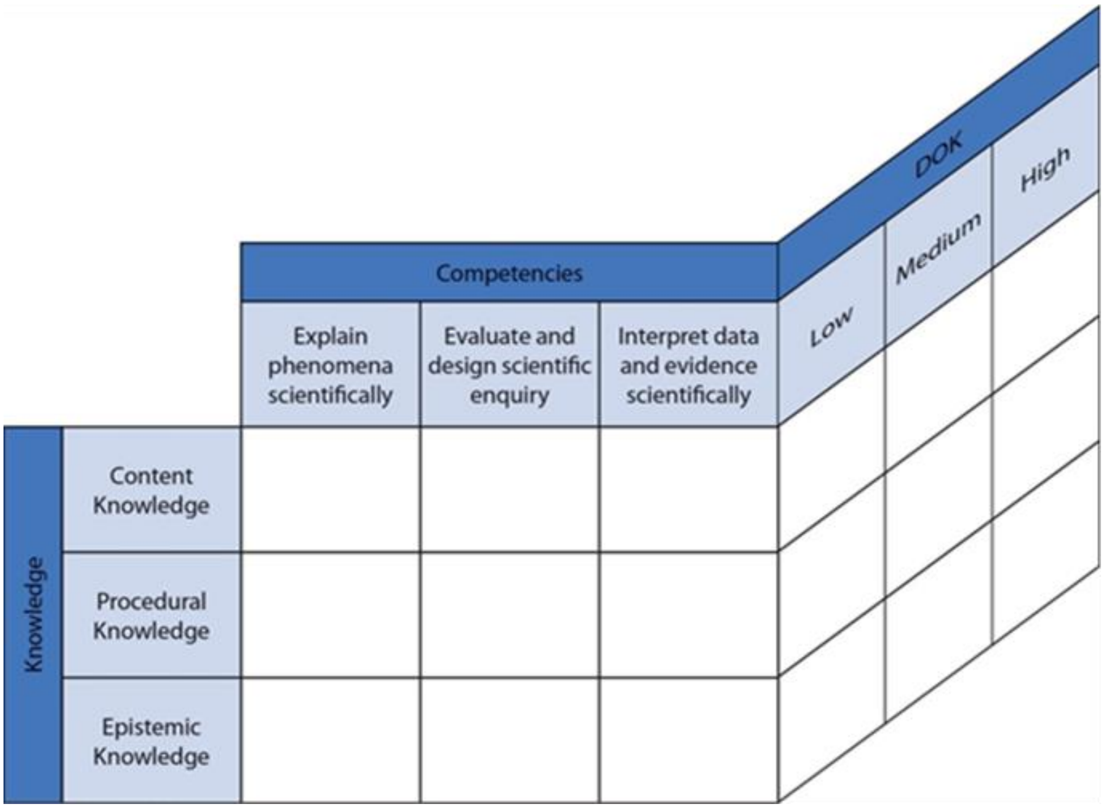
Webb's Depth of Knowledge (1997[22]) was specifically developed to address the disparity between assessments and the expectations of student learning. Webb's levels of depth are determined by the complexity of both the content and the task required. His framework consists of four levels: level 1 (recall), level 2 (using skills and/or conceptual knowledge), level 3 (strategic thinking) and level 4 (extended thinking). Each level is defined by a large number of verbs (some of which appear in more than one level) that describe cognitive processes. This framework offers a more holistic view of learning and assessment tasks and requires an analysis of both the content and cognitive process demanded by any task.

All the frameworks described above have helped to classify knowledge and competencies in the PISA 2018 science framework. In drawing up such a framework, it was recognised that there were challenges in developing test items based on a cognitive hierarchy. The three main challenges were that:

1. Too much effort would be made to fit test items into particular cognitive frameworks, which could lead to poorly developed items;
2. The intended and actual cognitive demand might not have align, with frameworks defining rigorous, cognitively demanding goals, but items operationalising the standard in a much less cognitively demanding way; and
3. Without a well-defined and understood cognitive framework, item writing and development might often focus on item difficulty and thus use only a limited range of cognitive processes and knowledge types. These would then only be described and interpreted *post hoc*, rather than being built from a theory of increasing competency.

The PISA 2018 science framework uses an adapted version of Webb's Depth of Knowledge grid (Webb, 1997[22]) alongside the desired scientific knowledge and competencies. As the competencies are the central feature of the framework, the cognitive framework needs to assess and report on them across the range of student abilities. Webb's Depth of Knowledge levels offer a taxonomy for cognitive demand that identifies both the cognitive demand from the verbal cues that are used (e.g., analyse, arrange or compare) and the expected depth of knowledge required.

Figure 4.2. PISA 2015/2018 framework for cognitive demand



The grid above (Figure 4.2) provides a framework for mapping items against the dimensions of knowledge and competencies. In addition, each item can also be mapped onto a third dimension based on depth of knowledge, which categorises cognitive demand into the following levels:

- **Low (L)**

  Carrying out a one-step procedure, such as recalling a fact, term, principle or concept or locating a single point of information from a graph or table.

- **Medium (M)**

  Using and applying conceptual knowledge to describe or explain phenomena; selecting appropriate procedures involving two or more steps; organising or displaying data; or interpreting or using simple data sets or graphs.

- **High (H)**

  Analysing complex information or data; synthesising or evaluating evidence; justifying; reasoning given various sources; developing a plan or sequence of steps to approach a problem.

Thus items that merely require the recollection of one piece of information make low cognitive demands, even if the knowledge itself might be quite complex. In contrast, items that require the recollection of more than one piece of knowledge and require a comparison and evaluation made of the competing merits of their relevance would be seen as having high cognitive demand, even if the knowledge itself is relatively simple. The difficulty of any item is therefore a combination of both the complexity and the range of knowledge it requires and the cognitive operations that are required to process this knowledge and thus resolve the item.

Therefore, the major factors that determine the difficulty of items assessing science achievement are:

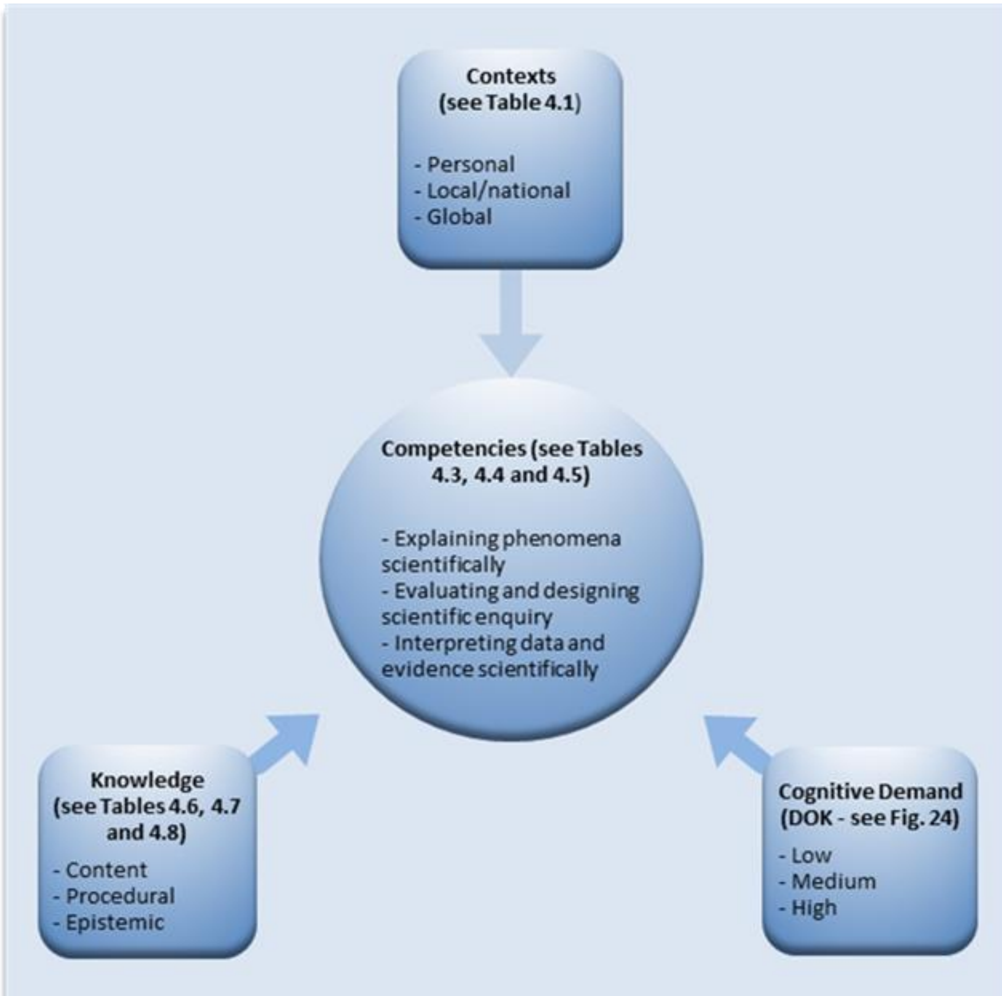
- The number and the degree of complexity of the elements of knowledge demanded by the item;
- The level of familiarity and prior knowledge that students may have of the content, procedural and epistemic knowledge involved;
- The cognitive operation required by the item (e.g., recall, analysis, evaluation); and
- The extent to which forming a response depends on models or abstract scientific ideas.

This four-factor approach allows for a broad measure of scientific literacy across a wide range of student abilities. It is relatively simple, therefore hopefully minimising the problems encountered in its application. This cognitive framework will also facilitate the development of an a priori definition of the descriptive parameters of the reporting proficiency scale (see Table 4.11).

Test Characteristics

Figure 4.3 relates the basic components of the PISA 2018 framework for the scientific literacy assessment to the structure and the content of assessment units (cf. Figure 4.1). As a starting point to construct assessment units, it shows the need to consider the contexts that will serve as stimulus material, the competencies required to respond to the questions or issues, the knowledge central to the units and the cognitive demand.

Figure 4.3. A tool for constructing and analysing assessment units and items



A test unit is introduced by specific stimulus material, which may be a brief written passage, or text accompanying a table, chart, graph or diagram. In units newly created for PISA 2015 (and reused in PISA 2018), the stimulus material may also include non-static stimulus material, such as animations and interactive simulations. The items within a unit are independently scored. Sample units can be found at www.oecd.org/pisa/test.

PISA groups items into units in order to use contexts that are as realistic as possible and that reflect the complexity of real-world situations, while making efficient use of testing time. Using situations about which several questions can be posed, rather than asking separate questions about a larger number of different situations, reduces the overall time required for a student to become familiar with the material in each question. However, score points (i.e. items) within a unit must remain independent of one another. Furthermore, because this approach reduces the number of different assessment contexts, it is important to ensure that there is an adequate range of contexts in order to minimise bias due to the choice of contexts.

PISA 2018 test units will require the use of all three scientific competencies and draw on all three forms of science knowledge. In most cases, each test unit will assess multiple competencies and knowledge categories. Individual items, however, will primarily assess only one form of knowledge and one scientific competency.

Students need to read the stimulus material and questions in the PISA 2018 science literacy assessment, therefore raising the issue that a certain level of reading literacy will be required to display science literacy. To address these concerns, stimulus material and questions will use language that is as clear, simple, brief and syntactically simple as possible while still conveying the appropriate meaning. The number of concepts introduced per paragraph will be limited. Questions within the domain of science that specifically assess reading or mathematical literacy will be avoided.

Item response formats

Three classes of items will be used to assess the competencies and scientific knowledge identified in the framework. The items will be divided approximately equally into these three classes:

Simple multiple-choice: Items calling for

- The selection of a single response from four options; or
- The selection of a "hot spot", or an answer that is a selectable element within a graphic or text.

Complex multiple-choice: Items calling for

- Responses to a series of related "Yes/No" questions that are treated as a single item for scoring purposes (the typical format in 2006);
- The selection of more than one response from a list;
- The completion of a sentence by selecting drop-down choices to fill multiple blanks; or
- "Drag-and-drop" responses, allowing students to move elements on screen to complete a task requiring matching, ordering or categorising.

Constructed response: Items calling for written or drawn responses. Constructed response items in the scientific literacy assessment typically call for a written responses ranging from a phrase to a short paragraph (i.e., two to four sentences of explanation). A small number of constructed response items call for the drawing of, for example, a graph or diagram. In the computer-based assessment, any such items will be supported by simple drawing editors that are specific to the response required.

Also, in PISA 2018, some responses will be captured by interactive tasks, such as a student's choices when manipulating variables in a simulated scientific enquiry. Responses to these interactive tasks will be scored as complex multiple-choice items. Some responses to interactive tasks are sufficiently open-ended that they are considered to be constructed responses.

Assessment Structure

Computer-based assessment will again be the primary mode of delivery for all domains, including scientific literacy, in PISA 2018. Science literacy items that were newly developed for the computer-based delivery of PISA 2015 will only be available in the computer-based assessment in PISA 2018. However, a paper-based assessment instrument (with a smaller selection of items) will be provided for countries choosing not to test their students on the computer.

PISA units are organised into 30-minute sections called "clusters." Each cluster includes either only units new to PISA 2015 or only units that have been used in previous PISA cycles, known as "trend units".

Each student will be assigned one two-hour test form. A test form is composed of four clusters, each designed to occupy thirty minutes of testing time. The clusters are placed in multiple computer–based test forms, according to a rotated test design.

Each student will spend a total of one hour on two clusters of reading literacy, with the remaining time assigned to either one or two of the additional domains of science, mathematics, and global competence. While the paper-based assessment will be limited to trend items and will not include any newly developed material, the computer-based instrument will include both newly developed items and trend items. Care will be taken when transposing paper-based trend items to an on-screen format so that the presentation, response format and cognitive demand remain comparable.

The desired score-point balance between the three types of knowledge (content, procedural and epistemic) and the three content knowledge categories is shown in Table 4.9. These weightings are broadly consistent with the previous framework and reflect a consensus view amongst the experts consulted in the writing of this framework.

Table 4.9. Target distribution of score points by knowledge type and content knowledge category

	Systems			
Knowledge types	Physical	Living	Earth & Space	Total over systems
Content	20-24%	20-24%	14-18%	**54-66%**
Procedural	7-11%	7-11%	5-9%	**19-31%**
Epistemic	4-8%	4-8%	2-6%	**10-22%**
Total over knowledge types	**36%**	**36%**	**28%**	**100%**

The target score-point balance for the scientific competencies is given in Table 4.10. These weightings have been chosen so that the assessment is evenly split between items which draw predominantly on content knowledge and items that draw predominantly on procedural or epistemic knowledge.

Table 4.10. Target distribution of score points for scientific competencies

Scientific Competencies	% of score points
Explaining phenomena scientifically	40-50%
Evaluating and designing scientific enquiry	20-30%
Interpreting data and evidence scientifically	30-40%
TOTAL	**100%**

Item contexts will be spread across personal, local/national and global settings roughly in the ratio of 1:2:1, as was the case in 2006 when science was first the major domain of assessment. A wide variety of areas of application will be selected, subject to the constraints imposed by the distribution of score points shown in Table 4.9 and Table 4.10.

Reporting scales

The development of scales of student achievement – or describing what students at different levels of attainment can do – is essential to report on and compare student achievement across the world. The 2015 framework (upon which this framework is largely based) explicitly defined the parameters of increasing competence and progression, allowing item developers to design items representing this growth in ability (Kane, 2006[23]; Mislevy and Haertel, 2006[24]). The scale has been extended down to Level "1b", which specifically addresses and provides a description of students at the lowest level of ability. These students demonstrate very minimal evidence of scientific literacy and would previously not have been included in the reporting scales.

Table 4.11. Summary description of the seven levels of proficiency in science in PISA 2015

Level	Characteristics of tasks
6	At Level 6, students can draw on a range of interrelated scientific ideas and concepts from the physical, life and earth and space sciences and use content, procedural and epistemic knowledge in order to offer explanatory hypotheses of novel scientific phenomena, events and processes or to make predictions. In interpreting data and evidence, they are able to discriminate between relevant and irrelevant information and can draw on knowledge external to the normal school curriculum. They can distinguish between arguments that are based on scientific evidence and theory and those based on other considerations. Level 6 students can evaluate competing designs of complex experiments, field studies or simulations and justify their choices.
5	At Level 5, students can use abstract scientific ideas or concepts to explain unfamiliar and more complex phenomena, events and processes involving multiple causal links. They are able to apply more sophisticated epistemic knowledge to evaluate alternative experimental designs and justify their choices and use theoretical knowledge to interpret information or make predictions. Level 5 students can evaluate ways of exploring a given question scientifically and identify limitations in interpretations of data sets including sources and the effects of uncertainty in scientific data.
4	At Level 4, students can use more complex or more abstract content knowledge, which is either provided or recalled, to construct explanations of more complex or less familiar events and processes. They can conduct experiments involving two or more independent variables in a constrained context. They are able to justify an experimental design, drawing on elements of procedural and epistemic knowledge. Level 4 students can interpret data drawn from a moderately complex data set or less familiar context, draw appropriate conclusions that go beyond the data and provide justifications for their choices.
3	At Level 3, students can draw upon moderately complex content knowledge to identify or construct explanations of familiar phenomena. In less familiar or more complex situations, they can construct explanations with relevant cueing or support. They can draw on elements of procedural or epistemic knowledge to carry out a simple experiment in a constrained context. Level 3 students are able to distinguish between scientific and non-scientific issues and identify the evidence supporting a scientific claim.
2	At Level 2, students are able to draw on everyday content knowledge and basic procedural knowledge to identify an appropriate scientific explanation, interpret data, and identify the question being addressed in a simple experimental design. They can use basic or everyday scientific knowledge to identify a valid conclusion from a simple data set. Level 2 students demonstrate basic epistemic knowledge by being able to identify questions that can be investigated scientifically.
1a	At Level 1a, students are able to use basic or everyday content and procedural knowledge to recognise or identify explanations of simple scientific phenomenon. With support, they can undertake structured scientific enquiries with no more than two variables. They are able to identify simple causal or correlational relationships and interpret graphical and visual data that require a low level of cognitive demand. Level 1a students can select the best scientific explanation for given data in familiar personal, local and global contexts.
1b	At Level 1b, students can use basic or everyday scientific knowledge to recognise aspects of familiar or simple phenomenon. They are able to identify simple patterns in data, recognise basic scientific terms and follow explicit instructions to carry out a scientific procedure.

Source: OECD (2016), *PISA 2015 Results (Volume I): Excellence and Equity in Education*, PISA, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264266490-en

References

American Association for the Advancement of Science (1989), *Science for All Americans*, Oxford University Press, New York, http://www.project2061.org/publications/sfaa/online/sfaatoc.htm. [5]

Bloom, B. (ed.) (1956), *Taxonomy of Educational Objectives, Book 1: Cognitive Domain*, Longmans Publishing. [21]

Brennan, R. (ed.) (2006), *Validation*, Praeger Publishers and the American Council on Education. [23]

Brookhart, S. and A. Nitko (2011), "Strategies for constructing assessments of higher order thinking skills", *Assessment of Higher Order Thinking Skills*, pp. 327-359. [20]

COSCE (2011), *Informe ENCIENDE, Enseñanza de las Ciencias en la Didáctica Escolar para edades tempranas en España*, Confederación de Sociedades Científicas de España, Madrid, https://www.cosce.org/pdf/Informe_ENCIENDE.pdf. [6]

Davis, S. and C. Buckendahl (2011), "Incorporating cognitive demand in credentialing examinations", *Assessment of Higher Order Thinking Skills*, pp. 327-359. [19]

Duschl, R. (2008), "Science Education in Three-Part Harmony: Balancing Conceptual, Epistemic, and Social Learning Goals", *Review of Research in Education*, Vol. 32/1, pp. 268-291, http://dx.doi.org/10.3102/0091732x07309371. [14]

Fensham, P. (1985), "Science for all: A reflective essay", *Journal of Curriculum Studies*, Vol. 17/4, pp. 415-435, http://dx.doi.org/10.1080/0022027850170407. [7]

Kuhn, D. (2010), "Teaching and learning science as argument", *Science Education*, Vol. 94/5, pp. 810-824, http://dx.doi.org/10.1002/sce.20395. [17]

Longino, H. (1990), *Science as Social Knowledge*, Princeton University Press, Princeton. [16]

Millar, R. et al. (1994), "Investigating in the school science laboratory: conceptual and procedural knowledge and their influence on performance", *Research Papers in Education*, Vol. 9/2, pp. 207-248, http://dx.doi.org/10.1080/0267152940090205. [12]

Millar, R. and J. Osborne (eds.) (1998), *Beyond 2000: Science Education for the Future*, School of Education, King's College London, http://www.nuffieldfoundation.org/sites/default/files/Beyond%202000.pdf. [8]

Ministry of Education, Chinese Taipei (1999), *General Senior High School Curriculum*, Ministry of Education website, http://www.k12ea.gov.tw. [11]

Mislevy, R. and G. Haertel (2006), "Implications of evidence-centered design for educational testing", *Educational Measurement: Issues and Practice*, Vol. 25/4, pp. 6-20. [24]

National Research Council (2012), *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas*, The National Academies Press, Washington, D.C., https://doi.org/10.17226/13165. [9]

OECD (2006), *Assessing Scientific, Reading and Mathematical Literacy: A Framework for PISA 2006*, PISA, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264026407-en. [3]

OECD (2003), *The PISA 2003 Assessment Framework: Mathematics, Reading, Science and Problem Solving Knowledge and Skills*, PISA, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264101739-en. [2]

OECD (1999), *Measuring Student Knowledge and Skills: A New Framework for Assessment*, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264173125-en. [1]

Osborne, J. (2010), "Arguing to Learn in Science: The Role of Collaborative, Critical Discourse", *Science*, Vol. 328/5977, pp. 463-466, http://dx.doi.org/10.1126/science.1183944. [18]

Roberts, R., R. Gott and J. Glaesser (2010), "Students' approaches to open-ended science investigation: the importance of substantive and procedural understanding", *Research Papers in Education*, Vol. 25/4, pp. 377-407, http://dx.doi.org/10.1080/02671520902980680. [13]

Rychen, D. and L. Salganik (eds.) (2001), *The Definition and Selection of Key Competencies*, OECD website, http://www.oecd.org/pisa/35070367.pdf. [4]

Ständige Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland (2005), *Beschlüsse der Kultusministerkonferenz: Bildungsstandards im Fach Biologie für den Mittleren Schulabschluss (Jahrgangsstufe 10)*, Wolters Kluwer Deutschland GmbH, München, Neuwied, https://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/2004/2004_12_16-Bildungsstandards-Biologie.pdf. [10]

Webb, N. (1997), *Criteria for Alignment of Expectations and Assessments in Mathematics and Science Education*, National Institute for Science Education, Washington, D.C. [22]

Ziman, J. (1978), *Reliable Knowledge: An Exploration of the Grounds for Belief in Science*, Cambridge University Press, Cambridge. [15]

5. PISA 2018 Financial Literacy Framework

PISA 2018 offers an optional assessment of financial literacy for the third time. This section presents the framework of this assessment and is based on the framework developed for the 2012 exercise, which was the first large-scale international study to assess the financial literacy of young people. It defines financial literacy as it pertains to youth, and it is organised around the content, processes and contexts that are relevant for the assessment of 15-year-old students. These content areas, processes and contexts are illustrated with several items drawn from the PISA 2018 field trial and previous assessments. In addition, the framework discusses the relationship between financial literacy and non-cognitive skills and between financial literacy and both reading and mathematics literacy.

Introduction

Policy interest in financial literacy

In recent years, developed and emerging economies have become increasingly aware of the importance of ensuring that their citizens are financially literate. This has stemmed in particular from shrinking public and private support systems, shifting demographic profiles including population aging, and wide-ranging developments in the financial marketplace. A lack of financial literacy contributes to ill-informed financial decisions and these decisions could, in turn, have tremendous adverse effects on both personal and, ultimately, global financial resilience (OECD, 2009[1]). As a result, financial literacy is now globally acknowledged as an essential life skill and targeted financial education policy is considered to be an important element of economic and financial stability and development. This is reflected in the G20 endorsement of the OECD/INFE (International Network on Financial Education) High-level Principles on National Strategies for Financial Education (G20, 2012[2]; OECD/INFE, 2012[3]) and the OECD/INFE Policy Handbook on National Strategies for Financial Education (OECD, 2015[4]). G20 leaders also recognise that this requires lifelong learning that starts in childhood, as indicated by their call for core competencies on financial literacy for young people and adults (OECD, 2015[5]; OECD, 2016[6]), and their statement supporting the widespread use of instruments to measure youth financial literacy, including the PISA financial literacy assessment (G20, 2013[7]).

A series of tangible trends underpin the rising global interest in financial literacy as a key life skill. These are summarised below.

Demographic and cultural shifts

Longevity is increasing in most countries, and in many the birth rate is falling. At the same time, women's participation in the labour force and the proportion of people entering higher education are both increasing, and grown-up children are less likely to continue to live in close proximity to older family members than they were in previous generations. The likely outcome of these shifts will be a greater need for financial security in retirement and professional care in old age, resulting in additional government expenditure (Colombo et al., 2011[8]). Working-age adults may be expected to shoulder the tax burden to finance this expenditure whilst at the same time also saving for their own retirement, potentially repaying their own student loans, and managing increasingly varied working-life trajectories which may include periods of inactivity, self-employment or retraining.

Risk shift and increased individual responsibility

There has been a widespread transfer of risk from both governments and employers to individuals, meaning that many people now face the financial risks associated with longevity, investment, credit, out-of-pocket healthcare and long-term care. The number of financial decisions that individuals have to make, and the significance of these decisions, is increasing as a consequence of changes in the market and the economy. For instance, longer life expectancy means that individuals need to ensure that they accumulate savings to cover much longer periods of retirement than previous generations, despite the steadily rising age of retirement in many countries. Younger adults may also face greater financial challenges than previous generations in saving for their own retirement whilst at the same time covering the heightened long-term health care needs of elderly relatives. Traditional pay-as-you-go (PAYG) public pension schemes are supplemented by privately funded schemes in which individuals may be responsible for making their own investment decisions, including the contribution rate, the investment allocation and the type of payout product. Moreover, defined-contribution pension plans are quickly replacing defined-benefit pension plans for new entrants, shifting onto workers the risks of uncertain investment performance and of longer life expectancy.

A financially literate consumer will know when to seek professional help to make a sound financial plan, but professional advisors are not an alternative to financial education. Even when individuals use the services of financial intermediaries and advisors, they need to understand what is being offered or advised, and they need the skills and knowledge to manage the products they choose. They should also be aware that some advisors are not independent and may face a conflict of interest as they provide advice and at the same time sell products or receive commissions. Depending on the national legal framework for financial advice, individuals may be fully responsible for the financial product they decide to purchase, facing all the direct consequences of their choice.

Surveys show that a majority of workers are unaware of the risks they now have to face, and even if they are aware of them, have neither the sufficient financial knowledge nor the skills to manage such risks adequately (Barrett, Mosca and Whelan, 2013[9]; Money and Pension Panel, 2013[10]; OECD, 2016[11]).

Increased supply of a wide range of financial products and services

Greater financial inclusion in emerging economies, as well as worldwide developments in technology and deregulation, have resulted in widening access to all kinds of financial products, from current accounts and remittance products to revolving credit and equity portfolios.[1] Growing numbers of consumers therefore have access to financial products and services from a variety of established and new providers delivered through traditional and digital channels. Whilst many of the products available bring advantages and help to improve financial well-being, many are also complex. Individuals are required to compare across a number of factors such as the fees charged, interest rates paid or received, length of contract and exposure to risk. They must also identify appropriate providers and delivery channels from the vast array of possibilities, including community groups, traditional financial institutions, online banks and mobile phone companies.

Increased demand for financial products and services

Economic and technological developments have brought greater global connectedness and massive changes in both the methods and frequency of communications and financial transactions, as well as in social interactions and consumer behaviour. Such changes have made it more important that individuals be able to interact with a wide range of financial providers and intermediaries. In particular, consumers often need access to financial services (including banks and other providers such as post offices) in order to make and receive electronic payments like income and remittances, and to perform online transactions. Together, these trends have enlarged the options for the majority of the population (including new financial consumers) and increased the level of complexity they face. Against this backdrop, individuals are expected to be sufficiently financially literate to take the necessary steps to protect themselves and their relatives and ensure their financial well-being.

New risks from financial products and services

As consumers attempt to access and use financial products, they also become vulnerable to certain associated risks. The spread of digital financial services means that consumers may face new types of risks, such as risks connected with high-cost short-term online credit. Consumers may also be exposed to new crimes, such as data and identity theft and fraud. Legitimate use of consumer data by a range of financial and non-financial companies to create consumers' digital profiles may also make it more costly or difficult to access certain types of financial products or services, as financial service providers seek to segment their consumer base and price or market their products accordingly (OECD, 2017[12]).

Expected benefits of financial education and improved levels of financial literacy

Existing empirical evidence shows that young people and adults in both developed and emerging economies who have been exposed to good quality financial education are subsequently more likely than others to plan ahead, save and engage in other responsible financial behaviours (Amagir et al., 2018[13]; Atkinson et al., 2015[14]; Bruhn et al., 2016[15]; Kaiser and Menkhoff, 2016[16]; Miller et al., 2014[17]). This evidence suggests a possible causal link between financial education and outcomes and indicates that improved levels of financial literacy can lead to positive changes in behaviour.

Other research indicates a number of potential benefits of being financially literate. There is evidence that in developed countries, those with higher financial literacy are better able to manage their money, participate in the stock market and perform better on their portfolio choices, and that they are more likely to choose mutual funds with lower fees (Clark, Lusardi and Mitchell, 2017[18]; Gaudecker, 2015[19]; Hastings and Tejeda-Ashton, 2008[20]; van Rooij, Lusardi and Alessie, 2011[21]). In emerging economies, financial literacy is shown to be correlated with holding basic financial products, like bank accounts, and with insurance take-up (Grohmann, Kluhs and Menkhoff, 2017[22]; Xu and Zia, 2012[23]). Similarly, 15-year-old students with bank accounts have higher levels of financial literacy than those without bank accounts on average across the OECD countries participating in the 2012 and 2015 PISA exercise (OECD, 2017[24]; OECD, 2014[25]). Moreover, adults who have greater financial knowledge are more likely to accumulate higher amounts of wealth (Behrman et al., 2012[26]; van Rooij, Lusardi and Alessie, 2012[27]).

Higher levels of financial literacy have been found to be related not only to the accumulation of assets but also to debt management, with more financially literate individuals opting for less costly mortgages and avoiding high interest payments and additional fees (Disney and Gathergood, 2013[28]; Lusardi and Tufano, 2015[29]).

In addition to the benefits identified for individuals, widespread financial literacy can be expected to improve economic and financial stability for a number of reasons (OECD, 2006[30]). Financially literate consumers can make more informed decisions and demand higher quality services, which can, in turn, encourage competition and innovation in the market. As they can protect themselves to a greater extent against the negative consequences of income or expenditure shocks and are less likely to default on credit commitments, macro-level shocks are likely to have a lower impact on financially literate populations. Financially literate consumers are also less likely to react to market conditions in unpredictable ways, less likely to make unfounded complaints and more likely to take appropriate steps to manage the risks transferred to them. All of these factors can lead to a more efficient financial services sector. They can also ultimately help in reducing government aid (and taxation) aimed at assisting those who have taken unwise financial decisions or no decision at all.

Box 5.1. OECD activities in relation to financial education

In 2002, the OECD initiated a far-reaching financial education project to address governments' emerging concerns about the potential consequences of low levels of financial literacy. This project is serviced by the OECD Committee on Financial Markets and the Insurance and Private Pensions Committee in coordination with other relevant bodies, including the PISA Governing Board and the Education Policy Committee on issues related to schools. The project takes a holistic approach to financial-consumer issues that highlights how, alongside improved financial access, adequate consumer protection and regulatory frameworks, financial education has a complementary role to play in promoting the outcome of financial literacy.

One of the first milestones of the financial education project was the adoption of the *Recommendation on Principles and Good Practices for Financial Education and Awareness* by the OECD Council (OECD, 2005[31]). Recognising the increasingly global nature of financial literacy and education issues, in 2008, the OECD created the International Network on Financial Education (INFE) to benefit from and encompass the experience and expertise of developed and emerging economies. More than 200 public institutions from more than 110 countries and economies are members of the INFE as of 2018. Members meet twice yearly to discuss the latest developments in their countries, share their expertise and collect evidence, as well as to develop analytical and comparative studies, methodologies, good practice, policy instruments and practical guidance on key priority areas.

Financial education for youth and in schools

The 2005 OECD Recommendation advised that "financial education should start at school. People should be educated about financial matters as early as possible in their lives" (OECD, 2005[31]). Two main reasons underpin the OECD Recommendation: the importance of focusing on youth in order to provide them with key life skills before they become active financial consumers, and the relative efficiency of providing financial education in schools rather than attempting remedial actions in adulthood.

At the time when the OECD Recommendation was published, there was a lack of guidance on how to implement financial education initiatives for youth and in schools. The OECD/INFE therefore subsequently developed a dedicated publication, *Financial Education for Youth: The Role of Schools* (OECD, 2014[32]). The publication includes case studies and guidelines on financial education learning frameworks and on introducing financial education in curricular teaching. It was supported by the Ministers of Finance of the Asia-Pacific Economic Cooperation (APEC) in 2012 (APEC, 2012[33]) and welcomed by G20 leaders in 2013.

Following a call by the G20 in 2013, the OECD/INFE also developed a *Core Competencies Framework on Financial Literacy for Youth*, which describes the financial literacy outcomes that are likely to be important for 15-18 year-olds and provides a tool for policy makers to develop national learning and assessment frameworks. The lessons learned from developing the PISA 2012 assessment framework and analysing the data thereby collected contributed to the development of this core competencies framework (OECD, 2015[5]; OECD, 2013[34]).

The two volumes collecting the results of the PISA 2012 and 2015 financial literacy assessment provide not only international evidence on the distribution of financial literacy among 15-year-old students within and across countries, but also policy suggestions on how policy makers in finance and education can improve it (OECD, 2017[24]; OECD, 2014[25]).

Focus on youth

People form habits and behaviours starting at a young age, learning from their parents and others around them, which indicates the importance of early interventions to help shape beneficial behaviours and attitudes (Whitebread and Bingham, 2013[35]). Furthermore, young people need financial knowledge and understanding from an early age in order to operate within the complex financial landscape they are likely to find themselves in, often even before reaching adulthood. Younger generations are not only likely to face more complex financial products, services and markets but as noted above, they are more likely than their parents to have to bear more financial risks in adulthood. In particular, as the previous discussion illustrates, they are likely to bear more responsibility for the planning of their own retirement savings and investments, and the coverage of their own healthcare needs.

Young people may learn beneficial behaviours from their friends and family, such as prioritising their expenditure or putting money aside for a rainy day, but recent changes in the financial marketplace and in social welfare systems mean it is unlikely that they can gain sufficient knowledge or information from their friends and family unless they work in related fields.[2] The majority of young people will have to apply their skills to search for financial information and solve problems, and will need to know when to make informed use of professional financial advice. Efforts to improve financial knowledge in the workplace or in other settings can be severely limited by a lack of early exposure to financial education and by a lack of awareness of the benefits of continuing financial education. It is therefore important to provide early opportunities for establishing the foundations of financial literacy.

In addition to preparing young people for their adult life, financial education for youth and in schools can also address the immediate financial issues facing young people. Children are often consumers of financial services. The results of the 2012 and 2015 PISA financial literacy assessments revealed that many 15-year-old students have a bank account (OECD, 2017[24]; OECD, 2014[25]). Moreover, it is not uncommon for them to have accounts with access to online payment facilities or to use mobile phones (with various payment options) even before they become teenagers. As both young people and their families are often unfamiliar with many emerging digital financial services, financial literacy skills would clearly be of benefit to young consumers when using such products. Before leaving school, they may also need to make decisions about issues such as scooter or car insurance, savings products and overdrafts.

In many countries, at around the ages of 15 to 18, young people (and their parents) face one of their most important financial decisions: whether or not to invest in tertiary education. The gap in wages between college and non-college educated workers has widened in many economies (OECD, 2016[36]). At the same time, the education costs borne by students and their families have increased, often resulting in large student loans to repay, reducing students' ability to save and potentially leading to a reliance on credit (Dolphin, 2012[37]; OECD, 2016[36]; Ratcliffe and McKernan, 2013[38]).

Efficiency of providing financial education in schools

Research suggests that, in developed countries, there is a link between financial literacy and family economic and educational background: those who are more financially literate disproportionately come from highly educated families that hold a wide range of financial products (Lusardi, Mitchell and Curto, 2010[39]). Results of the 2012 and 2015 PISA financial literacy assessments show that a remarkable proportion of the variation in student performance in financial literacy within each country and economy is associated with their family economic, social and cultural status, and that students with at least one parent with tertiary-level education have higher scores, on average, than other students (OECD, 2017[24]; OECD, 2014[25]). In order to provide equality of opportunity, it is important to offer financial education to those who would not otherwise have access to it through their families. Schools are well positioned to advance financial literacy among all demographic groups, thereby reducing financial literacy gaps and inequalities.

Recognising both the importance of financial literacy for youth and the unique potential to improve the knowledge and skills of future generations, an increasing number of countries have embarked on the development of financial education programmes for children and young people. These include efforts to introduce financial literacy topics into existing curricular subjects, such as mathematics, social sciences or citizenship, as well as extracurricular activities, such as national awareness events and educational games.

The need for data

Policy makers, educators and researchers need high-quality data on their students' levels of financial literacy in order to inform financial education strategies and the implementation of financial education programmes in schools, by identifying priorities and measuring change across time.

Several countries have undertaken national surveys of financial literacy across their adult population. Indeed, the OECD has developed a questionnaire designed to capture levels of financial literacy amongst adults at an international level, which was used for several international comparative studies (OECD, 2017[40]; OECD, 2016[41]). However, until financial literacy was included in the PISA 2012 assessment, there were few data collection efforts on the levels of financial literacy amongst young people under the age of 18, and none that could be compared across countries.

At the national level, a robust measure of financial literacy amongst young people can help identify issues to be addressed through schools or extracurricular programmes. This measure of financial literacy can also be used as a baseline from which to gauge success and review financial education programmes in future years.

An international study provides additional benefits to policy makers and other stakeholders. Comparing levels of financial literacy across countries makes it possible to see which countries have the highest levels of financial literacy and to possibly identify particularly effective national strategies and practices. It also makes it possible to recognise common challenges and explore the possibility of finding international solutions to the issues faced.

Against this backdrop, the collection of robust and internationally comparable financial literacy data in the student population provides policy makers, educators, curriculum and resource developers, researchers and others with:

- International evidence on how young people are distributed across the financial literacy proficiency scale, which can be used to inform the development of more targeted programmes and policies;
- An opportunity to compare financial education strategies across countries and explore good practice; and, ultimately,
- Comparable data over time to track trends in financial literacy and potentially assess the association between financial literacy and the availability of financial education in schools.

The measurement of financial literacy in PISA

PISA assesses the readiness of students for their life beyond compulsory schooling and, in particular, their capacity to use their knowledge and skills, by collecting and analysing cognitive and other information from 15-year-olds in many countries and economies.

The PISA financial literacy assessment provides a rich set of comparative data that policy makers and other stakeholders can use to make evidence-based decisions about financial education. International comparative data on financial literacy can answer questions such as "How well are young people prepared for the new financial systems that are becoming more global and more complex?" and "In which countries and economies do students show high levels of financial literacy?"

As with the core PISA domains of reading, mathematics and science, the main focus of the financial literacy assessment in PISA is the proficiency of 15-year-old students in demonstrating and applying knowledge and skills. And like other PISA domains, financial literacy is assessed using an instrument designed to provide data that are valid, reliable and interpretable.

The PISA financial literacy assessment framework developed in 2012 (OECD, 2013[34]) provided a first step in constructing an assessment that satisfies these three broad criteria. It also provided national authorities with the first detailed guidance about the scope and operational definition of financial literacy, which contributed to the development of national and international frameworks, including the OECD/INFE core competencies framework on financial literacy for youth, (OECD, 2015[5]).

The main benefit of constructing an assessment framework is improved measurement, as it provides an articulated plan for developing the individual items and designing the instrument that will be used to assess the domain. A further benefit is that it provides a common language for discussion of the domain, and thereby increases understanding of what is being measured. It also promotes an analysis of the kinds of knowledge and skills associated with competency in the domain, thus providing the groundwork for building descriptions of students' proficiency at different levels that can be used to interpret the results.

The development of the PISA frameworks, for financial literacy as for the other domains, can be described as a sequence of the following six steps:

- Developing a definition for the domain and a description of the assumptions that underlie that definition;
- Identifying a set of key characteristics that should be taken into account when constructing assessment tasks for international use;
- Operationalising the set of key characteristics that will be used in test construction, with definitions based on existing literature and experience in conducting other large-scale assessments;
- Evaluating how to organise the set of tasks constructed in order to report to policy makers and researchers on the achievement in each assessment domain for 15-year-old students in participating countries;
- Validating the variables and assessing the contribution each makes to understanding task difficulty across the various participating countries; and
- Preparing a described proficiency scale for the results.

The 2018 framework maintains the definition for the financial literacy domain whilst slightly updating the operationalisation of the domain to ensure it is in line with recent developments in financial markets and the latest research findings.

Defining financial literacy

In developing a working definition of financial literacy that can be used to lay the groundwork for designing an international financial literacy assessment, the Financial Literacy Expert Group (FEG) looked both to existing definitions of literacies in the other domains assessed by PISA, and to the nature of financial education.

PISA conceives of literacy as the capacity of students to apply knowledge and skills in key subject areas and to analyse, reason and communicate effectively as they pose, solve and interpret problems in a variety of situations. PISA is forward looking, focusing on young people's ability to use their knowledge and skills to meet real-life challenges, rather than merely on the extent to which they have mastered specific curricular content (OECD, 2009[42]).

In its *Recommendation on Principles and Good Practices for Financial Education and Awareness*, the OECD defined financial education as "the process by which financial consumers/investors improve their understanding of financial products, concepts and risks and, through information, instruction and/or objective advice, develop the skills and confidence to become more aware of financial risks and opportunities, to make informed choices, to know where to go for help, and to take other effective actions to improve their financial well-being" (OECD, 2005[31]).

The FEG agreed that "understanding", "skills" and the notion of applying understanding and skills ("effective actions") were key elements of this definition. It was recognised, however, that the definition of financial education describes a process – education – rather than an outcome. What was required for the assessment framework was a definition encapsulating the outcome of that process in terms of competency or literacy.

The definition of financial literacy for PISA is as follows:

> *Financial literacy is the knowledge and understanding of financial concepts and risks, and the skills, motivation and confidence to apply such knowledge and understanding in order to make effective decisions across a range of financial contexts, to improve the financial well-being of individuals and society, and to enable participation in economic life.*

This definition, like other PISA domain definitions, has two parts. The first part refers to the kinds of thinking and behaviour that characterise the domain. The second part refers to the purposes for developing the particular literacy.

In the following paragraphs, each part of the definition of financial literacy is considered in turn to help clarify its meaning in relation to the assessment.

Financial literacy…

Literacy is viewed as an expanding set of knowledge, skills and strategies on which individuals build throughout life, rather than a line to be crossed, with illiteracy on one side and literacy on the other. Literacy involves more than the reproduction of accumulated knowledge; instead, it involves a mobilisation of cognitive and practical skills, and other resources such as attitudes, motivation and values. The PISA assessment of financial literacy draws on a range of knowledge and skills associated with the development of the capacity to deal with the financial demands of everyday life and uncertain futures within contemporary society.

…is knowledge and understanding of financial concepts and risks…

Financial literacy is thus contingent on some knowledge and understanding of the fundamental elements of the financial world, including key financial concepts as well as the purpose and basic features of financial products. This also includes risks that may threaten financial well-being as well as insurance policies and pensions. It can be assumed that 15-year-olds are beginning to acquire this knowledge and gain experience of the financial environment that they and their families inhabit and the main risks they face. All of them are likely to have been shopping to buy household goods or personal items; some will have taken part in family discussions about money and whether what is wanted is actually needed or affordable; and a sizeable proportion of them will have already begun to earn and save money. Some students already have experience of financial products and commitments through a bank account or a mobile phone contract. A grasp of concepts such as interest, inflation and value for money are soon going to be, if they are not already, important for their financial well-being.

…and the skills…

These skills include generic cognitive processes such as accessing information, comparing and contrasting, extrapolating and evaluating, but applied in a financial context. They include basic skills in mathematical literacy such as performing basic calculations, computing a percentage, or converting from one currency to another, and language skills such as the capacity to read and interpret advertising and contractual texts.

…motivation and confidence…

Financial literacy involves not only the knowledge, understanding and skills to deal with financial issues, but also non-cognitive attributes: the motivation to seek information and advice in order to engage in financial activities, the confidence to do so and the ability to manage emotional and psychological factors that influence financial decision-making. These attributes are considered to be a goal of financial education, as well as being instrumental in building financial knowledge and skills.

…to apply such knowledge and understanding in order to make effective decisions…

PISA focuses on the ability to activate and apply knowledge and understanding in real-life situations rather than the reproduction of knowledge. In assessing financial literacy, this translates into a measure of young people's ability to transfer and apply what they have learnt about personal finance into effective decision-making. The term "effective decisions" refers to informed and responsible decisions that satisfy a given need.

…across a range of financial contexts…

Effective financial decisions can refer to a range of financial contexts that relate to young people's current daily lives and experiences, but also to steps they are likely to take in the near future as adults. For example, young people may currently make relatively simple decisions such as how they will use their pocket money or which mobile phone contract they will choose, but they may soon be faced with more significant decisions about education and work options with long-term financial consequences.

…to improve the financial well-being of individuals and society…

Financial literacy in PISA is primarily conceived of as literacy around personal or household finance and is distinguished from economic literacy, which includes concepts such as the theories of supply and demand, and the structure of markets. Financial literacy is concerned with how individuals understand, manage and plan their own and their households' – which often means their families' – financial affairs. It is recognised, however, that good financial understanding, management and planning on the part of individuals has some collective impact on the wider society, in contributing to national and even global stability, productivity and development.

…and to enable participation in economic life.

Like the other definitions of literacy in PISA, the definition of financial literacy emphasises the importance of the individual's role as a thoughtful and engaged member of society. Individuals with a high level of financial literacy are better equipped to make decisions that are of benefit to themselves, and also to constructively support and critique the economic world in which they live.

Organising the domain

How the domain is represented and organised determines how the assessment is designed, including how items are developed, and, ultimately, what evidence about student proficiencies can be collected and reported. Many elements are part of the concept of financial literacy, not all of which can be taken into account and varied in an assessment such as PISA. It is necessary to select the elements that will best ensure the construction of an assessment comprising tasks with an appropriate range of difficulty and a broad coverage of the domain.

A review of approaches and rationales adopted in previous large-scale studies, and particularly in PISA, shows that most consider what content, processes and contexts are relevant for assessment as they specify what they wish to assess. Content, processes and contexts can be thought of as three different perspectives on the area to be assessed:

- *Content* comprises the areas of knowledge and understanding that are essential in the area of literacy in question;
- *Processes* describes the mental strategies or approaches that are called upon to negotiate the material; and

- *Contexts* refers to the situations in which the knowledge, skills and understandings of the domain are applied, ranging from the personal to the global.

The development of the assessment starts by identifying and weighting the different categories within each perspective, and then ensuring that the tasks adequately reflect these categories. These steps ensure the coverage and validity of the assessment. The three perspectives are also helpful in thinking about how achievement in the area is to be reported.

The following section presents a discussion of each of the three perspectives and the categories into which they are divided. For each perspective, the framework presents lists of sub-topics and examples of what students should be able to understand and do; however, these examples should not be interpreted as a checklist of tasks to be included in any one assessment. Given that only one hour of financial literacy assessment material is administered in PISA, there is not enough space to cover every detail of each perspective.

The section includes examples of items drawn from the PISA 2018 field trial and previous assessments in order to illustrate these perspectives and categories. While they are representative of those used in the 2018 main survey, these particular items are not used in the 2018 assessment instrument; only secure, unpublished items are used for this purpose, to protect the integrity of the data that is collected to measure student proficiency.

Content

The content of financial literacy is conceived of as the areas of knowledge and understanding that must be drawn upon in order to perform a particular task. A review of the content of existing financial literacy learning frameworks indicated that there is some consensus on the financial literacy content areas (OECD, 2014[32]). The review showed that the content of financial education in schools around the world was – albeit with cultural differences – relatively similar, and that it was possible to identify a series of topics commonly included in these frameworks. These form the four content areas for financial literacy in PISA: *money and transactions*, *planning and managing finances*, *risk and reward*, and *financial landscape*. The work undertaken by the OECD/INFE to develop a core competencies framework on financial literacy for youth provides additional guidance on how these content areas map to desired financial literacy outcomes (OECD, 2015[5]).

Money and transactions

This content area includes awareness of the different forms and purposes of money and managing monetary transactions, which may include being aware of national, foreign and digital currencies; making payments using a variety of available tools including mobile or online ones, taking into account value for money; and using bank cards, cheques and bank accounts. It also covers practices such as taking care of cash and other valuables, calculating value for money, and filing documents and receipts, including those received electronically.

Tasks in this content area can, for example, ask students to show that they:

- Are aware of the different forms and purposes of money:

    - Recognise bank notes and coins;
    - Understand that money can be exchanged for goods and services;
    - Identify different ways to pay for items purchased in person or at a distance (e.g. on line);
    - Recognise that there are various ways of receiving money from other people and transferring money between people or organisations, such as cash, cheques, card payments in person or on line, or electronic transfers on line or via SMS; and
    - Understand that money can be borrowed or lent, and the purpose of interest (in this respect, the assessment takes into account that the payment and receipt of interest is forbidden in some religions).

- Are confident and capable at handling and monitoring transactions. Students can show that they know how to:
    - Use cash, cards and other payment methods to purchase items;
    - Use cash machines to withdraw cash or obtain an account balance;
    - Calculate the correct change;
    - Work out which of two consumer items of different sizes would give better value for money, taking into account the individual's specific needs and circumstances;
    - Use common tools, such as spreadsheets, online platforms or mobile applications, to monitor their transactions and perform budget calculations; and;
    - Check transactions listed on a bank statement provided on paper or digitally, and note any irregularities.

The following example from the unit BANK STATEMENT illustrates a task that requires students to understand a common financial document. In this question, and in many others, the unit of currency is the imaginary Zed. PISA questions often refer to situations that take place in the fictional country of Zedland, where the Zed is the unit of currency. This artifice (about which students are informed at the beginning of the testing session) has been introduced to enhance comparability across countries.

Figure 5.1. Illustrative PISA Financial literacy item 1 – BANK STATEMENT

Each week, Mrs Citizen transfers 130 zeds into her son's bank account.

In Zedland, banks charge a fee for each transfer.

Mrs Citizen received this statement from her bank in November 2011.

ZEDBANK				
Statement for: Mrs Citizen		**Account type** Current		
Month: November 2011		**Account number:** Z0005689		
Date	**Transaction details**	**Credit**	**Debit**	**Balance**
1-Nov	Opening balance			1780.25
5-Nov	Wages	575.00		2355.25
5-Nov	Transfer		130.00	2225.25
5-Nov	Transfer fee		1.50	2223.75
12-Nov	Wages	575.00		2798.75
12-Nov	Transfer		130.00	2668.75
12-Nov	Transfer fee		1.50	2667.25
13-Nov	Withdrawal		165.00	2502.25
19-Nov	Wages	575.00		3077.25
19-Nov	Transfer		130.00	2947.25
19-Nov	Transfer fee		1.50	2945.75
26-Nov	Wages	575.00		3520.75
26-Nov	Transfer		130.00	3390.75
26-Nov	Transfer fee		1.50	3389.25
27-Nov	Withdrawal		180.00	3209.25
27-Nov	Withdrawal (Rent)		1200.00	2009.25
30-Nov	Interest	6.10		2015.35

QUESTION:

What were the total fees charged by the bank in November?

Total bank fees in zeds: …………….

This question asks students to interpret a financial document, in this case, a bank statement. Students are required to identify bank fees from the statement and to perform a basic calculation (addition or multiplication). The purpose of the question is to test whether students can find the information on the statement and notice that it is not presented as a total, but as individual transactions. Such skills are fundamental to properly understanding the information received from financial service providers. The correct answer is 6.00.

Planning and managing finances

Income, expenditure and wealth need planning and managing over both the short and long term. This content area therefore reflects the process of managing, planning and monitoring income and expenses and understanding ways of enhancing wealth and financial well-being. It includes content related to credit use as well as savings and wealth creation.

This content area includes:

- The knowledge and ability to monitor and control income and expenses:
    - Identifying various types of income (e.g. allowances, salary, commission, benefits) and ways of discussing income (such as hourly wage and gross or net annual income) and
    - Drawing up a budget to plan regular spending and saving and staying within it.
- The knowledge and ability to make use of income and other available resources in both the short and long term to enhance financial well-being:
    - Understanding how to manipulate various elements of a budget, such as identifying priorities if income does not meet planned expenses, or finding ways to increase savings, such as reducing expenses or increasing income;
    - Assessing the impact of different spending plans and the ability to set spending priorities in both the short and long term;
    - Planning ahead to pay future expenses: for example, working out how much money needs to be saved each month to make a particular purchase or pay a bill;
    - Understanding the purposes of accessing credit and the ways in which expenditure can be smoothed over time through borrowing or saving;
    - Understanding the idea of building wealth, the impact of compound interest on savings, and the advantages and disadvantages of investment products;
    - Understanding the benefits of saving for long-term goals or anticipated changes in circumstances (such as living independently); and
    - Understanding how government taxes and benefits impact personal and household finances.

The examples MUSIC SYSTEM and ZCYCLE presented below illustrate items addressing *planning and managing finances* in contexts that are relevant to 15-year-olds as they think about their lives in the near future.

Figure 5.2. Illustrative PISA Financial literacy item 2 – MUSIC SYSTEM

Kelly asks her bank to lend her 2000 zeds to buy a music system.

Kelly has the choice to repay the loan over two years or over three years. The annual interest rate on the loan is the same in each case.

The table shows the repayment conditions for borrowing 2000 zeds over two years.

Repayment period	Monthly repayment (zeds)	Total repayment (zeds)	Total interest paid (zeds)
Two years	91.67	2200.08	200.08

QUESTION:

How will the repayment conditions for borrowing 2000 zeds over three years be different to the repayment conditions over two years?

Circle "True" or "False" for each statement.

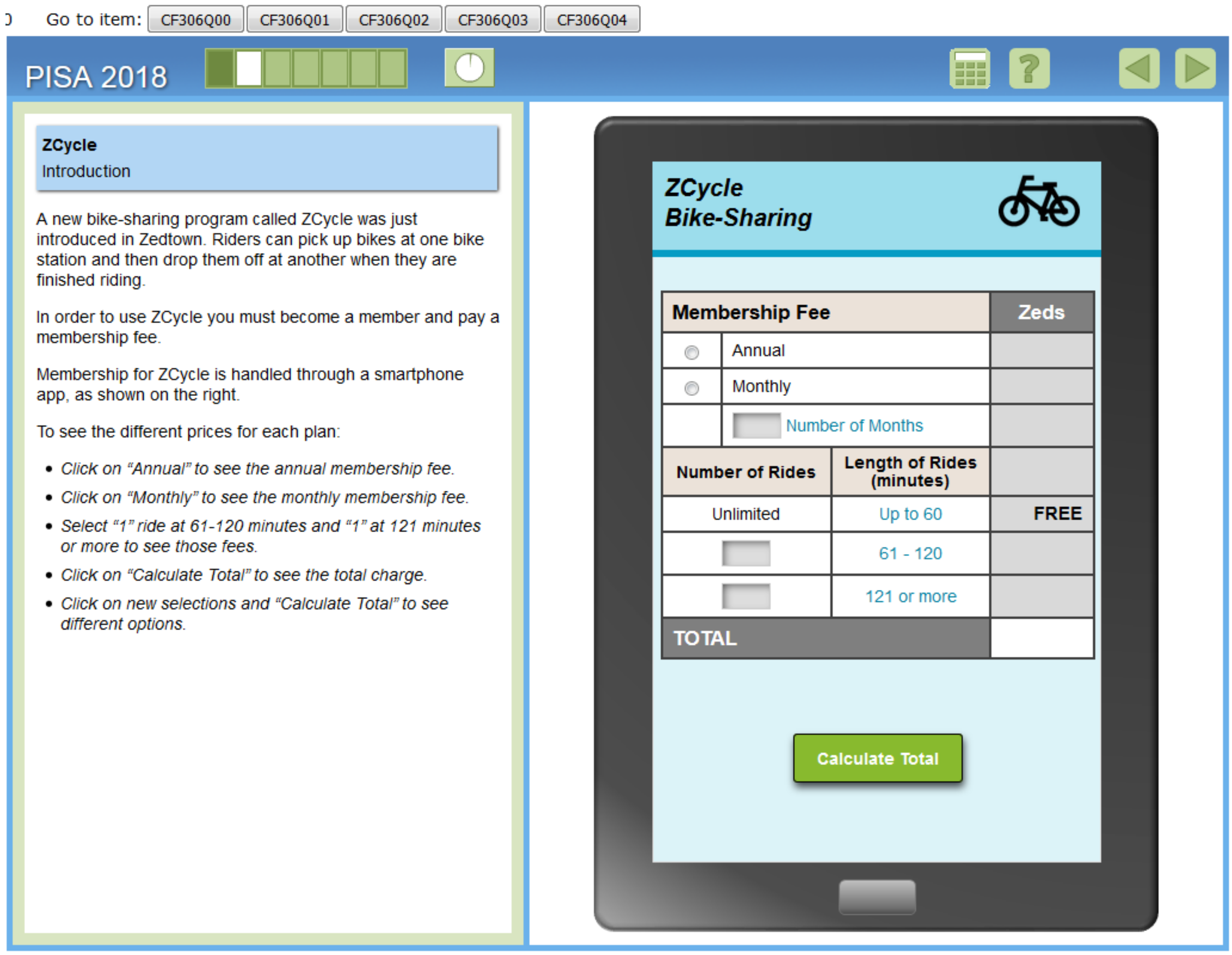
Statement	Is the statement true or false?
The monthly repayments will be larger for a loan over three years.	True / False
The total interest paid will be larger for a loan over three years.	True / False

The question MUSIC SYSTEM asks students to determine the effects of extending the loan repayment period from two to three years on the monthly interest payments and on the total interest paid when the annual interest rate does not change. As credit is widely available to young people and may be offered as an option when making a purchase in some countries, it is important that they understand how loans work so that they can make an informed decision about what is the best option for them. Students may be confronted with such a decision in the near future, for example, if they look to buy equipment to start a business or durable goods to furnish a home. The question requires anticipating the future consequences of choosing loans with different durations, without having to perform any calculations. Full credit for this question is gained by replying False and True in that order.

The unit ZCYCLE provides an example of another task that falls within the *planning and managing finances* content area. ZCYCLE is also an example of an interactive item where students use a hypothetical mobile application to find relevant information and support their reasoning.

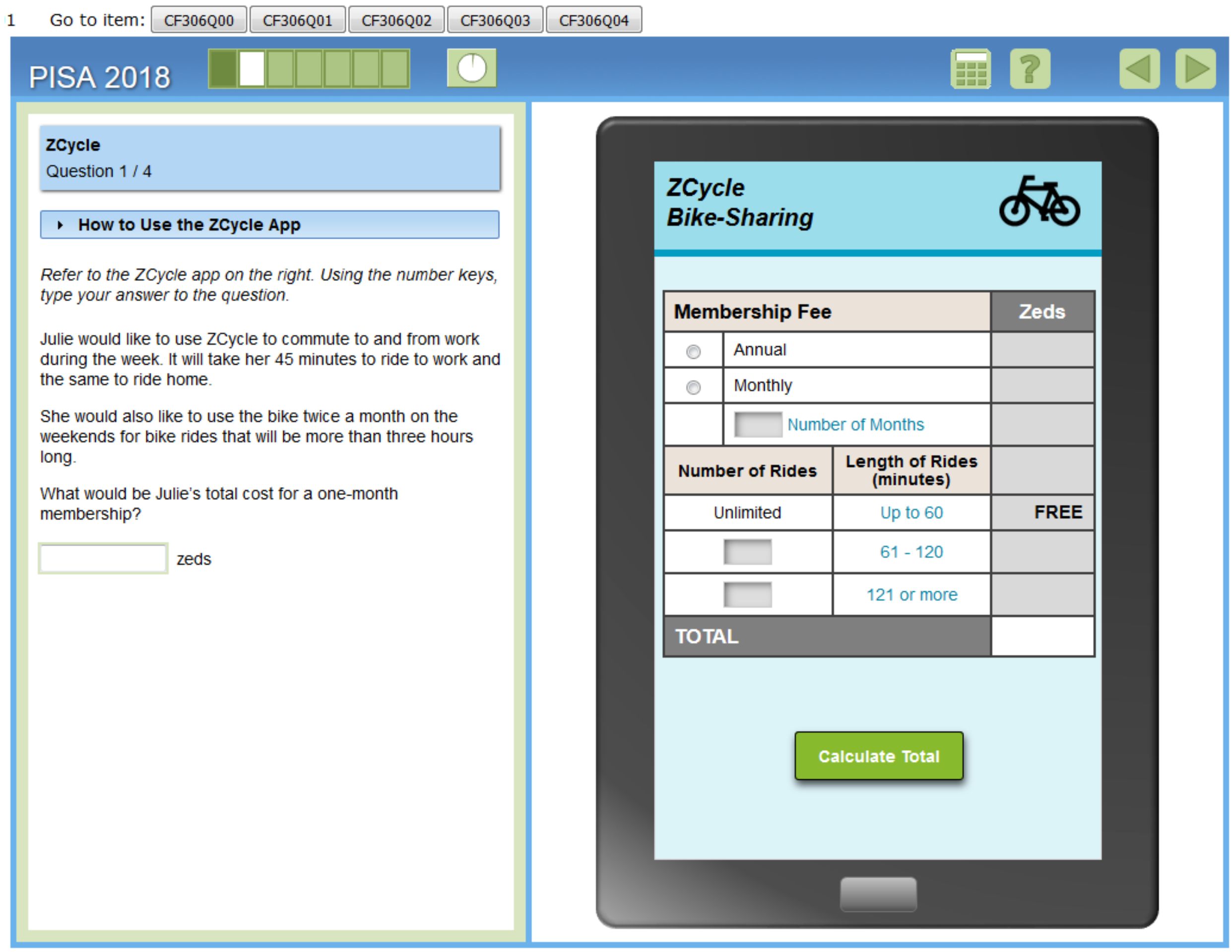
The first screenshot provides students with an introduction to a bike-sharing application that can be used to manage membership and fees.

Figure 5.3. Illustrative PISA Financial literacy item 3 – ZCYCLE



The following screenshot presents the question.

Figure 5.4. Illustrative PISA Financial literacy item 3 – ZCYCLE – Question



In this question students are asked to use the application to figure out how much membership in the bike-sharing scheme would cost given that Julie would like to use the bike for relatively short rides during the week and two longer rides during the weekend. This question falls into the *planning and managing finances content* area because students need to demonstrate the ability to put together different pieces of information on the relevant fees to choose among different options and plans. The correct response is 32 (the monthly fee is 20 zeds and each ride of at least 121 minutes costs 6 zeds).

Risk and reward

Risk and reward is a key area of financial literacy, incorporating the ability to identify ways of balancing and covering risks and managing finances in uncertainty and an understanding of the potential for financial gains or losses across a range of financial contexts. Two types of risk are of particular importance in this domain. The first relates to the risk of financial losses that an individual cannot cover using personal resources, such as those caused by catastrophic events. The second is the risk inherent in financial products, such as the risk of facing an increase in repayments on a credit agreement with variable interest rates, or the risk of loss or insufficient returns on investment products. This content area therefore includes knowledge of the types of products that may help people to protect themselves from the consequences of negative outcomes, such as insurance and savings, as well as being able to make a general assessment of the level of risk and reward related to different products, purchases, behaviours or external factors.

This content category includes:

- Recognising that certain financial products (including insurance) and processes (such as saving) can be used to manage and offset various risks (depending on different needs and circumstances):
    - Knowing how to assess whether certain insurance policies may be of benefit, and the level of cover needed.
- Applying knowledge of the benefits of contingency planning and diversification, and of the dangers of defaulting on bill and loan payments to decisions about:
    - Limiting the risk to personal capital;
    - Various types of investment and savings vehicles, including formal financial products and insurance products, where relevant; and
    - Various forms of credit, including informal and formal credit, unsecured and secured, rotating and fixed term, and those with fixed or variable interest rates.
- Knowing about and managing the risks and rewards associated with life events, the economy and other external factors, such as the potential impact of:
    - Theft or loss of personal items, job loss, birth or adoption of a child, and deteriorating health or mobility;
    - Fluctuations in interest rates and exchange rates; and
    - Other market changes.
- Knowing about the risks and rewards associated with substitutes for financial products, in particular:
    - Saving in cash, or buying property, livestock or gold as a store of wealth; and
    - Taking credit or borrowing money from informal lenders.
- Knowing that there may be unidentified risks and rewards associated with new financial products (such as mobile payment products and online credit).

An illustration from the *risk and reward* content category is provided in the example MOTORBIKE INSURANCE.

Figure 5.5. Illustrative PISA Financial literacy item 4 – MOTORBIKE INSURANCE

Last year, Steve's motorbike was insured with the PINSURA insurance company.

The insurance policy covered damage to the motorbike from accidents and theft of the motorbike.

QUESTION:

Steve plans to renew his insurance with PINSURA this year, but a number of factors in Steve's life have changed since last year.

How is each of the factors in the table likely to affect the cost of Steve's motorbike insurance this year?

Circle "Increases cost", "Reduces cost" or "Has no effect on cost" for each factor.

Factor	How is the factor likely to affect the cost of Steve's insurance?
Steve replaced his old motorbike with a much more powerful motorbike	Increases cost / Reduces cost / Has no effect on cost
Steve has painted his motorbike a different colour	Increases cost / Reduces cost / Has no effect on cost
Steve was responsible for two road accidents last year	Increases cost / Reduces cost / Has no effect on cost

Motorbike insurance falls under the content area of *risk and reward* because insurance is a product designed specifically to protect individuals against risks and financial losses that they would not otherwise be able to bear. Whilst insurance companies can provide many different products with different pricing options, they apply basic actuarial principles when calculating risk. The question tests whether students understand that the higher their risk exposure is with regards to measurable criteria, the more it will cost them to buy the same level of insurance cover. The correct answers are "Increases cost", "Has no effect on cost", and "Increases cost", in that order.

Financial landscape

This content area relates to the character and features of the financial world. It covers an awareness of the role of regulation and protection for financial consumers, knowing the rights and responsibilities of consumers in the financial marketplace and within the general financial environment, and the main implications of financial contracts that they may enter into in the near future, either with parental consent or alone. The financial landscape also takes into account the wide variety of information available on financial matters, from education to advertising. In its broadest sense, *financial landscape* also incorporates an understanding of the consequences of changes in economic conditions and public policies, such as changes in interest rates, inflation, taxation, sustainability and environmental targets or welfare benefits for individuals, households and society. The content in this area includes:

- Awareness of the role of regulation and consumer protection.
- Knowledge of rights and responsibilities, such as:
    - Understanding that buyers and sellers have rights, such as being able to apply for redress;
    - Understanding that buyers and sellers have responsibilities, such as:
        - For consumers and investors, giving accurate information when applying for financial products;
        - For providers, disclosing all material facts; and
        - For consumers and investors, being aware of the implications of one of the parties not doing so.
    - Recognising the importance of the legal documentation provided when purchasing financial products or services and the importance of understanding the content therein.
- Knowledge and understanding of the financial environment, including:
    - Identifying which providers are trustworthy, and which products and services are protected through regulation or consumer protection laws;
    - Identifying whom to ask for advice when choosing financial products, and where to go for help or guidance in relation to financial matters;
    - Awareness of existing financial crimes such as identity theft, data theft, online fraud and other scams;
    - Knowledge of how to take appropriate precautions to protect personal data and avoid other scams, and knowledge of their rights and responsibilities in the event that they are a victim; and
    - Awareness of the potential for new forms of financial crime and alertness to the risks.
- Awareness of the financial risks and implications of sharing personal data, and awareness that personal data may be used to create a person's digital profile, which can be used by companies to offer targeted products and services.
- Knowledge and understanding of the impact of their own financial decisions on themselves and others, and on the environment:
    - Understanding that individuals have choices in spending and saving, and that each action can have consequences for the individual and for society; and
    - Recognising how personal financial habits, actions and decisions have an impact at the individual, community, national and international levels.
- Knowledge of the influence of economic and external factors:

- Awareness of the economic climate and understanding of the impact of policy changes such as reforms related to the funding of post-secondary education or compulsory savings for retirement;
- Understanding how the ability to build wealth or access credit depends on economic factors such as interest rates, inflation and credit scores; and
- Understanding that a range of external factors, such as advertising and peer pressure, can affect individuals' financial choices and outcomes.

The item MOBILE PHONE CONTRACT provides an example of a question about the *financial landscape*.

Figure 5.6. Illustrative PISA Financial literacy item 5 – MOBILE PHONE CONTRACT

Alan wants a mobile phone but he is not old enough to sign the contract.

His mother buys the phone for Alan and signs a one-year contract.

Alan agrees to pay the monthly bill for the phone.

After 6 weeks, Alan’s mother discovers that the bill has not been paid.

QUESTION:

Is each statement about the mobile phone bill true or false?

Circle “True” or “False” for each statement.

Statement	Is the statement about the mobile phone bill true or false?
Alan’s mother is legally responsible for paying the bill.	True / False
The mobile phone shop must pay the bill if Alan and his mother do not.	True / False
The bill does not have to be paid if Alan returns the mobile phone to the shop.	True / False

To answer this question correctly, students should understand the legal implications of financial contracts and recognise the potential financial consequences on others (Alan's mother) if a contract is not honoured (if Alan does not pay the phone bill). Even if they cannot sign contracts at 15, students will soon be confronted with legal obligations and their financial consequences. In order to get full credit, students should answer True, False and False, in that order.

Processes

The process categories relate to cognitive processes. They are used to describe students' ability to recognise and apply concepts relevant to the domain, and to understand, analyse, reason about, evaluate and suggest solutions. Four process categories have been defined in PISA's financial literacy domain: *identify financial information*, *analyse information in a financial context*, *evaluate financial issues* and *apply financial knowledge and understanding*. While the verbs used here bear some resemblance to those in Bloom's taxonomy of educational objectives (Bloom, 1956[43]), an important distinction is that the processes in the financial literacy construct are not operationalised as a hierarchy of skills. They are, instead, parallel cognitive approaches, all of which are part of the financially literate individual's repertoire. The order in which the processes are presented here relates to a typical sequence of thought processes and actions, rather than to an order of difficulty or challenge. At the same time, financial thinking, decisions and actions are most often dependent on a recursive and interactive blend of the processes described in this section. For the purposes of this assessment, each task is identified with the process that is judged most central to its completion.

Identify financial information

This process is engaged when the individual searches and accesses sources of financial information, and identifies or recognises its relevance. In PISA 2018, the information is in the form of texts such as contracts, advertisements, charts, tables, forms and instructions displayed on screen. A typical task might ask students to identify the features of a purchase invoice, or recognise the balance on a bank statement. A more difficult task might involve searching through a contract that uses complex legal language to locate information that explains the consequences of defaulting on loan repayments. This process category is also reflected in tasks that involve recognising financial terminology, such as identifying "inflation" as the term used to describe increasing prices over time.

Example 6, PAY SLIP, shows an item that focuses on identifying and interpreting financial information.

Figure 5.7. Illustrative PISA Financial literacy item 6 – PAY SLIP

Each month, Jane's employer pays money into Jane's bank account.

This is Jane's pay slip for July.

EMPLOYEE PAY SLIP	**Jane Citizen**	
Position	Manager	1 July to 31 July
Gross salary	2 800 zeds	
Deductions	300 zeds	
Net salary	2 500 zeds	
Gross salary to date this year	19 600 zeds	

QUESTION:

How much money did Jane's employer pay into Jane's bank account on 31 July?

A. 300 zeds

B. 2 500 zeds

C. 2 800 zeds

D. 19 600 zeds

Students are asked to identify financial information in a simple pay slip and to indicate that the correct answer is 2 500 zeds.

The question BANK STATEMENT presented previously also belongs to the category *identifying financial information* as it requires the student to identify bank fees in a commonly used financial document, in that case, a bank statement.

Analyse information in a financial context

This process covers a wide range of cognitive activities undertaken in financial contexts, including interpreting, comparing and contrasting, synthesising, and extrapolating from information that is provided. Essentially, it involves recognising something that is not explicit: identifying the underlying assumptions or implications of an issue in a financial context. For example, a task may involve comparing the terms offered by different mobile phone contracts or working out whether an advertisement for a loan is likely to include unstated conditions. An example in this process category is provided below, in the unit PHONE PLANS.

Figure 5.8. Illustrative PISA Financial literacy item 7 – PHONE PLANS

Ben lives in Zedland and has a mobile phone. In Zedland, there are two different kinds of phone plans available.

Plan 1

- You pay the phone bill at the end of the month.
- The bill is the cost of the calls you make **plus** a monthly fee.

Plan 2

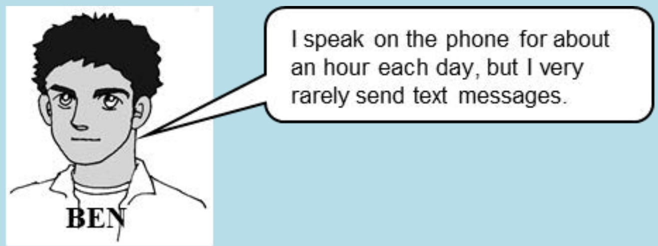
- You buy credit for the phone in advance.
- The credit lasts for a maximum of one month or until all credit has been used.

Ben decides to use Plan 1. He must now choose which phone company to use.

The table below shows the details of the four different phone companies that offer Plan 1. All costs are shown in zeds.

	Company 1	Company 2	Company 3	Company 4
Monthly fee (zeds)	20	20	30	30
Cost of call per minute (zeds)	0.27	0.25	0.30	0.25
Number of free minutes per month	90	90	60	60
Cost of text message (zeds)	0.02	0.02	free	0.01
Number of free text messages per month	200	100	unlimited	200

QUESTION:



Which company offers the best financial deal for Ben?

- ❑ Company 1
- ❑ Company 2
- ❑ Company 3
- ❑ Company 4

This question in the unit PHONE PLANS illustrates the process of *analysing information in a financial context* by looking at students' ability to select the most suitable telephone plan for a particular individual. Students are required to compare the conditions offered by different mobile phone companies by looking at multiple dimensions, such as flat fees, cost of calls and cost of messages, select the ones that are most relevant, and find the best offer for a given need. To get full credit, students should indicate that Company 2 offers the best deal for Ben's needs.

The item MUSIC SYSTEM is another example of *analysing information in a financial context* as students are asked to identify the implications of changing the duration of a loan on the total interest paid and the monthly repayments. The item MOTORBIKE INSURANCE is also an example of a question requiring students to *analyse information in a financial context*, as students have to show an understanding of the implications of different factors on the cost of insurance.

Evaluate financial issues

In this process the focus is on recognising or constructing financial justifications and explanations, by applying financial knowledge and understanding to specific contexts. It involves such cognitive activities as explaining, assessing and generalising. Critical thinking is brought into play in this process, when students must draw on knowledge, logic and plausible reasoning to make sense of and form a view about a finance-related problem. The information that is required to deal with such a problem may be partly provided in the stimulus of the task, but students will need to connect such information with their own prior financial knowledge and understandings. In the PISA context, any information that is required to understand the problem is intended to be within the expected range of experiences of a 15-year-old – either direct experiences or those that can be readily imagined and understood. For example, it is assumed that 15-year-olds are likely to be able identify with the experience of wanting something that is not essential (such as a music player or games console). A task based on this scenario could ask about the factors that might be considered in deciding on the relative financial merits of making a purchase or deferring it, given specified financial circumstances.

ONLINE SHOPPING provides an example of a demanding task that falls within the *evaluate financial issues* category.

Figure 5.9. Illustrative PISA Financial literacy item 8 – ONLINE SHOPPING

QUESTION:

Kevin is using a computer at an Internet café. He visits an online shopping website that sells sports equipment. He enters his bank card details to pay for a football.

The security of financial information is important when buying goods on line.

What is one thing Kevin could have done to increase security when he paid for the football on line?

The question ONLINE SHOPPING asks students to reflect on the potential risks of conducting financial transactions on line using computers in public places and to evaluate those risks. Internet cafés are less widespread in developed countries than they were when the item was designed, but young people may still be sharing computers with friends, and may make payments on line in public places or use public Wi-Fi to access personal data. In the case of ONLINE SHOPPING, all of the necessary information is provided in the question, but to gain credit students need to identify what is relevant and reflect on the consequences of taking a particular action. Various responses are awarded full credit, such as referring to using a secure computer rather than one in a public place, using a more secure or safer method of online payment, or using a trusted website.

The item MOBILE PHONE CONTRACT is also an example of evaluating financial issues, because students should use their critical thinking to recognise the implications of a contract.

Apply financial knowledge and understanding

The fourth process picks up a term from the definition of financial literacy: "to apply such [financial] knowledge and understanding". It focuses on taking effective action in a financial setting by using knowledge of financial products and contexts, and understanding of financial concepts. This process is reflected in tasks that involve performing calculations and solving problems, often taking into account multiple conditions. An example of this kind of task is calculating the interest on a loan over two years. This process is also reflected in tasks that require recognition of the relevance of prior knowledge in a specific context. For example, a task might require the student to work out whether purchasing power will increase or decrease over time when prices are changing at a given rate. In this case, knowledge about inflation needs to be applied.

The following example, RINGTONES, falls into the process category *apply financial knowledge and understanding*.

Figure 5.10. Illustrative PISA Financial literacy item 9 – RINGTONES

Colin sees this advertisement in a magazine for teenagers.

Get Cheeky Monkey™ ringtones for your phone.

Your phone will make a monkey noise when your friends call you.

Get one NOW for only 3 zeds*

Text the word MONK to 13 45 67

* Each ring-tone costs 3 zeds. By texting MONK to 13 45 67 customer agrees to receive a different Cheeky Monkey™ ring-tone every day. Customer can cancel contract at any time by texting STOP to 13 45 67. Cancellation fee is 5 zeds.

QUESTION:

Colin has 30 zeds credit on his phone.

He texts the word MONK to 13 45 67.

Colin does **not** use his phone again to make calls or send texts. He does **not** add any more credit.

How much credit will Colin have on his phone exactly one week later?

Credit in zeds: ………………

This question asks students to pay attention and interpret the small print to understand the terms and conditions of buying a service, and then to calculate the implications for the true cost. When developed as a test item for the 2012 assessment, this question presented a widely relevant situation; while ads for ringtones may have changed in the meantime in some countries, students continue to receive advertisements, as for purchases through digital games and apps, in a similar format. The question falls in the category *Apply financial knowledge and understanding* because it asks students to perform basic calculations (multiplication and subtraction) taking into account multiple elements that are not immediately evident (by buying one ringtone the user agrees to receiving – and pays for – a ringtone every day). This item also highlights a wider issue that young people face when starting to make financial decisions and budget their own money. An impulse decision to make a purchase of 3 zeds without first reading the small print would cost the student a minimum of 8 zeds even if they recognised their error immediately. The correct response is 9 or 6, recognising the potential ambiguity as to when the first or last download occurs.

Contexts

In building a framework, and developing and selecting assessment items based on this framework, attention is given to the breadth of contexts in which the domain literacy is exercised. Decisions about financial issues are often dependent on the contexts or situations in which they are presented. By situating tasks in a variety of contexts, the assessment offers the possibility of connecting with the broadest possible range of situations in which individuals need to function in the 21st century.

Certain situations will be more familiar to 15-year-olds than others. In PISA, assessment tasks are set in general, everyday situations, which may take place in but are not confined to the school. The focus may be on the individual, family or peer group, on the wider community, or even more widely on a global scale.

The contexts identified for the PISA financial literacy assessment are *education and work, home and family, individual* and *societal*.

Education and work

The context of *education and work* is of great importance to young people. Virtually all 15-year-olds will be starting to think about financial matters related to both education and work, whether they are spending existing earnings, considering future education options or planning their working life.

The educational context is obviously relevant to students sitting the PISA assessment, since they are by definition a sample of the school-based population; indeed, many of them will continue in education or training for some time. However, many 15-year-old students are also already engaged in some form of paid work outside school hours making the work context equally valid. Furthermore, many will move from education into some form of employment, including self-employment, before reaching their twenties.

Typical tasks within this context include understanding payslips, planning to save for tertiary education, investigating the benefits and risks of taking out a student loan, and participating in workplace savings schemes.

Item PAY SLIP illustrates a task designed to fall into the *education and work* context category. This question asks students to address problems related to earning income and identifying information on a payslip, which is a situation that students will soon encounter as they grow up.

Home and family

*Home and famil*y includes financial issues relating to the costs involved in running a household. Family is the most likely household circumstance for 15-year-olds; however, this category also encompasses households that are not based on family relationships, such as the kind of shared accommodation that young people often use shortly after leaving the family home. Tasks within this context include buying household items or family groceries, keeping records of family spending and making plans for family events. Decisions about budgeting and prioritising spending may also be framed within this context.

Both the items MOBILE PHONE CONTRACT and BANK STATEMENT discussed previously provide an illustration of the *Home and family* context, as both look at the interaction between a parent and her child and present typical situations that may happen in a family.

Individual

The context of the *individual* is important within personal finance since there are many decisions that a person takes entirely for personal benefit or gratification, and many risks and responsibilities that must be borne by individuals. These decisions span essential personal needs, as well as leisure and recreation. They include choosing personal products and services such as clothing, toiletries or haircuts; buying consumer goods such as electronic or sports equipment; as well as subscriptions for season tickets or gym membership. They also cover the process of making personal decisions and the importance of ensuring individual financial security, such as keeping personal information safe and being cautious about unfamiliar products.

Although the decisions made by an individual may be influenced by the family and society (and may impact society), when it comes to opening a bank account, buying shares or getting a loan it is typically the individual who has the legal responsibility and ownership associated with the operation. The *individual* context therefore includes contractual issues around events such as opening a bank account, purchasing consumer goods, paying for recreational activities, and dealing with relevant financial services that are often associated with larger consumption items, such as credit and insurance.

Items from the *individual* context category include MUSIC SYSTEM, ZCYCLE, MOTORBIKE INSURANCE, PHONE PLANS and RINGTONES. All of them illustrate decisions that impact the individual, like choosing a loan or a phone plan, planning monthly expenses, renewing insurance, and paying attention to hidden costs.

Societal

The environment young people are living in is characterised by change, complexity and interdependence. Globalisation is creating new forms of interdependence where actions are subject to economic influences and their consequences stretch well beyond the individual and the local community. While the core of the financial literacy domain is focused on personal finances, the *societal* context recognises that individual financial well-being cannot be entirely separated from the rest of society. Personal financial well-being affects and is affected by the local community, the nation and even global activities. Financial literacy within this context includes matters such as being informed about consumer rights and responsibilities, understanding the purpose of taxes and local government charges, being aware of business interests, and taking into account the role of consumer purchasing power. It extends also to considering financial choices such as donating to non-profit organisations and charities.

The task ONLINE SHOPPING shown earlier is categorised as falling within the *societal* context, since it relates to the protection of financial data and the risk of fraudulent behaviour targeted across society.

Non-cognitive factors

The PISA working definition of financial literacy includes the non-cognitive terms *motivation* and *confidence*, attitudes which, according to some, have an influence on money management behaviour (Mandell and Klein, 2009[44]; Arellano, Cámara and Tuesta, 2014[45]; Palameta et al., 2016[46]). PISA conceives of both financial attitudes and behaviour as aspects of financial literacy in their own right. Attitudes and behaviour are also of interest in terms of how they interact with the cognitive elements of financial literacy. Information collected about the financial attitudes and behaviour of 15-year-olds might constitute useful baseline data for any longitudinal investigation of the financial literacy of adults, including their financial behaviours.

The FEG identified four non-cognitive factors for inclusion in the framework: *access to information and education, access to money and financial products, attitudes towards and confidence about financial matters,* and *spending and saving behaviour*.

Access to information and education

There are various sources of financial information and education that may be available to students, including informal discussion with friends, parents or other family members, information from the financial sector, as well as formal school education. The literature in this area often refers to the process of ‘financial socialisation’, which can be seen as the process of acquiring financial literacy. Parents have a major role in the financial socialisation of children but, as discussed above, they may not have experience with all the financial contexts and decisions that their children face (Gudmunson and Danes, 2011[47]; Otto, 2013[48]). Copying and discussing financial behaviours with friends is another important source of socialisation, but this also may vary in terms of quality and reliability, with research from the UK indicating that money is rarely talked about honestly (Money Advice Service, 2014[49]). Moreover, the amount and quality of formal education and training about money and personal finance received by students varies within and across countries (OECD, 2014[32]).

Data about students’ access to financial information and education can be collected through both the student questionnaire and the questionnaire for school principals. In the student questionnaire, students can be asked about the typical sources of information that they access in order to analyse the extent to which each source is correlated with financial literacy. This is intended to provide a description of students’ main sources of financial socialisation, rather than assessing whether they understand the importance of using appropriate sources of information or advice, which is covered in the cognitive assessment. The 2018 student questionnaire also asks students whether they have heard of or learnt about specific financial concepts during school lessons and whether they have encountered some types of tasks about money matters at school.

In addition, the school questionnaire can ask principals about the availability and quality of financial education in their schools. Evidence about the extent to which there is a link between levels of financial literacy and financial education inside and outside schools is likely to be particularly useful in shaping education programmes for improving financial literacy.

Access to money and financial products

The results of the 2012 and 2015 PISA financial literacy exercise showed that in some countries, students with a bank account scored higher in financial literacy than students with similar socio-economic status who did not hold a bank account (OECD, 2017[24]; OECD, 2014[25]). Whilst this does not indicate a causal relationship, it is plausible to assume that real-life experiences of financial products may influence young people’s financial literacy and vice versa. Personal experience may come, for example, from using financial products such as payment cards, from dealing with the banking system, or from occasional working activities outside of school hours. In order to further understand the potential role of learning through experience, the 2018 non-cognitive student questionnaire collects evidence on a range of practical financial experiences, such as making payments using a mobile phone or making online purchases.

Students who have had more personal experience of dealing with financial matters from earning money or receiving an allowance might also be expected to perform better on the cognitive assessment than those without such experience; however, a recent review suggests that the key factor may not be experience, but the extent to which parents are involved in the spending decisions made by young people, with higher financial literacy associated with parents who are more involved (Drever et al., 2015[50]). The 2018 framework therefore recognises the importance of knowing whether students have access to money, through which channels, and to what extent spending and savings decisions are discussed with parents.

Attitudes towards and confidence about financial matters

The PISA definition of financial literacy highlights the important role of attitudes. Individual preferences can be related to financial behaviour and the ways in which financial knowledge is used. PISA 2012 showed that students' perseverance and openness to problem solving were strongly associated to their financial literacy scores (OECD, 2014[25]). PISA 2015 showed a positive association between students' financial literacy and their motivation to achieve (OECD, 2017[24]). In addition to this, the extent to which students believe that they are in control of their future and their preference for current consumption may influence their financial decisions, their independence, and their propensity to learn how to make plans for their own financial security (Golsteyn, Grönqvist and Lindahl, 2014[51]; Lee and Mortimer, 2009[52]; Meier and Sprenger, 2013[53]). Moreover, confidence in one's own ability to make a financial decision may make it more likely that a student will work through complex financial problems or carefully make choices across several possible products. At the same time, however, confidence may turn into over-confidence, leading to mistakes and overly risky decisions. The 2018 framework therefore recognises the importance of investigating students' perception of their own financial skills and asks them about their confidence in dealing with various financial matters, from understanding a bank statement to using digital devices to make payments.

Spending and saving behaviour

While items on the cognitive assessment test students' ability to make particular spending and savings decisions, it is also useful to have some measure of what their actual (reported) behaviour is, that is, how students save and spend in practice. The PISA financial literacy assessment provides the opportunity to look at the potential relationship between 15-year-olds' spending and saving behaviour and their results on the cognitive financial literacy assessment. In particular, the PISA 2018 assessment explores how students make spending decisions, such as whether they compare prices, check change or buy items that cost more that they intended to spend, and whether decisions are made alone or with the guidance or recommendation of a trusted adult.

Assessing financial literacy

The structure of the assessment

In 2012, the PISA financial literacy assessment was developed as a one-hour pen-and-paper exercise, to be completed alongside one hour of material from other cognitive domains. The financial literacy assessment was comprised of 40 items divided into two clusters, chosen from 75 tasks initially administered during the field trial. The choice of items was made based on their psychometric properties, ensuring that each item discriminated between high- and low-scoring students.

In 2015, items were transferred to a computer-based delivery platform, and additional items were developed for this form of delivery in order to replace items that had been released to the public in the report of the 2012 results and were therefore no longer valid for testing purposes. The 2015 financial literacy assessment was developed as a one-hour exercise, comprising 43 items divided into two clusters. All students sitting the financial literacy test also sat the standard PISA test of science, reading and mathematics.

New items were developed for the 2018 assessment. These incorporate specially developed interactive elements, in order to provide additional reality and interest for students. For instance, some interactive items require the student to actively seek more information by clicking links, rather than relying solely on the information presented on the first screen. Others include graphs that can be manipulated to see a variety of potential outcomes. Such items allow the student to test different scenarios and explain why certain outcomes occur, while at the same time eliminating the need to make calculations and allowing students to focus on financial decisions.

Twenty new interactive items were designed for the 2018 field trial in order to cover all dimensions of the framework and the different levels of difficulty. Out of these, 14 items were retained for the main 2018 survey. These are used alongside non-interactive items developed for the 2012 and 2015 assessments in order to ensure that the overall set of items continues to provide the necessary links across waves of data collection, and to provide the necessary breadth of coverage across the framework. Overall, the PISA 2018 financial literacy assessment consists of 43 items for a total of a one-hour financial literacy exercise.

As with other PISA assessment domains, computer-based financial literacy items are grouped in units comprising one or more items based around a common stimulus. The selection includes financially-focused stimulus material in diverse formats, including prose, diagrams, tables, charts and illustrations. All financial literacy assessments comprise a broad sample of items covering a range of difficulty that allow the strengths and weaknesses of students and key subgroups to be measured and described.

Response formats and coding

Some PISA items require short descriptive responses, others require more direct responses of one or two sentences or a calculation, and some can be answered by checking a box. Decisions about the form in which the data are collected – the response formats of the items – are based on what is considered appropriate given the kind of evidence that is being collected, and also on technical and pragmatic considerations. In the financial literacy assessment as in other PISA assessments, two broad types of items are used: constructed-response items and selected-response items.

Constructed-response items require students to generate their own answers. The format of the answer may be a single word or figure, or may be longer: a few sentences or a worked calculation. Constructed-response items that require a more extended answer are ideal for collecting information about students' capacity to explain decisions or demonstrate a process of analysis.

The second broad type of item in terms of format and coding is selected response. This kind of item requires students to choose one or more alternatives from a given set of options. The most common type in this category is the simple multiple-choice item, which requires the selection of one from a set of (usually) four options. A second type of selected-response item is the complex multiple-choice item, in which students respond to a series of "Yes/No"-type questions. Selected-response items are typically regarded as most suitable for assessing whether students can identify and recognise information, but they are also a useful way of measuring students' understanding of higher-order concepts that they may not easily be able to express.

Although particular item formats lend themselves to specific types of questions, care needs to be taken that the format of the item does not affect the interpretation of the results. Research suggests that different groups (for example, boys and girls, or students in different countries) respond differently to the various item formats. Several research studies on response format effects based on PISA data suggest that there are strong arguments for retaining a mixture of multiple-choice and constructed-response items. In their comparison of the PISA reading literacy assessment and the IEA Reading Literacy Study (IEARLS), Lafontaine and Monseur (2006[54]) found that response format had a significant impact on the performance of the different genders. In another study, countries were found to show differential item difficulties in the PISA reading assessment on items in different formats (Grisay and Monseur, 2007[55]). This finding may relate to the fact that students in different countries are not equally familiar with the particular formats. In summary, the PISA financial literacy option includes items in a variety of formats to minimise the possibility that the item format influences overall student performance. Such an influence would detract from the intended object of measurement, in this case, financial literacy.

The resources available to code students' responses and the equity issues discussed in the preceding paragraphs must both be weighed when considering the distribution of item formats. Selected-response items have predefined correct answers that can be computer-coded, therefore demanding fewer resources. While some of the constructed-response items are automatically coded by computer, some elicit a wider variety of responses that cannot be categorised in advance, thus requiring human coding from expert judges. The proportions of constructed- and selected-response items are determined taking into account all of these considerations. The majority of the items selected for the PISA 2018 main survey were automatically coded. Only 13 out of 43 items were human-coded.

Most items are coded dichotomously (full credit or no credit), but where appropriate, an item's coding scheme allows for partial credit. Partial credit makes possible a more nuanced scoring of items, to take into account that some answers, even though incomplete, are better than others.

Distribution of score points

In this section, we outline the distribution of score points across the categories of the three main framework characteristics discussed previously. The term "score points" is used in preference to "items", as some partial credit items are included. The distributions are expressed in terms of ranges, indicating the approximate weighting of the various categories. The assessment contains a mix of original items, developed for the 2012 assessment, and those items developed for the 2015 and 2018 assessments. In particular, care was taken to ensure that the interactive items cover most of the framework perspectives discussed above (content areas, processes and contexts).

While each PISA financial literacy item is categorised according to a single content area, a single process and a single context category, it is often the case that elements of more than one category are present in a task. In such cases, the item is identified with the category judged most integral to responding successfully to the task.

The target distribution of score points by financial literacy content areas is shown in Table 5.1. The distribution reflects that *money and transactions* is considered to be to the most relevant content area for 15-year-olds.

Table 5.2 shows the target distribution of score points over the four processes. The weighting shows that greater importance was attributed to *evaluating financial issues* and *applying financial knowledge and understanding*.

Table 5.3 shows the target distribution of score points over the four contexts. Consistent with an assessment of the personal financial literacy of 15-year-olds, there is a clear emphasis on scenarios focussing on the *individual*, but also a weighting towards the financial interests of the household or family unit. *Education and work* and *societal* contexts are given less emphasis, but included in the scheme as they are important elements of financial experience.

Table 5.1. Approximate distribution of score points in financial literacy, by content

Money and transactions	Planning and managing finances	Risk and reward	Financial landscape	Total
30% - 40%	25% - 35%	15% - 25%	10% - 20%	100%

Table 5.2. Approximate distribution of score points in financial literacy, by process

Identify financial information	Analyse information in a financial context	Evaluate financial issues	Apply financial knowledge and understanding	Total
15% - 25%	15% - 25%	25% - 35%	25% - 35%	100%

Table 5.3. Approximate distribution of score points in financial literacy, by contexts

Education and work	Home and family	Individual	Societal	Total
10% - 20%	30% - 40%	35% - 45%	5% -15%	100%

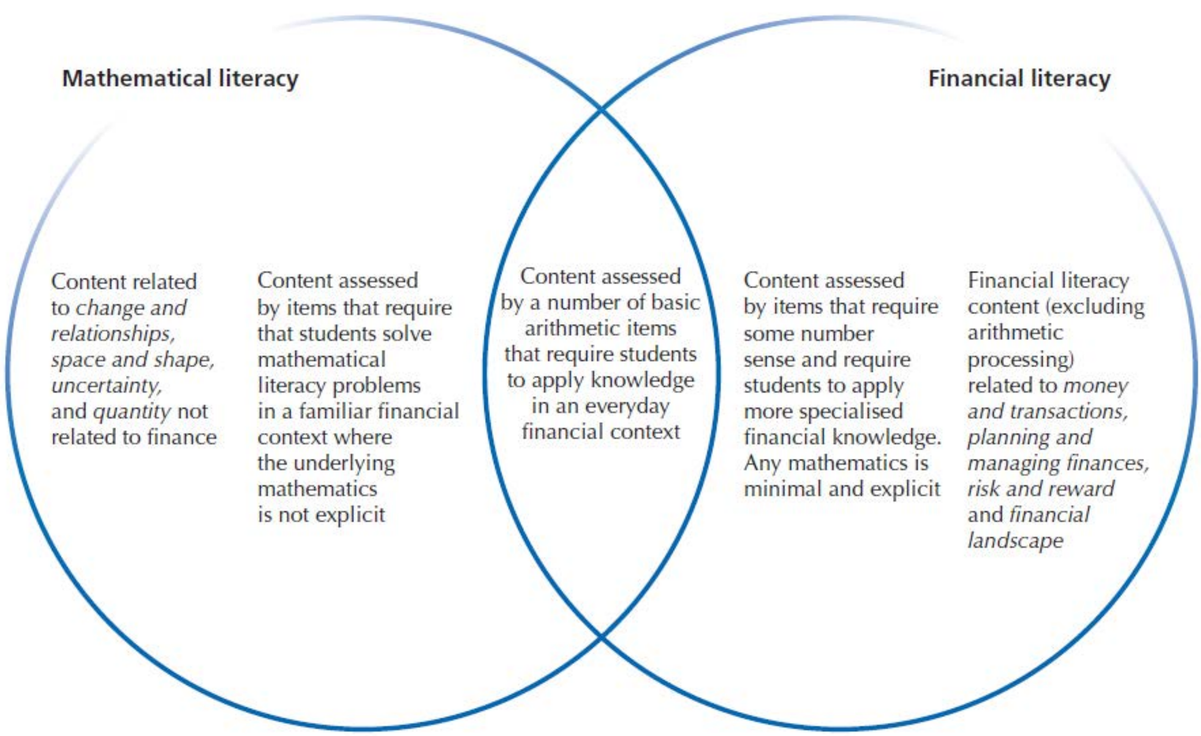
The impact of knowledge and skills in other domains on financial literacy

A certain level of numeracy (or mathematical literacy) is regarded as a necessary condition of financial literacy. Houston (2010[56]) argues that "if an individual struggles with arithmetic skills, this will certainly impact his/her financial literacy. However, available tools (e.g. calculators) can compensate for these deficiencies; thus, information directly related to successfully navigating personal finances is a more appropriate focus than numeracy skills for a financial literacy measure".

Mathematically-related proficiencies such as number sense, familiarity with multiple representations of numbers, and skills in mental calculation, estimation and the assessment of the reasonableness of results are intrinsic to some aspects of financial literacy.

On the other hand, there are large areas where the content of mathematical literacy and financial literacy do not intersect. As defined in the PISA 2012 mathematics literacy framework, mathematical literacy incorporates four content areas: *change and relationships, space and shape, quantity* and *uncertainty*. Of these, only *quantity* directly intersects with the content of the PISA financial literacy assessment. Unlike the mathematical literacy content area *uncertainty*, which requires students to apply probability measures and statistics, the financial literacy content area *risk and reward* requires an understanding that there will be a risk of losing money and (sometimes) a possibility of gains in a particular situation or for a particular financial product. This is a non-numerical appreciation of the way financial well-being can be affected by chance and an awareness of the related products and actions to protect against loss. In the financial literacy assessment, quantity-related proficiencies are applied to problems requiring more financial knowledge than can be expected in the mathematical literacy assessment. Figure 5.11 represents the relationship between the content of mathematical literacy and financial literacy in PISA.

Figure 5.11. Relationship between the content of financial literacy and mathematical literacy in PISA



Operationally, there are few items populating the portion of the diagram where the two circles intersect. In the financial literacy assessment, the nature of the mathematical literacy expected is basic arithmetic: the four operations (addition, subtraction, multiplication and division) with whole numbers, decimals and common percentages. Such arithmetic occurs as an intrinsic part of the financial literacy context and enables financial literacy knowledge to be applied and demonstrated. Use of financial formulae (requiring capability with algebra) is not considered appropriate for the assessment. The assessment minimises the need for substantial or repetitive calculation. The calculators used by students in their classrooms and on the PISA mathematics assessment will also be available in the financial literacy assessment, but success in the items will not depend on calculator use.

A similar reasoning holds for reading skills. It is assumed that all students taking part in the financial literacy assessment will have some basic reading proficiency, even while it is known from previous PISA surveys that reading proficiency varies widely both within and across countries (OECD, 2010[57]). To minimise the level of reading literacy required, stimulus material and task statements are generally designed to be as clear, simple and brief as possible. In some cases, however, stimulus may deliberately present complex or somewhat technical language: the capacity to read and interpret the language of financial documents or pseudo-financial documents is regarded as part of financial literacy.

Highly technical terminology relating to financial matters is avoided. The FEG has advised on terms that it judges reasonable to expect 15-year-olds to understand. Some of these terms may be the focus of assessment tasks.

In practice, the results of the 2012 and 2015 PISA financial literacy assessments gave a more precise measure of how students' performance in financial literacy was related to their mathematics and reading performance. In 2015, around 38% of the financial literacy score reflected factors that are uniquely captured by the financial literacy assessment (25% in 2012), while the remaining 62% of the financial literacy score reflected skills measured in the mathematics and/or reading assessments (75% in 2012) (OECD, 2017[24]; OECD, 2014[25]). The association between financial literacy and the other domains indicates that, in general, students who perform at higher levels in mathematics and/or reading also perform well in financial literacy. There were, however, wide variations in financial literacy performance for any given level of performance in mathematics and reading (OECD, 2017[24]; OECD, 2014[25]).

Reporting financial literacy

The data from the 2012 and 2015 financial literacy assessments are available at http://www.oecd.org/pisa/data/. The databases include financial literacy, reading and mathematics scores (as well as science scores in 2015), behavioural data from the short questionnaire on financial literacy, and data from the general student questionnaire and school questionnaire (a section on financial education was included in the school questionnaire only in 2012).

In each PISA cycle, financial literacy is discussed as an independent result, and in relation to performance in other domains, financial behaviour, and some background variables, such as gender, socioeconomic status and immigrant status. The data also allow the development of further work under the aegis of the OECD Project on Financial Education.

The financial literacy cognitive data is scaled in a similar way to data in other PISA domains. A comprehensive description of the modelling technique used for scaling can be found in the *PISA Technical Report* (OECD, 2014[58]; OECD, 2017[59]).

Each item is associated with a particular location on the PISA financial literacy scale of difficulty, and each student's performance is associated with a particular location on the same scale that indicates the student's estimated proficiency.

As with the other PISA domains, the relative difficulty of tasks in a test is estimated by considering the proportion of test takers getting each task correct. The relative proficiency of each student is estimated by considering the proportion of test items that they answer correctly and the difficulty of those items. A single continuous scale showing the relationship between the difficulty of items and the proficiency of students is constructed. Following PISA practice, a scale is constructed with a mean of 500 and a standard deviation of 100 among participating OECD countries.

The scale was divided into levels according to a set of statistical principles, following which descriptions of each level were generated based on the tasks located within each level. These descriptions encapsulate the kinds of skills and knowledge needed to successfully complete those tasks. The scale and set of descriptions are known as a described proficiency scale. The described proficiency scale helps in interpreting what students' financial literacy scores mean in substantive terms.

Five levels of proficiency in financial literacy were described in the 2012 assessment (OECD, 2014[25]). The same descriptions of the proficiency levels are used in the 2015 and 2018 financial literacy assessment.

Notes

[1] Financial inclusion has increased from 51% of the adult population with an account at a financial institution or mobile money service in 2011, to 62% in 2014. However, two billion adults remain unbanked (Demirguc-Kunt et al., 2015[60]).

[2] PISA 2012 indicates that students with a parent working in the financial services sector have higher levels of financial literacy on average, although data are only available for a limited number of countries.

References

Amagir, A. et al. (2018), “A review of financial-literacy education programs for children and adolescents”, *Citizenship, Social and Economics Education*, Vol. 17/1, pp. 56-80, http://dx.doi.org/10.1177/2047173417719555. [13]

APEC (2012), *2012 APEC Finance Ministers Statement*, https://www.apec.org/Meeting-Papers/Sectoral-Ministerial-Meetings/Finance/2012_finance (accessed on 5 March 2018). [33]

Arellano, A., N. Cámara and D. Tuesta (2014), “The Effect of Self-Confidence on Financial Literacy”*, BBVA Research Working Paper*, No. 14/28, BBVA Research, Madrid, https://www.bbvaresearch.com/wp-content/uploads/2014/10/WP14-28The-effect-of-self-confidence-on-financial-literacy.pdf. [45]

Atkinson, A. et al. (2015), “Financial Education for Long-term Savings and Investments: Review of Research and Literature”*, OECD Working Papers on Finance, Insurance and Private Pensions*, No. 39, OECD Publishing, Paris, http://dx.doi.org/10.1787/5jrtgzfl6g9w-en. [14]

Barrett, A., I. Mosca and B. Whelan (2013), “(Lack of) Pension Knowledge”*, IZA DP*, No. 7596, http://ftp.iza.org/dp7596.pdf (accessed on 5 March 2018). [9]

Behrman, J. et al. (2012), “How Financial Literacy Affects Household Wealth Accumulation.”, *The American economic review*, Vol. 102/3, pp. 300-304, http://www.ncbi.nlm.nih.gov/pubmed/23355747 (accessed on 5 March 2018). [26]

Bloom, B. (1956), *Taxonomy of educational objectives: The classification of educational goals.*, Prentice Hall. [43]

Bruhn, M. et al. (2016), “The Impact of High School Financial Education: Evidence from a Large-Scale Evaluation in Brazil”, *American Economic Journal: Applied Economics*, Vol. 8/4, pp. 256-295, http://dx.doi.org/10.1257/app.20150149. [15]

Clark, R., A. Lusardi and O. Mitchell (2017), “Financial knowledge and 401(k) investment performance: a case study”, *Journal of Pension Economics and Finance*, Vol. 16/03, pp. 324-347, http://dx.doi.org/10.1017/S1474747215000384. [18]

Colombo, F. et al. (2011), *Help Wanted?: Providing and Paying for Long-Term Care*, OECD Health Policy Studies, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264097759-en. [8]

Demirguc-Kunt, A. et al. (2015), *The Global Findex Database 2014 : measuring financial inclusion around the world*, http://documents.worldbank.org/curated/en/187761468179367706/The-Global-Findex-Database-2014-measuring-financial-inclusion-around-the-world (accessed on 5 March 2018). [60]

Disney, R. and J. Gathergood (2013), “Financial literacy and consumer credit portfolios”, *Journal of Banking & Finance*, Vol. 37/7, pp. 2246-2254, http://dx.doi.org/10.1016/J.JBANKFIN.2013.01.013. [28]

Dolphin, T. (2012), *Young people and savings*, Institute for Public Policy Research, https://www.ippr.org/files/images/media/files/publication/2012/11/young-people-savings_Nov2012_9849.pdf?noredirect=1 (accessed on 5 March 2018). [37]

Drever, A. et al. (2015), "Foundations of Financial Well-Being: Insights into the Role of Executive Function, Financial Socialization, and Experience-Based Learning in Childhood and Youth", *Journal of Consumer Affairs*, Vol. 49/1, pp. 13-38, http://dx.doi.org/10.1111/joca.12068. [50]

G20 (2013), *2013 St Petersburg G20 Leaders' Declaration*, http://www.g20.utoronto.ca/2013/2013-0906-declaration.html (accessed on 5 March 2018). [7]

G20 (2012), *G20 Los Cabos 2012: G20 Leaders Declaration*, http://www.g20.utoronto.ca/2012/2012-0619-loscabos.html (accessed on 5 March 2018). [2]

Gaudecker, H. (2015), "How Does Household Portfolio Diversification Vary with Financial Literacy and Financial Advice?", *The Journal of Finance*, Vol. 70/2, pp. 489-507, http://dx.doi.org/10.1111/jofi.12231. [19]

Golsteyn, B., H. Grönqvist and L. Lindahl (2014), "Adolescent Time Preferences Predict Lifetime Outcomes", *The Economic Journal*, Vol. 124/580, pp. F739-F761, http://dx.doi.org/10.1111/ecoj.12095. [51]

Grisay, A. and C. Monseur (2007), "MEASURING THE EQUIVALENCE OF ITEM DIFFICULTY IN THE VARIOUS VERSIONS OF AN INTERNATIONAL TEST", *Studies in Educational Evaluation*, Vol. 33/1, pp. 69-86, http://dx.doi.org/10.1016/J.STUEDUC.2007.01.006. [55]

Grohmann, A., T. Kluhs and L. Menkhoff (2017), "Does Financial Literacy Improve Financial Inclusion? Cross Country Evidence"*, Discussion paper*, No. 1682, DIW Berlin, http://dx.doi.org/10.2139/ssrn.3034178. [22]

Gudmunson, C. and S. Danes (2011), "Family Financial Socialization: Theory and Critical Review", *Journal of Family and Economic Issues*, Vol. 32/4, pp. 644-667, http://dx.doi.org/10.1007/s10834-011-9275-y. [47]

Hastings, J. and L. Tejeda-Ashton (2008), "Financial Literacy, Information, and Demand Elasticity: Survey and Experimental Evidence from Mexico"*, Working Paper*, No. 14538, National Bureau of Economic Research, Cambridge, MA, http://dx.doi.org/10.3386/w14538. [20]

Huston, S. (2010), "Measuring Financial Literacy", *Journal of Consumer Affairs*, Vol. 44/2, pp. 296-316, http://dx.doi.org/10.1111/j.1745-6606.2010.01170.x. [56]

Kaiser, T. and L. Menkhoff (2016), "Does Financial Education Impact Financial Literacy and Financial Behavior, and if So, When?"*, Discussion paper*, No. 1562, DIW Berlin, https://www.diw.de/documents/publikationen/73/diw_01.c.529454.de/dp1562.pdf (accessed on 5 March 2018). [16]

Lafontaine, D. and C. Monseur (2006), *Impact of test characteristics on gender equity indicators in the assessment of reading comprehension*, https://www.acer.org/files/lafontainemonseur.pdf (accessed on 5 March 2018). [54]

Lee, J. and J. Mortimer (2009), "Family Socialization, Economic Self-Efficacy, and the Attainment of Financial Independence in Early Adulthood.", *Longitudinal and life course studies*, Vol. 1/1, pp. 45-62, http://www.ncbi.nlm.nih.gov/pubmed/22025928 (accessed on 5 March 2018). [52]

Lusardi, A., O. Mitchell and V. Curto (2010), *Financial Literacy among the Young*, Wiley, http://dx.doi.org/10.2307/23859796. [39]

Lusardi, A. and P. Tufano (2015), "Debt literacy, financial experiences, and overindebtedness", *Journal of Pension Economics and Finance*, Vol. 14/04, pp. 332-368, http://dx.doi.org/10.1017/S1474747215000232. [29]

Mandell, L. and L. Klein (2009), "The Impact of Financial Literacy Education on Subsequent Financial Behavior", *Journal of Financial Counseling and Planning*, Vol. 20/1, pp. 15-24. [44]

Meier, S. and C. Sprenger (2013), "Discounting financial literacy: Time preferences and participation in financial education programs", *Journal of Economic Behavior & Organization*, Vol. 95, pp. 159-174, http://dx.doi.org/10.1016/J.JEBO.2012.02.024. [53]

Miller, M. et al. (2014), *Can you help someone become financially capable ? a meta-analysis of the literature*, http://documents.worldbank.org/curated/en/297931468327387954/Can-you-help-someone-become-financially-capable-a-meta-analysis-of-the-literature (accessed on 5 March 2018). [17]

Money Advice Service (2014), *It's time to talk: young people and money regrets*, Money Advice Service, https://mascdn.azureedge.net/cms/mas_money_regrets_online.pdf (accessed on 5 March 2018). [49]

Money and Pension Panel (2013), *Basic pension: The default option for labour-market pensions*, Report by the Committee of the Money and Pension Panel, https://www.raadtilpenge.dk/~/media/PPP/Pensionsrapport/Basic-pension_full-report.pdf?la=da (accessed on 5 March 2018). [10]

OECD (2017), *G20/OECD INFE Report on Adult Financial Literacy in G20 Countries*, OECD, http://www.oecd.org/daf/fin/financial-education/G20-OECD-INFE-report-adult-financial-literacy-in-G20-countries.pdf (accessed on 9 January 2018). [40]

OECD (2017), *G20/OECD INFE Report on Ensuring Financial Education and Consumer Protection for All in the Digital Age*, OECD, http://www.oecd.org/daf/fin/financial-education/G20-OECD-INFE-Report-Financial-Education-Consumer-Protection-Digital-Age.pdf (accessed on 9 January 2018). [12]

OECD (2017), *PISA 2015 Results (Volume V): Collaborative Problem Solving*, PISA, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264285521-en. [24]

OECD (2017), *PISA 2015 Technical Report*, OECD Publishing, http://www.oecd.org/pisa/data/2015-technical-report/ (accessed on 5 March 2018). [59]

OECD (2016), *Education at a Glance 2016 : OECD Indicators.*, OECD. [36]

OECD (2016), *G20/OECD INFE Core Competencies Framework on Financial Literacy for Adults*, OECD, http://www.oecd.org/daf/fin/financial-education/Core-Competencies-Framework-Adults.pdf (accessed on 9 January 2018). [6]

OECD (2016), *OECD/INFE International Survey of Adult Financial Literacy Competencies*, OECD, http://www.oecd.org/daf/fin/financial-education/OECD-INFE-International-Survey-of-Adult-FInancial-Literacy-Competencies.pdf (accessed on 9 January 2018). [41]

OECD (2016), "The role of financial education in supporting decision-making for retirement", in *OECD Pensions Outlook 2016*, OECD Publishing, Paris, http://dx.doi.org/10.1787/pens_outlook-2016-8-en. [11]

OECD (2015), *OECD/INFE Core Competencies Framework on Financial Literacy for Youth*, OECD, http://www.oecd.org/daf/fin/financial-education/Core-Competencies-Framework-Youth.pdf (accessed on 9 January 2018). [5]

OECD (2015), *OECD/INFE Policy Handbook on National Strategies for Financial Education*, OECD, http://www.oecd.org/daf/fin/financial-education/National-Strategies-Financial-Education-Policy-Handbook.pdf (accessed on 9 January 2018). [4]

OECD (2014), *Financial Education for Youth: the Role of Schools*, OECD Publishing, http://dx.doi.org/10.1787/9789264174825. [32]

OECD (2014), *PISA 2012 Results: Students and Money (Volume VI): Financial Literacy Skills for the 21st Century*, PISA, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264208094-en. [25]

OECD (2014), *PISA 2012 Technical report*, OECD Publishing, http://www.oecd.org/pisa/pisaproducts/pisa2012technicalreport.htm (accessed on 5 March 2018). [58]

OECD (2013), "Financial Literacy Framework", in *PISA 2012 Assessment and Analytical Framework: Mathematics, Reading, Science, Problem Solving and Financial Literacy*, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264190511-7-en. [34]

OECD (2010), *PISA 2009 Results: What Students Know and Can Do: Student Performance in Reading, Mathematics and Science (Volume I)*, PISA, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264091450-en. [57]

OECD (2009), *Financial Literacy and Consumer Protection: Overlooked Aspects of the Crisis*, OECD Publishing, http://www.oecd.org/finance/financial-markets/43138294.pdf (accessed on 5 March 2018). [1]

OECD (2009), *PISA 2009 assessment framework : key competencies in reading, mathematics and science.*, OECD, http://www.oecd.org/education/school/programmeforinternationalstudentassessmentpisa/pisa2009assessmentframework-keycompetenciesinreadingmathematicsandscience.htm (accessed on 5 March 2018). [42]

OECD (2006), "Improving Financial Literacy: Analysis of Issues and Policies", *Financial Market Trends*, Vol. 2005/2, http://dx.doi.org/10.1787/fmt-v2005-art11-en. [30]

OECD (2005), *Recommendation on Principles and Good Practices for Financial Education and Awareness*, http://www.oecd.org/finance/financial-education/35108560.pdf (accessed on 5 March 2018). [31]

OECD/INFE (2012), *OECD/INFE High-level Principles on National Strategies for Financial Education*, OECD, http://www.oecd.org/daf/fin/financial-education/OECD-INFE-Principles-National-Strategies-Financial-Education.pdf (accessed on 5 March 2018). [3]

Otto, A. (2013), "Saving in childhood and adolescence: Insights from developmental psychology", *Economics of Education Review*, Vol. 33, pp. 8-18, http://dx.doi.org/10.1016/J.ECONEDUREV.2012.09.005. [48]

Palameta, B. et al. (2016), *The Link Between Financial Confidence and Financial Outcomes Among Working-Aged Canadians*, The Social Research and Demonstration Corporation, Ottawa, http://www.srdc.org/media/199920/fcac-full-report-on-financial-confidence-en.pdf. [46]

Ratcliffe, C. and S. McKernan (2013), *Who Has Student Loan Debt?*, Urban Institute, https://www.urban.org/sites/default/files/publication/23736/412849-Forever-in-Your-Debt-Who-Has-Student-Loan-Debt-and-Who-s-Worried-.PDF?RSSFeed=UI_Education.xml (accessed on 5 March 2018). [38]

van Rooij, M., A. Lusardi and R. Alessie (2012), "Financial Literacy, Retirement Planning and Household Wealth", *The Economic Journal*, Vol. 122/560, pp. 449-478, http://dx.doi.org/10.1111/j.1468-0297.2012.02501.x. [27]

van Rooij, M., A. Lusardi and R. Alessie (2011), "Financial literacy and stock market participation", *Journal of Financial Economics*, Vol. 101/2, pp. 449-472, http://dx.doi.org/10.1016/J.JFINECO.2011.03.006. [21]

Whitebread, D. and S. Bingham (2013), *Habit Formation and Learning in Young Children Habit Formation and Learning in Young Children Section 2*, Money Advice Service, https://www.moneyadviceservice.org.uk/en/corporate/habit-formation-and-learning-in-young-children (accessed on 5 March 2018). [35]

Xu, L. and B. Zia (2012), *Financial literacy around the world : an overview of the evidence with practical suggestions for the way forward*, http://documents.worldbank.org/curated/en/264001468340889422/Financial-literacy-around-the-world-an-overview-of-the-evidence-with-practical-suggestions-for-the-way-forward (accessed on 5 March 2018). [23]

6. PISA 2018 Global Competence Framework

Twenty-first century students live in an interconnected, diverse and rapidly changing world. Emerging economic, digital, cultural, demographic and environmental forces are shaping young people's lives around the planet and increasing their intercultural encounters on a daily basis. This complex environment presents both an opportunity and a challenge. Young people today must not only learn to participate in a more interconnected world but also appreciate and benefit from cultural differences. Developing a global and intercultural outlook is a process – a lifelong process – that education can shape (Barrett et al., 2014[1]; Boix Mansilla and Jackson, 2011[2]; Deardorff, 2009[3]; UNESCO, 2013[4]; 2014[5]; 2016[6]). This section presents the framework for how the Programme for International Student Assessment (PISA) measures global competence, or students' ability to interact with the wider world around them.

Introduction: The importance of an international global competence assessment

What is global competence?

Global competence is a multidimensional, life-long learning goal. Globally competent individuals can examine local, global and intercultural issues, understand and appreciate different perspectives and worldviews, interact successfully and respectfully with others, and take responsible action toward sustainability and collective well-being.

Can schools promote global competence?

Schools play a crucial role in helping young people to develop global competence. They can provide opportunities for young people to critically examine global developments that are significant to both the world at large and to their own lives. They can teach students how to critically, effectively and responsibly use digital information and social media platforms. Schools can encourage intercultural sensitivity and respect by allowing students to engage in experiences that foster an appreciation for diverse peoples, languages and cultures (Bennett, 1993[7]; Sinicrope, Norris and Watanabe, 2007[8]). Schools are also uniquely positioned to enhance young people's ability to understand their place in the community and the world and improve their ability to make judgements and take action (Hanvey, 1975[9]).

Why do we need global competence?

To live harmoniously in multicultural communities

Education for global competence can promote cultural awareness and respectful interactions in increasingly diverse societies. Since the end of the Cold War, ethno-cultural conflicts have become the most common source of political violence in the world, and they show no sign of abating (Kymlicka, 1995[10]; Sen, 2007[11]; Brubaker and Laitin, 1998[12]). The many episodes of indiscriminate violence in the name of a religious or ethnic affiliation challenge the belief that people with diverse cultures are able to live peacefully in close proximity, accept differences, find common solutions and resolve disagreements. With the high influx of immigrants in numerous countries, communities have to redefine their identity and local culture. Contemporary societies call for complex forms of belonging and citizenship where individuals must interact with distant regions, people and ideas while also deepening their understanding of their local environment and the diversity within their own communities. By appreciating the differences in the communities to which they belong – the nation, the region, the city, the neighbourhood, the school – young people can learn to live together as global citizens (Delors, 1996[13]; UNESCO, 2014[14]). While education cannot bear the sole responsibility for ending racism and discrimination, it can teach young people the importance of challenging cultural biases and stereotypes.

To thrive in a changing labour market

Educating for global competence can boost employability. Effective communication and appropriate behaviour within diverse teams are keys to success in many jobs, and will remain so even more as technology continues to make it easier for people to connect across the globe. Employers increasingly seek to attract learners who easily adapt and are able to apply and transfer their skills and knowledge to new contexts. Work readiness in an interconnected world requires young people to understand the complex dynamics of globalisation, be open to people from different cultural backgrounds, build trust in diverse teams and demonstrate respect for others (British Council, 2013[15]).

To use media platforms effectively and responsibly

Over the past two decades, radical transformations in digital technologies have shaped young people's outlook on the world, their interactions with others and their perception of themselves. Online networks, social media and interactive technologies are giving rise to new types of learning, where young people exercise greater control over what and how they learn. At the same time, young people's digital lives can cause them to disconnect from themselves and the world, and ignore the impact that their actions may have on others. Moreover, while technology helps people to easily connect around the world, online behaviour suggests that young people tend to "flock together" (Zuckerman, 2013[16]) favouring interactions with a small set of people with whom they have much in common. Likewise, access to an unlimited amount of information is often paired with insufficient media literacy, meaning that young people are easily fooled by partisan, biased or fake news. In this context, cultivating students' global competence can help them to capitalise on digital spaces, better understand the world in which they live and responsibly express their voice online.

To support the Sustainable Development Goals

Finally, educating for global competence can help form new generations who care about global issues and engage in tackling social, political, economic and environmental challenges. The 2030 Agenda for Sustainable Development recognises the critical role of education in reaching sustainability goals. For example, in Target 4.7, it calls on all countries "to ensure, by 2030, that all learners acquire the knowledge and skills needed to promote sustainable development, including, among others, through education for sustainable development and sustainable lifestyles, human rights, gender equality, promotion of a culture of peace and non-violence, global citizenship and appreciation of cultural diversity and of culture's contribution to sustainable development" (UNESCO, 2016, p. 19[17]).

Should we assess global competence?

Every school should encourage its students to try and make sense of the most pressing issues defining our times. The high demands placed on schools to help their students cope and succeed in an increasingly interconnected environment can only be met if education systems define new learning objectives based on a solid framework, and use different types of assessment to reflect on the effectiveness of their initiatives and teaching practices. In this context, the PISA assessment of global competence aims to provide a comprehensive overview of education systems' efforts to create learning environments that invite young people to understand the world beyond their immediate environment, interact with others with respect for their rights and dignity, and take action towards building sustainable and thriving communities. A fundamental goal of this work is to support evidence-based decisions on how to improve curricula, teaching, assessments and schools' responses to cultural diversity in order to prepare young people to become global citizens.

How do we assess global competence?

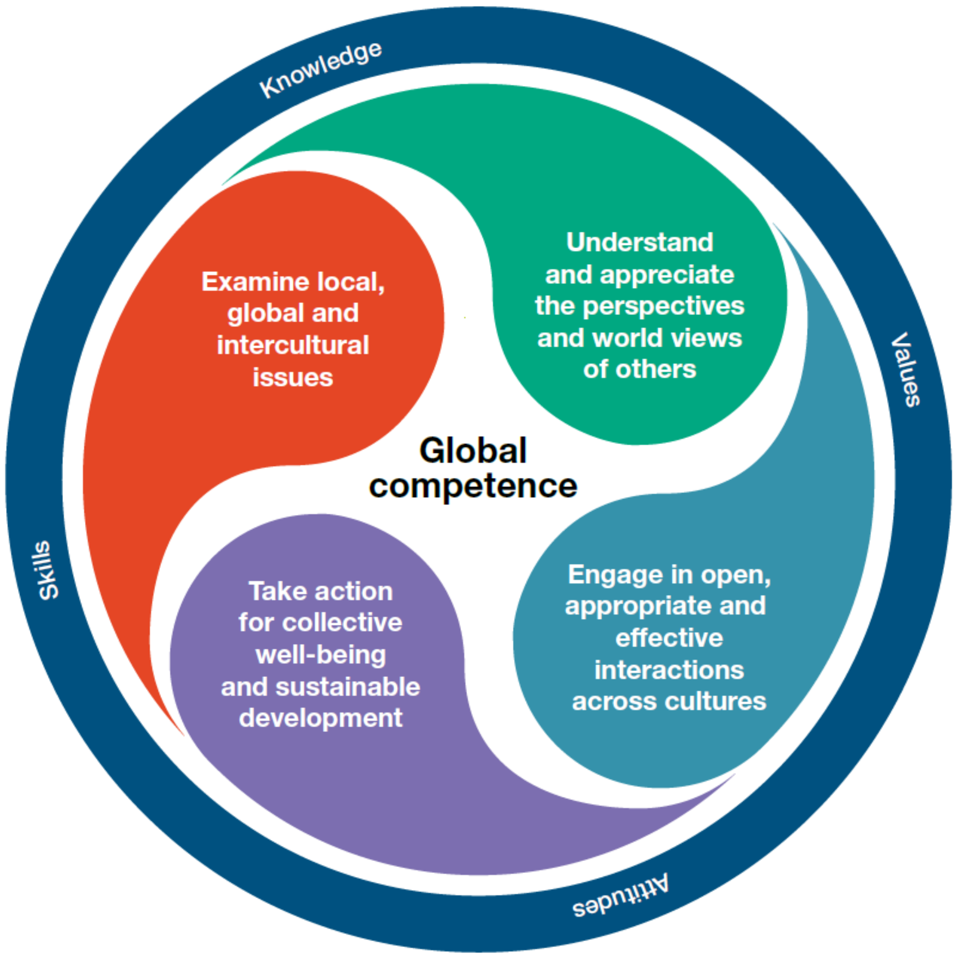
The global competence assessment in PISA 2018 is composed of two parts: a cognitive assessment and a background questionnaire. The cognitive assessment is designed to elicit students' capacities to critically examine global issues; recognise outside influences on perspectives and world views; understand how to communicate with others in intercultural contexts; and identify and compare different courses of action to address global and intercultural issues.

In the background questionnaire, students will be asked to report how familiar they are with global issues; how developed their linguistic and communication skills are; to what extent they hold certain attitudes, such as respect for people from different cultural backgrounds; and what opportunities they have at school to develop global competence. Answers to the school and teacher questionnaires will provide a comparative picture of how education systems are integrating global, international and intercultural perspectives throughout the curriculum and in classroom activities.

Taken together, the cognitive assessment and the background questionnaire address the following educational policy questions:

- To what degree are students able to critically examine contemporary issues of local, global and intercultural significance?
- To what degree are students able to understand and appreciate multiple cultural perspectives (including their own) and manage differences and conflicts?
- To what degree are students prepared to interact respectfully across cultural differences?
- To what degree do students care about the world and take action to make a positive difference in other peoples' lives and to safeguard the environment?
- What inequalities exist in access to education for global competence between and within countries?
- What approaches to multicultural, intercultural and global education are most commonly used in school systems around the world?
- How are teachers being prepared to develop students' global competence?

Figure 6.1. The dimensions of global competence



The building blocks of global competence – knowledge, skills, attitudes and values

The four dimensions of global competence[1] are supported by four inseparable factors: **knowledge, skills, attitudes and values**. For example, examining a global issue (dimension 1) requires *knowledge* of a particular issue, the *skills* to transform this awareness into a deeper understanding, and the *attitudes* and *values* to reflect on the issue from multiple cultural perspectives, keeping in mind the interest of all parties involved.

Effective education for global competence gives students the opportunity to mobilise and use their knowledge, skills, attitudes and values together while exchanging ideas on a global issue in or outside of school or while interacting with people from different cultural backgrounds (for example, engaging in a debate, questioning viewpoints, asking for explanations or identifying directions for deeper exploration and action).

A school community that wishes to nurture global competence should focus on clear and manageable learning goals. This means engaging all educators to reflect on teaching topics that are globally significant, the types of skills that foster a deeper understanding of the world and facilitate respectful interactions in multicultural contexts, and the attitudes and values that drive autonomous learning and inspire responsible action.

This section provides a general description of the content knowledge, skills, attitudes and values that individuals need in order to be globally competent. Policy makers, school leaders and teachers can refer to this section as they define strategies for teaching and assessing global competence. However, this description does not pretend to be conclusive or omni-comprehensive (other perspectives on global competence might put more emphasis on other important skills or attitudes, such as problem framing or emotional self-management). The definition and targeting of relevant skills and attitudes should also be adapted to the context in which the school operates.

Knowledge about the world and other cultures

Global competence is supported by knowledge of the global issues that affect lives locally and around the globe as well as by intercultural knowledge, that is, knowledge about the similarities, differences and relationships between cultures. This knowledge helps people to challenge misinformation and stereotypes about other countries and people, and thus counters intolerance and oversimplified representations of the world.

Global issues are those that affect all individuals, regardless of their nation or social group. They range from trade to poverty, human rights, geopolitics and the environment. Global issues reveal how different regions around the world are interconnected by shedding light on the diversity and commonality of their experiences (Boix Mansilla and Jackson, 2011[2]). For example, pollution in one place affects the ozone layer somewhere else; floods in agricultural areas not only ruin the local environment and economy, but also affect markets worldwide and drive waves of migration. Global issues are also local issues: they are global in their reach but local communities experience them in very diverse ways.

As global issues emerge when ecological and socio-economic interests cross borders, **intercultural issues** (situations) arise from the interaction of people with different cultural backgrounds. In this interaction, each party's way of thinking, believing, feeling and acting are interpreted by the other. This process can be smooth if there are not extreme differences between cultures and individuals are open to learning about and accepting those differences. But intercultural interactions can also face miscommunication and misunderstanding. In the worst cases, these misunderstandings degenerate into negative stereotypes, discrimination and violent conflict.

More than in other domains of knowledge, global competence requires engaging with controversial issues. Schools can provide a safe space in which students can explore complex and controversial global issues that they encounter through the media and their own experiences.

The list of relevant global or intercultural issues that can be introduced to children and young people in school is a long one. There have been recent attempts to systematise these complex sets of issues into a coherent sequence of lessons and learning materials at all curriculum levels (IBO, 2012[18]; Oxfam, 2015[19]; Reimers, 2017[20]). A curriculum should pay attention to the following four knowledge domains**: culture and intercultural relations**; **socio-economic development and interdependence**; **environmental sustainability**; and **global institutions, conflicts and human rights**. Teaching these four domains should highlight differences in opinions and perspectives, questioning concepts such as "truth" and "information". For example, while examining inequalities in economic development across the world, the teacher can explain that there are different interpretations of what development means and implies, inciting students to measure development according to different metrics.

The first key domain of knowledge for global competence relates to the manifold expressions **of culture and intercultural relations**, such as languages, arts, knowledge, traditions and norms. Acquiring knowledge in this domain can help young people become more aware of their own cultural identity, help them understand differences and similarities among and within cultures, and encourage them to value the importance of protecting cultural differences and diversity. As they engage in learning about other cultures and individual differences, students start to recognise multiple, complex identities and avoid categorising people through single markers of identity (e.g. black, white, man, woman, poor, rich). Students can acquire knowledge in this domain by reflecting on their own cultural identity and that of their peers, by analysing common stereotypes towards people in their community, or by studying illustrative cases of conflict or successful integration between cultural groups.

The domain of **socio-economic development and interdependence** refers to the study of development patterns in different regions of the world, with a focus on the links and interdependences between societies and economies. Students can analyse, at different levels of complexity and in developmentally appropriate ways, the many forms of globalisation, such as international migration, transnational production, global brands and technologies. By doing so, students can start to make sense of how local, national and global processes jointly shape the development patterns of countries, and the inequalities in opportunities available to individuals.

Students need a solid foundation in environmental issues in order to promote and support sustainability. Learning activities in the domain of **environmental sustainability** help students understand the complex systems and policies surrounding the demand for and use of natural resources.

The fourth knowledge domain of global competence focuses on formal and informal **institutions** that support peaceful relationships between people and the respect of fundamental human rights. Students can learn how global institutions such as the United Nations were established, can reflect on the contested nature of global governance in a world with highly unbalanced power relationships, review causes of and solutions for current and historical conflicts between countries, ethnic or social groups, and examine spaces and opportunities for young people to play an active role in society, take responsibility and exercise their rights. Acquiring deep knowledge in this domain is instrumental for young people to develop values such as peace, non-discrimination, equality, justice, non-violence, tolerance and respect.

Box 6.1. Integrating global and intercultural issues in the curriculum

Research on global education tends to focus on social studies and foreign language classes, often in the upper grade levels (Gaudelli, 2006[21]; Karaman and Tochon, 2007[22]; Merryfield, 2008[23]; Myers, 2006[24]; Rapoport, 2010[25]; Suarez, 2003[26]). However the local, global and intercultural issues that students should learn about, in order to take responsibility for and act upon them, cut across education levels and academic disciplines (Gaudelli, 2003[27]; O'Connor and Zeichner, 2011[28]). For global education to move from abstraction to action, many advocates recommend integrating global issues and topics into existing subjects (Klein, 2013[29]; UNESCO, 2014[5]). In practice, several countries are pursuing a dual approach, where content knowledge related to global competence is both integrated into the existing curriculum and also taught in specific subjects or courses (e.g. human rights education). Students can come to understand local, global and intercultural issues across ages, beginning in early childhood when such issues are presented in developmentally appropriate ways (Boix Mansilla and Jackson, 2011[2]; UNESCO, 2015[30]).

The way that a teacher frames a topic in the curriculum can significantly shape its contribution to global competence. When framing a topic to explore with students, teachers may consider the ways in which this topic addresses local and global dynamics, and how it can enable students to understand broad global patterns and the impact on their local environment. For instance, a mathematics teacher might invite students to decide whether linear or exponential functions best fit the data on world population growth, or a music teacher may explore how today's hip hop is expressed differently around the world.

In order to avoid the risk that global education becomes a catch-all curriculum where everything fits, teachers must have clear ideas about the global and intercultural issues that they want students to reflect upon. Teachers need to collaboratively research topics and carefully plan the curriculum, giving students multiple opportunities to learn about a core set of issues that increase in complexity throughout their education (Gaudelli, 2006[21]). Professional learning communities can be highly effective to engage all teachers and to facilitate collaboration and peer learning. For example, Lee et al. (2017[31]) document how highly motivated teachers in Thailand followed a training course on global competence promoted by the Ministry of Education, and then created professional learning communities in their school to engage other teachers, helping them integrate global and intercultural topics in their courses and promoting school-wide projects.

Teaching about minority cultures in different subject areas requires accurate content about and comprehensive portrayals of ethnically and racially diverse groups and experiences. Curricula should promote the integration of knowledge of other people, places and perspectives into everyday activities in the classroom throughout the year (UNESCO, 2014[5]), rather than using a "tourist approach", giving students a superficial glimpse of life in different countries every now and then.

Textbooks and other instructional materials can also distort cultural and ethnic differences (Gay, 2013[32]). Teachers and their students should thus critically analyse their textbook and teaching resources, and compensate for inadequacies when necessary.

Connecting global and intercultural topics to the reality, contexts and needs of the learning group is an effective methodological approach to make them relevant to adolescents (North-South Centre of the Council of Europe, 2012[33]). People learn better and become more engaged when the content relates to them, and when they can see the parallels between many global issues and their immediate environment. For example, students can become aware of the risks related to climate change by studying the effects that natural phenomena (hurricanes, floods) have on their own community. Capitalising on local expertise and the experience of young people in culturally responsive ways is particularly relevant when teaching less privileged or immigrant youth (Suárez-Orozco, Suárez-Orozco and Todorova, 2008[34]).

Skills to understand the world and to take action

Global competence also builds on specific cognitive, communication and socio-emotional "skills". Skills are defined as the capacity to carry out a complex and well-organised pattern of thinking (in the case of a cognitive skill) or behaviour (in the case of a behavioural skill) in order to achieve a particular goal. Global competence requires numerous skills, including **reasoning with information**, **communication skills in intercultural contexts**, **perspective taking**, **conflict resolution skills** and **adaptability**.

Globally competent students are able to **reason with information** from different sources, e.g. textbooks, peers, influential adults, traditional and digital media. They can autonomously identify their information needs, and select sources purposefully on the basis of their relevance and reliability. They use a logical, systematic and sequential approach to examine information in a piece of text or any other form of media, examining connections and discrepancies. They can evaluate the worth, validity and reliability of any material on the basis of its internal consistency, and its consistency with evidence and with one’s own knowledge and experience. Competent students question and reflect on the source author's motives; purposes and points of view; the techniques used to attract attention; the use of image, sound and language to convey meaning; and the range of different interpretations which are likely for different individuals.

Competent students are able to **communicate effectively and respectfully** with people who are perceived to have different cultural backgrounds. Effective communication requires being able to express oneself clearly, confidently and without anger, even when expressing a fundamental disagreement. Respectful communication requires understanding the expectations and perspectives of diverse audiences, and applying that understanding to meet the audience’s needs. Respectful communicators also check and clarify the meanings of words and phrases when they engage in an intercultural dialogue. Speaking more than one language is a clear asset for effective intercultural communication. Smooth communication in intercultural contexts is also facilitated by active listening, which means looking for not only what is being said but also how it is being said, through the use of voice and accompanying body language. Competent students are capable speakers who can use their body language and voice effectively when they discuss and debate global issues, express and justify a personal opinion or persuade others to pursue a particular course of action.

Perspective taking refers to the cognitive and social skills individuals need in order to understand how other people think and feel. It is the capacity to identify and take on often conflicting points of view or “stepping into someone else’s shoes”. Perspective taking does not only involve imagining another person’s point of view but also entails understanding how various perspectives are related to one another. Understanding others' perspectives facilitates more mature and tolerant interpretations of differences among groups.

Competent students approach conflicts in a constructive manner, recognising that conflict is a process to be managed rather than seeking to negate it. Taking an active part in **conflict management and resolution** requires listening and seeking common solutions. Possible ways to address conflict include: analysing key issues, needs and interests (e.g. power, recognition of merit, division of work, equity); identifying the origins of the conflict and the perspectives of those involved in the conflict; recognising that the parties might differ in status or power; identifying areas of agreement and disagreement; reframing the conflict; managing and regulating emotions, interpreting changes in one’s own and others' underlying emotions and motivation; dealing with stress, anxiety and insecurity, both in oneself and in others; prioritising needs and goals; and deciding on possible compromises and the circumstances under which to reach them (Rychen and Salganik, 2003[35]) (However, approaches to managing and resolving conflict may vary by societal expectations, so not all adhere to the steps outlined here).

Adaptability refers to the ability to adapt one's thinking and behaviours to the prevailing cultural environment, or to novel situations and contexts that might present new demands or challenges. Individuals who acquire this skill are able to handle the feelings of "culture shock", such as frustration, stress and alienation in ambiguous situations caused by new environments. Adaptable learners can more easily develop long-term interpersonal relationships with people from other cultures, and remain resilient in changing circumstances.

Box 6.2. Pedagogies for promoting global competence

Various student-centred pedagogies can help students develop critical thinking with regards to global issues, respectful communication, conflict management skills, perspective taking and adaptability.

Group-based co-operative project work can improve reasoning and collaborative skills. It involves topic- or theme-based tasks suitable for various levels and ages, in which goals and content are negotiated by all participants, and learners can create their own learning materials that they present and evaluate together. In order to co-operate effectively, learners need to feel safe and comfortable, and the task and its goals must be clearly set for them. Learners participating in co-operative tasks soon realise that in order to be efficient, they need to be respectful, attentive, honest and empathetic (Barrett et al., 2014[1]). Project work can effectively connect students within and across borders. For example, Global Cities has created a digital exchange program (Global Scholar) through which students in 26 countries are given the opportunity to work in classrooms across the world (Global Cities, 2017[36]).

Students can voice their differences, biases and culturally determined beliefs through organised discussions in the classroom. In order to stimulate discussion, a teacher typically uses a thought-provoking video clip, image or text (Costa and Kallick, 2013[37]). Students can then present supporting evidence, comment and express their differing points of view. Class discussion is, by nature, an interactive endeavour, and reflective dialogue engenders proactive listening and responding to ideas expressed by one's peers. By exchanging views in the classroom, students learn that there is not always a single right answer to a problem to be memorised and presented; they learn to understand the reasons why others hold different views and are able to reflect on the origins of their own beliefs.

Structured debates constitute a specific format of class discussion that is increasingly used in secondary and higher education as a way to raise students' awareness about global and intercultural issues, and to let them practice their communication and argumentation skills[2]. In this format, students are given instructions to join a team either supporting or opposing a polemic point of view, such as "the Internet should be censored" or "hosting the Olympics is a good investment". It is often helpful for students to articulate views that may be different from their own.

Service learning is another tool that can help students develop multiple global skills through real-world experience. This requires learners to participate in organised activities that are based on what has been learnt in the classroom and that benefit their communities. After the activities, learners are required to reflect critically on their service experience to gain further understanding of course content and enhance their sense of their role in society with regard to civic, social, economic and political issues (Bringle et al., 2016[38]). Service learning is strongly tied to the curriculum and differs from other types of educational experiences in the community and from volunteering. Through service learning, students not only "serve to learn," which is applied learning, but also "learn to serve" (Bringle et al., 2016[38]).

The Story Circle approach has been used in numerous classrooms around the world to let students practice key intercultural skills, including respect, cultural self-awareness and empathy (Deardorff, 2019, forthcoming[39]). The students, in groups of 5 or 6, take turns sharing a 3-minute story from their own experience based on specific prompts such as "Tell us about your first experience when you encountered someone who was different from you". After all students in the group have shared their personal stories, students then take turns briefly sharing the most memorable point from each story in a "flashback" activity. Other types of intercultural engagement involve simulations, interviews, role play and online games.[3]

Attitudes of openness, respect for people from different cultural backgrounds and global mindedness

Global competence embodies and is propelled by key dispositions or attitudes. Attitudes refer to the mind-set that an individual adopts towards a person, a group, an institution, an issue, a behaviour or a symbol. This mind-set integrates beliefs, evaluations, feelings and tendencies to behave in a particular way. Globally competent behaviour requires an attitude of **openness** towards people from other cultural backgrounds, an attitude of **respect** for cultural differences and an attitude of **global mindedness** (i.e. that one is a citizen of the world with commitments and obligations toward the planet and others, irrespective of their particular cultural or national background). Such attitudes can be fostered explicitly, through participatory and learner-centred teaching, as well as implicitly through a curriculum characterised by fair practices and a welcoming school climate for all students.

Openness toward people from other cultural backgrounds involves sensitivity toward, curiosity about and a willingness to engage with other people and other perspectives on the world (Byram, 2008[40]; Council of Europe, 2016[41]). It requires an active willingness to seek out and embrace opportunities to engage with people from other cultural backgrounds, to discover and learn about their cultural perspectives and how they interpret familiar and unfamiliar phenomena, and to learn about their linguistic and behavioural conventions. Another important characteristic of open learners is their willingness to suspend their own cultural values, beliefs and behaviours when interacting with others, and not to assume that their own values, beliefs and behaviours are the only possible correct ones. The attitude of openness towards cultural otherness needs to be distinguished from only being interested in collecting "exotic" experiences merely for one's own personal enjoyment or benefit. Rather, intercultural openness is demonstrated through a willingness to engage, cooperate and interact with those who are perceived to have cultural affiliations that differ from one's own, on an equal footing.

Respect consists of positive regard and esteem for someone or something based on the judgement that they have intrinsic worth. In this framework, respect assumes the dignity of all human beings and their inalienable right to choose their own affiliations, beliefs, opinions or practices. Being respectful of cultural differences does not require minimising or ignoring significant and profound differences that might exist between oneself and others, nor does it require agreeing with, adopting or converting to others' beliefs. Respect for others also has certain limits that are set by the principle of human dignity. For example, respect should not be accorded to beliefs and opinions or to lifestyles and practices which undermine or violate the dignity of others (Council of Europe, 2016[41]).

The concept of respect should be distinguished from the concept of tolerance. Tolerance may, in some contexts, simply mean enduring difference. Respect is a less ambiguous and more positive concept. It is based on recognition of the dignity, rights and freedoms of the other in a relationship of equality.

Global mindedness is defined by Hett, as cited in Hansen (2010[42]), as “a worldview in which one sees oneself as connected to the world community and feels a sense of responsibility for its members”. A globally-minded person has concerns for other people in other parts of the world, as well as feelings of moral responsibility to try to improve others' conditions irrespective of distance and cultural differences. Globally-minded people care about future generations, and so act to preserve the environmental integrity of the planet. Globally-minded individuals exercise agency and voice with a critical awareness of the fact that other people might have a different vision of what humanity needs, and are open to reflecting on and changing their vision as they learn about these different perspectives. Rather than believing that all differences can be eliminated, globally-minded people strive to create space for different ways of living with dignity.

Valuing human dignity and diversity

Values go beyond attitudes: they transcend specific objects or situations. They are more general beliefs about the desirable goals that individuals strive for in life, reflecting modes of conduct or states of being that an individual finds preferable to all other alternatives. In this way, values serve as standards and criteria that people use both consciously and unconsciously in their judgements. They have a normative, prescriptive quality about what ought to be done or thought in different situations. Values therefore motivate certain behaviours and attitudes. For example, people for whom independence is an important value are triggered if their independence is threatened, feel despair when they are helpless to protect it, and are happy when they can enjoy it (Schwartz, 2012[43]).

Valuing human dignity and **valuing cultural diversity** contribute to global competence because they constitute critical filters through which individuals process information about other cultures and decide how to engage with others and the world. Individuals who cultivate these values become more aware of themselves and their surroundings, and are strongly motivated to fight against exclusion, ignorance, violence, oppression and war.

Education has a deep influence on the values of individuals. During their time at school, young citizens form habits of mind, beliefs and principles that will stay with them throughout their lives. This is why it is crucial to reflect on the type of education that best "cultivates humanity" (Nussbaum, 1997[44]). An education that encourages valuing dignity, human rights and diversity emphasises shared commonalities that unite people around the world, rather than the issues that divide them; provides learning experiences so that students see the world from many different perspectives, enabling them to examine their own thoughts and beliefs, and their society's norms and traditions; encourages people to understand the significance of another person's sufferings; and emphasises the importance of reasoning, careful argument, logical analysis, self-questioning, the pursuit of truth and objectivity.

While most people would agree that education should help students develop into human beings who care for and respect others (Delors, 1996[13]), deciding which values education systems around the world should promote is subject to debate. It is not easy to identify a core set of rights that are universally valid and interpreted in the same way everywhere and in every circumstance, as morals and social institutions vary across cultures and historical contexts (Donnelly, 2007[45]).

Article 1 of the Universal Declaration of Human Rights describes the constitutive elements of a minimum core of rights that can guide education around the world: *"All human beings are born free and equal in dignity and rights. They are endowed with reason and conscience and should act towards one another in a spirit of brotherhood"*. The article defines two basic foundations of human dignity: the first is that every human being possesses an intrinsic worth, merely by being human; the second is that this intrinsic worth should be recognised and respected by others, and certain forms of treatment by others are inconsistent with respect for this intrinsic worth. Individuals have a distinct moral obligation to treat each other in ways that are constrained by certain inviolable limits. Embracing this value often means helping others to protect what is most important to them in life.

The concept of respecting the fundamental right of human dignity is often associated with protection from discrimination. Clapham (2006[46]) has suggested that valuing the equality of core rights and dignity is comprised of four aspects: (1) the prohibition of all types of inhuman treatment, humiliation or degradation by one person over another; (2) the assurance of the possibility for individual choice and the conditions for each individual's self-fulfilment, autonomy or self-realisation; (3) the recognition that the protection of group identity and culture may be essential for the protection of personal dignity; and (4) the creation of the necessary conditions for each individual to have their essential needs satisfied. Nussbaum (1997[44]) has argued that a minimally just society must endeavour to nurture and support a core set of basic "capabilities", defined as opportunities for choice and action (e.g. being secure against violent assault; being able to imagine, to think and to reason; being able to love, to grieve, to experience longing, gratitude and justified anger, etc.). People from many different traditions, with many different conceptions of "good", can agree on these core capabilities as the necessary basis for pursuing a good life.

A controversial issue relates to the Western roots of the concept of human dignity and to the Western dominance in the discussion and definitions of human rights. However, deep reflections on human dignity can be found in several different countries and cultures. For example, the indigenous African concept of Ubuntu has a strong connection with the conceptualisation of human dignity in Western philosophy. Ubuntu generally translates as humaneness, and its spirit emphasises respect for human dignity, marking a shift from confrontation to conciliation (Mokgoro, 1995[47]). Box 6.3 discusses this further.

Box 6.3. Perspectives on global competence from different cultures

The literature, theories and frameworks on intercultural competence, global competence and global citizenship emerge predominantly from a Western, Euro-American context. However, related concepts exist in many countries and cultures around the world. One interesting perspective on global competence comes from South Africa and involves the concept of Ubuntu. There is much literature written about Ubuntu (Nwosu, 2009[48]; Khoza, 2011[49]; Khoza, 2011[49]), the core idea of which can be summarised by the Zulu proverb "Umuntu Ngumuntu Ngabantu", meaning that a person is a person because of others. Ubuntu can be used to illustrate a collective identity, as well as connectedness, compassion, empathy and humility. There are concepts similar to Ubuntu found in different cultures around the world, including in indigenous cultures in the Andes and in Malaysia. Collective identity, relationships and context (as affected by historical, social, economic and political realities) are all major emphases in other cultural discourses on global competence. In summarising some central themes across different cultures with regard to global competence, Deardorff (2011[50]) noted the following key elements: respect, listening, adaptation, relationship-building, seeing from multiple perspectives, self-awareness and cultural humility.

Even if the cultural context varies, the common core values of respecting human dignity is sufficiently robust to challenge the legitimacy of a wide array of systems that abuse their power against individuals and groups.[4] Abuses of power against vulnerable individuals are not unique to war-torn regions or fragile states. They can happen anywhere: neighbourhoods, offices or schools. Schools, in particular, are places where human dignity takes on a concrete meaning, because every student deserves equal justice, equal opportunity and equal dignity. Discrimination at school can be overtly displayed through xenophobic comments, bullying, name-calling, segregation and physical altercations. Discrimination can also be less apparent but still present in stereotypes, fear of others and unconscious reactions to or intentional avoidance of certain groups. Teaching youth to use human rights as a frame of reference for their behaviour can allow them to break down stereotypes, biases and discrimination, thereby improving the school environment and social relationships in the communities that schools serve.

Respecting human beings' core rights and dignity is, in most cases, compatible with respecting and valuing cultural diversity. Globally competent learners should not only have a positive attitude towards cultural diversity (the attitudes of "openness" and "respect" defined above), but should also value cultural diversity as an asset for societies and a desirable goal for the future. However, valuing cultural diversity has certain limits that are determined by the inviolability of human dignity (UNESCO, 2001[51]). The possible tension between valuing cultural diversity and valuing human rights can be solved by establishing a normative hierarchy between the two: valuing core human rights is more important than valuing cultural diversity, in cases where the two values are in conflict with each other.

Promoting the value of cultural diversity in education practice involves encouraging students to take actions to safeguard both tangible and intangible cultural heritage around the world, as well as actions to promote the rights of all people to embrace their own perspectives, views, beliefs and opinions (UNESCO, 2009[52]). It also means conveying the message to all students that their cultural heritage is important and enriches society.

Evaluating how much students care about and cherish the values of human dignity and cultural diversity is complex and therefore beyond the scope of this PISA assessment of global competence. However, the inclusion of values in this framework hopes to stimulate a productive debate on how education can shape children's development of an ethical decision-making framework grounded on human rights, while fully preserving the value of diverse opinions and beliefs. Acknowledging the importance of values in education does not mean promoting a uniform and fixed way to interpret the world; it rather implies giving students some essential references to navigate a world where not everyone holds their views, but everyone has a duty to uphold the principles that allow different people to co-exist and to prosper.

Box 6.4. Teaching attitudes and values related to global competence

Allocating teaching time to a specific subject dealing with human rights issues and non-discrimination is an important first step in cultivating values for global competence. But even more can be achieved by mainstreaming the principle of respect for human dignity and for cultural diversity across all subjects. For example, teachers can use multi-ethnic and multicultural examples to illuminate general principles and concepts, or emphasise the contributions of people from different ethnic groups to our collective knowledge and quality of life. Teachers thus need to develop repertoires of culturally diverse examples, and the skills and confidence to use them fluidly and routinely in classroom instruction.

Values and attitudes are partly communicated through the formal curriculum but also through the ways in which educators and students interact, how discipline is encouraged and the types of opinions and behaviour that are validated in the classroom. For example, a history lesson on the American Civil War may emphasise valuing racial equality; however, if the teacher disciplines minority students more severely, he or she communicates a contradictory value system. It is likely that students will assimilate the culture of the classroom more readily than they will learn the curriculum. Therefore, recognising the school and classroom environments' influence on developing students' values can help educators become more aware of the effects that their teaching has on students. For example, a teacher might reconsider the seating plan of the classroom if he is hoping to promote racial and gender integration among his students.

Teachers can be instrumental in replacing stereotypes of minority and disadvantaged students with more positive ones. However, teachers often find it difficult to engage in open discussions about diversity and discrimination. Part of the problem is a lack of experience with people who are different, and the assumption that conversations about discrimination and ethics will always be contentious. Consequently, teachers may concentrate only on "safe" topics about cultural diversity, such as cross-group similarities, ethnic customs, cuisines, costumes and celebrations, while neglecting more troubling issues such as inequities, injustices and oppression (Gay, 2013[32]).

These difficulties can be overcome by giving educators access to continual professional development. Specific training programmes and modules can help teachers to acquire a critical awareness of the role that different subject and teaching approaches can play in the struggle against racism and discrimination; the skills to acknowledge and take into account the diversity of learners' needs, especially those of minority groups; and a command of basic methods and techniques of observation, listening and intercultural communication (UNESCO, 2006[53]).

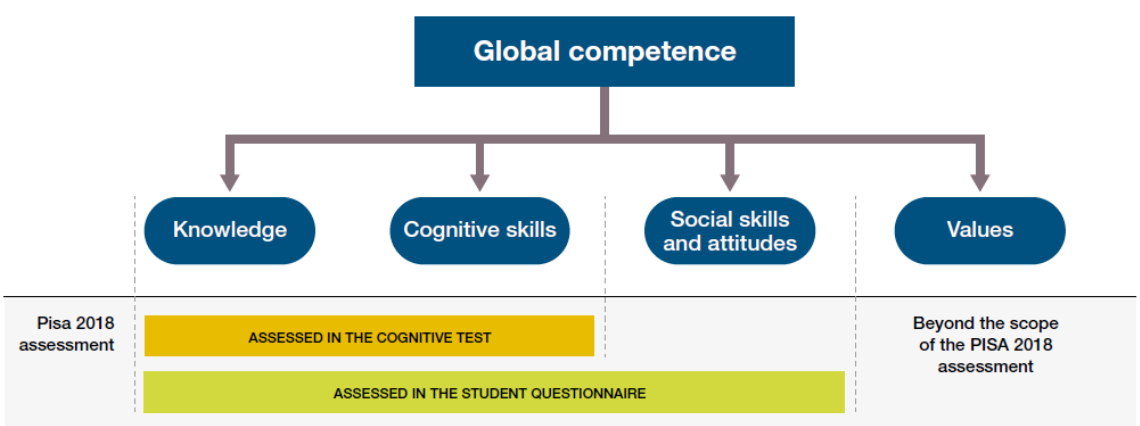
The assessment of global competence in PISA

The assessment strategy

Assessing global competence in all of its complexity requires a multi-method, multi-perspective approach. The PISA 2018 assessment of global competence is a development in this direction, although clear challenges and limitations remain. The most salient challenge for the PISA assessment is that – through a single international instrument – it needs to account for the large variety of geographic and cultural contexts represented in participating countries. Students who perform well on a question assessing their reasoning about a global issue are likely to have some prior knowledge of the issue, and the type of knowledge students already have of global issues is influenced by their experiences within their unique social context. On the one hand, cultural variability in the tested population requires that the test material cannot be too biased towards a particular perspective, for example, the perspective of a student in a rich country who thinks about a problem in a poor country. Similarly, the test units should focus on issues that are relevant for 15-year-old students in all countries. On the other hand, leaning too much towards "cultural neutrality" in the design of scenarios and questions reduces the authenticity and relevance of the tasks. The test design is further limited by the time constraints of the assessment and the narrow availability of internationally-valid instruments that measure the behavioural elements of global competence.

Accounting for these limitations and challenges, the PISA 2018 global competence assessment has two components: 1) **a cognitive test** exclusively focused on the construct of *"global understanding"*, defined as the combination of background knowledge and cognitive skills required to solve problems related to global and intercultural issues; and 2) **a set of questionnaire items** collecting self-reported information on students' awareness of global issues and cultures, skills (both cognitive and social) and attitudes, and information from schools and teachers on activities to promote global competence.

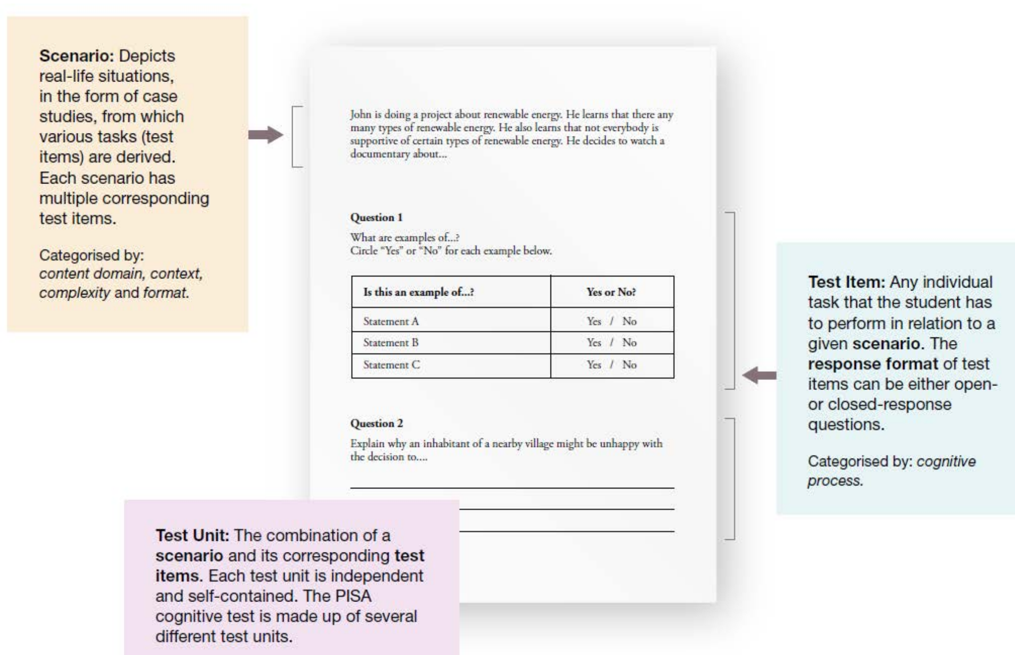
Figure 6.2. The PISA approach to assessing global competence



The reporting of the results will reflect the differences between these two assessment components. Students' answers to the questions in the cognitive test can be objectively scored as right (or partially right) or wrong, and can thus be presented on a scale. Given that the capacity to understand global or intercultural issues and situations can be developed at school, the PISA proficiency scale is expected to yield results that can be interpreted in educational policy terms. For some of the questions measuring attitudinal or socio-emotional traits (e.g. "openness"), however, defining right or wrong answers is more controversial because the development of these traits and their contribution towards global competence might be non-linear (beyond a certain threshold, more "openness" may not necessarily be better). Measurement issues are also more acute in self-reported items, so ranking students or countries on the basis of students' responses to the questionnaire risks errors of misrepresentation and misinterpretation. For example, people from some cultural backgrounds tend to exaggerate their responses to typical questionnaire items based on a Likert-type scale (e.g. questions asking students whether they strongly disagree, disagree, agree or strongly agree with a statement), whereas others tend to take a middle ground (Harzing, 2006[54]). The responses to the questionnaire items will thus *not* be used to position countries and students on a scale. Instead, they will only be used to illustrate general patterns and differences within countries in the development of the skills and attitudes that contribute to global competence among 15-year-old students, as well as to analyse the relationship between those skills and attitudes and students' results on the cognitive test.

Global understanding is assessed in the PISA cognitive test by asking students to complete several test units. Each test unit is composed of one scenario (or case study) and various scenario-based tasks (see Figure 6.3). In a typical test unit, students read about a case and respond to questions (otherwise referred to as test items) that evaluate their capacity to understand its complexity and the multiple perspectives of the diverse actors involved. Each scenario will expose students to a range of different situations and test their capacity to apply their background knowledge and cognitive skills in order to analyse the situation and suggest solutions.

Figure 6.3. Elements of a typical PISA 2018 global competence test unit



The cognitive skills demanded by global understanding are relevant measures of all four dimensions of students' global competence. Test items asking students to critically analyse statements and information will provide relevant information about students' capacity to "examine global and intercultural issues" (dimension 1). "Understanding perspectives" (dimension 2) can be assessed through test items examining students' capacity to recognise different perspectives while being aware of one's own cultural lens and biases, as well as those of other people; consider the contexts (cultural, religious, regional) that influence these perspectives; and find possible connections or "common ground" across perspectives. Elsewhere, "engage in appropriate and effective interactions" (dimension 3) can be assessed through items testing students' capacity to understand communicative contexts and the norms of respectful dialogue. Finally, "take action for sustainability and well-being" (dimension 4) can be assessed by students' capacity to consider possible actions to global problems and weigh direct and indirect consequences.

The student questionnaire will provide complementary information on the attitudes, knowledge and skills that people need to navigate everyday life in globally and culturally competent ways, but whose measurement goes beyond the parameters of the PISA cognitive test. Self-reported skills and attitudes will be measured through Likert-type scales that have been selected on the basis of a review of empirical studies.

The cognitive test on global understanding

A short review of cognitive assessments in this area

Research in this area has predominantly been based on student self-reports, and only a few examples of cognitive assessments exist. In the Global Understanding Survey (Barrows et al., 1981[55]), the authors define global understanding as a sum of four components: (a) knowledge; (b) attitudes and perceptions; (c) general background correlations and (d) language proficiency. The knowledge domain in the Global Understanding Survey consisted of 101 multiple-choice questions that addressed international institutions, major historical events and trends, and legal and policy frameworks associated with 13 global themes.

Test items in the Global Understanding Survey addressed real-world issues. Students who reported regular news consumption scored higher on the test. However, the authors found only weak relationships between students' educational experiences – coursework, language study or study abroad – and their levels of international knowledge. The final report also recognised that the assessment provided only limited insights on the nature and development of global understanding.

The IEA Studies on Civic Education (Amadeo et al., 2002[56]) and the International Civic and Citizenship Study (Schulz et al., 2010[57]; Schulz et al., 2018[58]) are other relevant examples that could guide item development in PISA. The key research questions in the ICCS concern student achievement in civic and citizenship education and their disposition to engage with such issues. The ICCS measures the cognitive processes of knowledge, reasoning and analysis across four content domains: (a) civic society and systems, (b) civic principles, (c) civic participation, and (d) civic identities (Schulz et al., 2010[57]; Torney-Purta et al., 2015[59]). The item format combines multiple-choice and open-ended questions.

Some of the items in the ICCS measure students' ability to analyse and reason. Reasoning asks students to apply their knowledge and understanding of familiar concrete situations in order to reach conclusions about complex, multifaceted, unfamiliar and abstract situations (Schulz et al., 2010[57]).

Outside the context of global and civic education, an increasing number of assessments have attempted to measure students' capacity to evaluate information and think critically about problems.[5] In many of these tests, students read a short text and decide whether a series of statements related to the text are likely to be true or false. Some of these tests also include constructed response questions, where students need to develop logical arguments or explain how someone else's conclusions could be verified or strengthened. All these assessments emphasise reasoning, analysis, argumentation and evaluation (Liu, Frankel and Roohr, 2014[60]). These tests treat those skills as generic, while PISA will look at the application of these capacities in the specific context of global and intercultural issues.

The Global Integrated Scenario-Based Assessment of Reading (GISA), is another relevant reference for the PISA test (O'Reilly and Sabatini, 2013[61]; Sabatini et al., 2014[62]; Sabatini et al., 2015[63]). GISA assesses students' "global reading literacy ability", a multidimensional competence that requires students to not only use and process texts but also to employ other cognitive, language and social reasoning skills, as well as call upon their own knowledge, strategies and dispositions. Unlike traditional reading assessments that present students with a set of unrelated texts and no purpose for reading them, GISA uses a scenario-based approach with a carefully structured sequence of tasks. By employing scenarios that provide authentic contexts and purposes for reading, the assessment better reflects the cognitive processes that students engage in when confronted with real learning activities.

The GISA assessments also include collaborative activities. For example, test takers "interact" with simulated peers to identify errors, correct misconceptions and provide feedback. The members of the simulated interactions can state facts, present incorrect information, give their opinions and go off topic, just as people do in real life. Performance moderators such as background knowledge, self-regulatory strategies and motivation are also measured in GISA and are used to interpret the reading score.

Relatively few assessments of perspective-taking skills exist. One relevant example for the PISA test is the perspective-taking measure developed within the Catalyzing Comprehension through Discussion and Debate (CCDD) initiative.[6] This assessment is designed to assess students' ability to acknowledge, articulate, position and interpret the perspectives of multiple stakeholders in a social conflict, and provide solutions that consider and integrate their respective different positions. The assessment puts students in the shoes of an "advisor", who needs to address social conflicts that can occur in different contexts. In a sample assessment unit, test takers read a story about a student named Casey who is a victim of bullying, and are asked what they would recommend Casey should do, why, and what might go wrong with the recommendation. Students have to provide answers to these questions in the form of short, open responses.

Defining the construct of global understanding

Global understanding is a process that involves awareness of global issues and intercultural experiences. Access to information from around the globe and opportunities for intercultural encounters have greatly increased over the last decade, meaning that the majority of PISA students are exposed to a wide range of perspectives on global issues and intercultural experiences even if they do not actively search for them. However, access to information about the world and other cultures does not always go together with understanding. The oversimplification of complex knowledge is a significant contributing factor to deficiencies in learning (Spiro et al., 1988[64]), and is particularly frequent in the domain of global and cultural issues. Although misconceptions often arise from a lack of information, they are compounded by the fact that initial and deeply held beliefs about how the world works are difficult to subsequently change. Given that humans learn by creating classification systems, a lack of new knowledge or experiences can lead to oversimplified categorisations and generalisations which, in turn, can result in prejudice and stereotyping. However, misconceptions also arise even when students are exposed to appropriate information but absorb this information in a passive way, without reflecting on its deeper meaning or using the information to adjust their prior beliefs.

Students need to use knowledge and skills simultaneously in order to develop global understanding (Figure 6.4). If a student does not know much about a certain issue, they will find it difficult to identify flaws in texts, consider multiple perspectives, communicate in rich ways or consider the consequences of actions related to the issue in question (Willingham, 2008[65]). However, knowledge alone of intercultural and global issues without understanding adds little value. One can know, and continue to judge and dismiss superficially (Williams-Gualandi, 2015[66]). Understanding is the ability to use knowledge to find meaning and connections between different pieces of information and perspectives.

The cognitive processes that support global understanding

For analytical and assessment purposes, this framework distinguishes four interrelated cognitive processes that globally competent students need to use in order to fully understand global or intercultural issues and situations:

1. The capacity to **evaluate information, formulate arguments and explain complex situations and problems** by using and connecting evidence, identifying biases and gaps in information and managing conflicting arguments.
2. The capacity to **identify and analyse multiple perspectives and world views**, positioning and connecting their own and others' perspectives on the world.
3. The capacity to **understand differences in communication,** recognising the importance of socially appropriate communication conventions and adapting communication to the demands of diverse cultural contexts.
4. The capacity to **evaluate actions and consequences** by identifying and comparing different courses of action and weighing these actions against one another on the basis of short- and long-term consequences.

Globally competent students should thus be able to perform a wide variety of tasks utilising different cognitive processes. The first of these cognitive processes requires students to be able to reason with evidence about an issue or situation of local, global and intercultural significance; search effectively for useful sources of information; evaluate information on the basis of its relevance and reliability; synthesise information in order to describe the main ideas in an argumentative text or the salient passages of a conversation; and combine their background knowledge, new information and critical reasoning to build multi-causal explanations of global or intercultural issues.

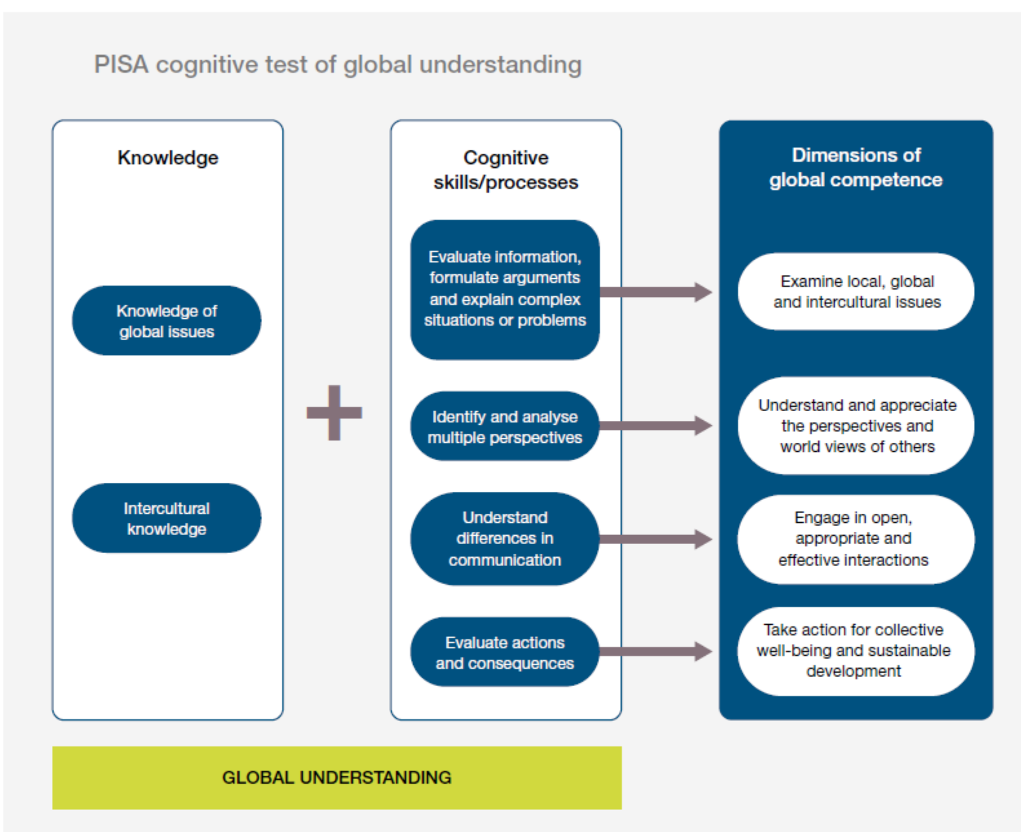
Furthermore, a solid understanding of a global or intercultural problem also requires recognising that one's beliefs and judgements are always contingent upon one's own cultural affiliations and perspectives. Students should therefore be able to recognise the perspectives of other people or groups and the factors that might influence them, including their access to information and resources. Students need to be able to explain how perspectives and contexts shape human interactions and interpretations of events, issues or phenomena.

Globally competent students should also identify ways to manage conflicts that emerge from communication problems, by analysing communicative contexts and conventions and recognising markers of respect.

Finally, students demonstrate their level of global understanding when they can evaluate different courses of action, propose solutions and consider the immediate and indirect implications of actions. The last constitutive cognitive process of global understanding therefore involves the ability to draw sound conclusions from the information one possesses and acquires.

Different types of tasks can test students' level of proficiency in applying each of these interrelated cognitive processes to a global or intercultural issue. For example, students can be asked to select the most reliable among a selection of different sources of information about an issue; they can evaluate whether a statement is valid and based on evidence; they can be asked to summarise and explain an issue or situation, or choose among possible summaries; they can be asked to identify passages of a media message transmitting negative stereotypes or making hasty generalisations; they can identify the different stakeholders in a case and list the possible contextual and cultural drivers of their respective positions; they can identify what passages in a conversation demonstrate a clear ignorance of intercultural communication approaches; or they can be asked to list or select the possible consequences of a proposed course of action for solving a problem.

Figure 6.4. The relationship between the cognitive test of global understanding and the dimensions of global competence



While all four cognitive processes are important indicators of a globally competent individual's skills, the test items in the PISA 2018 global competence assessment are not expected to cover all four cognitive processes in a balanced way. In particular, creating test items that validly measure students' understanding of communication norms and differences (process 3) is especially complex and might require a longer period of development and validation. This cognitive process is thus expected to be less represented than the other three in the 2018 PISA test.

Table 6.1 describes students' abilities at basic, intermediate and advanced levels of development of the four typologies of cognitive processes that constitute global understanding, the cognitive facet of global competence.

Table 6.1. Typologies of cognitive processes by level in the PISA 2018 test of global competence

Cognitive process	Sub-category	Basic	Intermediate	Advanced
1. Evaluate information, formulate arguments and explain complex situations or problems	**Selecting sources (range)**	The student prefers using sources stemming from her own cultural context without having an apparent strategy to search, select or differentiate between sources.	The student searches for and selects sources stemming from geographic and cultural contexts (region, language, perspective) beyond her own. She can also search for and select more than one source type (e.g. newspapers, publications, personal testimonies, government reports). However, no concrete strategy beyond a commitment to use different sources is apparent.	The student is able to frame the search systematically in a way which enables identifying the nature and extent of information needed to address the issue. She selects sources purposefully drawing on contexts and types that will inform her understanding of the issue at hand.
	Weighing sources (reliability and relevance)	The student takes the information at face value without considering contextual factors (author, geo-perspective, cultural context), or source kind. She cannot yet detect clear bias or inconsistencies. The student does not weigh the sources' relevance vis-à-vis the topic or claim at hand.	The student weighs sources for their relevance vis-à-vis the topic or claim at hand. The student also considers contextual factors that can inform her evaluation of a source's reliability. She can detect clear biases and inconsistencies, yet she shows a rather binary view of reliability ("biased"-"non-biased").	The student pays attention to contextual factors to establish a source's reliability and relevance. She understands the significance of different sources' perspectives, can distinguish between the communicative intentions of sources and claims (facts, opinions, propaganda), evaluate whether the assumptions or premises are reasonable and well-grounded in evidence, and identify assumptions or claims that reveal stereotypes.
	Employing sources (reasoning with evidence)	The student views the use of sources as a simple, unproblematic matter of copying and pasting information into an argument.	The student understands the need for multiple sources but uses a mechanistic approach when including sources in an argument (e.g. two "pro"- two "against" sources).	The student recognises the provisional nature of evidence and that multiple arguments can stem from similar sources. The student can consider

Cognitive process	Sub-category	Basic	Intermediate	Advanced
				evidence to explore and meet counterarguments. She can also address conflicting claims or sources.
	Describing and explaining complex situations or problems	The student can produce short summaries of information or perspectives. Summaries read as a string of information with little substantive organisation. The student is not yet capable of classifying the information.	The student can describe the issue/situation at hand in ways that connect larger concepts (e.g. culture, identity, migration) and simple examples. She can order content in a way that supports others' understanding of the issues.	The student can describe the issue/situation at hand in ways that connect larger concepts (e.g. culture, identity, migration) and relevant examples. She can develop and express clear, sound and effective arguments synthesising and connecting information provided in the task and information she acquired in or outside of school.
2. Identify and analyse multiple perspectives and world views	**Recognising perspectives and world views**	The student has a simplistic view of perspectives: one person, one perspective. She cannot yet explain the source of the perspective. The student views context as either irrelevant or as deterministic ("context as destiny"). She views perspectives (cultural, religious, linguistic) as relatively fixed, bounded or impermeable markers of a person's identity and world view. The student views an individual's identity as predominantly one category (such as nationality or religion). The student does not consider herself as having a distinct cultural perspective or world view and rather believes that what she knows is "the norm".	The student can identify different actors and points of view on an issue. The student begins to recognise that differences in perspectives or world views are rooted in cultural, religious, socio-economic, regional and other backgrounds, and that she also holds a particular view of the world. The student cannot yet articulate how multiple perspectives relate to one another. Differences in perspectives or world view start to be seen as rooted in cultural, religious, socio-economic, regional and other backgrounds. .	The student can describe and interpret multiple perspectives and world views. The student understands that perspectives are rooted in cultural, religious, socio-economic, regional and other backgrounds, and she understands how someone's geographic and cultural context can shape how that person sees the world. She also understands that an individual's identity is complex (one can be at once a girl, a daughter, a farmer, and a citizen). She can articulate relationships among perspectives, placing the perspectives in a broader encompassing frame (e.g. when the student sees that two classmates from different ethnic groups fight because of cultural prejudices, she understands that their relationship reflects broader tensions in today's society).

Cognitive process	Sub-category	Basic	Intermediate	Advanced
				The student views herself as holding perspectives and blind spots. She understands that her perspective is informed by her context and experiences and that others may perceive her in ways that may differ from the way she sees herself.
	Identifying connections	The student does not recognise connections among human beings apart from physical connotations and evident cultural markers. The student does not recognise the impact that actions have on others and sees individuals from different cultures or contexts as distant or exotic, who think and behave differently and who do not share similar rights or needs.	The student recognises that people from different cultures share most basic human rights and needs (e.g. food, shelter, work, education, happiness). She understands the meaning of these rights or needs and some of the ways in which they can be met.	The student appreciates common human rights and needs and reflects on individual, cultural or contextual differences critically, understanding the obstacles that individuals and societies may confront (economic inequality, unequal power relations, violence or unsustainable conduct) in affirming their rights to diversity and well-being. She also understands that universal human rights leave considerable space for national, regional and cultural individuality and other forms of diversity, and that they allow individuals and groups to pursue their own vision of what consists of a good life as long as their choices do not impede others' core human rights.
3. Understand differences in communication	**Understanding communicative contexts and respectful dialogue**	The student does not yet understand how to effectively and appropriately communicate based on audience and context. Specifically, she does not recognise cultural norms, interactive styles, expectations or levels of formality in a given social and cultural context and audience. The student is not yet able to observe, listen actively or interpret social and contextual clues such as body language, tone, diction, physical	The student is aware of her way of communicating and attempts to make that communication fit the context. The student can identify some interactive styles, expectations, or levels of formality in a given social and cultural context but cannot yet calibrate her language and communication choices accordingly. The student can respond to breakdowns in communication (for example, by requesting	The student is aware of her own styles of communication and understands that effective and appropriate communication must be adapted to audience, purpose, and context. Specifically, she is sensitive to nuances in cultural norms, interactive styles, expectations and levels of formality of a given social and cultural context and audience. She listens actively, observes carefully and gathers insight, including social and cultural

Cognitive process	Sub-category	Basic	Intermediate	Advanced
		interactions, dress code or silences. The student is surprised by any breakdowns in communication and lacks a communicative repertoire that can resolve or prevent such breakdowns.	repetitions or reformulations) but does so very tentatively.	clues that inform her communicative choices. The student can break down his or her messages, providing re-statements, revisions or simplifications of her own communication. She employs linguistic devices such as avoiding categorical claims, connecting to what others said, sharing questions and puzzles, and acknowledging contributions in ways that advance civil and reciprocal dialogue.
4. Evaluate actions and consequences	**Considering actions**	The student considers one course of action as obvious and unproblematic. For example, when presented with a problem about industrial pollution, her immediate conclusion would be "just close all polluting factories".	The student understands that multiple courses of action are possible and necessary to address an issue/situation or contribute to the well-being of individuals and societies. She can identify directions for future investigations if the available evidence is insufficient for reaching conclusions about the best course of action.	The student demonstrates an ability to identify and evaluate different courses of action to solve an issue/situation. She weighs these actions against one another, for example, by looking at precedents, considering and evaluating available evidence, and assessing the conditions that may make actions possible.
	Assessing consequences and implications	The student understands the implications of simple actions in linear terms without weighing multiple actions and implications or considering unintended consequences.	The student understands the most likely immediate consequences of a given position or course of action, and can assess how these consequences compare with available alternative positions/views.	The student considers the immediate and indirect consequences or implications of different possible actions and decisions. She can weigh short- and long-term consequences as well as short-range and spatially distant consequences. The student also considers the possibility of unintended consequences as a result of actions.

Content of the test units

A typical test unit is based on a scenario that focuses on one global or intercultural issue and presents different perspectives on the issue. Scenarios are often used as teaching tools, and their use in the test units can yield useful evidence for education policy and teachers as they encourage students to think logically and systematically.

A scenario-based design in an international assessment assumes that it is possible to identify a set of "big issues" that all young people should learn about, regardless of where they live or their socio-cultural background. However, an exact delimitation of relevant content for the scenarios is difficult because global and intercultural issues are in constant evolution. Nonetheless, Table 6.2 outlines four content domains and their related subdomains, which can be considered relevant for all students. Every scenario in the PISA cognitive test can therefore be categorised according to one of these content (sub-) domains.

Table 6.2. Content domains and subdomains of the scenarios

Content Domain 1: Culture and intercultural relations
Subdomain 1.1: Identity formation in multicultural societies
Subdomain 1.2: Cultural expressions and cultural exchanges
Subdomain 1.3: Intercultural communication
Subdomain 1.4: Perspective taking, stereotypes, discrimination and intolerance
Content Domain 2: Socio-economic development and interdependence
Subdomain 2.1: Economic interactions and interdependence
Subdomain 2.2: Human capital, development and inequality
Content Domain 3: Environmental sustainability
Subdomain 3.1: Natural resources and environmental risks
Subdomain 3.2: Policies, practices and behaviours for environmental sustainability
Content Domain 4: Institutions, conflicts and human rights
Subdomain 4.1: Prevention of conflicts and hate crimes
Subdomain 4.2: Universal human rights and local traditions
Subdomain 4.3: Political participation and global engagement

Test developers should aim for a balanced coverage of the four content domains across the different units that constitute each 1-hour cognitive test, favouring scenarios that cut across multiple content domains. The test units should privilege stimulus material that is familiar and relevant to 15-year-olds in order to facilitate students' engagement with the task. The risk associated with sensitive topics (e.g. a case study on hate violence against minorities may be sensitive for a student from a minority group) should be carefully assessed and minimised during the design of the scenarios and related test items. The combination of appropriate media, such as texts, comic strips and photography, can increase the quality and relevance of the scenario for students, reducing the reading load and increasing students' engagement with the tasks. It is also important to avoid scenarios which present a stereotypical representation of certain identities or cultural groups, and could thus further contribute to single stories and prejudice.

As well as varying by content, the scenarios in each test unit can vary by context. For example, they can refer to the personal context of the student (situations relating to the self, family and peer groups), to their local context (wider social networks, neighbourhood, city or country) or to a global context (life across the world, as experienced through exposure to the media and participation in social networks). For example, in the personal context of student interaction within a multicultural classroom – whereby such a classroom encompasses not only differences in national backgrounds but also in gender, religion, socio-economic background and so on – students can be assessed on their intercultural communication and understanding skills (cognitive processes 2 and 3, and content domain 1). Scenarios that incorporate histories of conflicts or positive cultural exchanges in multicultural neighbourhoods (local context) can serve as useful background for test items assessing students' understanding of the challenges of social integration with their local community; scenarios in which students are required to analyse global news in or work remotely on a project with students in a different country can tap into a wide variety of content domains and cognitive processes.

Complexity of the test units

The effective use of the assessed cognitive processes (described in Table 6.1) is intimately tied to the content knowledge the students have of the issue or situation they are asked to work on. While the cognitive skills of analysing and evaluating information are intrinsically general in nature, global and intercultural issues present their own specific challenges that require knowledge of the world and of cultural differences. For example, only those students who have some degree of knowledge of the consequences of climate change can fully understand conflicting positions in a debate on the reduction of carbon emission in cities. Similarly, if a student does not know anything about an issue, they will find it difficult to consider the issue from multiple perspectives. Background content knowledge is considered, in this framework, as an important facilitator of the cognitive processes that students activate when asked to reflect about the particular case study presented in a test unit.

When students read a piece of text or follow a conversation presented in the scenario of each test unit, their understanding is constrained by, on one hand, the content and complexity of the material in the scenario, and on the other hand, by the development of the cognitive processes necessary for global understanding. The cognitive demand of individual test units is therefore defined both by the level of content knowledge and cognitive skills that students need to activate in order to solve the tasks. In more demanding test units, the student must generally contribute information from his or her own knowledge about the content domain that is not explicitly stated in the scenario.

Table 6.3 sets out the categorisations of complexity of the test units, according to the level of content knowledge and general reading skills required by the scenario and test items. Although general language decoding and comprehension skills are not integral components of global competence, the language used in the test scenarios and items will inescapably influence the difficulty of test units. Highly complex language thus needs to be avoided to reduce the risk that the test results become heavily influenced by differences in text decoding and language comprehension skills. As for domain-specific content knowledge, the requirement of prior exposure to relevant information and intercultural situations is an important driver of a test unit's difficulty, and thus of students' performance on the cognitive test.

Table 6.3. Dimensions and levels of complexity of the scenarios

Levels of complexity	Domain-specific knowledge	Percentage of scenarios	General knowledge (text and language)	Percentage of scenarios
Low	The topic analysed in the test unit is familiar to the vast majority of students. Very limited prior knowledge of the topic/issue is required from students to understand what the unit requires.	Around 40%	The scenario is framed in very simple language, without technical words or expressions that are typical of a certain socio-cultural or demographic group.	Around 60%
Medium	Most students regularly hear about the topic/issue but they are not necessarily familiar with all its aspects. Students who have had some exposure to the	Around 40%	The language in the scenario is familiar to the majority of 15-year-old students. The choice of words is typical of communication addressed to non-specialist audiences. Differences in communication styles across groups are minimised whenever fictional conversations	Around 30%

Levels of complexity	Domain-specific knowledge	Percentage of scenarios	General knowledge (text and language)	Percentage of scenarios
	topic/issue in or outside of school can be expected to perform better on the unit.		are used as scenarios. Single texts are internally coherent and multiple texts are clearly connected.	
High	Most students have heard about the topic/issue but, given its complexity, only a minority of students can be expected to be familiar with the content of the unit. Students who have had some exposure to the topic/issue in or outside of school can more easily engage with the test unit and are expected to perform significantly better.	Around 20%	The scenario is framed in more complex language that is typical of formal writing or professional conversation, and can include a limited amount of content-specific or technical vocabulary. Communication between actors in the scenario can reflect differences in communication styles among groups, although most students are expected to be able to follow the conversation and understand its overall meaning (no jargon or convoluted phrasing is used).	Around 10%

International asymmetries in a student's opportunity to learn the subject matter are probably more important in an assessment of global competence than in assessments of more traditional subjects, such as science or mathematics. This is because only a minority of schools already consciously include global education in their curriculum, and the content of global education varies significantly across countries. Moreover, the learning process of global competence takes place within a context that extends far beyond the classroom: an important factor determining the extent to which students know about global issues and other cultures can be traced to the varying socio-cultural environments in which they live and learn. Learning for global competence is a cultural activity, not just because it is partly acquired through social interactions but also because the process is influenced by the way in which specific cultural groups interpret the world and transmit information.

These asymmetries in content knowledge are expected to matter for performance in the test. However, the design of the test makes the PISA cognitive assessment fundamentally different from a knowledge quiz. Firstly, no test item directly assesses factual knowledge (for example, an item would not ask students to specify the increase in global temperature reported in the last International Panel on Climate Change report). Secondly, only a minority of test units will require students to have a high level of background knowledge of global and intercultural issues (Table 6.2). While background content knowledge assists students' understanding of the scenario, performance on the test should mainly reflect students' capacities to use their reasoning and perspective-taking skills to connect their general knowledge of global issues to new and unanticipated problems and situations. The test design mitigates international asymmetries in students' opportunity to learn background content knowledge because it asks students to work on several short test units in different content domains. Test takers from a given socio-cultural context will thus likely have more background knowledge on some areas, but not on others.

Format of the scenarios

The scenarios used in the test should reflect the variety of contexts and roles in which students can learn about global issues or explore the complexity of intercultural interactions. The authenticity and relevance of the tasks are critically important to stimulate a sufficient level of engagement with the test. The scenarios can be designed using the following four formats that assign a particular role to the student, providing a clear purpose to engage in the task:

1. Students as researchers
2. Students as reporters
3. Students as mediators or team-members
4. Students as debaters

In the first format – students as researchers – the test takers are asked to imagine that they are enrolled in a course at their school and that they need to submit a collaborative research paper with other fellow students at the end of the school term. In this scenario, the student has to examine information from web searches or from inputs from other students on the team. This format tests multiple types of cognitive processes: students' capacity to select information can be assessed by presenting them with multiple results from web queries and asking them to select the one that is most appropriate to the research; students' perspective-taking abilities can be assessed by asking them to examine the causes of a misunderstanding or conflict between two members on the research team.

The second format presents performance tasks that students should solve by acting as reporters: the scenario asks students to put themselves in the shoes of a journalist who wants to write an article about a piece of news he or she has heard. The text in this type of scenario typically takes the form of an extract from a newspaper or from social media where the main elements of a case are presented. A first question or set of questions typically verifies whether the students understand the message, can assess the quality and credibility of information reported in the source, and can reason beyond the text by questioning possible motivations and subjective interpretations of the information by the author. The scenario then develops as students are asked to search for their own information and sources, such as by asking students to identify which stakeholders they would like to interview and/or selecting relevant questions to ask different actors in order to better understand their actions and perspectives. This type of scenario can assess all the cognitive processes in the framework, and works particularly well for assessing students' capacity to select, use information and assess the validity of information. The investigative nature of the tasks should be sufficiently stimulating and realistic for most students.

'Students as mediators/team-members' scenarios ask students what they would suggest to moderate or solve a conflict in their school or neighbourhood. The text typically takes the form of a conversation, where two or more actors have a conflict over an issue. The questions ask students to identify who is involved in the situation, how the different stakeholders are likely to feel, think and react, and why they think and react in this way, based on the relationships between characters and their social and cultural characteristics. The test-taker can also be asked to generate or identify possible solutions that consider the interests of all or most parties. This type of scenario can effectively test students' ability to acknowledge, articulate, position and interpret multiple stakeholders' perspectives in a given social conflict, and provide solutions that consider and integrate these different positions.

Finally the 'student as debater' scenarios require test takers to develop arguments and compare different perspectives on an issue in the form of a debate. The scenario typically provides some background information on the issue that students can use for their responses. The questions in the scenario ask the students to develop (or select) arguments for their side, and address and rebut the arguments put forth by their opponent's side. If properly transposed to an assessment format, the debate format can stimulate students' engagement and give them the opportunity to demonstrate their grasp of thinking and communication skills.

This description of scenario formats is not exhaustive, and other types of scenarios can be explored during the test development process.

Response format

The form in which the evidence is collected – the response format – varies according to the cognitive process that is assessed and the chosen format of the scenario. Various response formats can require different skills. For example, closed and multiple-choice response items depend more on decoding skills, because readers have to eliminate incorrect responses, when compared to open-constructed response items (Cain and Oakhill, 2006[67]).

As in any large-scale assessment, the range of feasible item formats is limited to some combination of open and closed response questions. However, contextualised open-response items are particularly relevant for this assessment as they ask the learner to assemble relevant, abstract, conceptual and case-specific knowledge components for a problem-solving task (Spiro et al., 1988[64]). Open-response items were already used and validated in the ICCS's International Cognitive Test (Schulz et al., 2010[57]), NAEP Civics (National Assessment Governing Board, 2010[68]) and in the United Kingdom's GCSE examination in Citizenship Studies (UK Government Department for Education, 2014[69]). The open-response items are scored using rubrics, or scoring guidelines that include detailed qualitative descriptions of performance standards (Andrade, 2005[70]; Popham, 1997[71]; Popp, Ryan and Thompson, 2009[72]; Stellmack et al., 2009[73]; Thaler, Kazemi and Huscher, 2009[74]).[7] Most units in the test should include at least one question with an open-response format.

Moderators of performance: reading comprehension, attitudes and values

Certain individual factors that are not explicitly assessed in the PISA cognitive test may nonetheless moderate students' performance. In the 2018 iteration of the test, the scenarios are mostly based on written texts, despite efforts to efficiently integrate texts and images. The capacities that students need in order to perform well on the global competence test therefore overlap to a certain extent with those required for reading literacy, because the PISA definition of reading literacy has progressively put more emphasis on students' capacities to analyse, synthesise, integrate and interpret multiple texts (OECD, 2016[75]). However, this framework identifies a set of perspective-taking and reasoning abilities that clearly go beyond reading proficiency, and focuses on the application of these abilities to specific content areas (global and intercultural issues). The specificities of global issues and intercultural relations contribute to defining and determining the cognitive processes and skills employed in the tasks.

It will be possible to measure and partially account for the correlation between reading skills and global understanding as students tested in global competence in 2018 will also be tested in reading. Thus, individual students' and countries' results on the assessment could be compared before and after accounting for their performance in reading.

Attitudes can facilitate global and intercultural understanding at the affective level, and can thus act as moderators of performance in the cognitive test. Some examples of attitudes that support the practice and development of cognitive skills with respect to global competence are a curiosity about other cultures; inquisitiveness with regard to a wide range of global issues; conscious efforts to remain well-informed about current events at the local and global level; a positive and respectful regard of cultural differences; and a desire to do something about global problems that threaten the needs and freedoms of current and future generations (global mindedness). These attitudes will not be measured directly in the cognitive test. However in the contextual PISA questionnaire, students will report the extent to which they agree with a series of statements related to such attitudes (see Section 3.3). The triangulation of the results of the cognitive test and the self-reported information from the questionnaire will provide relevant evidence on how attitudes support global and intercultural understanding.

Arguably, the most complex issue for the operationalisation of this assessment framework relates to a clear definition of the way in which values affect global and intercultural understanding. While values are an integral part of global competence, the PISA cognitive test does not assess values. The proposed test asks students to reflect on the validity and consequences of statements, and to elaborate their own conclusions about a specific issue or situation.

This issue requires a careful choice of the test questions that can be included in the international cognitive assessment. Students could be asked to evaluate statements that are clearly right or wrong on the basis of objective criteria, because they adhere to or contradict agreed scientific or historical evidence. However, the questions in the cognitive test should not aim at assessing students on their ethics and opinions, but rather on their capacity to recognize and explain the complexity of a case and the multiplicity of possible positions. For example, in a hypothetical scenario describing the case of a father who steals in order to feed his starving children, the students would not be asked to conclude whether or not the action deserves a given punishment; instead, students would be asked to demonstrate an understanding that the law may in some cases and under certain perspectives collide with basic human needs, and to identify/explain the possible risks and uncertainties of establishing ad-hoc exceptions to the law.

Self-reported information in the student questionnaire

In addition to the results of the cognitive assessment, the reporting on global competence in PISA 2018 will include country- or sub-population-level information on students', school principals', teachers' and parents' responses to questionnaire items.

For socio-emotional skills and attitudes, finding the right method of assessment is arguably more a stumbling block than deciding what to assess. It is essentially not possible to define scales for self-reported attitudes and skills that are always completely valid. The strategy adopted in PISA 2018 has privileged the use and adaptation of scales that have already been validated in other empirical assessments.

The most common problem with assessing self-reported skills and attitudes is that of social desirability. Attitudes, in particular, are related to self-image and social acceptance. In order to preserve a positive self-image, students may be tempted to answer questionnaire items in a way that they believe is socially acceptable. Self-reported scales that measure attitudes towards race, religion, sex, etc. are particularly affected by social desirability bias. Respondents who harbour a negative attitude towards a particular group may not wish to admit, even to themselves, that they have these feelings. In a study of attitudes towards refugees, Schweitzer et al. (2005[76]) found that social desirability bias accounted for 8% of the variance in attitudes.

A large number of Likert-type scales appear in the literature on civic and democratic attitudes and a number of them are related to global competence as defined in PISA.[8] The Global-Mindedness Scale, for example, was developed in order to "measure attitudes of students related to their sense of connection to, interest in and responsibility for the global community and the behaviours associated with this perspective" (Hett, 1993[77]). The items in the scale addressed both beliefs and behaviours: for example, students were asked to report the extent to which they agreed with the statement "I tend to judge the values of others based on my own value system".

Following this literature, the student questionnaire in PISA 2018 includes multi-statement items using Likert-type methods. These items are based, as much as possible, on pre-existing work, taking into account issues of testing time and question sensitivity and adapted as best as possible to the reality of 15-year-old students. The questions and items on global competence that were included in the PISA 2018 student questionnaire are available on line (www.oecd.org/pisa). These questions are a subset of a larger set of material that was field trialled across all countries participating in PISA. In the transition from the field trial to the main study, some questions were deleted and some scales were shortened in order to save testing time, all the while still ensuring the proper coverage of this framework and preserving the psychometric validity of the scales. The longer questionnaire tested in the PISA field trial and the field trial analysis of the psychometric quality of the material are both available upon request.

The analysis of the responses to these items is expected to support the future development of questions on attitudes and behavioural or emotional skills that might be included in future rounds of PISA. Future work beyond 2018 might also consider integrating other methods for measuring attitudes and "soft skills" that are less prone to social-desirability bias.

Self-reported knowledge and skills

Self-reported knowledge of global and intercultural issues

A first set of questions in the student questionnaire covers the dimension of knowledge of global and intercultural issues. One question in the PISA 2018 questionnaire asks students to report how easily they could perform a series of tasks relating to global issues, such as explaining how carbon dioxide emissions affect global climate change. Another question asks students to report how familiar they are with different global issues, such as climate change and global warming, global health and migration.

Self-reported ability to communicate in multicultural contexts

A second set of questions refers to the linguistic, communication and behavioural skills that are required to communicate with other people, to manage breakdowns in communication and to mediate between speakers of different languages or cultures. Students' progression in this component can be evaluated according to their proficiency in a foreign language and through their self-reported ability to handle communication with people from other cultural backgrounds and in unfamiliar contexts.

Self-reported data on foreign language proficiency can be used to examine the relationships between acquiring a second language and measured levels of global understanding or positive dispositions toward other countries and cultures. Such an investigation could have several relevant policy implications for both language teaching efforts and curricular programmes aimed at increasing the level of students' understanding of global issues.

The student questionnaire for PISA 2018 reports how many languages students and their parents speak well enough to be able to converse with others. The questionnaire also includes one question asking the students the extent to which they would explain things very carefully, check understanding or adapt their language when talking in their native language with people whose native language is different.

Self-reported adaptability

Research on intercultural communication has developed and validated several items and scales on adaptability and flexibility. For example, the Intercultural Effectiveness Scale of Portalla and Chen (2010[78]) includes self-reported measures of behavioural flexibility, such as the level of agreement with the statement "I often act like a very different person when interacting with people from different cultures". The PISA 2018 question includes one multi-statement question on adaptability, asking students how they deal with challenging interactions with people from other cultural backgrounds. The six items in the question were adapted from validated scales in Martin and Rubin (1995[79]) and Dennis and Vander Wal (2010[80]).

Self-reported perspective taking

As in the case of adaptability, there are several scales on perspective taking and on empathy that have been specifically designed for adolescents and reviewed for the PISA questionnaire. These include the Index of Empathy for Children and Adolescents (Bryant, 1982[81]), the empathy subscale from the Children's Behaviour Questionnaire (Rothbart, Ahadi and Hershey, 1994[82]), the Interpersonal Reactivity Index (Davis, 1980[83]), the Basic Empathy Scale (Jolliffe and Farrington, 2006[84]), and the Adolescent Measure of Empathy and Sympathy (Vossen, Piotrowski and Valkenburg, 2015[85]). In the PISA student questionnaire, one question comprised of five items assesses perspective taking. The five items have been adapted from Davis (1983[86]) and are expected to form a unidimensional construct.

Self-reported attitudes

Self-reported openness toward people from other cultural backgrounds

The PISA questionnaire includes one question assessing students' "interest in learning about other cultures". The question assesses a student's desire or willingness to learn about other countries, religions and cultures. The four items included in the question have been adapted from different sources, such as Chen et al. (2016[87]) and Mahon and Cushner (2014[88]).

Self-reported respect for people from other cultural backgrounds

One question in the PISA questionnaire asks the students to report to what extent they feel they respect and value other people as equal human beings, no matter their cultural background. The five items were adapted from the Council of Europe (2016[89]), Munroe and Pearson (2006[90]), Lázár (2012[91]), and Fritz et al. (2002[92]).

Self-reported global mindedness

The PISA questionnaire includes one question on global mindedness. The six items in the question are expected to assess the following facets of global mindedness: sense of world citizenship (item 1), responsibility for others in the world (items 2, 4 and 6), sense of interconnectedness (item 3) and global self-efficacy (item 5).

Questionnaire items on strategies, pedagogies and attitudes to teach global competence

The PISA 2018 questionnaire will provide information on innovations in curricula and teaching methods aimed at preparing students for global citizenship. Two questions focus on the curriculum. One question asks principals and teachers whether the curriculum includes global topics such as climate change and global warming, global health or migration. Another question asks principals and teachers whether the formal curriculum refers to global competence skills and dispositions, such as communicating with people from different cultural backgrounds or countries, or openness to intercultural experiences.

A second set of questions focuses on educators' beliefs and practices. One question asks principals to report on their teachers' general beliefs about how the school should handle ethnic diversity. A second enquires about specific practices for multicultural learning at the school level, such as teaching about the beliefs, customs or arts of diverse cultural groups that live in the country, or encouraging students to communicate with people from other cultures via the internet and social media.

Two questions in the PISA teacher questionnaire enquire about teachers' level of preparation to respond to different student communities, potentially through different teaching strategies. One question provides information as to whether a teacher has studied intercultural issues or received training in pedagogical methods to teach effectively in multicultural environments. Another question will provide information about teachers' self-efficacy in coping with the challenges of a multicultural classroom and adapting their teaching to the cultural diversity of students.

The student questionnaire also provides information on teachers' behaviours from the perspective of the students. One question, in particular, asks students to report whether they think that their teachers treat students from all cultural groups with equal respect.

Conclusions

How education systems respond to growing economic interdependence, cultural divides, new digital opportunities and calls for sustainability will have a significant impact on the well-being of all members of the communities they serve. All people, in both diverse and homogenous communities, are called upon to challenge cultural stereotypes, reflect on the causes of racial, religious and hate violence, and participate in the creation of respectful, integrated and sustainable societies.

Achieving global competence through education will require significant changes in the classroom: changes concerning what students learn about the world and other cultures, the opportunities they have to practice what they learn, and how teachers support this learning by working with diverse students. Some national curricula now put an emphasis on education for sustainable development and intercultural education. Many teachers also encourage students to analyse and reflect on the root causes of global issues, and share ideas on possible solutions. However, progress has been uneven and good practices have not been shared sufficiently at the international level.

The conceptual framework for global competence, and the approach that PISA will take for its first international assessment in 2018, described in this document, will offer the first comprehensive overview of education systems' success in equipping young people to address global developments and collaborate productively across cultural differences in their everyday lives. The data will provide insights on which policy approaches to global education are most commonly used in school systems around the world, and on how teachers are being prepared to promote global competence. Education systems will thus learn from each other about how to best adapt curricula, promote innovative teaching methods and adjust teachers' initial education and training so as to facilitate the development of global competence.

The results of the assessment can also stimulate innovation at the level of individual schools as they seek effective approaches to enhance their students' global competence. A broad range of learning activities in the classroom can in fact influence students' global competence and involve teachers in all subject areas, although to differing degrees. These may include role-playing activities that allow students to take on different perspectives, discussions on prejudice and discrimination, or project-based activities that encourage students to analyse and reflect on the root causes of global issues.

No single assessment can fully account for the complexity of global competence as a learning goal. Importantly, the PISA approach reflects the needs and the constraints of a large-scale international assessment. It is thus no substitute for formative assessments of global competence at the classroom and school level. More efforts, beyond 2018, will be needed to build on the lessons learnt from this initiative to further improve the measurement of the constructs defined in this framework. The most challenging, but perhaps most urgent endeavour will be to experiment with and evaluate new methods to further improve the measurement of the socio-emotional, attitudinal and value dimensions of global competence.

Notes

[1] The discussion of knowledge, attitudes, skills and values in this section draws upon the conceptualisation of these components provided by the Council of Europe (2016[41]) which was developed through an extensive process. It involved auditing 101 existing conceptual global, intercultural and civic competence schemes. The basic values, attitudes, skills, knowledge and understanding throughout the schemes were then identified, and a set of criteria identifying the core values, attitudes, skills, knowledge and understanding was established. Next, a first draft of the resulting model was produced and academic experts, education practitioners and policy makers reviewed and endorsed the model. It was then fine-tuned and finalised, taking into account the experts' feedback. Full details of the development process can be found in Council of Europe (2016[41]).

[2] See the web platform "idebate.org" and Schuster and Meany (2005[93]) for resources on debates in school education.

[3] For examples of specific activities to use in the classroom, see (Anna Lindh Foundation, 2017[95]; Berardo and Deardorff, 2012[96]; Council of Europe, 2015[97]; Fantini, 1997[98]; Seelye, 1996[99]; Storti, 2017[100]; Stringer and Cassiday, 2009[106]).

[4] Here system is used in a broad sense to include not just states and markets, but also husbands, parents, officials, landowners, social authorities etc. In other words, all those who have power and can use it to control or interfere in people's lives.

[5] Measurement instruments of critical thinking include the Ennis–Weir Critical Thinking Essay Test (Ennis and Weir, 1985[101]), Cornell Critical Thinking Test (Ennis, Millman and Tomko, 1985[102]), ETS HEIghten™ Critical Thinking Assessment (Liu, Frankel and Roohr, 2014[60]; Liu et al., 2016[103]) and the Halpern Critical Thinking Assessment (Halpern, 2010[104]).

[6] See http://ccdd.serpmedia.org/ for more information.

[7] (Doscher, 2012[105]) explores the validity and reliability of two rubrics for the Global Learning Initiative at Florida International University (FIU). The rubrics referred to two case studies measuring university students' global awareness and perspectives. The rubrics yielded scores that reliably measured students' global learning outcomes. Students who attended global learning courses scored significantly higher on the performance tasks than students who did not attend such courses.

[8] Likert scales involve a series of statements to which respondents indicate agreement or disagreement on, for example, a 4- or 5-point response scale.

References

Amadeo, J. et al. (2002), *Civic Knowledge and Engagement. An IEA Study of Upper Secondary Students in Sixteen Countries (CIVED)*, IEA, Amsterdam, http://pub.iea.nl/fileadmin/user_upload/Publications/Electronic_versions/CIVED_Phase2_Upper_Secondary.pdf. [56]

Andrade, H. (2005), "Teaching with Rubrics: The Good, the Bad, and the Ugly", *College Teaching*, Vol. 53/1, pp. 27-30, http://www.jstor.org/stable/27559213 . [70]

Anna Lindh Foundation (2017), *The Anna Lindh Education Handbook: Intercultural Citizenship in the Euro-Mediterranean Region*. [95]

Barrows, T. et al. (1981), *College Students' Knowledge and Beliefs: A Survey of Global Understanding*, Change Magazine Press, New Rochelle, New Jersey, https://files.eric.ed.gov/fulltext/ED215653.pdf. [55]

Bennett, J. (1993), "Towards ethnorelativism: A developmental model of intercultural sensitivity", in Paige, M. (ed.), *Education for the intercultural experience*, Intercultural Press, Yarmouth, ME, https://www.amazon.com/Education-Intercultural-Experience-Paige/dp/1877864250. [7]

Berardo, K. and F. Trompenaars (eds.) (2012), *Building Cultural Competence: Innovative Activities and Models*, Stylus Publishing, Sterling, VA, https://www.amazon.com/Building-Cultural-Competence-Innovative-Activities/dp/1579228046. [96]

Bringle, R. and P. Clayton (2012), "Civic Education through Service Learning: What, How, and Why?", in McIlrath L., Lyons A. and Munck R. (eds.), *Higher Education and Civic Engagement*, Palgrave Macmillan, New York, https://doi.org/10.1057/9781137074829_7. [94]

Bringle, R. et al. (2016), *Service Learning in Psychology: Enhancing Undergraduate Education for the Public Good*, American Psychological Association, Washington D.C. [38]

British Council (2013), *Culture At Work: The Value of Intercultural Skills in the Workplace*, British Council, United Kingdom, http://www.britishcouncil.org. [15]

Brubaker, R. and D. Laitin (1998), "Ethnic and Nationalist Violence", *Annual Review of Sociology*, Vol. 24, pp. 423-452, https://www.jstor.org/stable/223488. [12]

Bryant, B. (1982), "An Index of Empathy for Children and Adolescents", *Child Development*, Vol. 53/2, pp. 413-425, http://dx.doi.org/10.2307/1128984. [81]

Byram, M. (2008), *From Foreign Language Education to Education for Intercultural Citizenship*, Multilingual Matters, Clevedon. [40]

Cain, K. and J. Oakhill (2006), "Assessment Matters: Issues in the Measurement of Reading Comprehension", *British Journal of Educational Psychology*, Vol. 76/4, pp. 697-708, http://dx.doi.org/10.1348/000709905x69807. [67]

Chen, S. et al. (2016), “Conceptualizing psychological processes in response to globalization: Components, antecedents, and consequences of global orientations”, *Journal of Personality and Social Psychology*, Vol. 110/2, pp. 302-331, http://dx.doi.org/10.1037/a0039647. [87]

Clapham, A. (2006), *Human Rights Obligations of Non-State Actors*, Oxford University Press, Oxford, http://dx.doi.org/10.1093/acprof:oso/9780199288465.001.0001. [46]

Costa, A. and B. Kallick (2013), *Dispositions: Reframing Teaching and Learning*, Corwin Press, Thousand Oaks, CA. [37]

Council of Europe (2016), *Competences for Democratic Culture (CDC): 559 Validated Descriptors*. [89]

Council of Europe (2016), *Competences for Democratic Culture: Living Together as Equals in Culturally Diverse Democratic Societies*, Council of Europe Publishing, Strasbourg, http://book.coe.int. [41]

Davis, M. (1983), “Measuring Individual Differences in Empathy: Evidence for a Multidimensional Approach”, *Journal of Personality and Social Psychology*, Vol. 44, pp. 113-26, http://dx.doi.org/10.1037/0022-3514.44.1.113. [86]

Davis, M. (1980), “A Multidimensional Approach to Individual Differences in Empathy”, *JSAS Catalog of Selected Documents in Psychology*, Vol. 10/85. [83]

Deardorff, D. (2011), *Promoting Understanding and Development of Intercultural Dialogue and Peace: A Comparative Analysis and Global Perspective of Regional Studies on Intercultural Competence*, UNESCO Division of Cultural Policies and Intercultural Dialogue, Paris. [50]

Deardorff, D. (2009), “Implementing Intercultural Competence Assessment”, in Deardorff, D. (ed.), *The SAGE Handbook of Intercultural Competence*, Sage Publications, Thousand Oaks, CA. [3]

Deardorff, D. (2019, forthcoming), *UNESCO Manual on Developing Intercultural Competence Based on Human Rights*, UNESCO, Paris. [39]

Delors, J. (1996), *Learning: The Treasure Within*, UNESCO, Paris. [13]

Dennis, J. and J. Vander Wal (2010), “The Cognitive Flexibility Inventory: Instrument Development and Estimates of Reliability and Validity”, *Cognitive Therapy and Research*, Vol. 34/3, pp. 241-253, http://dx.doi.org/10.1007/s10608-009-9276-4. [80]

Donnelly, J. (2007), “The Relative Universality of Human Rights”, *Human Rights Quarterly*, Vol. 29/2, pp. 281-306, http://dx.doi.org/10.1016/0925-4005(93)85399-U. [45]

Doscher, S. (2012), *The Development of Rubrics to Measure Undergraduate Students’ Global Awareness and Global Perspective: A Validity Study*, Florida International University, http://dx.doi.org/10.25148/etd.FI12050125. [105]

Ennis, R., J. Millman and N. Tomko (1985), *Cornell Critical Thinking Tests*, Midwest Publications, Pacific Grove, CA. [102]

Ennis, R. and E. Weir (1985), *The Ennis–Weir Critical Thinking Essay Test*, Midwest Publications, Pacific Grove, CA, http://www.academia.edu/1847582/The_Ennis-Weir_Critical_Thinking_Essay_Test_An_Instrument_for_Teaching_and_Testing. [101]

Fantini, A. (ed.) (1997), *New Ways in Teaching Culture*, TESOL, Alexandria, VA, https://eric.ed.gov/?id=ED421001. [98]

Fritz, W., A. Mollenberg and G. Chen (2002), "Measuring intercultural sensitivity in different cultural contexts", *Intercultural Communication Studies*, Vol. 11, pp. 165-176. [92]

Gaudelli, W. (2006), "Convergence of Technology and Diversity: Experiences of Two Beginning Teachers in Web-Based Distance Learning for Global/Multicultural Education", *Teacher Education Quarterly*, Vol. 33/1, pp. 97-116, https://www.jstor.org/stable/23478737. [21]

Gaudelli, W. (2003), *World Class: Teaching and Learning in Global Times*, Lawrence Erlbaum Associates Publishers. , Mahwah, New Jersey. [27]

Gay, G. (2013), "Teaching To and Through Cultural Diversity", *Curriculum Inquiry*, Vol. 43/1, pp. 48-70, http://dx.doi.org/10.1111/curi.12002. [32]

Global Cities (2017), *A Framework for Evaluating Student Outcomes in Global Digital Education*, Global Cities Inc., New York. [36]

Halpern, D. (2010), *Halpern Critical Thinking Assessment Manual*, Schuhfried GmbH, Vienna, https://sites.google.com/site/dianehalperncmc/home/research/halpern-critical-thinking-assessment. [104]

Hansen, R. (2010), *Impact of Study Abroad on Ethnocultural Empathy and Global-Mindedness"*, Unpublished Doctoral Dissertation, Ball State University. [42]

Hanvey, R. (1975), *An Attainable Global Perspective*, Center for War/Peace Studies, New York. [9]

Harzing, A. (2006), "Response Styles in Cross-national Survey Research", *International Journal of Cross Cultural Management*, Vol. 6/2, pp. 243-266, http://dx.doi.org/10.1177/1470595806066332. [54]

Hett, E. (1993), *The Development of an Instrument to Measure Global Mindedness*, Unpublished Doctoral Dissertation, University of San Diego. [77]

Huber, J. and C. Reynolds (eds.) (2014), *Developing intercultural competence through education*, Council of Europe Publishing, Strasbourg, http://book.coe.int. [1]

IBO (2012), *What is an IB Education? (brochure)*, IBO, Cardiff. [18]

Jolliffe, D. and D. Farrington (2006), "Development and validation of the Basic Empathy Scale", *Journal of Adolescence*, Vol. 29/4, pp. 589-611, http://dx.doi.org/10.1016/j.adolescence.2005.08.010. [84]

Karaman, A. and F. Tochon (2007), "International Student Teaching in World Language Education: Critical Criteria for Global Teacherhood", *Critical Inquiry in Language Studies*, Vol. 4/2-3, pp. 237-264, http://dx.doi.org/10.1080/15427580701389714. [22]

Khoza, R. (2011), *Attuned Leadership: African Humanism as Compass*, Penguin Global, Johannesburg, https://www.amazon.co.uk/Attuned-Leadership-African-Humanism-Compass/dp/014352819X. [49]

Klein, J. (2013), "Making Meaning in a Standards-Based World: Negotiating Tensions in Global Education", *Educational Forum*, Vol. 77/4, pp. 481-490, http://dx.doi.org/10.1080/00131725.2013.822044. [29]

Kymlicka, W. (1995), *Multicultural Citizenship: A Liberal Theory of Minority Rights*, Oxford University Press, Oxford, http://dx.doi.org/10.1093/0198290918.001.0001. [10]

Lázár, I. (2012), *Recognising Intercultural Competence: What Shows That I Am Interculturally Competent?*, Council of Europe, Strasbourg. [91]

Lee, H. (2017), *An International Collaborative Study of GCED Policy and Practice (I)*, Korea Institute for Curriculum and Evaluation, Seoul. Research Report RRO 2017-1. [31]

Liu, O., L. Frankel and K. Roohr (2014), *Assessing Critical Thinking in Higher Education: Current State and Directions for Next-Generation Assessment*, Educational Testing Service, Princeton, NJ, http://dx.doi.org/10.1002/ets2.12009. [60]

Liu, O. et al. (2016), "Assessing Critical Thinking in Higher Education: The HEIghten™ Approach and Preliminary Validity Evidence", *Assessment and Evaluation in Higher Education*, Vol. 41/5, pp. 677-694, http://dx.doi.org/10.1080/02602938.2016.1168358. [103]

Mahon, J. and K. Cushner (2014), "Revising and updating the inventory of cross-cultural sensitivity", *Intercultural Education*, Vol. 25/6, pp. 484-496, http://dx.doi.org/10.1080/14675986.2014.990232. [88]

Martin, M. and R. Rubin (1995), "A New Measure of Cognitive Flexibility", *Psychological Reports*, Vol. 76/2, pp. 623-626, https://doi.org/10.2466/pr0.1995.76.2.623. [79]

Merryfield, M. (2008), "Scaffolding Social Studies for Global Awareness", *Social Education*, Vol. 72/7, pp. 363-366. [23]

Mokgoro, J. (1995), *State v. Makwanyane and Mchunu*, (6) BCLR 665 (CC), para 308. [47]

Mompoint-Gaillard, P., I. Lázár and F. Arató (eds.) (2015), *TASKs for democracy; 60 activities to learn and assess Transversal Attitudes, Skills, and Knowledge*, Council of Europe Publishing, Strasbourg, http://www.coe.int. [97]

Munroe, A. and C. Pearson (2006), "The Munroe Multicultural Attitude Scale Questionnaire: A new instrument for multicultural studies", *Educational and Psychological Measurement*, Vol. 66/5, pp. 819-834, http://dx.doi.org/10.1177/0013164405285542. [90]

Myers, J. (2006), "Rethinking the Social Studies Curriculum in the Context of Globalization: Education for Global Citizenship in the U.S", *Theory and Research in Social Education*, Vol. 34/3, pp. 370-394, http://dx.doi.org/10.1080/00933104.2006.10473313. [24]

National Assessment Governing Board (2010), *Civics Framework for the 2010 National Assessment of Edutional Progress*, U.S. Department of Education, Washington D.C., https://www.nagb.gov/assets/documents/publications/frameworks/civics/2010-civics-framework.pdf. [68]

North-South Centre of the Council of Europe (2012), *Global Education Guidelines: Concepts and Methodologies on Global Education for Educators and Policy Makers*. [33]

Nussbaum, M. (1997), *Cultivating Humanity: A Classical Defense of Reform in Liberal Education*, Harvard University Press, Cambridge, MA. [44]

Nwosu, P. (2009), "Understanding Africans' conceptualizations of intercultural competence", in Deardorff, D. (ed.), *The SAGE Handbook of Intercultural Competence*, Sage, Thousand Oaks, CA. [48]

O'Connor, K. and K. Zeichner (2011), "Preparing US teachers for critical global education", *Globalisation, Societies and Education*, http://dx.doi.org/10.1080/14767724.2011.605333. [28]

OECD (2016), *PISA 2018 Draft Reading Framework*, OECD, Paris. [75]

Omerso, E. (ed.) (2011), *Educating for Global Competence: Preparing Our Youth to Engage the World*, Asia Society, Council of Chief State School Officers, https://asiasociety.org/files/book-globalcompetence.pdf. [2]

O'Reilly, T. and J. Sabatini (2013), *Reading for Understanding: How Performance Moderators and Scenarios Impact Assessment Design*, ETS Research Report RR-13-31, http://www.ets.org/Media/Research/pdf/RR-13-31.pdf. [61]

Oxfam (2015), *Education for Global Citizenship A Guide for Schools*. [19]

Patel, V. (ed.) (1988), *Cognitive Flexibility Theory: Advanced Knowledge Acquisition in Ill-Structured Domains*, Lawrence Erlbaum , Hillsdale, NJ. [64]

Popham, W. (1997), "What's Wrong—and What's Right—with Rubrics", *Educational Leadership*, Vol. 55/2, pp. 72-75. [71]

Popp, S., J. Ryan and M. Thompson (2009), "The critical role of anchor paper selection in writing assessment", *Applied Measurement in Education*, Vol. 22/3, pp. 255-271, http://dx.doi.org/10.1080/08957340902984026. [72]

Portalla, T. and G. Chen (2010), "The Development and Validation of the Intercultural Effectiveness Scale", *Intercultural Communication Studies*, Vol. 19/3, pp. 21-37, https://web.uri.edu/iaics/files/02TamraPortallaGuo-MingChen.pdf. [78]

Rapoport, A. (2010), “We Cannot Teach What We Don’t Know: Indiana Teachers Talk About Global Citizenship Education”, *Education, Citizenship and Social Justice*, Vol. 5/3, pp. 179-190, http://dx.doi.org/10.1177/1746197910382256. [25]

Reimers, F. (2017), *Empowering Students to Improve the World in Sixty Lessons. Version 1.0*, CreateSpace Independent Publishing Platform. [20]

Rothbart, M., S. Ahadi and K. Hershey (1994), “Temperament and social behavior in childhood”, *Merrill-Palmer Quarterly: Journal of Developmental Psychology*, Vol. 40, pp. 21-39. [82]

Rychen, D. and L. Salganik (2003), *Key Competencies for a Successful Life and a Well-Functioning Society*, Hogrefe and Huber. [35]

Sabatini, J. et al. (2014), “Broadening the Scope of Reading Comprehension Using Scenario-Based Assessments: Preliminary findings and challenges”, *L’Année Psychologique*, Vol. 114/4, pp. 693-723, http://dx.doi.org/10.4074/S0003503314004059. [62]

Sabatini, J. et al. (2015), “Improving Comprehension Assessment for Middle and High School Students: Challenges and Opportunities”, in *Improving Reading Comprehension of Middle and High School Students*, Springer, New York. [63]

Schulz, W. et al. (2010), *ICCS 2009 International Report: Civic Knowledge, Attitudes, and Engagement Among Lower-Secondary School Students in 38 Countries*, IEA, http://www.iea.nl. [57]

Schulz, W. et al. (2018), *Becoming Citizens in a Changing World IEA International Civic and Citizenship Education Study 2016 International Report*, International Association for the Evaluation of Educational Achievement (IEA), https://doi.org/10.1007/978-3-319-73963-2. [58]

Schuster, K. and J. Meany (2005), *Speak Out! Debate and Public Speaking in the Middle Grades*, Interntational Debate Education Association, New York. [93]

Schwartz, S. (2012), “An Overview of the Schwartz Theory of Basic Values”, *Online Readings in Psychology and Culture*, Vol. 2/1, http://dx.doi.org/10.9707/2307-0919.1116. [43]

Schweitzer, R. et al. (2005), “Attitudes Towards Refugees: The Dark Side of Prejudice in Australia”, *Australian Journal of Psychology*, Vol. 57/3, pp. 170-179, http://dx.doi.org/10.1080/00049530500125199. [76]

Seelye, H. (ed.) (1996), *Experiential Activities for Intercultural Learning*, Intercultural Press, Yarmouth, ME. [99]

Sen, A. (2007), *Identity and Violence: The Illusion of Destiny*, Penguin Books India, India. [11]

Sinicrope, C., J. Norris and Y. Watanabe (2007), “Understanding and Assessing Intercultural Competence: A Summary of Theory, Research and Practice, Technical Report for the Foreign Language Program Evaluation Project”, *Second Language Studies*, Vol. 26/1, pp. 1-58, http://citeseerx.ist.psu.edu/viewdoc/download;jsessionid=D8D4BDBE0C5918D895EFC04B36C738E4?doi=10.1.1.517.1193&rep=rep1&type=pdf. [8]

Stellmack, M. et al. (2009), "An Assessment of Reliability and Validity of a Rubric for Grading APA-Style Introductions", *Teaching of Psychology*, Vol. 36/2, pp. 102-107, http://dx.doi.org/10.1080/00986280902739776. [73]

Storti, C. (2017), *Cross-Cultural Dialogues: 74 Brief Encounters with Cultural Difference*, Intercultural Press, Yarmouth, ME. [100]

Stringer, D. and P. Cassiday (2009), *52 Activities for Improving Cross-Cultural Communication*, Intercultural Press, Yarmouth, ME. [106]

Suarez, D. (2003), "The Development of Empathetic Dispositions through Global Experiences", *Educational Horizons*, Vol. 81/4, pp. 180-182, https://www.jstor.org/stable/42925459. [26]

Suárez-Orozco, C., M. Suárez-Orozco and I. Todorova (2008), *Learning a New Land: Immigrant Students in American Society*, Harvard University Press, Cambridge, MA. [34]

Thaler, N., E. Kazemi and C. Huscher (2009), "Developing a Rubric to Assess Student Learning Outcomes Using a Class Assignment", *Teaching of Psychology*, Vol. 36/2, pp. 113-116, http://dx.doi.org/10.1080/00986280902739305. [74]

Torney-Purta, J. et al. (2015), "Assessing Civic Competency and Engagement in Higher Education: Research Background, Frameworks, and Directions for Next-Generation Assessment", *ETS Research Report Series*, http://dx.doi.org/10.1002/ets2.12081. [59]

UK Government Department for Education (2014), *Citizenship studies: Draft GCSE subject content (DFE-00582-2014)*, Department of Education. [69]

UNESCO (2016), *Education 2030: Incheon Declaration and Framework for Action for the implementation of Sustainable Development Goal 4: Ensure inclusive and equitable quality education and promote lifelong learning opportunities for all*. [17]

UNESCO (2016), *Global Education Monitoring Report*, UNESCO, Paris, https://unesdoc.unesco.org/ark:/48223/pf0000246230. [6]

UNESCO (2015), *Global Citizenship Education: Topics and Learning Objectives*, UNESCO, Paris. [30]

UNESCO (2014), *Global Citizenship Education: Preparing Learners for the Challenges of the 21st Century*, UNESCO, Paris, https://unesdoc.unesco.org/ark:/48223/pf0000227729?posInSet=1&queryId=50e44b1f-6384-4b61-a20a-f07f8cd63538. [5]

UNESCO (2014), *Learning to Live Together: Education Policies and Realities in the Asia-Pacific*, UNESCO, Paris, http://www. [14]

UNESCO (2013), *Intercultural Competences: Conceptual and Operational Framework*, UNESCO, Paris, http://www.unesco.org/new/en/bureau-of-strategic-planning/themes/culture-of-peace-and-non-violence/. [4]

UNESCO (2009), *Investing in Cultural Diversity and Intercultural Dialogue: UNESCO World Report*, UNESCO, Paris, https://unesdoc.unesco.org/ark:/48223/pf0000185202. [52]

UNESCO (2006), *UNESCO Guidelines on Intercultural Education*, UNESCO, Paris, https://unesdoc.unesco.org/ark:/48223/pf0000147878. [53]

UNESCO (2001), *Universal Declaration of Cultural Diversity*, UNESCO, Paris, https://unesdoc.unesco.org/ark:/48223/pf0000127160. [51]

Vossen, H., J. Piotrowski and P. Valkenburg (2015), "Development of the Adolescent Measure of Empathy and Sympathy (AMES)", *Personality and Individual Differences*, Vol. 74, pp. 66-71, http://dx.doi.org/10.1016%2Fj.paid.2014.09.040. [85]

Williams-Gualandi, D. (2015), "Intercultural Understanding: What Are We Looking for and How do we Assess What We Find?", *International and Global Issues for Research*, No. 2015/7, University of Bath, Bath, http://bathedresearch.weborium.net. [66]

Willingham, D. (2008), "Critical Thinking: Why Is It So Hard to Teach?", *Arts Education Policy Review*, Vol. 109/4, pp. 21-32, http://dx.doi.org/10.3200/AEPR.109.4.21-32. [65]

Zuckerman, E. (2013), *Rewire: Digital Cosmopolitans in the Age of Connection*, W. W. Norton & Company. [16]

Annex 6.A. Illustrative examples of scenarios for the cognitive assessment of global understanding

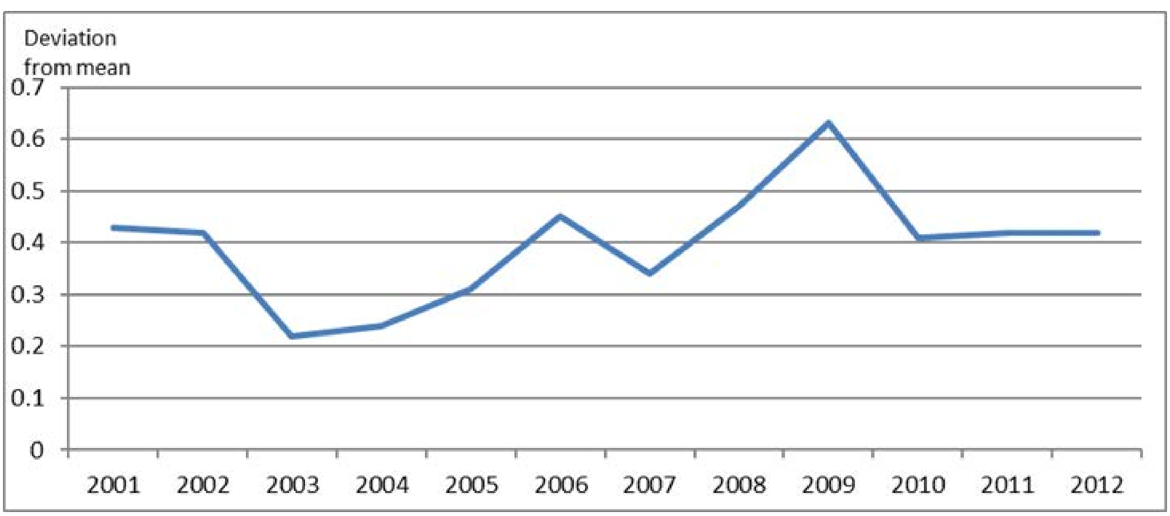
This section presents several examples of tasks that students might be asked to complete in the PISA global competence assessment. They are provided for illustrative purposes only. They have not been prepared nor verified by the professional test developers who are responsible for developing the cognitive instruments for PISA 2018. No fully developed test items are included among these examples. The questions and their answer keys below were meant to guide the development of both multiple choice and open response items.

Are global temperatures rising?

Content domain: 3. Environmental Sustainability (3.1: Natural resources and environmental risks)

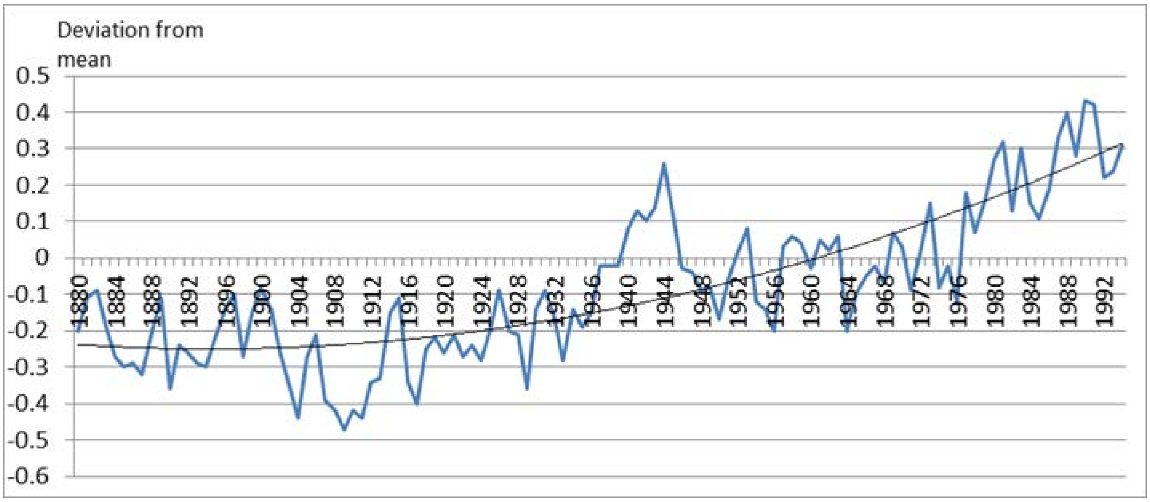
In her science class, Mei reads a research article that was featured in the daily press. The author of the article uses the following graph to argue that popular claims about a rise in global temperatures are not supported by the data. In fact, global temperatures were lower in 2011 and 2012 than in 2008 and 2009.

Annex Figure 6.A.1. Global temperature



Mei's teacher asks the class to have a look at another chart she produced from the same source of data in the article.

Annex Figure 6.A.2. Global temperature



Question: What can you infer about the validity of the article's claim by comparing the two charts?

***Question classification**: 1. Evaluate information, formulate arguments and explain complex situations or problems (1.2: weighing sources).*

***Answer key**: The author's claim is not based on solid evidence. The author should have considered a longer time frame to analyse* changes in global temperatures.

The teacher tells the class that the research in the article was financed by a major oil corporation. She also explains that some companies that hire researchers to perform a study require the researchers to sign a nondisclosure agreement before they are funded, by which researchers waive their right to release any results independently.

Question: What is a possible consequence of allowing unregulated sponsoring of scientific research by industrial companies?

***Question classification**: 4. Evaluate actions and consequences (4.2 Assessing consequences and implications)*

***Answer key**: If not properly regulated, some financing might result in a "funding bias", due to the fact that a researcher might be induced to support the interests of the sponsor.*

A talented player

Content domain: 1. Culture and intercultural relations (1.4 Perspective taking, stereotypes, discrimination and intolerance).

Last week-end your team lost because a foreign-born player decided to walk away from the game after putting up with racial insults by the visiting team's fans for almost one hour, forcing your team to play 10 against 11. One of your friends was at the stadium, and told you that the player should have gone on with the game, and not have let the insults get to him.

Question: What could have prevented the player leaving and destabilising his team?

***Question classification**: 4. Evaluate actions and consequences (4.1 Considering actions)*

***Answer key**: Clear regulations enforced by the referee in which he or she suspends a match whenever he/she hears racial insults, disqualifying the team whose supporters perpetrate racist acts.*

As you keep talking about the player who left the game, you realise that both you and your friend have never used his real name but always referred to him as "the Animal". This is the nickname he got from the press after his first game with your team. The captain of your team, who is also the captain of your national team, is nicknamed “the Brain".

Question: What is a possible consequence of the choice of nicknames?

***Question classification**: 4. Evaluate actions and consequences (4.2 Assessing consequences and implications)*

***Answer key**: It can reinforce a belief that national players are smart, hardworking, team players while foreign players are athletes who get by on their natural gifts.*

A song in Quechua

Content domain: 1. Culture and Intercultural Relations (1.1 Identity formation in multi-cultural societies)/ 4. Institutions, conflicts and human rights (4.3 Political participation and global engagement)

In a YouTube video that reached over 2 million viewers, Renata Flores sings in Quechua, her native language, to Michael Jackson's "The Way You Make Me Feel” against the backdrop of ancient Inca ruins. Renata is an activist participating in a project called *‘Las juventudes tambien hablamos Quechua‘* (The youth, we speak Quechua too).

Question: What messages do you think Renata is trying to convey?

***Question classification**: 2. Identify and analyse multiple perspectives (2.1 Recognising perspectives and contexts)*

***Answer keys**: She wants to combat young people's perceptions of the indigenous language as unhip and backwards. She wants to revive her culture and combat uniformity.*

Several other initiatives are trying to revive disappearing languages. For example, one of the top Internet search engines has launched a version in Quechua and the New South Wales Government of Australia has proposed legislation for protecting and reviving Aboriginal languages. However, keeping alive a disappearing language is not an easy task.

Question: Which factors, among the following, can contribute to the disappearance of languages?

***Question classification**: 1. Evaluate information, formulate arguments and explain complex situations or problems (1.4 Describing and explaining complex situations or problems)*

***Answer keys**: Young people from minority groups who think that speaking their heritage language is not cool; lack of Aboriginal and indigenous language teachers; few disappearing languages have written grammar and dictionaries that people can use to learn them.*

Annex 6.B. Description of possible topics for the scenarios of the cognitive test

This annex lists global and intercultural issues that can be used as reference topics to develop scenarios in the cognitive test. It is implied in this list that these complex topics have to be developmentally appropriate for 15-year-old and sufficiently engaging.

Culture and intercultural relations

This content domain relates to the manifold expressions of cultural diversity, such as languages, arts, knowledge, traditions and norms. Acquiring knowledge in this domain can help young people recognise that perspectives are shaped by multiple cultural influences, better understand differences among cultures, and value the importance of protecting cultural differences.

Identity formation in multi-cultural societies

This subdomain focuses on how young people develop their cultural identity in multicultural communities and interconnected societies. Scenarios in this content area can describe: situations where minority individuals and/or migrants must navigate between minority ethnic (home) culture and majority national (peer group and school-academic) cultures; young citizens' rights and responsibilities in different societies; complex views of identity (national, gender, religious); ideas of culture as fixed and determined versus dynamic and permeable; expectations of how adolescents should behave in and outside of school; causes of supportive and conflicting relationships between teachers and students in multicultural classes; relationships with parents, family and community networks in different cultures; tensions between cultural celebrations and attempts to affirm larger cultural identities; understanding of power and privilege within a society; distinction between collective and individual cultural orientations and the different value judgements which can arise from these. Scenarios may also address how young people construct and respond to digital identities. It will be important for these scenarios to address the multiple, complex identities held by individuals so that they do not perpetuation the "single story" identity.

Cultural expressions and cultural exchanges

This subdomain focuses on issues related to preserving the world's cultural capital (e.g. language, arts and traditions) and the relationships between dominant and non-dominant cultures. Scenarios in this content area can describe: expression of different cultures in a globalised world; significance of cultural diversity; public policies to protect and promote the diversity of language and other cultural expressions; school initiatives to encourage learning and appreciating different cultural traditions; different perspectives on what development means and on how countries should support other countries' development; designing art and cultural education programmes in schools; new technologies' role in providing access to cultural expressions; diversity of public media (access, content and language); convergence of people's *habits* and *consumption* patterns and how transnational ideas (e.g. hip hop, meditation) are culturally appropriated in local contexts and/or fused with other cultural practices to form hybrid cultures. Scenarios could include recognising cultural elements or messages within such expressions.

Intercultural communication

This subdomain focuses on what students can learn about the complexity of communicative processes involving individuals from different cultural backgrounds. Scenarios in this area can represent situations where: diverse audiences interpret different meanings from the same information; two or more people fail to understand each other because they follow different communication norms; individuals explore the idea that languages sometimes encode meanings which can be difficult to access in other languages; multicultural settings such as schools, community organisations, or workplaces become more effective as colleagues/peers adjust their communication styles; people fail to understand each other because of different non-verbal communication styles (especially given that more is often communicated nonverbally than through spoken word); individuals adapt (or fail to adapt) their communication style to different contexts (academic/informal neighbourhood/online settings); individuals seek to communicate while not sharing a language. These situations can be within informal contexts to which 15-year-olds may better be able to relate such as conflict on a sports team, within a friend group, in welcoming a new student (even from within the same country but different background), and so on.

Perspective taking, stereotypes, discrimination and intolerance

This subdomain refers to what students can learn about social/cultural understanding and perspective taking as well as the nature, manifestations and impact of cultural prejudices and ways to combat these. Scenarios in this subdomain can reproduce texts, media messages or conversations that: exhibit some explicit or implicit cultural bias against some groups; describe how individuals adjust and suffer as a result of cultural prejudices; show how people correct their stereotypes as they acquire new information about the others. Common expressions of prejudice and oversimplification include: gender or socioeconomic-based stereotyping about what students can achieve in different subjects; gender or racial biases while selecting applicants for a job; perceptions about certain groups' predispositions to violence and crime; stereotypes about indigenous cultures; intolerance towards sexual inclinations; religious stereotypes. The scenarios may invite students to identify, articulate, explain and position different cultural perspectives. They may ask students to engage with these discrimination cases and manage dilemmas associated with conflicting value systems. Specifically, this could be a conversational exchange in which a biased remark is made and the respondent must determine how to respond.

Socio-economic development and interdependence

This domain focuses on economic links between local, regional and worldwide levels and looks at how these links influence opportunities around the globe and across social or cultural groups. Students who acquire an advanced level of knowledge in this domain more easily understand how people, places and economies are strongly interrelated, and are aware that economic policies and choices made at any level have consequences at all levels, individual to global.

Economic interactions and interdependence

This subdomain focuses on the connections and interdependencies of economic systems at multiple levels. Some examples of scenario topics framed in this subdomain are: transnational production of everyday goods (cell phones, clothing); financial liberalisation, contagion and crisis; capital flow directions and instability; the emergence of global corporations; impacts of low-cost travel and shipping on local economic systems; technological investments and technology exchanges; wage differences and foreign investments; and the impact of job migration on countries.

Human capital, development and inequality

This subdomain focuses on the relationship between economic integration and social development. Examples of topics in this subdomain include: inequality in education, trends in income inequalities between and within countries; economic integration and reducing poverty; developing sustainable tourism; changes in employment opportunities in the face of global automated production and computerisation; impact of immigration on employment and wages; education mobility and brain drain.

Environmental sustainability

This content domain focuses on the complex systems surrounding the demand for and use of natural resources. Students who are more exposed to this area learn about the main drivers that deplete the planet's natural environment, and better understand how improving the quality of life should be pursued without damaging the planet for future generations.

Natural resources and environmental risks

In this subdomain students learn about the main environmental risks facing our planet and about the ecological interdependence of the natural world. The environmental risks considered in this subdomain are widespread, concerning both developed and developing countries, and cause harm to people who have not voluntarily chosen to suffer their consequences, requiring public authority regulation. In most cases, these risks cannot be assessed precisely, and can be evaluated differently in different contexts and social terms. A partial list of these risks include: climate change; air pollution and related health risks; pollution and over acidification of the oceans; soil degradation; desertification and drought; population growth and unsustainable urbanisation; natural disasters; glacier mass balance; contamination from pesticide residues; loss of biodiversity on the planet; access to clean, fresh water; overfishing, clearing of forests. With any of these topics, it will be important to select ones that are most relevant to 15-year-olds, such as clean water and air pollution, since everyone needs to drink water and go outside.

Policies, practices and behaviours for environmental sustainability

This subdomain focuses on what policy makers and individuals can do to reduce resource depletion and better manage environmental risks. Scenarios in this subdomain can ask students to reflect on tools and instruments (e.g. standards, taxes, subsidies, communications campaigns, education) put in place to encourage sustainable consumption and production; how environmental risks are communicated in the media; how governments weigh the risks of the depletion of natural resources when making choices of economic policy; what role non-government organisations have in forming the public opinion about environmental issues and changing policies; trade-offs between development and environmental concerns and differences in how sustainable development is understood and political responsibilities allocated in different countries and contexts.

Institutions, conflicts and human rights

This content domain focuses on the formal and informal institutions supporting peaceful relationships between people and the respect of fundamental human rights. Students can learn how global institutions such as the United Nations have developed, can be asked to reflect on the contested nature of global governance in a world with highly asymmetrical power relationships, review factors and solution of current and historical conflicts between countries, ethnic or social groups, and examine spaces and opportunities for young people to play an active part in society and exercise their rights and responsibilities.

Prevention of conflicts and hate crimes

This subdomain relates to institutions and strategies for managing, resolving and preventing violent conflicts. Relevant conflicts include international wars, civil wars, ethnic or religious conflicts and hate crimes against particular groups. Scenarios in this area can expose students to different interpretations about the causes of a particular violent conflict; present different historical reconstruction of conflicts driven by competition over scarce natural resources or by economic competition between countries; encourage them to analyse strategies for managing, resolving and preventing conflicts; ask them to think about why some conflicts are more difficult to resolve than others; let them reflect on the psychological preconditions that might be necessary for reconciliation between conflicting parties (e.g. willingness to admit that one's own group has perpetrated unacceptable acts, etc.); make them examine the role of non-violent protests in social and political change, conflicting definitions of social justice, and contrasting arguments about the conditions for lasting peace and greater social cohesion.

Universal human rights and local traditions

This subdomain includes human right education and scenarios can refer to key documents such as the Universal Declaration of Human Rights or the UN Convention on the Rights of the Child. students might be asked to reflect on the reasons why some people's rights are denied (e.g. why gender inequalities in access to education persist); enquire about the political, legal, socio-cultural, religious and economic factors that can undermine human rights in particular contexts; analyse opposing arguments and evidence about the universality or relativity of human rights; reflect on the obligations of states in relation to human rights and/or on the means to protect oneself which are available to citizens; reflect on rights that are in conflict with one another and how to resolve such conflicts.

Political participation and global engagement

This subdomain refers to the opportunities young people across the world have to express their voice and make a difference in local or global contexts. Scenarios in this area can describe real experiences of young people who have taken action to improve peoples' living conditions in their own or other communities, or who are evaluating which actions they can take on a social, civic or political issue. The situations presented in the scenarios can also describe practical difficulties young people face when they start volunteering, such as lack of knowledge about the people they wish to help, recognising their limits in taking action as an individual, backlash, discouragement and fatigue. This subdomain also includes issues related to how young people are exposed to political propaganda and develop their political opinions.

7. PISA 2018 Questionnaire Framework

This document presents the framework for the background questionnaires for the 2018 cycle of the Programme for International Student Assessment (PISA). These questionnaires are, to a large extent, developed from ones that were used in previous cycles of PISA and, as such, permit the monitoring of trends in student-, school-, and system-level factors that may be related to student outcomes. A variety of constructs are discussed, including student background constructs, schooling constructs, and non-cognitive/metacognitive constructs. The document also makes explicit the link to reading literacy, which is the major domain of this cycle of PISA. The relevance of each of the constructs to policy issues and student outcomes is also highlighted.

Introduction

Outline of the framework

This document provides the framework for the development of the background questionnaire. It defines all major constructs, meaning what is measured by one or more items of a scale, that will be assessed in the background questionnaires of PISA 2018, including student background constructs, school-level constructs and non-cognitive and metacognitive constructs. The framework also discusses the current literature on these constructs. However, the framework focusses on the why and how of assessing constructs rather than on the relationships between the constructs, which has been addressed in many previous publications of PISA results. A wide variety of theoretical models have been proposed to link systemic or curriculum variables (e.g. instructional approaches and educational expenditure) to student achievement. However, a "yield study" such as PISA in which a set of tests is administered to a group of students only once without any follow-up has only a limited scope for causal analysis. Therefore, the emphasis in the present framework is on identifying constructs that are of interest in either previous PISA cycles or the current literature, rather than on specifying their links.

The document is organised into two main parts: (1) defining the core content of the PISA questionnaires and elaborating its modular structure, and (2) explaining the policy issues that the questionnaires cover. Detailed references to current research are provided throughout the document.

The first part of this document links the current framework to the overarching (cross-cycle) structure of previous PISA assessments and questionnaires, as set out in the PISA 2012 and 2015 frameworks (Klieme et al., 2013[1]; Klieme and Kuger, 2014[2]; OECD, 2013, p. 168[3]). The constructs that need to be covered for monitoring trends in education are discussed in the context of research into the effectiveness of education systems. These measures have been used previously in PISA reports, as international indicators (published in *Education at a Glance*) and in secondary analyses.

The second part of this document explores the in-depth policy issues covered in PISA 2018. These issues are organized by modules, which comprise one or more related constructs (assessed by items or scales); for example, the module on domain-general student attitudes and behaviours is composed of various scales, such as self-related beliefs and attitudes towards school, well-being and the utilisation of information and communications technology (ICT). Additionally, this part of the document explains how the modules were implemented in the PISA 2018.

Some newly developed questions, spread across various modules, will be tested in the 2018 field trial, providing a broad set of measures that can be used in the PISA 2018 main study and/or in later cycles. Modules, constructs, questions and items will be selected for inclusion in the PISA 2018 main study based on the results from the field trial.

Defining the questionnaire core in PISA 2018

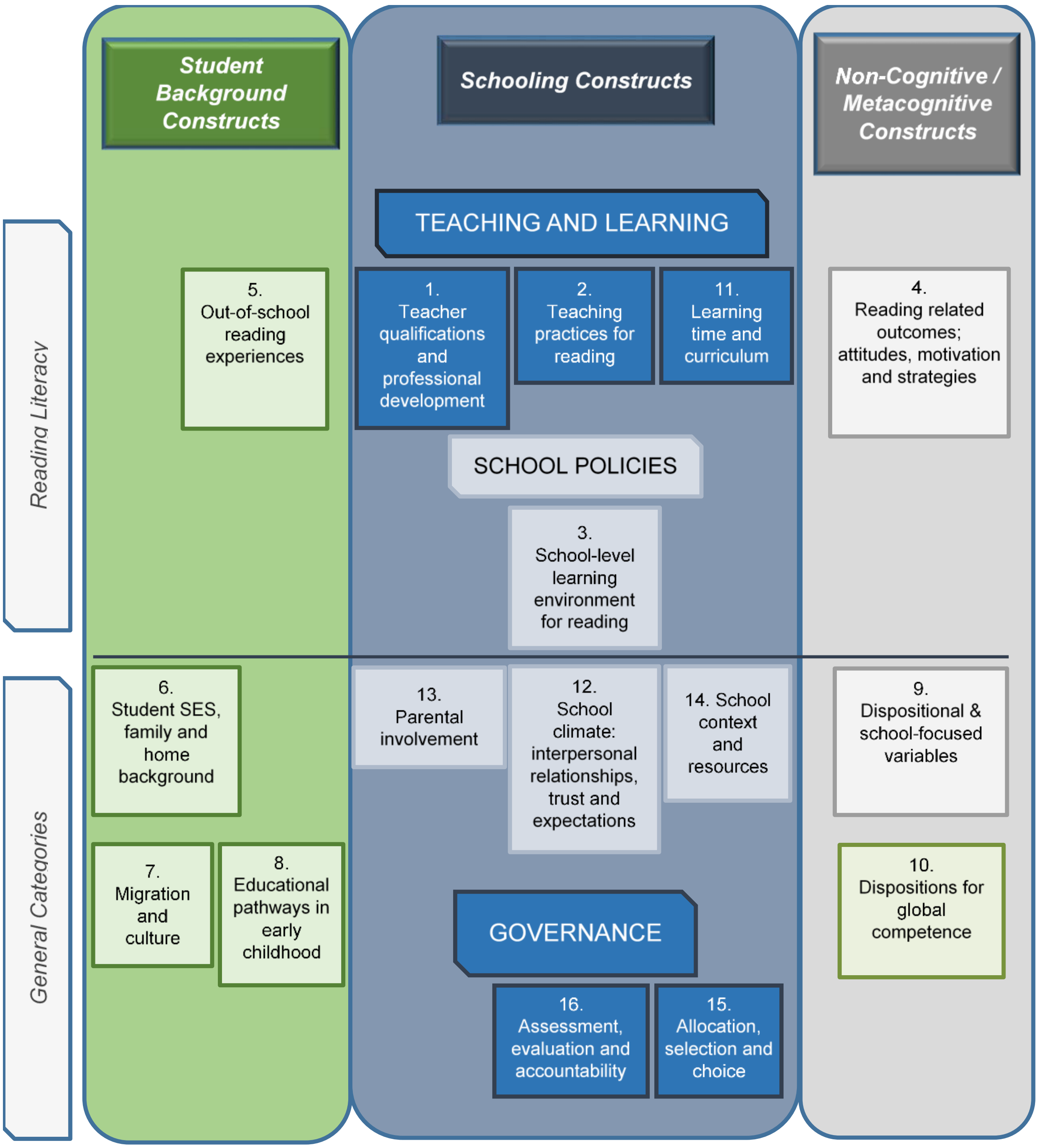
One of the major features of the implementation of PISA is the cyclical change in focus of the cognitive assessment: reading literacy was the major domain of assessment in PISA 2000 and 2009 and is so again in PISA 2018, whilst mathematics was the major domain of PISA 2003 and 2012 and science in PISA 2006 and 2015. The major domain of cognitive assessment is also the focus of domain-specific context assessment in the associated questionnaire – in other words, various reading-related constructs are assessed in the 2018 questionnaire since reading is the major domain. However, there is also a need for stability in measures administered in different waves in order to gauge and understand trends in education. Stability has to be considered at two levels: across waves of three years (various questions in the questionnaires tend to recur in every cycle) and in subject-specific constructs across waves of nine years (reading-specific constructs assessed in the 2009 wave could be reused in 2018).

The questionnaire framework first established for PISA 2012 and continued for PISA 2015 identifies core questionnaire content that should be kept comparable across cycles (OECD, 2013, p. 189[3]) to allow for the continuous monitoring of education systems and the establishment of valid and reliable trends. This includes both domain-specific and cross-domain measures that assess the conditions, processes and outcomes of education, at the level of both the student and the school.

This overarching framework, which specifies the constructs and measures in more detail and provides arguments that support the choice of core content for PISA 2018, is described below.

Outline of the content covered in the questionnaires

Figure 7.1. PISA 2018 Questionnaire modules



The green columns on the left-hand side of Figure 7.1 (student background constructs; modules 5 to 8) summarise students' family background and the education they have received to date. The items associated with these columns are typically asked of students or parents. The constructs in the blue columns in the middle of Figure 7.1 (modules 1 to 3 and 11 to 16), refer to educational processes on different levels (system, school and classroom). Most of the questions associated with these columns are answered by schools but some may be answered by students or parents. The last grey column on the right of Figure 7.1 (modules 4, 9 and 10) asks students about various non-cognitive and metacognitive (strategy awareness) constructs. The upper half mainly deals with domain-specific (in this case, reading-related) topics, while the lower half of the figure deals with general topics not focusing on a specific domain or those domains other than reading.

Every module represents a focus of policy making, and the set of 16 modules covers a wide and comprehensive array of policy issues that are relevant across countries. Indeed, most topics treated by Sykes, Schneider, and Plank (2009[4]) and by the OECD (2015[5]) in their reviews of educational policy research are covered here. This framework first discusses non-cognitive and metacognitive constructs, followed by student background constructs, teaching and learning constructs, and finally school policy and governance constructs.

PISA treats the mandatory core questionnaires (school questionnaire and student questionnaire) separately from the optional questionnaires, which countries may opt out of. The framework attempts to make the connections among the questionnaires as transparent as possible. It also describes what is conceptually covered in each questionnaire, which constructs are examined at the student and at the school levels, and who responds to each individual question.

Reading as the major domain

A new reading framework has been developed for PISA 2018. While this new framework shares many similarities with the 2000 and 2009 frameworks, it has been reconceptualised to address the main differences between print and online reading (Afflerbach and Cho, 2010[6]). In online reading, the text is not given: the reader has to build his/her own text, choosing which paths to follow and which ones to dismiss, in a context where the reader is offered many more options and opportunities in which to get lost. The PISA 2018 framework for reading literacy aims to address the additional complexities linked to online reading comprehension as defined by Coiro and Dobler (2007[7]), such as additional sources of prior knowledge (knowledge about search engines, website structures), a higher incidence of multilevel forward inferential reasoning (predicting what is behind a link), and new dimensions of self-regulated reading (the integration of physical actions such as clicking and scrolling down with cognitive processes such as predicting, assessing and evaluating the quality of information). All the reading-related constructs and questions should therefore cover both print and online reading.

Moreover, the new reading framework highlights the importance of metacognitive task management processes, such as setting goals and plans, monitoring and regulation.

Global competence

PISA 2018 introduces the new domain of global competence. This domain is seen as being critically important because our learning, working and living environments are becoming more global, interconnected and interdependent. Young people will encounter, actively engage with and help shape these environments, no matter where they are born, attend school, work or live. They therefore need to leave school equipped with the necessary knowledge, skills and attitudes that will enable them to learn, work and live in a globalised world, and with the tools to further develop these attributes as they move through life. In particular, this involves a knowledge of and interest in engaging with the world around them; a growing confidence and a spirit of curiosity, adventure, flexibility and resilience; and the communication and interaction skills necessary to make the most of the opportunities and challenges that fast-changing, interconnected and interdependent environments bring. Classrooms and schools should foster the value of and embrace the diversity of peoples, languages and cultures. Schools should also encourage intercultural sensitivity and help students move away from ethnocentric world views and beyond tolerance to acceptance, respect and appreciation. Students can be given the chance to engage in experiences that facilitate international and intercultural relationships, exchanges and conversations and should then reflect upon what they have learned from such experiences (Bennett, 1993[8]; Sinicrope, Norris and Watanabe, 2007[9]).

Several authors have signalled that because of the increasing speed of change in society (such as developments in information and communications technology (ICT) and the rise of cross-border working), schools need to adapt their curricula to account more precisely for what students will need in their future lives (Fisch and McLeod, 2009[10]). Furthermore, Anderson (2008[11]) points out that the knowledge and skills required to prosper in the 21st century go far beyond the traditional literacies. He identifies knowledge construction, adaptability, finding, organising and retrieving information, information management, critical thinking and teamwork as the skills demanded by modern societies. Meanwhile, Binkley et al. (2012[12]) maintain that achieving competence in 21st century skills, such as global competence, requires specific knowledge, skills, attitudes, values and ethics.

Constructs to be covered in the questionnaires

Since PISA began in 2000, the background questionnaires have served two interrelated purposes. The first purpose has been to provide a context through which to interpret scores from the cognitive assessment (both within and across education systems). The second purpose has been to provide reliable and valid non-cognitive outcomes, which can inform policy and research in their own right. Over the six cycles of PISA to date, new non-cognitive outcomes have emerged for both domain-specific and cross-domain features of education. The background questionnaire has also tracked developments in psychometric theory and survey research methodology so as to provide increasingly reliable and valid measures of non-cognitive constructs that are not sensitive to cultural differences in response style. These developments have taken place while maintaining the ability to report trends across PISA cycles.

PISA 2018 is the seventh cycle of PISA and the third cycle where reading is the major domain of assessment. In addition, PISA 2018 will also introduce a new domain – global competence – while diminishing the distinctions among the major and minor domains. These goals will require additional questionnaire delivery time, and may therefore prompt a change in how the questionnaires are developed and designed.

This section of the questionnaire framework presents the constructs for PISA 2018 and is organised around: (1) non-cognitive and metacognitive constructs; (2) student background constructs; (3) teaching and learning constructs; and (4) school policies and governance constructs.

Non-cognitive and metacognitive constructs

PISA measures and documents the outcomes of education attained at the age of 15. Educating a student means *fostering* his or her *individual development* as a unique, self-determined, knowledgeable person who gradually gains in his or her ability to participate in society. As each PISA cycle is a cross-sectional study, it does not capture developmental processes in the same way that longitudinal studies can; rather, PISA serves as a snapshot of students' developmental status at the age of 15. This includes achievement in cognitive areas such as reading, mathematics and science, all of which are measured in PISA;

however, other outcomes are also important. Success in school – and in life – also depends on being committed to learning, respecting and understanding others, being motivated to learn and being able to regulate one's own behaviour. These constructs can be perceived as prerequisites to learning, but they may themselves also be judged as goals of education, as elaborated in the OECD project *Defining and Selecting Key Competencies* (DeSeCo) (Rychen and Salganik, 2003[13]). Educational research has shown that non-cognitive factors are very important for individual development as well as for success in life and well-being, and thus have an impact on individuals and society alike (Almlund et al., 2011[14]; Heckman, Stixrud and Urzua, 2006[15]).

Given the increasing importance of non-cognitive outcomes, PISA complements the assessment of cognitive, learning-related behaviour (e.g. self-regulation, strategies and invested time) with non-cognitive and metacognitive outcomes, such as attitudes, beliefs, motivation and aspirations, as measured primarily in the student questionnaire (but also in the school questionnaire). These outcomes may be of a general nature, such as the achievement motivation and well-being of students and the drop-out rates of schools, or related to the domains of the cognitive assessment, such as reading engagement, interest in mathematics or enjoyment of science. Domain-specific non-cognitive outcomes are also mentioned in the respective definitions of literacy, so this array of constructs serves as a link between the cognitive frameworks and the questionnaire framework.

Student background

In order to understand learning outcomes, educational trajectories and equity issues within and across countries, one must take into account family background variables, such as socio-economic status (SES) and ethnic background.

PISA has become well known for its detailed, theory-based assessment of family background, SES and immigration background. Much effort has gone into the definition and operationalisation of individual student background indicators, leading to the establishment of a composite indicator for economic, social and cultural status, known as the ESCS (Willms, 2006[16]). The components of this indicator need to be assessed in as stable a way as possible across PISA cycles.

Furthermore, PISA gathers retrospective and prospective information about educational pathways. In recent years, researchers and the public debate in many countries have stressed the importance of early childhood education (Blau and Currie, 2006[17]; Cunha et al., 2006[18]). Therefore, PISA intends to collect at least some information on students' participation in primary and pre-primary education, bearing in mind that, for the most part, this would be solicited from 15-year-olds, which could challenge the validity of the reports.

Beyond individual student background, the social, ethnic and academic composition of the school has an impact on students' learning processes and outcomes. Therefore, PISA aggregates student data at the school level to characterise schools' background factors. These are used in combination with structural factors, such as the location and size of a school.

Teaching and learning

School-based instruction is the core process of formal education. Therefore, policy makers need information on the teaching and learning that takes place in schools. To increase the explanatory power of the study, the assessment of teaching and learning will focus on the *major domain of assessment*, which in 2018 is reading, as well as on the innovative domain for 2018, global competence. Research on education effectiveness identifies the following core factors as possibly affecting students' reading literacy: teachers' qualifications, teaching practices, classroom climate, learning time and learning opportunities provided both within and outside of school (Creemers and Kyriakides, 2008[19]; Scheerens and Bosker, 1997[20]). Teaching practices are comprised of three basic dimensions (Klieme, Pauli and Reusser, 2009[21]): (i) structure and classroom management; (ii) teacher support; and (iii) cognitive challenge.

One challenge in addressing teacher- and teaching-related factors in PISA is that sampling is performed by age rather than by grade or class. Another challenge is linked to the reading domain itself. When students are 15 years old, reading is no longer taught as a standalone subject in the same way that mathematics and science are. However, reading literacy is still improved by teaching practices, and reading strategies are taught or learned through not only language arts and literature courses in the test language, but also through foreign language courses and social and natural science courses, known in their entirety as "content literacy" (McKenna and Robinson, 1990[22]; Shanahan and Shanahan, 2008[23]). While questions about teaching and learning mathematics and science can be, to a great extent, limited to solely mathematics and science lessons, there is clear evidence that rich and valuable information about reading (especially online reading) cannot be obtained solely from test language instruction lessons. Indeed, one of the most striking differences between countries in their reading curriculum is their emphasis on and time dedicated to content literacy, including the teaching reading in other subjects, (Lafontaine et al., 2017[24]). Consequently, any teacher questionnaire implemented in PISA 2018 investigating the teaching of reading literacy should be administered to a sample of teachers across domains, rather than only to test language teachers.

School policies and governance

Policy makers have only a limited direct impact on teaching and learning processes. Instead, much of their impact takes place via their influence on school-level factors that directly affect schools, and thus indirectly affect student learning. It is hence important to gather information on these school-level factors. As with teacher and teaching variables (Barile et al., 2012[25]), research has shown that "essential supports" at the level of the school promote school effectiveness (Bryk et al., 2010[26]; Chapman et al., 2011[27]). These essential supports comprise professional capacity with a focus on professional development; a well-organised curriculum; leadership and school management; parental involvement; an ambitious but nurturing school climate (clear norms and shared values, high achievement expectations, truthful and mutually supportive interactions between stakeholders); and the use of assessment and evaluation for improvement. These factors will be addressed within the PISA questionnaires as cross-domain processes on the school level. In addition, the questionnaires will cover school-level support for teaching the major domain, such as the provision of libraries, ICT equipment and a school curriculum for reading literacy, including multimodal aspects of reading in a digital era.

The PISA 2018 questionnaires also need to address issues related to governance at the system level (Hanushek and Woessmann, 2010[28]; Woessmann et al., 2007[29]). Student allocation, selection and evaluation are the basic processes that policy makers and/or school administrators use to control school quality and to monitor and foster school improvement. Some of this information can be obtained from other sources (as documented in OECD's *Education at a Glance*), while other information can be assessed through the PISA school questionnaire.

Previous use of contextual data from PISA: Measures that have been and will be important for analysis and reporting

Comparability and trend items

In previous cycles, statistical analyses – in particular, exploratory and confirmatory factor analyses – were conducted to address whether identical underlying constructs were measured across all participating countries and whether scores could be compared across countries (OECD, 2012[30]; OECD, 2014[31]). In other words, these analyses attempted to determine whether constructs and scores were invariant across countries. Invariance issues will continue to be important issue in the analysis of questionnaire data.

An important asset of the PISA study is its use of trend items, or those that have been used in at least one previous round. As many trend items as possible will be retained in the 2018 questionnaire to compare 2018 data with those from previous rounds and to thereby conduct trend analyses.

The PISA 2009 report

The OECD combines PISA data from both the cognitive assessments and the questionnaires to create its various reports, and the use of this data in previous cycles helped decide which variables should be included in the 2018 study. This section describes the background variables used in the PISA 2009 Initial Report, so chosen as it was the last cycle during which reading was the major domain (as it is in 2018). In addition to student achievement, non-cognitive outcomes, such as student engagement, cognitive strategies and metacognitive strategies, were studied in detail, and the impact of background variables and classroom-, school- and system-level factors was reported. Most of these were gathered through the student and school questionnaires. In more detail, the six volumes describing the PISA 2009 results used the following questionnaire data:

Volume I: What Students Know and Can Do: Student Performance in Reading, Mathematics and Science

- Student background: gender

Volume II: Overcoming Social Background: Equity in Learning Opportunities and Outcomes

- Student background: economic, social and cultural status (ESCS), gender, immigration status, language spoken at home, age of arrival, country of origin
- Support for students assessed through parent questionnaire: parental support (at the beginning of primary education and at age 15), pre-primary education (attendance, quality)

Volume III: Learning to Learn: Student Engagement, Strategies, and Practices

- Student background: ESCS, gender, immigration status, language spoken at home
- Outcomes: enjoyment of reading, time and material used for reading, metacognition (awareness of strategies), self-reported use of reading strategies (memorisation, elaboration, control)

Volume IV: What Makes a School Successful? Resources, Policies, and Practices

- Student background: socio-economic status

- Student-reported factors: learning time (previous education, learning time at school, enrichment/remedial education, after-school lessons), teacher-student relationships, disciplinary climate, teacher's stimulation of reading engagement
- School- and system-level factors (as reported by the principal): type of school (public/private), number of programmes, class size, educational resources (e.g. ICT, library), school responsibility for assessment and curriculum and for resource allocation, extra-curricular activities available, age of school entry, grade repetition, school admittance/grouping/transfer policies, assessment practices/purposes, use of achievement data, school accountability, methods for monitoring teachers, teacher and student behaviour, parent involvement and expectations, leadership, school climate

Volume V: Learning Trends: Changes in Student Performance since 2000

- Trends in student background variables: socio-economic status, immigration status, language spoken at home
- Trends in non-cognitive outcomes and schooling constructs: reading attitudes and practices (reading for pleasure, diversity of texts read, reading engagement, reading fiction), school climate indicators (teachers-student relationships, disciplinary climate)

Volume VI: Students On Line. Digital Technologies and Performance.

- ICT familiarity (optional questionnaire): access to ICT at home and at school, use of ICT at home and at school, students' attitudes towards and self-confidence in using computers, self-confidence in performing ICT tasks and activities, navigation indices extracted from log-file data (number of pages visited, number of relevant pages visited)

In PISA 2000, in addition to the main international report, an in-depth thematic report was dedicated to reading (Kirsch et al., 2002[32]).

As will be outlined below in more detail, most measures that were described in the PISA 2009 Initial Report are included among the 2018 instruments, thereby ensuring the opportunity to compare findings between 2009 and 2018.

Research publications

Numerous scientific research papers using PISA data can be found in the literature. Many papers discuss non-cognitive, domain-specific outcomes: re-scaling the questionnaire items, studying the structure of indices based on questionnaire items within and across countries, analysing outcomes across subgroups and across countries, examining the impact of student and family background and identifying and explaining school-level factors.

Coverage of policy issues in PISA 2018

A balance needed to be struck between the need for trend items and the need for new or changed constructs in PISA 2018. Where possible and sensible, constructs and modules were carried forward intact or with only minor changes. If measures were outdated, redundant, or did not comply with psychometric criteria (e.g. due to low internal consistency), they were recommended for deletion. Finally, two types of constructs were added: (1) those found throughout the research literature that have not previously been covered and (2) those relevant to the new domain of global competence. (Existing constructs were also extended to cover global competence or the new dimensions in the reading framework – specifically, online reading).

This section is divided into four subsections that group the modules into the larger constructs of: (1) assessing non-cognitive and metacognitive constructs; (2) assessing student background; (3) assessing teaching and learning processes; and (4) assessing school policies and governance. The subsection on assessing non-cognitive and metacognitive constructs contains the largest number of changes for 2018, as it includes the new domain of global competence in addition to reading-specific variables.

Assessing non-cognitive and metacognitive constructs

This subsection summarises the conceptual foundations for modules 4 (reading-related outcomes: attitudes, motivation, attitudes and strategies), 9 (dispositional and school-focused variables), and 10 (dispositions for global competence, Figure 7.1).

PISA has traditionally considered only the results from cognitive achievement tests as student outcomes. Students' motivations, attitudes, beliefs and behaviours were seen to be important precursors and predictors of scholastic performance, educational attainment and labour market success. However, non-cognitive outcomes are increasingly considered to be important in their own right from the standpoint of both educational policy and labour market policy, because they are instrumental for personal growth, individual success, long-term achievement and society as a whole (Marsh et al., 2006[33]). Research has demonstrated the considerable power of non-cognitive outcomes for success in secondary education, higher education and the workforce (Heckman, Stixrud and Urzua, 2006[15]; Lindqvist and Vestman, 2011[34]; Poropat, 2009[35]; Richardson, Abraham and Bond, 2012[36]; Shiner et al., 2007[37]). By collecting information on non-cognitive outcomes, PISA can investigate the complex relationships (e.g. moderation or mediation) between non-cognitive outcomes and achievement at the individual, school and country levels.

Previous PISA cycles have focused on domain-specific student attitudes and behaviours, such as interest in and motivation towards reading and mathematics; mathematics self-concept and mathematics anxiety; or knowledge of reading strategies (metacognition). Most of these attitudes and behaviours display robust relationships with student proficiency scores. Domain-specific student attitudes and behaviours are once again included in PISA, particularly, in module 4 (reading-related outcomes). In addition, the current framework includes a broader set of non-cognitive student factors, which will increase the policy relevance of the PISA 2018 database.

The questions in this section cover students' achievement-relevant dispositions, school-focused variables, and reading- and global competence-specific variables. The questionnaire thus adopts a hierarchical approach, investigating constructs that are specific to a domain and then constructs that cut across domains in order to understand and explain student achievement, engagement and behaviour (Elliot and Thrash, 2001[38]).

Reading-related outcomes (module 4)

Reading motivation, engagement and practices

Reading motivation, engagement and practices have been shown to be strongly linked with reading proficiency (Becker, McElvany and Kortenbruck, 2010[39]; Guthrie et al., 1999[40]; Klauda and Guthrie, 2015[41]; Mol and Bus, 2011[42]; Morgan and Fuchs, 2007[43]; Pfost, Dörfler and Artelt, 2013[44]; Schaffner, Philipp and Schiefele, 2016[45]; Schiefele et al., 2012[46]). In PISA 2000 and 2009, reading engagement (i.e. interest, intrinsic motivation, avoidance and practices) was a factor with one of the strongest relationships with reading proficiency; indeed, reading engagement was more strongly associated with reading proficiency than socio-economic status (OECD, 2010[47]; Kirsch et al., 2002[32]). It was shown that a high level of engagement compensated, to some extent, for a poor socio-economic background. It was also found that if boys were equally engaged in reading as girls, the gender gap would be reduced by two thirds (OECD, 2010[48]). In other studies, reading engagement has been shown to explain reading achievement more than any other variable besides previous reading achievement (Guthrie and Wigfield, 2000[49]). Thus, motivation and engagement are powerful variables, and are therefore possible levers on which one can act in order to enhance reading proficiency and reduce gaps between groups of students.

In the past, the components of motivation that PISA mainly targeted, particularly when reading was the major domain, were interest and intrinsic motivation. Other motivational constructs, such as self-efficacy and self-concept, were investigated when mathematics and science were the major domain. Here, self-efficacy (Bandura, 1997[50]; Ferla, Valcke and Cai, 2009[51]) refers to an individual's perceived capacity of doing specific tasks while self-concept is a general measure of an individual's own perceived abilities in a domain such as reading, mathematics or science, (Marsh and Craven, 1997[52]). Positive self-concept and self-efficacy are highly related to motivation, learning behaviour, general expectations for the future and performance (Baker and Wigfield, 1999[53]; Marsh and Craven, 2006[54]; Morgan and Fuchs, 2007[43]; Retelsdorf, Köller and Möller, 2011[55]; Solheim, 2011[56]; OECD, 2007[57]; OECD, 2016[58]) OECD, 2007. As a result, these constructs are measured in the PISA 2018 student questionnaire. Following Chapman and Tunmer's recommendations (1995[59]), the instruments should cover not only students' perceptions of their own *competence* in reading, but also perceptions of their *difficulty* with reading. Indeed, Klauda and Guthrie (2015[41]) have provided evidence that perceived difficulty with reading is a stronger predictor of reading achievement than self-efficacy. Similarly, they confirmed that students' avoidance of reading and their devaluation of reading (the belief that reading is not useful) are negatively correlated with growth in engagement and motivation indicators among grade 7 pupils, even if students show positive self-efficacy and engagement in reading tasks.

PISA has now made the identification of students who perceive themselves to be struggling readers a higher priority through emphasizing the importance of basic components of reading such as fluency, and through extending the lower end of the reading scale since PISA 2009.

Metacognition

Like engagement, metacognition is significantly correlated with reading proficiency and is responsive to teaching and learning (Artelt, Schiefele and Schneider, 2001[60]; Brown, Palincsar and Armbruster, 2004[61]). The prominent metacognitive reading strategies include setting reading goals, adapting one's reading strategies depending on these goals, knowing how to summarise a piece of text or remember essential information, monitoring comprehension and knowing how to repair comprehension problems. The new PISA 2018 reading literacy framework acknowledges the paramount importance played by these reading task management processes. They are now an integral part of the model of reading processing that organises the reading literacy framework (see Figure 2.2 in the reading literacy framework).

Explicit or formal instruction of metacognitive strategies leads to an improvement in understanding text and using information (National Reading Panel, 2000[62]). That is, when readers are given cognitive and metacognitive strategy instruction, they make more significant gains on measures of reading comprehension than students only trained with conventional instruction procedures (Baker and Carter-Beall, 2009[63]; Dole, Nokes and Drits, 2009[64]; Pressley, Graham and Harris, 2006[65]; Pressley et al., 1989[66]; Rosenshine and Meister, 1994[67]; Rosenshine, Meister and Chapman, 1996[68]; Waters and Schneider, 2010[69]).

PISA 2009 assessed students' metacognitive strategies by asking them how useful they thought two reading strategies – summarizing a piece of text and understanding and memorizing a piece of text – were in order to solve a reading task. Correlations of these two metacognitive strategies with reading performance were robust, with median correlations across OECD countries of, respectively, 0.46 and 0.39 (Artelt and Schneider, 2015[70]).

The growing importance of digital reading literacy in PISA 2018 makes the need to assess metacognition even more important. Coiro and Dobler (2007[7]) pointed out that in online reading, efficient and specific self-regulated strategies (such as selecting the most relevant links and pathways and avoiding distracting information) are crucial to facilitating reading goals and plans. A new questionnaire item will focus on another important aspect of online reading, namely assessing the quality and the credibility of sources.

Dispositional and school-focused variables (module 9)

Complementing the reading-related outcomes are dispositions towards achievement and school-focused variables common across domains. Dispositional variables are the personality-based context in which students approach or avoid learning; they are the result of a lifetime of socialisation from parents, teachers, coaches and one's cultural surroundings, and they capture how behaviour is energised over time. School-focused variables, or how students view and approach school, are influenced by both students' disposition and the particular situation in which they find themselves. These dispositional and school-focused variables are the best predictors of both the aforementioned domain-specific variables (e.g. reading-related outcomes) and achievement outcomes. Together, dispositional and school-focused variables provide important information on students' attitudes towards learning and achievement.

The following paragraphs provide an overview of some of the target dispositional and school-focused variables. These variables are relevant to all domains and focus on the non-cognitive components important to learning. The dispositional variables include the achievement motives of competitiveness, work mastery, and the fear of failure; incremental mind-set; perseverance; subjective well-being; and ICT motivation and practices. The school-focused variables include learning beliefs and attitudes towards school, and achievement goals.

Dispositional variables

Achievement motives - Competitiveness, Work mastery, and Fear of Failure: *Achieving motivation*[1], as assessed in the 2015 field trial, represented a combination of competitiveness and extrinsic motivation. However, the questionnaire expert group has replaced the construct of *achieving motivation* in PISA 2018 with the constructs of *competitiveness* and *work mastery*, which is how achievement motivation theorists have conceptualized their central construct over the past four decades (Helmreich et al., 1978[71];

Elliot and McGregor, 2001[72]). Here, competitiveness is defined as the dispositional desire to outperform others, while work mastery is defined as the dispositional desire to work hard to master tasks. Research shows that these two components of approach-oriented achievement motivation are linked to different sets of antecedents and consequences; so, when assessing achievement motivation, it is important to measure these constructs separately (Baranik, Barron and Finney, 2007[73]; Murayama and Elliot, 2012[74]; Spence and Helmreich, 1983[75]).

The questionnaire expert group has replaced the 2015 construct of general *test anxiety* with the construct of *fear of failure*. Test anxiety is worry about potential failure at the task- or domain-specific level of analysis (Hembree, 1988[76]). Fear of failure, however, is the more general tendency to self-protectively avoid potential mistakes and failures because they are experienced as shameful, which may be more predictive of cognitive achievement in real-life situations than test anxiety. Research has shown that fear of failure leads students to be self-protective and to avoid challenging situations and opportunities that are essential for learning and development (Covington, 1992[77]; Heckhausen, 1975[78]; Kaye, Conroy and Fifer, 2008[79]).

The optimal learner is high in work mastery and low in fear of failure. Competitiveness alone can be problematic, but the confluence of high competitiveness and work mastery appears to be beneficial. The positive and negative implications of competitiveness is a hot topic in the achievement literature, and data on this variable both within and across countries should prove valuable and garner considerable attention.

Incremental mind-set: Students with an incremental mind-set believe that ability is changeable rather than fixed, which is another core characteristic of an optimal learner. Having an incremental mind-set is related to perseverance, and has been found to be positively correlated with work mastery while negatively correlated with performance avoidance. This mind-set has also been shown to be correlated with both persistence in the face of failure and performance attainment (Aronson, Fried and Good, 2002[80]; Blackwell, Trzesniewski and Dweck, 2007[81]; Dweck, 2007[82]).

Perseverance: This construct was included in the 2012 student background questionnaire but was not included in the 2015 wave because of time constraints. Despite the six-year gap, examining perseverance is still valuable given that it has been shown to be an important predictor of achievement (Duckworth et al., 2007[83]; Ţuţu and Constantin, 2012[84]). Many different labels are used in the current literature for this construct, including "persistence" and "grit". The optimal learner is high in perseverance.

Subjective well-being: Subjective well-being can be defined as "good mental states, including all of the various evaluations, positive and negative, that people make of their lives and the affective reactions of people to their experiences" (OECD, 2013, p. 10[3]; OECD, 2011[85]; OECD, 2013[86]; OECD, 2017[87]). This definition encompasses three elements: *life evaluation* – one's reflective assessment of one's life (including the "general life satisfaction" question: "Overall, how satisfied are you with your life as a whole these days?"); *affect* – one's emotional state, typically at a particular point of time; and *eudaemonia* – a sense of meaning and purpose in life, which can increase one's sense of belonging. The recent growing interest from researchers and policy makers in this construct has resulted in recommendations to statistical agencies to "incorporate questions on subjective well-being in their standard surveys to capture people's life evaluations, hedonic experiences and life priorities" (Stiglitz, Sen and Fitoussi, 2009, p. 216[88]). The OECD (OECD, 2013[86]) has responded to this charge in providing guidelines on measuring subjective well-being. The QEG has included information on all three elements, *life* *evaluation*, *affect*, and *eudaemonia*, for PISA 2018. The optimal learner has a positive life evaluation, frequent positive affect and infrequent negative affect, and strong eudaemonia.

ICT motivation and practices: Module 9 also covers information and communications technology (ICT). ICT-related behaviours and motivational attributes can be regarded as domain-general student outcomes because ICT plays a role across all educational domains. Following the OECD's DeSeCo project and the 21st Century Skills Initiative, the optimal learner exhibits general skills related to information, media and technology above and beyond the traditional core subjects (OECD, 2005[89]; Partnership for 21st Century Skills, 2008[90]). The ICT familiarity questionnaire of PISA 2018 will assess students' interest in ICT, use of ICT, perceived competence and autonomy in using ICT, and the use of social media.

School-focused variables

Learning beliefs and attitudes towards school: Beliefs about one's own success or failure in school learning have been shown to be strong predictors for further effort and success, including for test scores in student assessments (Opdenakker and Van Damme, 2000[91]; Rumberger and Palardy, 2005[92]). PISA 2018 investigates several factors last examined in 2012, including students' school self-efficacy, their evaluation of their experience in school, and their attitudes toward school. The optimal learner has strong school self-efficacy and a positive inclination toward school.

Achievement goals: One important characteristic of optimal learners is that they are focused on improvement in the classroom; in other words, they pursue mastery-approach goals. Students who adopt mastery-approach goals have been shown to engage in deep learning, to persist upon failure and to show high levels of intrinsic motivation (Hulleman et al., 2010[93]; Kaplan and Maehr, 2007[94]; Middleton and Perks, 2014[95]). An equally important characteristic of an optimal learner is that he/she does *not* strive to avoid performing worse than other students, or in other words, he/she does not pursue performance-avoidance goals. Students who adopt such goals have been shown to engage in shallow learning, to give up in the face of failure, and to display low levels of both performance and intrinsic motivation (Hulleman et al., 2010[93]; Rawsthorne and Elliot, 1999[96]; Van Yperen, Blaga and Postmes, 2014[97]). In short, achievement goals – both mastery-approach and performance-avoiding goals – are key predictors of the two central outcomes indicative of sustainable student success: performance attainment (which shows that short-term learning has taken place) and intrinsic motivation (which shows that the motivation for continued, long-term learning is in place). Focusing on both the presence of the positive (mastery-approach goals) and the absence of the negative (performance-avoidance goals) is important, as both are essential for optimal learning to take place (Elliot, 2005[98]).

Dispositions for global competence (module 10)

Global competence is the new domain in PISA 2018. This domain is critically important because learning, working and living environments are becoming more global, interconnected and interdependent. Young people will encounter, actively engage with and help shape those environments during their lifetime no matter where they are born, educated, work or live. It is therefore important that students leave school equipped with the necessary knowledge, skills and attitudes that will enable them to learn, work and live in a globalised world.

In PISA 2018, global competence is defined as the capability and disposition to act and interact appropriately and effectively, both individually and collaboratively, when participating in an interconnected, interdependent and diverse world. The domain of global competence is comprised of the following four dimensions:

- **Communication and relationship management** refers to the willingness and capability to adapt one's communication and behaviour in order to interact appropriately and effectively with others holding diverse perspectives and in different contexts.
- **Knowledge of**[2] **and interest in global developments, challenges and trends** refers to a learner's interest in and knowledge of cultures, major issues, events and phenomena in the world, as well as the learner's ability to understand their global significance and their implications for adapting appropriately and effectively to learning, working, and living with others holding diverse perspectives and in different contexts.
- **Openness and flexibility** refers to being receptive to and understanding of new ideas, people and situations, as well as different perspectives and practices. It also refers to the ability to seek out and understand new and different perspectives and experiences and to appropriately and effectively adapt one's thinking, behaviour and actions to learning, working and living situations that involve others holding diverse perspectives and in different contexts.
- **Emotional strength and resilience** refers to the ability to deal appropriately with the ambiguity, changes, and challenges that different perspectives and experiences can present and to have the resilience to maintain one's identity and/or to develop personally despite or as a result of encountering different perspectives and experiences.

Generally, the questionnaire items related to global competence (construct-related and contextual information to inform the interpretation of outcomes) will focus on the two dimensions of *openness and flexibility* and *emotional strength and resilience*. Approximately ten questions will be asked for each of the following four scales: openness, flexibility, emotional strength and resilience. The questions for each scale will primarily be behaviour-based (e.g. "When I meet people who are different from me, I am interested in learning more about them"; "I like to eat in a variety of ethnic restaurants"), will be located in simple contexts (e.g. "When I travel abroad …"; "When I am at school …") and will be presented in a consistent format whenever possible (e.g. using a Likert scale).

Topics for the global competence questionnaire items could include (most measures comprise both construct and contextual components):

- Languages:
  - Languages spoken at home/understood by the student and by his/her mother/father/siblings
  - Languages taught in school; number of languages spoken/understood by the student
  - Number of languages to which the student has access or with which the student has experience (actual, virtual)
- Global developments/challenges/trends:

  - The student's engagement with others about global events/issues (e.g. on line chat);
  - Extent of exposure to/awareness of global developments/challenges/trends (e.g. via the news or other media)
- Migration/movement of student:
  - The student's own background; the student's experience living abroad
  - The possibility of the student working/studying abroad in the future
  - Demographics in his/her community
- Student interaction with or exposure to people from other countries/cultures; student's travel experience
- Student's degree of curiosity/motivation to travel

Existing questionnaire items (i.e. from the school and teacher questionnaires) could also be used or extended to inform the measures of global competence. For example:

- School climate:
  - Student demographics (at the school level)
  - School philosophy/values/policies related to global competence
  - Support given to students for language and/or culturally-based reasons (e.g. such items could be revised to focus on global competence-related challenges)
- Global competence in the curriculum
- Professional development: how/the degree to which teachers are prepared to manage multiculturalism and/or facilitate global competence (e.g. global trends, international events)
- The availability of resources related to global competence

Finally, the following items could be added to the teacher questionnaire in order to measure global competence:

- Teacher demographics/background (e.g. cultures represented)
- Number of teachers who have taught in another country
- Number of teachers who hold a certification from another country
- Exposure/experience/engagement with international and/or global events
- ICT literacy
- Awareness of and support for school policies that relate to/support global competence
- Teacher practices related to global competence, including cross-curricular practices

Most existing measures in the area of global competence have been developed for older learners (e.g. college students or adult employees) rather than for the 15-year-old students assessed in PISA 2018. Nonetheless, various measures for 15-year-old students can be derived from the extant literature (Deardorff, 2009[99]).

Assessing student background

This subsection covers module 6 (student SES, family and home background), module 7 (ethnicity and migration), and module 8 (educational pathways in early childhood). These topics require careful revisiting every cycle because they contain the basic information needed to calculate the index of economic, social and cultural status (ESCS), the proxy that PISA uses for socio-economic status.

Student SES, family and home background (module 6): PISA 2018 keeps questions regarding socio-economic status and other background variables basically unchanged in order to be able to observe trends in social, cultural and economic indicators. However, some minor changes have become necessary. For example, extensive developments in ICT mean that items that were once only sometimes found in students' homes (e.g. laptops or tablets) are now commonplace and are therefore less discriminatory as a marker of socio-economic status. Thus, the measures of home possessions will be updated to ensure better coverage of within- and cross-country variation of home possessions. These changes are expected not to have an impact on the important trend measures in this module.

Migration and culture (module 7): Many nations are home to several subpopulations with different languages and cultures. International migration perpetuates this diversity. On average across OECD countries, first and second generation immigrant students composed 12.5% of the student population in 2015, up from 9.4% in 2006 (OECD, 2016[58]). At the same time, students from ethnic minority groups and immigrant students often face particular challenges. In a number of education systems, immigrant students perform at significantly lower levels than their native peers in key school subjects (Stanat and Christensen, 2006[100]; OECD, 2016[58]), and both groups are often faced with overt or covert discrimination with potentially detrimental consequences for their psychological development and well-being. Thus, providing students from different linguistic and cultural backgrounds with equal opportunities is often considered one of the central challenges for education systems in the 21st century (OECD, 2010[47]).

PISA 2015 put a special focus on diversity-related aspects of the school climate. A new question developed for the PISA 2015 field trial asked students about their membership in a group that they believed to be discriminated against in their country. If they identified themselves as belonging to such a group, they were then asked whether they felt treated in a respectful and fair manner by their teachers and equal to their classmates, a factor that has been shown to be related to educational outcomes (Fisher, Wallace and Fenton, 2000[101]; Wong, Eccles and Sameroff, 2003[102]). Another new question, implemented in the optional parent questionnaire, assessed perceived barriers to parental involvement. Additionally, teachers and principals are asked about diversity-related assumptions among teachers in their school. The wording of the question is based on research on how multiculturalism is supported (van de Vijver, Breugelmans and Schalk-Soekar, 2008[103]). Additionally, PISA 2015 examined aspects of multicultural education practices and the extent to which multicultural educational practices were implemented in different schools. Altogether, findings from this module may help researchers better understand educational inequalities and can suggest ways to address these inequalities. However, these items have proved problematic and will need to be reworked if they are to be successful in measuring this important concept in 2018. Given the relevance of this module for global competence, such a reworking (presumably with more emphasis on opportunities for intercultural encounters) is worthwhile.

Educational pathways in early childhood (module 8): Children already have varying levels of ability in their language, pre-reading and early numeracy skills by the time they enter primary school, and these differences are often maintained throughout life. Promoting school readiness and better adjustment to school is hypothesised to be an efficient means of raising the achievement levels of all children, but especially of those children who experience a lack of parental support or who grow up in disadvantaged circumstances. It has been argued that investing in early education programmes will have large long-term monetary and non-monetary benefits (Heckman, 2006[104]).

According to UNESCO (2006[105]), Early Childhood Care and Education (ECCE) programmes are "programmes that, in addition to providing children with care, offer a structured and purposeful set of learning activities either in a formal institution (pre-primary or ISCED 0) or as part of a non-formal child development programme" (p. 348). PISA will also use this definition, as opposed to inquiring only about students' experiences in ISCED 0; most literature on early childcare – such as brain research, studies on domain-specific development and support, evaluation studies of model programmes, and longitudinal large-scale studies – do the same.

Yet not all ECCE programmes lead to long-term benefits. For example, the British EPPE study found short-term effects showing that pre-school attendance was beneficial for cognitive and socio-emotional development, in particular for children from disadvantaged backgrounds. However, in the long term, only those children who attended a high-quality pre-school centre showed lasting beneficial effects (Sammons et al., 2009[106]; Sylva et al., 2011[107]); cf. also (Valenti and Tracey, 2009[108]). A certain degree of intensity in terms of hours per week/months also seems to be a precondition for the beneficial effects of ECCE attendance (Logan et al., 2011[109]; Sylva et al., 2011[110]).

Thus, asking about early education experience in PISA only makes sense if specific aspects of quantity, quality and curriculum can be retrieved retrospectively, which is highly unlikely when the questions are asked of students (Fivush and Hudson, 1990[111]; Markowitsch and Welzer, 2009[112]). As a consequence, PISA 2018, whilst keeping a short question on ISCED 0 attendance in the student questionnaire, will continue to administer a series of questions on students' ECCE attendance in the parent questionnaire as parents are expected to be a more reliable source of information. Those countries administering the optional parent questionnaire will thus gain information on basic characteristics of the ECCE arrangements in their countries and the reasons for attending or not attending ECCE.

Assessing teaching and learning processes

This subsection summarises the conceptual foundations for module 1 (teacher qualifications and professional development), module 2 (teaching practices for reading), module 5 (out-of-school reading experience), and module 11 (learning time and curriculum).

Teaching and learning are at the heart of education. Most cognitive and non-cognitive, curricular and extra-curricular goals of school education are achieved or impeded by the way students and teachers interact in classrooms. Whilst teaching is the core process in schools, the curriculum determines its content and professional teachers are the force who implement the curriculum, orchestrate learning activities and thus arrange for quality learning time.

There is ample evidence that teaching and learning activities are very good predictors of student ability and PISA therefore needs to examine these activities if it is to inform educational policy making at the system and the school level. Module 2 describes reading education by broad sets of teaching and learning activities. In addition, this module investigates general dimensions of teaching quality, such as the structure of instruction, classroom management and support, and cognitive activation, as they are applied in reading education. Next, module 11 covers learning time, including non-mandatory, additional instruction within and out of school, as well as the coherence, focus and rigour of the reading curriculum. Module 5 examines students' engagement with reading outside of school. Finally, the initial education, beliefs and professional development of the teaching force will be described in module 1.

Teaching practices for reading (module 2) and learning time and curriculum (module 11)

Teaching practices and classroom support for reading growth and engagement in PISA

Research on reading shows that classroom practices, such as the direct teaching of reading strategies (Pressley, 2000[113]; Rosenshine and Meister, 1997[114]; Waters and Schneider, 2010[69]) and teachers' scaffolding and support for autonomy, competence and ownership (Guthrie, Klauda and Ho, 2013[115]; Guthrie, Wigfield and You, 2012[116]) are powerful ways of improving students' reading proficiency, awareness of strategies (metacognition) and engagement (motivation) in reading.

Hence, it is important that the 2018 questionnaire address teaching practices around reading literacy. Two broad theories inform the selection of the most relevant constructs:

- *Practices that support reading engagement and motivation* relate to self-determination theory (Deci and Ryan, 1985[117]; Reeve, 2012[118]; Vansteenkiste, Lens and Deci, 2006[119]). The pathway to students' self-determination in reading depends on support from "significant others" in their lives. A favourable family and school context leads to greater confidence in one's reading ability and expanded autonomy in directing one's own reading activities. This, in turn, results in intrinsically motivated and self-determined reading and finally to greater reading proficiency. The teacher is a "significant other" for reading literacy. Numerous studies show that teachers who improve students' sense of ownership and competence enable them to become active and competent readers. By contrast, teachers who neglect these instructional practices impede students' efforts to become autonomous, resulting in students who disengage from reading and fail to progress in reading achievement (Guthrie, 2008[120]).
- *Practices that enhance reading skills and metacognitive strategies* are based on direct instruction of reading strategies (Pressley, 2000[113]), and also correspond to the "cognitive activation" facet of the tridimensional model of quality of teaching (Klieme, Pauli and Reusser, 2009[21]), namely the model on which the selection of teaching constructs for PISA 2012 and 2015 has been built. According to this model, classroom management processes, teacher support and cognitive activation independently predict growth in students' mathematics, science and language skills; teacher support is the main predictor of motivation growth; and cognitive activation is the main predictor of cognitive gains at different grades (Klieme, Steinert and Hochweber, 2010[121]).

However, when students are 15 years old, reading is no longer taught as a subject matter in the same way that mathematics and science are. It is therefore a challenge to tailor questions capturing the classroom practices that students are exposed to and their opportunities to develop their reading skills. Reading is indeed not a part of test language lessons, but also of, for example, social science, natural science, foreign language, civic education and ICT lessons. There is even more uncertainty in where the new skills and processes related to digital reading are taught: they are learned mostly in non-formal contexts outside of school in some countries, taught in specific courses in others, and considered to be "transversal" and not taught at all in specific courses but incorporated throughout the curriculum in yet other countries.

There is an ongoing debate in the field of reading research between scholars supporting the view that "every teacher is a reading teacher" ("generic" or "intermediate" reading literacy) and advocates of the "content" or "disciplinary" literacy view, who argue that reading texts dealing with, for example, mathematics, chemistry or history require distinct reading skills linked to the domain that should be taught by content matter teachers (Shanahan and Shanahan, 2008[23]). Taking into account time limitations and the fact that teaching practices were not considered a priority for PISA 2018, the questionnaire will ask the students only a small number of questions about their experience and exposure to "generic" or "intermediate" literacy teaching practices in their classes in general, as took place in PISA 2009 (Lafontaine et al., 2015[122]).

Out-of-school experience (module 5)

In previous PISA cycles when reading was the major domain (2000 and 2009), reading practices were measured as *self-reported frequencies of reading activities with diverse content in various media*. The initial list of possible content included fiction and non-fiction books, comics, newspapers, magazines and e-mails. In 2009, new items about online reading practices were included. Some of these new items focused on social online practices (blogs, forums, e-mails), while others focused on searching for information (searching for news online, for information in order to learn about a topic, or for practical information). While students' online practices related to searching for information explained a significant, yet small, proportion of the variance of their online digital reading abilities, their social practices on line had no significant relationship with digital reading in most countries that participated in PISA 2009 (OECD, 2011[123]; Naumann, 2015[124]).

The list of online reading practices will be extended in PISA 2018 to take into account recent and emerging reading media (e.g. e-books, social networks). However, despite the growing importance of reading on digital devices, it should be kept in mind that traditional forms of reading (especially books) are still the most influential means for developing students' reading comprehension and vocabulary (Pfost, Dörfler and Artelt, 2013[44]).

As in the reading cognitive assessment, it is important to keep some of the 2009 items measuring reading practices in order to measure trends. This allowed PISA to say, for instance, that 15-year-olds read less in 2009 than in 2000, with a stronger decline among males. However, the same question might take on different meanings in 2009 and in 2018 because reading itself has changed: the term "book", for instance, might now refer to both printed books and e-books. The reading practices reported by students could change drastically depending on how questions are worded.

Due to the limitations of self-reported measures based on Likert scales (Allen, Cipielewski and Stanovich, 1992[125]; Pfost, Dörfler and Artelt, 2013[44]), alternative ways of capturing reading practices (forced-choice, behavioural or situational questions) will be tested during the field trial. If some of these prove to be more stable, more culturally invariant and better correlates of reading ability, they could replace or complement some of the trend items and potentially become new trend items themselves.

Teacher qualifications and professional development (module 1)

Many studies have demonstrated a clear influence of teacher-related factors on student learning and outcomes, and there has been increased focus in recent years on teacher-related policies. In addition to teachers' professional behaviour within the classroom, the age distribution and educational level of the teaching force, teachers' initial education and qualifications, their individual beliefs and competencies, and their professional practices on the school level (such as collaboration and professional development), and how these factors are related to student outcomes, are core topics in education policy. Basic information on these topics will be available from the PISA 2018 school as well as the optional teacher questionnaire.

Assessing school policies and governance

This subsection summarises the conceptual foundations for module 3 (school-level learning environment for reading) and modules 12-18.

School policies and approaches to educational governance

During the last two decades, research on educational effectiveness has largely been concerned with the impact of school-level factors on students' learning. Studies show that such school factors are related to student progress. It has been asserted that the school environment can influence the behaviour of teachers and students and thus – mostly indirectly – their success in teaching and learning. Both "soft" factors, such as school climate and parental involvement, and "hard" factors, such as school management activities and allocation policies, vary within and across countries and are related to student outcomes.

School-level learning environment for reading (module 3): Conceptually, this module overlaps to a considerable degree with other modules dealing with school-level factors, such as module 11 (learning time and curriculum), module 14 (school context and resources), and module 16 (assessment, evaluation, and accountability). The questionnaire expert group recommends that some questions in the school questionnaire focus directly on the status of reading education in the school and available resources dedicated to reading. Accordingly, a question has been developed to investigate the overall value of reading within the school ("Is reading education a shared priority for stakeholders?"), along with questions on resources available that are primarily dedicated to reading education: the size of the teaching staff, resources such as libraries and digital learning devices, and potentially cooperation with external partners.

School climate (module 12): The school climate encompasses a school's shared norms and values, the quality of the interpersonal relationships within the school, and its general atmosphere. The general consensus about both the mission of the school and the value of education, shared by school leaders, staff and parents, affects the norms of student peer groups and facilitates learning (Opdenakker and Van Damme, 2000[91]; Rumberger and Palardy, 2005[92]). In addition, an orderly learning atmosphere maximises the use of learning time. By contrast, disrespectfulness and an unruly environment are counterproductive for teachers and students alike and distract from the school's educational mission. As in previous PISA assessments, school climate will be assessed in both the student questionnaire (such as through questions on student-teacher relationships and achievement pressure) and the school questionnaire (such as through questions on teacher morale and behaviours that could affect school climate).

A number of items related to student-teacher relationships, as identified by Caldwell and Bradley (1984[126]) in their "home inventory", were added to the PISA 2015 questionnaires: emotional warmth; guidance and boundaries; stimulation/scaffolding and stability. Students report on their relationship with teachers and parents, whilst school principals and parents (in the optional questionnaires) are asked parallel items. All in all, an interesting picture of social relationships between students, school actors and parents could emerge in countries that administer both the parent and the school questionnaires. Finally, two scales in the student questionnaire that cover more problematic, often hidden aspects of school climate should be continued: bullying by peers and unfair treatment by teachers.

Parental involvement (module 13): Over the past years, the involvement of parents in educational processes has gained importance in the educational debate and relevance for educational policy. PISA has collected information on parental involvement in education since 2006, when the parent questionnaire was administered for the first time, directly addressing the parents of PISA students. In PISA 2015, specific aspects of parental involvement were added to all questionnaires, not only the parent questionnaire, focusing on (a) parent-school communication and collaboration and (b) parental support for learning; these scales will be continued for 2018. Although aspects of parental involvement can be found in other modules, the majority of items and topics regarding parental involvement are included in the parent questionnaire that will be administered as an international option.

The parent questionnaire will also ask questions specifically related to reading literacy, in particular regarding parental support during the development of their child's early literacy skills (at ISCED 0), and regarding parents' own interest in and motivation for reading. There is extensive evidence that parental support before the formal instruction of reading at school, such as the joint reading of books or playing language games, is critical to the development of early or "emergent" literacy (Neuman and Dickinson, 2003[127]). However, there is still much to be discovered about the relationship between parental support and adolescents' reading motivation and practices (Klauda, 2009[128]).

School context and resources (module 14): Information on school type (public vs. private) and class size has always been included in the school questionnaire. In addition to these trend questions, the PISA 2015 field trial expanded this module, discriminating between types of private schools (religious/denominational, not-for-profit and for-profit) and eliciting more advanced information on ICT use. All PISA cycles to date have included a question on the degree to which a school experiences problems due to a lack of resources or poor-quality resources. PISA 2018 now inquires about these issues in one set of coherent questions in the school questionnaire.

Allocation, selection, choice and grade repetition (module 15): The way students are channelled into educational pathways, schools, tracks or courses is a core issue of educational governance known as stratification, streaming or tracking. Selection and allocation procedures are also important aspects of school organisation. The learning environment in highly selective schools may differ from that in more comprehensive schools. For all of these reasons, questions on allocation, selection, choice and grade repetition answered by school administrators and parents have been retained from previous PISA questionnaires.

PISA 2015 asked students whether they had ever repeated a grade. Many longitudinal studies have demonstrated grade retention to have a negative relationship with individual careers and outcomes (Ou and Reynolds, 2010[129]; Griffith et al., 2010[130]), student behaviour, and well-being (Crothers et al., 2010[131]). Grade repetition is less common in secondary schools compared to primary schools, but the negative effects of late retention seem to be larger (Ou and Reynolds, 2010[129]). Greene and Winter (2009[132]) showed that once a test-based retention policy had been installed, those who were exempted from the policy performed worse. Babcock and Bedard (2011[133]) showed that a large number of students being retained could have a positive effect on the cohort (i.e. on all students, including those who were promoted). Kloosterman and De Graaf (2010[134]) argued that in highly tracked systems, such as in some European countries, grade repetition might serve as a preferred alternative to moving into a lower track; indeed, they found evidence that this strategy is preferred for low-performing students with higher SES. Thus, changing grade repetition policies might be a viable low-cost intervention (Binder, 2009[135]). Therefore, it is worthwhile to take a closer, comparative look at grade retention policies and their relationship with student outcomes (for both students who do and do not repeat a grade). PISA 2018 will explore grade-repetition related question.

Assessment, evaluation, and accountability (module 16)

Assessing students and evaluating schools is a common practice in most countries (Ozga, 2012[136]). Since the 1980s, policy instruments, such as performance standards, standard-based assessment, annual reports on student progress and school inspections, have been promoted and implemented across education systems. Reporting and sharing data from assessments and evaluations with different stakeholders provides multiple opportunities for monitoring, feedback and improvement. In recent years, there has been a growing interest in the use of assessment and evaluation results for quality management and improvement (OECD, 2010, p. 76[47]); formative assessments, also known as assessment for learning, have been one of the dominant movements in this domain (Baird et al., 2014[137]; Black, 2015[138]; Hattie, 2009[139]). Accountability systems based on these instruments are increasingly common in OECD countries (Rosenkvist, 2010[140]; Scheerens, 2002, p. 36[141]).

Previous PISA cycles have covered aspects of assessment, evaluation and accountability in the school questionnaire by identifying a variety of purposes for the assessment of students. School leaders have been asked: whether they use test results to make comparisons with other schools at the district or national level; or to improve teacher instruction by asking students for written feedback on lessons, teachers or resources. Relevant research on school evaluation and student assessment is summarised below to provide the rationale for questionnaire development in this module in PISA 2018.

Evaluation: The evaluation of schools is used as a means of assuring transparency, making judgements about systems, programmes, educational resources and processes, and guiding school development (Faubert, 2009[142]). In PISA 2018, the term evaluation will be used for processes at the school and system level, as was done in PISA 2015.

Evaluation can be either external or internal (Berkenmeyer and Müller, 2010[143]). In an external evaluation, the process is controlled and headed by an external body and the school does not define the areas that are judged. On the other hand, an internal evaluation is part of a process controlled by the school, where the school defines the areas that are judged; the evaluation may be conducted by members of the school (self-evaluation) or by persons/institutions commissioned by the school. Different evaluation practices generally co-exist and benefit from each other (Ryan, Chandler and Samuels, 2007[144]). External evaluations can expand the scope of internal evaluation and validate the results from an internal evaluation. Internal evaluations can improve the interpretation and increase the utilisation of results from an external evaluation. However, the improvement of schools seems to be more likely when an internal evaluation is undertaken. Moreover, country and school-specific context factors may influence the implementation of evaluation results as well as their effects on schools. In many countries, the individual evaluation of teachers and principals, as separate from a school-wide evaluation, is also common (Faubert, 2009[142]; Santiago and Benavides, 2009[145]); they are treated here as a separate type of evaluation.

Assessment: Many countries have implemented national standards to assess students' learning outcomes. Together with formative assessment practices, these summative assessment systems influence the way teachers teach and students learn. In particular, formative assessment practices can enhance students' achievement (Black and Wiliam, 1998[146]). However, there is large variation in the implementation of formative assessment practices, as reported in recent studies in Canada, Norway, Scotland, Singapore, Sweden and the United States among others (Wylie and Lyon, 2015[147]; DeLuca et al., 2015[148]; Jonsson, Lundahl and Holmgren, 2015[149]; Hayward, 2015[150]; Ratnam-Lim and Tan, 2015[151]; Hopfenbeck and Stobart, 2015[152]). PISA 2018 therefore aims to assess both the formative and summative aspects of student assessment through questions in both the student and school questionnaires.

Dealing with response bias

The analysis of response bias plays an important role in the analysis of PISA data. Two types of approaches have been implemented to handle bias: adapted instruments (using novel constructs or response formats) and advanced statistical modelling. The field trial will enable experimentation with different question formats (e.g. using interactive features of the computer-based administration system), and with new content that could be subject to response bias (e.g., measures of students' physical and emotional well-being). The field trial will assess the psychometric properties of these instruments and whether new measures are can be compared across different education systems. A subset of measures will be selected for the main study based on findings from the field trial.

Well-known examples of design measures to correct for bias include overclaiming, anchoring vignettes and cognitive interviews to examine response styles. Overclaiming is a procedure in which students are asked about their knowledge of a number of concepts, some of which do not exist. Students who indicate a high knowledge of these non-existent concepts show strong response styles, which is somewhat akin to social desirability. Statistical correction for overclaiming had an impact on cross-national differences in scores. For example, correlations between motivation and performance are often computed at the individual and country levels as a test of bias correction procedures. These correlations are usually positive and significant yet small at the individual level, but strong and negative at the country level (Marsh and Hau, 2004[153]). It has been argued that cross-national differences in response styles are responsible for this change in correlation after aggregation. Correcting for overclaiming led to a sizable reduction of the negative correlation between motivation and performance at the country level (Kyllonen and Bertling, 2014[154]).

Anchoring vignettes present descriptions of hypothetical persons, usually with very high, medium and very low levels of a target construct, before assessing the target construct itself (King et al., 2004[155]). Students are asked how they would rate the motivation of these hypothetical persons on a scale. The response on the target item, how the students would rate their own reading motivation, is then re-scaled using the anchor scores of the three hypothetical persons. The results of the use of anchoring vignettes have been mixed. Very promising results were obtained in the 2012 field trial, where the country-level correlations between motivation and achievement were rather close to the individual-level correlations. However, the 2015 field trial yielded a more complex and less supportive set of results. Other problems with anchoring vignettes are the additional testing time they require and the relatively high reading load of the items.

Various statistical procedures to correct for cross-national bias have been proposed. One example is the correction for response styles, such as acquiescence, extremity and social desirability, in analyses of covariance, using response style indices as covariates (He and Van de Vijver, 2016[156]; Van de Vijver and He, 2014[157]). Such procedures do not tend to have much impact on the negative correlation between motivation and achievement at the country level (see above). Another example of such a procedure is the use of propensity score matching (Buckley, 2009[158]), which attempts to increase the comparability of samples obtained in different countries by matching them through relevant background characteristics, such as socio-economic status. Although still frequently used, there are indications that the negative correlation between motivation and achievement at country level is not strongly affected by propensity matching. A final example is the statistical modelling of a response style factor in a confirmatory factor analysis (Billiet and McClendon, 2000[159]). The applicability of this procedure may be limited, however, as it can only be used in balanced scales where some items are formulated in a positive direction and others formulated in the opposite direction (such as in measures of extroversion, where some items assess extroversion and some introversion). PISA background scales do not use such a balanced approach.

It can be concluded that there is no simple way to eliminate cross-cultural bias in the PISA 2018 background questionnaires. It is advisable to use cognitive interviewing to avoid response styles and other problematic aspects as much as possible, and to use data from the field trial to further investigate the feasibility of different statistical approaches (including the standard tests of invariance using confirmatory factor analysis). However, despite the aspects of the questionnaire design outlined above and the statistical analyses proposed, it may not be possible to eradicate all sources of bias in the PISA 2018 background questionnaire.

In summary, the field trial will test the psychometric properties of the new scales and question formats. Given the inconclusiveness about how to deal with cross-cultural bias and response styles, it is proposed to employ procedures already used in the past (e.g. confirmatory factor analysis and IRT modelling of response styles) instead of using the field trial for further development of these issues.

Notes

[1] Please note that "achievement motives" include both the PISA 2015 constructs of "achieving motivation" and "test anxiety".

[2] Refers to implicit and explicit, and to procedural and declarative knowledge.

References

Afflerbach, P. and B. Cho (2010), "Determining and Describing Reading Strategies: Internet and Traditional Forms of Reading", in Waters, H. and W. Schneider (eds.), *Metacognition, Strategy Use and Instruction*, Guilford Press, New York. [6]

Allen, L., J. Cipielewski and K. Stanovich (1992), "Multiple indicators of children's reading habits and attitudes: Construct validity and cognitive correlates", *Journal of Educational Psychology*, Vol. 84/4, pp. 489-503, http://dx.doi.org/10.1037/0022-0663.84.4.489. [125]

Almlund, M. et al. (2011), "Personality Psychology and Economics", *IZA Discussion Papers*, No. 5500, Institute for the Study of Labour (IZA), Bonn, http://ftp.iza.org/dp5500.pdf. [14]

Anderson, R. (2008), "Implications of the Information and Knowledge Society for Education", in Voogt, J. and G. Knezek (eds.), *International Handbook of Information Technology in Primary and Secondary Education*, Springer, New York. [11]

Aronson, J., C. Fried and C. Good (2002), "Reducing the Effects of Stereotype Threat on African American College Students by Shaping Theories of Intelligence", *Journal of Experimental Social Psychology*, Vol. 38/2, pp. 113-125, http://dx.doi.org/10.1006/jesp.2001.1491. [80]

Artelt, C., U. Schiefele and W. Schneider (2001), "Predictors of reading literacy", *European Journal of Psychology of Education*, Vol. XVI/3, pp. 363-383, http://dx.doi.org/10.1007/BF03173188. [60]

Artelt, C. and W. Schneider (2015), "Cross-Country Generalizability of the Role of Metacognitive Knowledge in Students' Strategy Use and Reading Competence", *Teachers College Record*, Vol. 117/1, pp. 1-32, https://www.tcrecord.org/content.asp?contentid=17695. [70]

Babcock, P. and K. Bedard (2011), "The Wages of Failure: New Evidence on School Retention and Long-Run Outcomes", *Education Finance and Policy*, Vol. 6/3, pp. 293-322, http://dx.doi.org/10.1162/EDFP_a_00037. [133]

Baird, J. et al. (2014), *Assessment and Learning: State of the Field Review*, Knowledge Center for Education, Oslo, https://www.forskningsradet.no/servlet/Satellite?c=Rapport&cid=1253996755700&lang=en&pagename=kunnskapssenter%2FHovedsidemal. [137]

Baker, L. and L. Carter-Beall (2009), "Metacognitive Processes and Reading Comprehension", in Israel, S. and G. Duffy (eds.), *Handbook of Research on Reading Comprehension*, Routledge, New York. [63]

Baker, L. and A. Wigfield (1999), "Dimensions of Children's Motivation for Reading and Their Relations to Reading Activity and Reading Achievement", *Reading Research Quarterly*, Vol. 34/4, pp. 452–477, http://dx.doi.org/10.1598/RRQ.34.4.4. [53]

Bandura, A. (1997), *Self-Efficacy: The Exercise of Control*, W. H. Freeman, New York. [50]

Baranik, L., K. Barron and S. Finney (2007), “Measuring Goal Orientation in a Work Domain: Construct Validity Evidence for the 2×2 Framework”, *Educational and Psychological Measurement*, Vol. 67/4, pp. 697-718, http://dx.doi.org/10.1177/0013164406292090. [73]

Barile, J. et al. (2012), “Teacher-Student Relationship Climate and School Outcomes: Implications for Educational Policy Initiatives”, *Journal of Youth and Adolescence*, Vol. 41, pp. 256–267, http://dx.doi.org/10.1007/s10964-011-9652-8. [25]

Becker, M., N. McElvany and M. Kortenbruck (2010), “Intrinsic and Extrinsic Reading Motivation as Predictors of Reading Literacy: A Longitudinal Study”, *Journal of Educational Psychology*, Vol. 102/4, pp. 773-785, http://dx.doi.org/10.1037/a0020084. [39]

Bennett, J. (1993), “Towards ethnorelativism: A developmental model of intercultural sensitivity”, in Paige, M. (ed.), *Education for the intercultural experience*, Intercultural Press, Yarmouth, ME, https://www.amazon.com/Education-Intercultural-Experience-Paige/dp/1877864250. [8]

Berkenmeyer, N. and S. Müller (2010), “Schulinterne Evaluation - nur ein Instrument zur Selbststeuerung von Schulen?”, in Altrichter, H. and K. Maag Merki (eds.), *Handbuch Neue Steuerung im Schulsystem*, VS Verlag für Sozialwissenschaften, Wiesbaden. [143]

Billiet, J. and M. McClendon (2000), “Modeling Acquiescence in Measurement Models for two Balanced Sets of Items”, *Structural Equation Modeling*, Vol. 7/4, pp. 608-628, http://dx.doi.org/10.1207/S15328007SEM0704_5. [159]

Binder, M. (2009), “Why Are Some Low-Income Countries Better at Providing Secondary Education?”, *Comparative Education Review*, Vol. 53/4, pp. 513-534, http://www.journals.uchicago.edu/t-and-c. [135]

Binkley, M. et al. (2012), “Defining Twenty-First Century Skills”, in Griffin, P., B. McGaw and E. Care (eds.), *Assessment and Teaching of 21st Century Skills*, Springer, New York. [12]

Black, P. (2015), “Formative Assessment – An Optimistic but Incomplete Vision”, *Assessment in Education: Principles, Policy and Practice*, Vol. 22/1, pp. 161-177, http://dx.doi.org/10.1080/0969594X.2014.999643. [138]

Black, P. and D. Wiliam (1998), “Assessment and Classroom Learning”, *Assessment in Education: Principles, Policy & Practice*, Vol. 5/1, pp. 7-74, http://dx.doi.org/10.1080/0969595980050102. [146]

Blackwell, L., K. Trzesniewski and C. Dweck (2007), “Implicit Theories of Intelligence Predict Achievement Across an Adolescent Transition: A Longitudinal Study and an Intervention”, *Child Development*, Vol. 78/1, pp. 246 – 263, http://dx.doi.org/10.1111/j.1467-8624.2007.00995.x. [81]

Blau, D. and J. Currie (2006), “Pre-School, Day Care, and After-School Care: Who’s Minding the Kids?”, in Hanushek, E. and F. Welch (eds.), *Handbook of the Economics of Education*, Elsevier, Amsterdam, http://dx.doi.org/10.1016/S1574-0692(06)02020-4. [17]

Brown, A., A. Palincsar and B. Armbruster (2004), "Instructing Comprehension-Fostering Activities in Interactive Learning Situations", in Ruddell, R. and N. Unrau (eds.), *Theoretical Models and Processes of Reading*, International Reading Association, Newark. [61]

Bryk, A. et al. (2010), *Organizing Schools for Improvement: Lessons from Chicago*, University of Chicago Press, Chicago. [26]

Buckley, J. (2009), *Cross-National Response Styles in International Educational Assessments: Evidence from PISA 2006*, New York University, New York, https://edsurveys.rti.org/PISA/documents/Buckley_PISAresponsestyle.pdf. [158]

Caldwell, B. and R. Bradley (1984), *Home Observation for Measurement of the Environment: Administration Manual, Revised Edition*, Unpublished manuscript. University of Arkansas, Little Rock. [126]

Chapman, C. et al. (2011), *School Effectiveness and Improvement Research, Policy and Practice: Challenging the Orthodoxy?*, Routledge, Abingdon. [27]

Chapman, J. and W. Tunmer (1995), "Development of Young Children's Reading Self-Concepts: An Examination of Emerging Subcomponents and Their Relationship With Reading Achievement", *Journal of Educational Psychology*, Vol. 87/1, pp. 154-167, http://dx.doi.org/10.1037/0022-0663.87.1.154. [59]

Coiro, J. and E. Dobler (2007), "Exploring the Online Reading Comprehension Strategies Used by Sixth-Grade Skilled Readers to Search for and Locate Information on the Internet", *Reading Research Quarterly*, Vol. 42/2, pp. 214-257, http://dx.doi.org/10.1598/RRQ.42.2.2. [7]

Covington, M. (1992), *Making the Grade A Self-Worth Perspective on Motivation and School Reform*, Cambridge University Press, New York, http://dx.doi.org/10.1017/CBO9781139173582. [77]

Creemers, B. and L. Kyriakides (2008), *The Dynamics of Educational Effectiveness: A Contribution to Policy, Practice and Theory in Contemporary Schools*, Routledge, Abingdon. [19]

Crothers, L. et al. (2010), "A Preliminary Study of Bully and Victim Behavior in Old-For-Grade Students: Another Potential Hidden Cost of Grade Retention or Delayed School Entry", *Journal of Applied School Psychology*, Vol. 26/4, pp. 327-338, http://dx.doi.org/10.1080/15377903.2010.518843. [131]

Cunha, F. et al. (2006), "Interpreting the Evidence on Life Cycle Skill Formation", in *Handbook of the Economics of Education, Volume 2*, Elsevier, Amsterdam. [18]

Deardorff, D. (2009), *The SAGE Handbook of Intercultural Competence*, Sage Publications, Thousand Oaks. [99]

Deci, E. and R. Ryan (1985), *Intrinsic Motivation and Self-Determination in Human Behavior*, Plenum Press, New York. [117]

DeLuca, C. et al. (2015), "Instructional Rounds as a Professional Learning Model for Systemic Implementation of Assessment for Learning", *Assessment in Education: Principles, Policy and Practice*, Vol. 22/1, pp. 122-139, http://dx.doi.org/10.1080/0969594X.2014.967168. [148]

Dole, J., J. Nokes and D. Drits (2009), "Cognitive Strategy Instruction", in Israel, S. and G. Duffy (eds.), *Handbook of Research on Reading Comprehension*, Routledge, New York. [64]

Duckworth, A. et al. (2007), "Grit: Perseverance and Passion for Long-Term Goals", *Journal of Personality and Social Psychology*, Vol. 92/6, pp. 1087-1101, http://dx.doi.org/10.1037/0022-3514.92.6.1087. [83]

Dweck, C. (2007), *Mindset: The New Psychology of Success*, Ballantine Books, New York. [82]

Elliot, A. (2005), "A Conceptual History of the Achievement Goal Construct", in Elliot, A. and C. Dweck (eds.), *Handbook of Competence and Motivation*, The Guilford Press, New York. [98]

Elliot, A. and H. McGregor (2001), "A 2×2 Achievement Goal Framework", *Journal of Personality and Social Psychology*, Vol. 80/3, pp. 501-519, http://dx.doi.org/10.1037/0022-3514.80.3.501. [72]

Elliot, A. and T. Thrash (2001), "Achievement goals and the hierarchical model of achievement motivation", *Educational Psychology Review*, Vol. 13/2, pp. 139-156, http://link.springer.com/article/10.1023/A:1009057102306. [38]

Faubert, V. (2009), "School Evaluation: Current Practices in OECD Countries and a Literature Review OECD Education Working Paper No. 42", *OECD Education Working Papers*, OECD, Paris, http://www.oecd.org/edu/workingpapers. [142]

Ferla, J., M. Valcke and Y. Cai (2009), "Academic Self-Efficacy and Academic Self-concept: Reconsidering Structural Relationships", *Learning and Individual Differences*, Vol. 19/4, pp. 499-505, http://dx.doi.org/10.1016/j.lindif.2009.05.004. [51]

Fisch, K. and S. McLeod (2009), *Did you know? 3.0 – 2009 Edition*, http://www.youtube.com/watch?v=PHmwZ96_Gos (accessed on 15 May 2009). [10]

Fisher, C., S. Wallace and R. Fenton (2000), "Discrimination Distress During Adolescence", *Journal of Youth and Adolescence*, Vol. 29/6, pp. 679–695, http://dx.doi.org/10.1023/A:1026455906512. [101]

Fivush, R. and J. Hudson (eds.) (1990), *Knowing and Remembering in Young Children*, Cambridge University Press, New York. [111]

Greene, J. and M. Winters (2009), "The Effects of Exemptions to Florida's Test-based Promotion Policy: Who is Retained?. Who Benefits Academically?", *Economics of Education Review*, Vol. 28/1, pp. 135-142, http://dx.doi.org/10.1016/j.econedurev.2008.02.002. [132]

Griffith, C. et al. (2010), "Grade Retention of Students During Grades K–8 Predicts Reading Achievement and Progress During Secondary Schooling", *Reading & Writing Quarterly*, Vol. 26/1, pp. 51-66, http://dx.doi.org/10.1080/10573560903396967. [130]

Guthrie, J. (2008), “Reading Motivation and Engagement in Middle and High School: Appraisal and Intervention”, in Guthrie, J. (ed.), *Engaging Adolescents in Reading*, Corwin, Thousand Oaks. [120]

Guthrie, J., S. Klauda and A. Ho (2013), “Modeling the Relationships Among Reading Instruction, Motivation, Engagement, and Achievement for Adolescents”, *Reading Research Quarterly*, Vol. 48/1, pp. 9–26, http://dx.doi.org/10.1002/rrq.035. [115]

Guthrie, J. and A. Wigfield (2000), “Engagement and Motivation in Reading”, in Kamil, M. et al. (eds.), *Handbook of Reading Research, Volume 3*, Lawrence Erlbaum Associates, Mahwah, NJ. [49]

Guthrie, J. et al. (1999), “Motivational and Cognitive Predictors of Text Comprehension and Reading Amount”, *Scientific Studies of Reading*, Vol. 3/3, pp. 231-256, http://dx.doi.org/10.1207/s1532799xssr0303_3. [40]

Guthrie, J., A. Wigfield and W. You (2012), “Instructional Contexts for Engagement and Achievement in Reading”, in Christenson, S., A. Reschly and C. Wylie (eds.), *Handbook of Research on Student Engagement*, Springer Science, New York, http://dx.doi.org/10.1007/978-1-4614-2018-7. [116]

Hanushek, E. and L. Woessmann (2010), “The Economics of International Differences in Educational Achievement”, in Hanushek, E., S. Machin and L. Woessmann (eds.), *Handbook of the Economics of Education, Volume 3*, Elsevier, Amsterdam. [28]

Hattie, J. (2009), *Visible Learning: A Synthesis of Over 800 Meta-Analyses Relating to Achievement*, Routledge, New York. [139]

Hayward, L. (2015), “Assessment is Learning: The Preposition Vanishes”, *Assessment in Education: Principles, Policy and Practice*, Vol. 22/1, pp. 27-43, http://dx.doi.org/10.1080/0969594X.2014.984656. [150]

Heckhausen, H. (1975), “Fear of Failure as a Self-Reinforcing Motive System”, in Sarason, I. and C. Spielberger (eds.), *Stress and anxiety*, Hemisphere, Washington, DC. [78]

Heckman, J. (2006), “Skill Formation and the Economics of Investing in Disadvantaged Children”, *Science*, Vol. 312, pp. 1900-1902, http://dx.doi.org/10.1126/science.1128898. [104]

Heckman, J., J. Stixrud and S. Urzua (2006), “The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior”, *Journal of Labor Economics*, Vol. 24/3, pp. 411-482, http://dx.doi.org/10.1086/504455. [15]

He, J. and F. Van de Vijver (2016), “The Motivation-Achievement Paradox in International Educational Achievement Tests: Toward a Better Understanding”, in King, R. and A. Bernardo (eds.), *The Psychology of Asian Learners: A Festschrift in honor of David Watkins*, Springer, Singapore, http://dx.doi.org/10.1007/978-981-287-576-1_16. [156]

Helmreich, R. et al. (1978), “Achievement Motivation and Scientific Attainment”, *Personality and Social Psychology Bulletin*, Vol. 4/2, pp. 222-226, http://dx.doi.org/10.1177/014616727800400209. [71]

Hembree, R. (1988), “Correlates, Causes, Effects, and Treatment of Test Anxiety”, *Review of Educational Research*, Vol. 58/1, pp. 47-77, http://dx.doi.org/10.3102/00346543058001047. [76]

Hopfenbeck, T. and G. Stobart (2015), “Large-Scale Implementation of Assessment for Learning”, *Assessment in Education: Principles, Policy and Practice*, Vol. 22/1, pp. 1-2, http://dx.doi.org/10.1080/0969594X.2014.1001566. [152]

Hulleman, C. et al. (2010), “A Meta-Analytic Review of Achievement Goal Measures: Different Labels for the Same Constructs or Different Constructs with Similar Labels?”, *Psychological Bulletin*, Vol. 136/3, pp. 422-449, http://dx.doi.org/10.1037/a0018947. [93]

Jonsson, A., C. Lundahl and A. Holmgren (2015), “Evaluating a Large-Scale Implementation of Assessment for Learning in Sweden”, *Assessment in Education: Principles, Policy and Practice*, Vol. 22/1, pp. 104-121, http://dx.doi.org/10.1080/0969594X.2014.970612. [149]

Kaplan, A. and M. Maehr (2007), “The Contributions and Prospects of Goal Orientation Theory”, *Educational Psychology Review*, Vol. 19/2, pp. 141–184, http://dx.doi.org/10.1007/s10648-006-9012-5. [94]

Kaye, M., D. Conroy and A. Fifer (2008), “Individual Differences in Incompetence Avoidance”, *Journal of Sport & Exercise Psychology*, Vol. 30/1, pp. 110-132. [79]

King, G. et al. (2004), “Enhancing the Validity and Cross-Cultural Comparability of Measurement in Survey Research”, *American Political Science Review*, Vol. 98/1, pp. 191-207, http://dx.doi.org/10.1017/S000305540400108X. [155]

Kirsch, I. et al. (2002), *Reading for Change Performance and Engagement Across Countries: Results From PISA 2000*, OECD, Paris. [32]

Klauda, S. (2009), “The Role of Parents in Adolescents’ Reading Motivation and Activity”, *Educational Psychology Review*, Vol. 21/4, pp. 325-363, http://dx.doi.org/10.1007/s10648-009-9112-0. [128]

Klauda, S. and J. Guthrie (2015), “Comparing Relations of Motivation, Engagement, and Achievement Among Struggling and Advanced Adolescent Readers”, *Reading and Writing*, Vol. 28/2, pp. 239–269, http://dx.doi.org/10.1007/s11145-014-9523-2. [41]

Klieme, E. et al. (2013), “PISA 2012 Context Questionnaires Framework”, in *PISA 2012 Assessment and Analytical Framework: Mathematics, Reading, Science, Problem Solving and Financial Literacy*, OECD, Paris. [1]

Klieme, E. and S. Kuger (2014), *PISA 2015 Draft Questionnaire Framework*, OECD, Paris, http://www.oecd.org/pisa/pisaproducts/PISA-2015-draft-questionnaire-framework.pdf. [2]

Klieme, E., C. Pauli and K. Reusser (2009), “The Pythagoras Study. Investigating Effects of Teaching and Learning in Swiss and German Mathematics Classrooms”, in Janík, T. (ed.), *The Power of Video Studies in Investigating Teaching and Learning in the Classroom*, Waxmann, Münster. [21]

Klieme, E., B. Steinert and J. Hochweber (2010), “Zur Bedeutung der Schulqualität für Unterricht und Lernergebnisse”, in Bos, W., E. Klieme and O. Köller (eds.), *Schulische Lerngelegenheiten und Kompetenzentwicklung*, Waxmann, Münster. [121]

Kloosterman, R. and P. de Graaf (2010), “Non-Promotion or Enrolment in a Lower Track? The Influence of Social Background on Choices in Secondary Education for Three Cohorts of Dutch Pupils”, *Oxford Review of Education*, Vol. 36/3, pp. 363-384, https://www.jstor.org/stable/25699590. [134]

Kyllonen, P. and J. Bertling (2014), “Innovative Questionnaire Assessment Methods to Increase Cross-Country Comparability”, in Rutkowski, L., M. von Davier and D. Rutkowski (eds.), *Handbook of International Large-Scale Assessment: Background, Technical Issues, and Methods of Data Analysis*, CRC Press, Boca Raton. [154]

Lafontaine, D. et al. (2015), “Equity in Opportunity-To-Learn and Achievement in Reading: A Secondary Analysis of PISA 2009 Data”, *Studies in Educational Evaluation*, Vol. 47, pp. 1-11, http://dx.doi.org/10.1016/j.stueduc.2015.05.001. [122]

Lafontaine, D. et al. (2017), “Pratiques d’enseignement et compétences en lecture des élèves : qu’évaluent les enquêtes internationales et que peuvent en apprendre les enseignants ?”, in Bianco, M. and L. Lima (eds.), *Comment enseigner la compréhension en lecture?*, Hatier, Paris. [24]

Lindqvist, E. and R. Vestman (2011), “The Labor Market Returns to Cognitive and Noncognitive Ability: Evidence from the Swedish Enlistment”, *American Economic Journal: Applied Economics*, Vol. 3, pp. 101-128, http://dx.doi.org/10.1257/app.3.1.101. [34]

Logan, J. et al. (2011), “Children’s Attendance Rates and Quality of Teacher-Child Interactions in At-Risk Preschool Classrooms: Contribution to Children’s Expressive Language Growth”, *Child and Youth Care Forum*, Vol. 40, pp. 457-477, http://dx.doi.org/10.1007/s10566-011-9142-x. [109]

Markowitsch, H. and H. Welzer (2009), *The Development of Autobiographical Memory*, Psychology Press, Hove. [112]

Marsh, H. and R. Craven (2006), “Reciprocal Effects of Self-Concept and Performance From a Multidimensional Perspective Beyond Seductive Pleasure and Unidimensional Perspectives”, *Perspectives on Psychological Science*, Vol. 1/2, p. 133, http://dx.doi.org/10.1111/j.1745-6916.2006.00010.x. [54]

Marsh, H. and R. Craven (1997), “Academic Self-Concept: Beyond the Dustbowl”, in Phye, G. (ed.), *Handbook of Classroom Assessment: Learning, Achievement, and Adjustment*, Academic Press, Orlando, FL. [52]

Marsh, H. and K. Hau (2004), “Explaining Paradoxical Relations Between Academic Self-Concepts and Achievements: Cross-Cultural Generalizability of the Internal/External Frame of Reference Predictions Across 26 Countries”, *Journal of Educational Psychology*, Vol. 96/1, pp. 56-67, http://dx.doi.org/10.1037/0022-0663.96.1.56. [153]

Marsh, H. et al. (2006), "OECD's Brief Self-Report Measure of Educational Psychology's Most Useful Affective Constructs: Cross-Cultural, Psychometric Comparisons Across 25 Countries", *International Journal of Testing*, Vol. 6/4, pp. 311-360, http://dx.doi.org/10.1207/s15327574ijt0604_1. [33]

McKenna, M. and R. Robinson (1990), "Content Literacy: A Definition and Implications", *Journal of Reading*, Vol. 34/3, pp. 184-186. [22]

Middleton, M. and K. Perks (2014), *Motivation to Learn: Transforming Classroom Culture to Support Student Achievement*, Sage Publishing, Thousand Oaks, http://dx.doi.org/10.4135/9781483394480. [95]

Mol, S. and A. Bus (2011), "To Read or not to Read: A Meta-Analysis of Print Exposure from Infancy to Early Adulthood", *Psychological Bulletin*, Vol. 137/2, pp. 267-296, http://dx.doi.org/10.1037/a0021890. [42]

Morgan, P. and D. Fuchs (2007), "Is There a Bidirectional Relationship Between Children's Reading Skills and Reading Motivation?", *Exceptional Children*, Vol. 73/2, pp. 165-183, http://dx.doi.org/10.1177/001440290707300203. [43]

Murayama, K. and A. Elliot (2012), "The Competition–Performance Relation: A Meta-Analytic Review and Test of the Opposing Processes Model of Competition and Performance", *Psychological Bulletin*, Vol. 138/6, pp. 1035-1070, http://dx.doi.org/10.1037/a0028324. [74]

National Reading Panel (2000), *Report of the National Reading Panel: Teaching Children to Read: An Evidence-Based Assessment of the Scientific Research Literature on Reading and Its Implications for Reading Instruction: Report of the Subgroups*, National Institutes of Health. [62]

Naumann, J. (2015), "A model of online reading engagement: Linking engagement, navigation, and performance in digital reading", *Computers in Human Behavior*, Vol. 53, pp. 263–277, http://dx.doi.org/10.1016/j.chb.2015.06.051. [124]

Neuman, S. and D. Dickinson (2003), *Handbook of Early Literacy Research, Volume 1*, Guilford Press, New York. [127]

OECD (2017), *PISA 2015 Assessment and Analytical Framework: Science, Reading, Mathematic, Financial Literacy and Collaborative Problem Solving*, PISA, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264281820-en. [87]

OECD (2016), *PISA 2015 Results (Volume I): Excellence and Equity in Education*, PISA, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264266490-en. [58]

OECD (2015), *Education Policy Outlook 2015: Making Reforms Happen*, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264225442-en. [5]

OECD (2014), *PISA 2012 Technical Report*, OECD, Paris. [31]

OECD (2013), *OECD Guidelines on Measuring Subjective Well-being*, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264191655-en. [86]

OECD (2013), *PISA 2012 Assessment and Analytical Framework: Mathematics, Reading, Science, Problem Solving and Financial Literacy*, PISA, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264190511-en. [3]

OECD (2012), *PISA 2009 Technical Report*, PISA, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264167872-en. [30]

OECD (2011), *How's Life?: Measuring Well-being*, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264121164-en. [85]

OECD (2011), *PISA 2009 Results: Students On Line: Digital Technologies and Performance (Volume VI)*, PISA, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264112995-en. [123]

OECD (2010), *PISA 2009 Results: Learning to Learn: Student Engagement, Strategies and Practices (Volume III)*, PISA, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264083943-en. [48]

OECD (2010), *PISA 2009 Results: What Students Know and Can Do: Student Performance in Reading, Mathematics and Science (Volume I)*, PISA, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264091450-en. [47]

OECD (2007), *PISA 2006: Science Competencies for Tomorrow's World: Volume 1: Analysis*, PISA, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264040014-en. [57]

OECD (2005), *The Definition and Selection of Key Competencies: Executive Summary*, OECD, Paris. [89]

Opdenakker, M. and J. Van Damme (2000), "Effects of Schools, Teaching Staff and Classes on Achievement and Well-Being in Secondary Education: Similarities and Differences Between School Outcomes", *School Effectiveness and School Improvement*, Vol. 11/2, pp. 165-196, http://dx.doi.org/10.1076/0924-3453(200006)11:2;1-q;ft165. [91]

Ou, S. and A. Reynolds (2010), "Grade Retention, Postsecondary Education, and Public Aid Receipt", *Educational Evaluation and Policy Analysis*, Vol. 32/1, pp. 215–229, http://dx.doi.org/10.3102/0162373709354334. [129]

Ozga, J. (2012), "Assessing PISA", *European Educational Research Journal*, Vol. 11/2, pp. 166-171, http://dx.doi.org/10.2304/eerj.2012.11.2.166. [136]

Partnership for 21st Century Skills (2008), *21st Century Skills, Education & Competitiveness: A Resource and Policy Guide*, Partnership for 21st Century Skills, Tucson, AZ, http://www.21stcenturyskills.org. [90]

Pfost, M., T. Dörfler and C. Artelt (2013), "Students' Extracurricular Reading Behavior and the Development of Vocabulary and Reading Comprehension", *Learning and Individual Differences*, Vol. 26, pp. 89-102, http://dx.doi.org/10.1016/j.lindif.2013.04.008. [44]

Poropat, A. (2009), "A Meta-Analysis of the Five-Factor Model of Personality and Academic Performance", *Psychological Bulletin*, Vol. 135/2, pp. 322–338, http://dx.doi.org/10.1037/a0014996. [35]

Pressley, M. (2000), "What Should Comprehension Instruction be the Instruction Of?", in Kamil, M. et al. (eds.), *Handbook of reading research, Volume III*, Lawrence Erlbaum Associates, Mahwah, NJ. [113]

Pressley, M., S. Graham and K. Harris (2006), "The State of Educational Intervention Research as Viewed Through the Lens of Literacy Intervention", *British Journal of Educational Psychology*, Vol. 76, pp. 1-19, http://dx.doi.org/10.1348/000709905X66035. [65]

Pressley, M. et al. (1989), "Strategies That Improve Children's Memory and Comprehension of Text", *The Elementary School Journal*, Vol. 90/1, pp. 3-32, http://dx.doi.org/10.1086/461599. [66]

Ratnam-Lim, C. and K. Tan (2015), "Large-Scale Implementation of Formative Assessment Practices in an Examination-Oriented Culture", *Assessment in Education: Principles, Policy and Practice*, Vol. 22/1, pp. 61-78, http://dx.doi.org/10.1080/0969594X.2014.1001319. [151]

Rawsthorne, L. and A. Elliot (1999), "Achievement Goals and Intrinsic Motivation: A Meta-Analytic Review", *Personality and Social Psychology Review*, Vol. 3/4, pp. 326-344, http://dx.doi.org/10.1207/s15327957pspr0304_3. [96]

Reeve, J. (2012), "A Self-Determination Theory Perspective on Student Engagement", in Christenson, S., A. Reschly and C. Wylie (eds.), *Handbook of Research on Student Engagement*, Springer, New York. [118]

Retelsdorf, J., O. Köller and J. Möller (2011), "On the Effects of Motivation on Reading Performance Growth in Secondary School", *Learning and Instruction*, Vol. 21/4, pp. 550-559, http://dx.doi.org/10.1016/j.learninstruc.2010.11.001. [55]

Richardson, M., C. Abraham and R. Bond (2012), "Psychological Correlates of University Students' Academic Performance: A Systematic Review and Meta-Analysis", *Psychological Bulletin*, Vol. 138/2, pp. 353-387, http://dx.doi.org/10.1037/a0026838. [36]

Rosenkvist, M. (2010), "Using Student Test Results for Accountability and Improvement: A Literature Review", *OECD Education Working Papers*, No. 54, OECD Publishing, Paris, https://dx.doi.org/10.1787/5km4htwzbv30-en. [140]

Rosenshine, B. and C. Meister (1997), "Cognitive strategy instruction in reading", in Stahl, S. and D. Hayes (eds.), *Instructional Models in Reading*, Lawrence Erlbaum Associates, Mahwah, NJ. [114]

Rosenshine, B. and C. Meister (1994), "Reciprocal Teaching: A Review of the Research", *Review of Educational Research*, Vol. 64/4, pp. 479-530, http://dx.doi.org/10.3102/00346543064004479. [67]

Rosenshine, B., C. Meister and S. Chapman (1996), "Teaching Students to Generate Questions: A Review of the Intervention Studies", *Review of Educational Research*, Vol. 66/2, pp. 181-221, http://dx.doi.org/10.3102/00346543066002181. [68]

Rumberger, R. and G. Palardy (2005), "Test Scores, Dropout Rates, and Transfer Rates as Alternative Indicators of High School Performance", *American Educational Research Journal*, Vol. 42/1, pp. 3-42, http://dx.doi.org/10.3102/00028312042001003. [92]

Ryan, K., M. Chandler and M. Samuels (2007), "What Should School-Based Evaluation Look Like?", *Studies in Educational Evaluation*, Vol. 33/3-4, pp. 197-212, http://dx.doi.org/10.1016/j.stueduc.2007.07.001. [144]

Rychen, D. and L. Salganik (2003), *Key Competencies for a Successful Life and a Well-Functioning Society*. [13]

Sammons, P. et al. (2009), "Children's Cognitive Attainment and Progress in English Primary Schools During Key Stage 2: Investigating the potential continuing influences of pre-school education", in Roßbach, H. and H. Blossfeld (eds.), *Frühpädagogische Förderung in Institutionen*, VS Verlag für Sozialwissenschaften, Wiesbaden, http://dx.doi.org/doi.org/10.1007/978-3-531-91452-7_12. [106]

Santiago, P. and F. Benavides (2009), *Teacher Evaluation A Conceptual Framework and Examples of Country Practices*, OECD, Paris. [145]

Schaffner, E., M. Philipp and U. Schiefele (2016), "Reciprocal Effects Between Intrinsic Reading Motivation and Reading Competence? A Cross-lagged Panel Model for Academic Track and Nonacademic Track Students", *Journal of Research in Reading*, Vol. 39/1, pp. 19–36, http://dx.doi.org/10.1111/1467-9817.12027. [45]

Scheerens, J. (2002), "School Self-Evaluation: Origins, Definition, Approaches, Methods and Implementation", in Nevo, D. (ed.), *School-Based Evaluation: An International Perspective*, JAI Press, Amsterdam. [141]

Scheerens, J. and R. Bosker (1997), *The Foundations of Educational Effectiveness*, Pergamon, Oxford. [20]

Schiefele, U. et al. (2012), "Dimensions of Reading Motivation and Their Relation to Reading Behavior and Competence", *Reading Research Quarterly*, Vol. 47/4, pp. 427–463, http://dx.doi.org/10.1002/RRQ.030. [46]

Shanahan, T. and C. Shanahan (2008), "Teaching Disciplinary Literacy to Adolescents: Rethinking Content- Area Literacy", *Harvard Educational Review*, Vol. 78/1, pp. 40-59, http://dx.doi.org/10.17763/haer.78.1.v62444321p602101. [23]

Shiner, R. et al. (2007), "The Power of Personality: The Comparative Validity of Personality Traits, Socioeconomic Status, and Cognitive Ability for Predicting Important Life Outcomes", *Perspectives on Psychological Science*, Vol. 2/4, pp. 313-345, http://dx.doi.org/10.1111/j.1745-6916.2007.00047.x. [37]

Sinicrope, C., J. Norris and Y. Watanabe (2007), "Understanding and Assessing Intercultural Competence: A Summary of Theory, Research, and Practice", *Second Language Studies*, Vol. 26/1, pp. 1-58. [9]

Solheim, O. (2011), "The Impact of Reading Self-efficacy and Task Value on Reading Comprehension Scores in Different Item Formats", *Reading Psychology*, Vol. 32/1, pp. 1-27, http://dx.doi.org/10.1080/02702710903256601. [56]

Spence, J. and R. Helmreich (1983), “Achievement-Related Motives and Behavior”, in Spence, J. (ed.), *Achievement and Achievement Motives: Psychological and sociological Approaches*, Freeman, San Francisco, CA. [75]

Stanat, P. and G. Christensen (2006), *Where Immigrant Students Succeed - A Comparative Review of Performance and Engagement in PISA 2003*, OECD, Paris. [100]

Stiglitz, J., A. Sen and J. Fitoussi (2009), *Report by the Commission on the Measurement of Economic Performance and Social Progress*, European Commission, Brussels, http://www.stiglitz-sen-fitoussi.fr. [88]

Sykes Gary, B. Schneider and D. Plank (eds.) (2009), *Handbook of Education Policy Research*, Routledge, New York. [4]

Sylva, K. et al. (2011), “Pre-School Quality and Educational Outcomes at Age 11: Low Quality Has Little Benefit”, *Journal of Early Childhood Research*, Vol. 9/2, pp. 109–124, http://dx.doi.org/10.1177/1476718X10387900. [107]

Sylva, K. et al. (2011), “Effects of Early Child-Care on Cognition, Language, and Task-Related Behaviours At 18 Months: An English Study”, *British Journal of Developmental Psychology*, Vol. 29, pp. 18-45, http://dx.doi.org/10.1348/026151010X533229. [110]

Ţuţu, A. and T. Constantin (2012), “Understanding Job Performance Through Persistence and Job Competency”, *Procedia - Social and Behavioral Sciences*, Vol. 33, pp. 612-616, http://dx.doi.org/10.1016/j.sbspro.2012.01.194. [84]

UNESCO (2006), *Education For All Global Monitoring Report 2007: Strong Foundations*, UNESCO, Paris, http://dx.doi.org/10.1007/978-94-007-0753-5_101111. [105]

Valenti, J. and D. Tracey (2009), “Full-Day, Half-Day, and No Preschool”, *Education and Urban Society*, Vol. 41/6, pp. 695-711, http://dx.doi.org/10.1177/0013124509336060. [108]

van de Vijver, F., S. Breugelmans and S. Schalk-Soekar (2008), “Multiculturalism: Construct Validity and Sstability”, *International Journal of Intercultural Relations*, Vol. 32/2, pp. 93-104, http://dx.doi.org/10.1016/j.ijintrel.2007.11.001. [103]

Van de Vijver, F. and J. He (2014), “Report on Social Desirability, Midpoint and Extreme Responding in TALIS 2013”*, OECD Education Working Papers*, No. 107, OECD Publishing, Paris, https://dx.doi.org/10.1787/5jxswcfwt76h-en. [157]

Van Yperen, N., M. Blaga and T. Postmes (2014), “A Meta-Analysis of Self-Reported Achievement Goals and Nonself-Report Performance Across Three Achievement Domains (Work, Sports, and Education)”, *PLoS ONE*, Vol. 9/4, pp. 1-16, http://dx.doi.org/10.1371/journal.pone.0093594. [97]

Vansteenkiste, M., W. Lens and E. Deci (2006), “Intrinsic Versus Extrinsic Goal Contents in Self-Determination Theory: Another Look at the Quality of Academic Motivation”, *Educational Psychologist*, Vol. 41/1, pp. 19-31, http://dx.doi.org/10.1207/s15326985ep4101_4. [119]

Waters, H. and W. Schneider (eds.) (2010), *Metacognition, Strategy Use, and Instruction*, Guilford Press, New York. [69]

Willms, J. (2006), *Learning Divides: Ten Policy Questions About the Performance and Equity of Schools and Schooling Systems*, UNESCO Institute of Statistics, Montreal, http://www.uis.unesco.org. [16]

Woessmann, L. et al. (2007), "School Accountability, Autonomy, Choice, and the Level of Student Achievement: International Evidence from PISA 2003", *OECD Education Working Papers*, No. 13, OECD Publishing, Paris, https://dx.doi.org/10.1787/246402531617. [29]

Wong, C., J. Eccles and A. Sameroff (2003), "The Influence of Ethnic Discrimination and Ethnic Identification on African American Adolescents' School and Socioemotional Adjustment", *Journal of Personality*, Vol. 71/6, pp. 1197-232, http://dx.doi.org/10.1111/1467-6494.7106012. [102]

Wylie, E. and C. Lyon (2015), "The Fidelity of Formative Assessment Implementation: Issues of Breadth and Quality", *Assessment in Education: Principles, Policy and Practice*, Vol. 22/1, pp. 140–160., http://dx.doi.org/10.1080/0969594X.2014.990416. [147]

8. PISA 2018 Well-being Framework

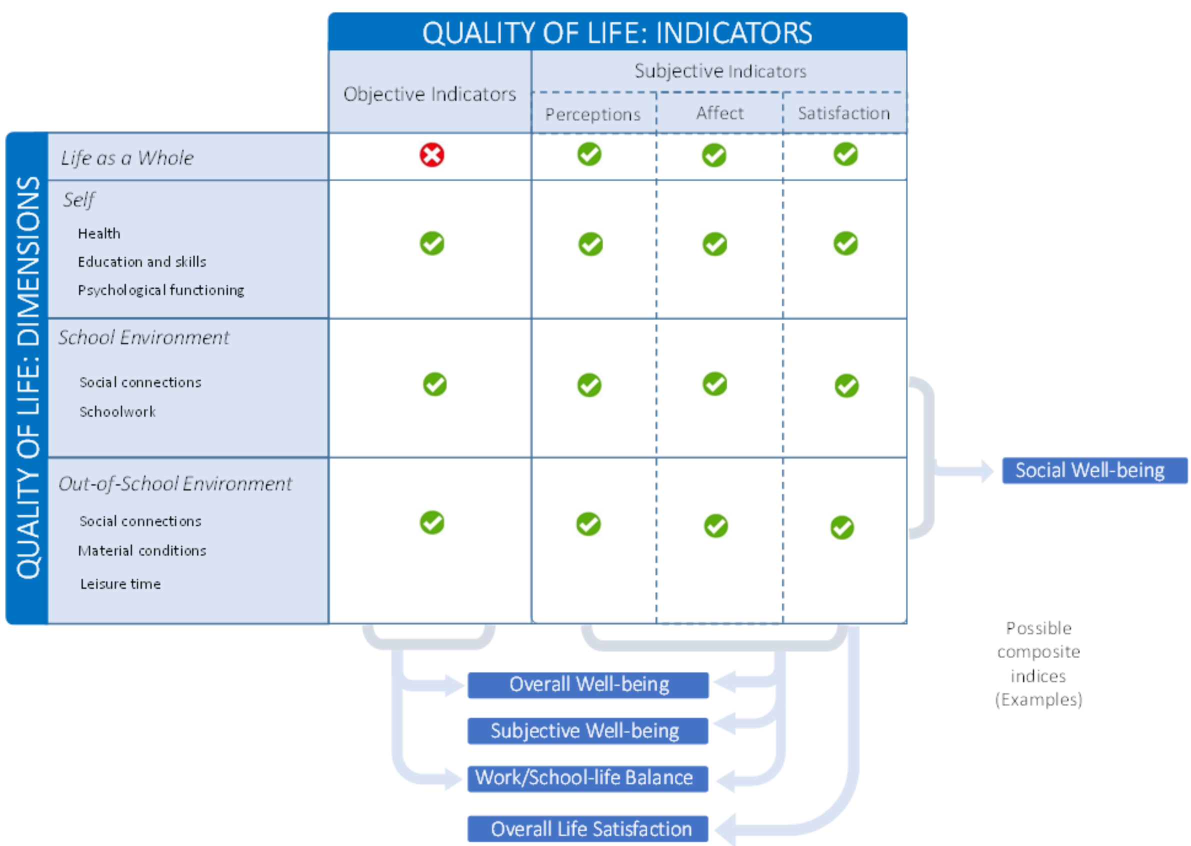
This section presents the theoretical framework for the way in which the 2018 cycle of the Programme for International Student Assessment (PISA) assesses student well-being. PISA was the first large-scale study to examine student well-being in its 2015 cycle. This framework discusses potential objective and subjective indicators of student well-being, grounding them in previous attempts from the literature. It also distinguishes between various dimensions of well-being, including life as a whole, self-related well-being, school-related well-being, and well-being out of school. Potential measurement issues are also presented. Potential composite indicators, combining responses to various questions into a single indicator, are suggested at the end.

Executive summary

Well-being, defined as the quality of people's lives and their standard of living, is of growing interest to policy makers and educators around the world. There seems to be a consensus that well-being is a multi-dimensional construct that comprises both objective, material components and subjective, psychological facets. While there is a growing body of research on the topic, only a few large-scale studies for adolescents have taken a comprehensive view on well-being. Besides PISA, no large-scale assessment directly links students' well-being to their educational achievement, and little has thus been established regarding the relationship between student learning and well-being.

By measuring well-being, PISA can create international benchmarks of student well-being across OECD and partner countries via a database of tremendous utility for educators, researchers, and policy makers. Research into well-being involves a variety of approaches used in public health, education, psychology, and economics. The framework outlined here aims to integrate different perspectives on well-being and to present a comprehensive model that covers different dimensions of well-being with a spectrum of indicators (both objective and subjective).

Figure 8.1. Framework overview



The proposed modular framework (Figure 8.1) distinguishes three main dimensions of well-being in addition to students' perceived quality of life as a whole:

- First, well-being in terms of how fit and healthy students are, the education and skills they have, and how they feel about themselves and their lives (*self*);
- Second, well-being in terms of the environment a student is exposed to at school (*school environment*); and
- Third, well-being in terms of the living environment and circumstances outside of school experienced by a student (*out-of-school environment*).

Several sub-dimensions under each of these broader dimensions can be directly mapped to the dimensions proposed in other frameworks.

Possible measurement approaches are presented and specific indicators are outlined for all framework components. Recommendations are informed by a review of the relevant literature as well as by the pragmatic considerations of space in the questionnaire, student burden, and available survey methods. Special consideration is given to issues of cross-cultural comparability and the age appropriateness of the proposed survey methods. In order to measure well-being in a brief and efficient manner, innovative survey methods drawing on the day reconstruction method are outlined, thereby further extending the approaches successfully implemented in PISA 2015.

The framework is modular in two ways (Figure 8.1). First, the framework can be broken down into modules by dimension (i.e., self, school environment, and out-of-school environment). Second, the framework can be broken down into modules by the type of indicator (i.e., objective well-being indicators and subjective perceptions, affect, and satisfaction). The different cells in the framework, which are themselves indicators, therefore also give rise to potential composite indicators that can be used as robust reporting elements in areas of key policy interest. These include, among others, indices of overall well-being, subjective well-being, social well-being and work/school-life balance.

Introduction

Well-being can be defined as the quality of people's lives and their standard of living. It is often quantified both via objective measures, such as household income, educational resources and health status, and via subjective indicators such as experienced affect (or emotions), perceptions of quality of life and life satisfaction (Casas, 2011[1]).

Economists have proposed several possible alternatives to using only gross domestic product (GDP) as an indicator of nations' well-being (Diener and Seligman, 2004[2]; Kahneman et al., 2004[3]; Stiglitz, Sen and Fitoussi, 2009[4]). For instance, Stiglitz et al. (2009, p. 58[4]) recommended that, "[s]tatistical offices should incorporate questions to capture people's life evaluations, hedonic experiences, and priorities in their own surveys". Several countries have started collecting data and reporting more comprehensive well-being metrics, including measures of subjective well-being (SWB) (Boarini, Kolev and Mcgregor, 2014[5]; Evans, Macrory and Randall, 2015[6]; Federal Interagency Forum on Child and Family Statistics, 2009[7]; The Foundation for Child Development (FCD), 2012[8]; UNICEF, 2007[9]; UNICEF, 2012[10]; Statham and Chase, 2010[11]; The Children's Society, 2015[12]). Numerous studies have identified important determinants for adult subjective well-being, often defined as how desirable people find their lives, following the definition proposed by Diener et al. (1999[13]). Among the most important determinants include health, employment-related factors (e.g. income and unemployment) and social contacts (Dolan, Peasgood and White, 2008[14]; Sacks, Stevenson and Wolfers, 2010[15]; Winkelmann and Winkelmann, 1998[16]; Helliwell, Layard and Sachs, 2015[17]). There is empirical evidence that SWB and objective measures of health are related to important work-related outcomes, with healthy individuals being more productive and making less use of health care services (Keyes and Grzywacz, 2005[18]). Longitudinal studies have shown that mental health is an important predictor of subsequent work performance (Wright, Bonett and Sweeney, 1993[19]). Diener and Chan (2011[20]) reported that people who are happier tended to report a better health status and have a higher life expectancy than individuals who frequently experience anger, depression or anxiety. Findings pointing to the importance of well-being for general life outcomes and workplace success have also increased interest in well-being among the business community (Beal, Rueda-Sabater and Ling Heng, 2015[21]).

Policy makers now increasingly call for information on their citizens' and workforce's well-being in addition to indicators of their knowledge and skills. Large international health surveys, such as the World Health Survey, Health Behavior in School-aged Children, WHO-5 (Topp et al., 2015[22]) and KIDSCREEN (Ravens-Sieberer et al., 2014[23]), and adult household surveys, such as the Gallup World Poll (Boarini et al., 2012[24]) already include measures of well-being. However, most international well-being assessments have so far focused on adult populations. Indeed, the 2015 Good Childhood Report states that, "[p]eople's subjective well-being has become a topic of widespread – and growing – interest. However, discussion of children's subjective well-being has been notable by its absence" (The Children's Society, 2015, p. 9[12]). While more studies that specifically focus on adolescent and child well-being now exist (Ben-Arieh, 2008[25]; Cummins and Lau, 2005[26]; Lippman, Moore and McIntosh, 2011[27]; Pollard and Lee, 2003[28]; Huebner, 2001[29]; Bradshaw et al., 2011[30]; Gilman and Huebner, 2003[31]; Huebner and Dew, 1996[32]; Saha et al., 2010[33]), many of these studies focus on specific subgroups rather than on the general child and adolescent population (Casas, 2011[1]).

As the first international large-scale assessment of students' well-being, the 2015 *Programme for International Student Assessment* (PISA) included a few questions on students' subjective well-being to its student questionnaire (OECD, 2017[34]). For the first time, indicators of students' well-being have been directly related to students' achievement across a large number of education systems (OECD., 2017[35]). However, the set of questions included in PISA 2015, and therefore the conclusions that could be drawn from these questions, were limited in scope. That might be changed in PISA 2018. A separate well-being questionnaire encompassing questions covering the entire well-being construct could be a building block for international benchmarks on adolescent well-being. The OECD has already established guidelines for the measurement of adult well-being through its *Better Life Initiative* (OECD, 2015[36]), and it now has the chance to do the same for adolescents.

It is important to monitor adolescent well-being as today's adolescents are tomorrow's workforce: how they fare today is directly related to how their countries will fare in an increasingly globalised and competitive economy. Media reports about extremely long school hours and rising suicide rates in some high-performing countries, or findings that large proportions of students report disliking school and show diminished school engagement (McGill et al., 2012[37]) and report high levels of anxiety and stress regarding school (Natsuaki, Biehl and Ge, 2009[38]) raise questions about the trade-offs between different educational and societal objectives. As Helliwell, Layard and Sachs (2015, p. 11[17]) remarked in the 2015 World Happiness Report, "if schools do not measure the well-being of their children, but do measure their intellectual development, the latter will always take precedence".

Schleicher (2015[39]) described three ways in which well-being is of direct policy relevance to PISA. First, adolescent well-being is intrinsically important as it is a part of governments' efforts to ensure all of their citizens' and residents' well-being. Second, adolescent well-being is an important determinant of adult well-being. Finally, adolescent well-being is a substantial driver of educational outcomes in the school system. Educators and policy makers are in need of valid and reliable information on student well-being so that they can evaluate the efficacy of policy interventions targeting child well-being, such as bullying prevention programs.

An initial step in defining well-being measures for adolescents is determining how the construct differs between adolescents and adults, the group for which the majority of research has so far been conducted. Some of the key components of adult well-being, such as job satisfaction, earnings or work-life balance, are conceptually rooted in adult life and must be adapted for younger populations. Also, adolescent well-being must take into account adolescents' priorities, their opportunities to spend leisure time and time with friends, and their relationships with parents, teachers and adults in general (The Children's Society, 2015[12]). Peer relationships, in particular, become more important in adolescence (Hardy, Bukowski and Sippola, 2002[40]; Way and Greene, 2006[41]; McGill et al., 2012[37]; Wang and Eccles, 2012[42]; Way, Reddy and Rhodes, 2007[43]). Indeed, when a large sample of 14 and 15-year-olds were asked about what having a good life meant to them, five of the six concepts they most commonly discussed were "friends", "family", "bullying", "parents", and "school"[1] (The Children's Society, 2015[12]). All of these are related to their relationships, not to material conditions.

The framework proposed herein for measuring well-being in PISA is based on other such frameworks that have been proposed, for both children and adults, and it integrates aspects that have so far often been separately treated. It aims especially to accomplish the following:

1. To recognise that well-being is a multi-dimensional construct and that its measurement requires covering different domains, not just overall life satisfaction;
2. To distinguish between overall well-being and subjective well-being. More specifically, the framework distinguishes between objective and subjective indicators of student well-being;
3. To focus on adolescent well-being, therefore placing special emphasis on the life environment of school-aged children. Indicators specific to adolescent well-being are included in addition to those used for adults that also apply to adolescents;
4. To focus on *individual* well-being, as that can be measured by PISA through the main student questionnaire or a supplementary well-being questionnaire. Indicators that might be collected at the system level are briefly mentioned but not elaborated on in detail. Such system-level indicators of well-being include aspects of environmental quality, crime or employment statistics, which might be inferred from available other data sources based on a school's location;
5. To consider measurement challenges such as the age-appropriateness of items and item formats, cross-cultural comparability and respondent burden. The framework will also propose solutions to addressing such challenges, including multi-method assessment strategies involving self-reported biodata and behaviours, subjective self-reports, and elements of the *day reconstruction method* or *event reconstruction method* (Kahneman and Krueger, 2006[44]);

6. To lay out the foundations of a well-being assessment plan for PISA 2018 that specifies which components of the framework are already covered by previous PISA questionnaires, and which components would need to be added.

Well-being as a multi-dimensional construct

Adolescent well-being, defined as the quality of students' lives and their standard of living, is of growing interest to policy makers and educators around the world. There seems to be a consensus that well-being is a multi-dimensional construct that comprises both objective, material components and subjective, psychological facets. While there is a growing body of research on the topic, only a few large-scale studies for adolescents have taken a comprehensive view on well-being. Some studies have focussed mainly on material well-being and health outcomes (e.g., Health Behaviour in School-Aged Children [HBSC]) and other studies have focussed more on subjective well-being (e.g., Children's Worlds and the Gallup Student Poll). However, none of these studies have directly linked well-being to students' educational achievement.

Despite sometimes being used interchangeably, it is important to differentiate between well-being and subjective well-being. Well-being is a multifaceted construct that includes subjective well-being but also objective well-being. Subjective well-being can be defined as "people's evaluations of their lives—the degree to which their thoughtful appraisals and affective reactions indicate that their lives are desirable and proceeding well" (Diener, Oishi and Lucas, 2015[45]). It includes both an affective component – both positive and negative emotions – and a cognitive component – one's judgment of one's overall life satisfaction or satisfaction with specific domains of one's life. This framework presents a comprehensive model that defines different dimensions of well-being and a variety of both objective and subjective indicators available for each dimension.

Figure 8.1 presents a graphical depiction of the overall framework. In addition to overall well-being, three main dimensions of well-being have been identified: well-being in term of how fit and healthy students are and how they feel about themselves and their lives (*self*); well-being in the school environment a student is exposed to (*school environment*); and well-being in the student's living environment and circumstances outside of school (*out-of-school environment*). Under each of these broader dimensions, several sub-dimensions (e.g., social connections or health) can be directly mapped to the 11 dimensions of the quality of life proposed by the OECD Better Life Initiative (OECD, 2013[46]), as well as to dimensions of other key frameworks described in the literature (Table 8.1).

Table 8.1. Dimensions of well-being in other frameworks

Framework	Dimensions of well-being as defined in each framework
Lippman et al. (2011)	- Self (physical health, development and safety; cognitive development and education; psychological/emotional development; social development and behaviour) - Relationships on different levels (family, peers, school, community and macrosystem) - Contexts (family, peers, school, community and macrosystem).
Adamson (2007)	- Material well-being, health and safety, educational well-being, interpersonal relationships, behaviour and risks, and subjective well-being
Land, Lamb and Mustillo (2001)	- Material well-being (poverty, employment and income) - Health (mortality rate and personal health) - Social relationships - Safety/behavioural concerns (e.g., smoking) - Educational attainment - One's place in the community (including enrolment and engagement) - Emotional well-being
Bradshaw, Hoelscher and Richardson (2007)	- Material situation (poverty, deprivation and parental joblessness) - Housing (overcrowding, quality of the local environment, housing problems) - Health (health at birth, immunization, health behaviour) - Subjective well-being (self-defined health, personal well-being and well-being at school) - Education (educational attainment, educational participation, youth labour market outcomes from education) - Relationships (family structure, relationships with parents, relationships with peers) - Civic participation (participation in civic activities, political interest) - Risk and safety (child mortality, risky behaviour, experiences of violence)
Moore et al. (2008)	- Child health and safety - Educational achievement and cognitive development - Social and emotional development - Family processes - Family demographics Also distinguishes between: • Child well-being (physical health, psychological health, social health and educational/intellectual development) • Contextual well-being (family, community and socio-demographic factors)

Unlike most of the frameworks in the literature, this framework intersects subjective well-being with all other dimensions of well-being rather than setting it apart as a separate independent dimension. Moreover, as Casas et al. (2012[47]) pointed out, assessing quality of life involves measuring both the material and non-material characteristics of life in large populations, and subjective measures should be utilized to add to objective measures rather than replace them. Indeed, early literature on well-being already investigated its subjective components: a paper from over four decades ago defined non-material quality of life as peoples' "perceptions, evaluations, and aspirations concerning their own lives and life conditions" (Campbell, Converse and Rodgers, 1976[48]).

The framework is modular in two ways. First, the framework can be broken down by dimension (i.e., life as a whole, self, school environment, and out-of-school environment). Second, the framework can be broken down into modules by the type of indicator (i.e., objective well-being indicators, subjective perceptions, affect, and satisfaction). The different cells in the framework therefore give rise to potential composite indicators that can be used as robust reporting elements in areas of key policy interest. In addition to the proposed composite indicators of overall well-being and subjective well-being, composite indicators of emotional well-being, social well-being, life satisfaction and work/life balance (school/life balance) are suggested for further consideration. A school/life balance index could, for instance, serve as a benchmark for how well students in different countries are able to integrate curricular demands and school life with time for personal activities, leisure and maintenance of a healthy lifestyle. These broader composite indices directly address the policy need for a smaller set of robust reporting elements following the model of PISA's index of economic, social, and cultural status (ESCS). Several authors have emphasised the value of creating composite well-being indices, particularly to facilitate the measurement of trends and the comparison of trends across sub-groups or regions (Ben-Arieh, 2008[25]; Fernandes, Mendes and Teixeira, 2012[49]; Land et al., 2007[50]).

Before describing each of the components of the framework in detail, the next section discusses important measurement challenges and methodological considerations.

Addressing measurement challenges

There are several challenges in measuring psychological, subjective, or non-cognitive constructs in PISA: robust measurement approaches are required yet student burden must remain low; questionnaire and survey items cannot be perceived as being too intrusive; and cross-national and cross-cultural comparability of the recorded responses must be maintained. This section presents five recommendations that directly address these measurement challenges and that provide a basis for the selection of the proposed measures.

Balance single-item measures with multi-item indices

Previous studies that measure well-being have often relied on single-item indicators or on a set of very few questions. While this approach is easy to administer, it is conceptually unsatisfactory and potentially invalid and unreliable for several reasons:

- There is no consensus in the literature as to the best single question for measuring well-being, and research evidence on the equivalence of approaches is insufficient;
- A single question or a small set of single-item indicators will overemphasise certain aspects of well-being while underrepresenting others and thereby fail to capture the construct in its entirety;
- Some well-being questions are likely to be more sensitive to cross-cultural norms and response styles than others. Creating a valid international well-being indicator requires sampling a larger number of questions and field-testing them in order to select the most appropriate questions for operational use; and
- Reporting elements based on one or a very few number of questions are less reliable when used as well-being indicators in cross-country comparisons.

Hence, in order to robustly measure well-being across nations and economies, it is crucial to rely on multiple indicators and a multi-item measurement approach for the construct at hand. These recommendations are consistent with Casas et al. (2012, p. 26[47]), who state that future research for cross-country comparability should collect data using more than one scale in a given area and that "we need much more data and from more countries to analyze in any real depth the qualities and possible weaknesses of each scale for the international comparison of adolescent populations."

At the same time, some components of the overall well-being construct require fewer questions than others to ensure valid and reliable measurement. In particular, some objective indicators might be captured directly as observable variables. However, not all variables of interest can be directly measured, therefore requiring the use of one or several proxies for a variable of interest. Creating multi-item indicators would be consistent with current practice in PISA and other large-scale testing programs. For example, TIMSS and PIRLS currently use a multi-item index approach and NAEP has recently moved to an index approach for more robust reporting (Bertling, 2014[51]).

Use a meaningful number of scale points and clearly differentiated scale labels

Most established well-being instruments have been developed and validated for adult populations and must be adapted to PISA's 15-year-old student population. These instruments tend to use response formats with substantially more response scale points (such as a 0-10 or a 1-10 scale where only the scale endpoints are labelled with a description) than current PISA practice (generally four to five scale points; see below). However, the observed frequency distributions for such instruments with ten or more scale points are highly skewed with large proportions of responses far above the scale mid-point. Given the very sparse frequencies on the lower end of the response scale, it is uncertain whether all scale points are conceptually meaningful and practically useful. Indeed, it may be that scales with fewer scale points might be equally or even more valid. This problem appears to be even more severe for younger respondents than for adults. For example, data from the 2015 Children's Worlds Survey show that across 15 countries, more than 80% of all student responses fell into categories 9 and 10 on a zero to 10 scale (Rees and Main, 2015[52]). Researchers have recommended using scales with fully-labelled response options whenever possible (Dillman, Smyth and Christian, 2014[53]; Gehlbach, 2015[54]; Krosnick and Fabrigar, 1997[55]). Furthermore, reducing the number of scale points below 11 could potentially improve the validity of PISA questionnaire and survey items.

However, the dominant question format in the PISA student questionnaires might be subject to the opposite problem: too few scale points. Most PISA items currently use a four-point Likert-type response format with the verbal anchors "strongly disagree", "disagree", "agree" and "strongly disagree". Several researchers have criticised both the low number of scale points and the nature of the written descriptions of the degree of agreement (Gehlbach, 2015[54]).

It is therefore essential to find the right balance between fewer versus more scale points and fully-labelled versus unlabelled response scales. The need to translate the survey into a large number of languages poses a further challenge to extending response scales beyond four scale points. If feasible, alternative versions of questionnaire items (such as fully labelled versus incompletely labelled response options) should be tested and compared in future PISA field trials.

Select measures that maximise cross-cultural comparability

Another challenge is developing well-being questions that allow for the comparison of the resulting data across cultural and national borders and across subgroups within a country – a challenge that has been well documented for PISA and other international surveys (Kyllonen and Bertling, 2014[56]). Classical measurement approaches based on self-reports often suffer from limited inter-individual comparability due to individual- or group-specific response styles. There is ample evidence that responses to even seemingly objective questions can often not be interpreted as objective indicators and display only limited comparability across countries, before accounting for differences in response style (Kim, Schimmack and Oishi, 2012[57]). For example, anchoring vignettes, or brief descriptions or anecdotes that define various points on a scale, have been successfully applied to increase consistency across respondents (Angelini et al., 2014[58]; Kristensen and Johansson, 2008[59]; Kyllonen P. C. and Bertling J. P., 2014[60]; Salomon, Tandon and Murray, 2004[61];

van Soest et al., 2011[62]). However, alternative survey methods to increase cross-country and inter-individual comparability (which also include situational judgment tests and forced choice) might be less valuable for well-being as many components of this construct explicitly involve a subjective component that, by definition, is influenced by cultural norms and the respondent's personality. (White, 2007[63]) identified 'culture' as a key influence on the way one's perception of well-being is constructed and therefore suggested that well-being should be understood as a process grounded in a specific time and place. Thus, although the well-being construct proposed herein will capture culture-specific aspects of student responses, PISA can maximize cross-cultural comparability by choosing clear, translatable, and where possible quantifiable response formats and, particularly when using anchoring vignettes, by including short definitions as part of the item stem whenever a question involves certain reference points that might limit cross-cultural comparability.

Consider item formats beyond traditional self-reports

Measuring well-being more comprehensively requires a survey approach that goes beyond the self-report questionnaires traditionally used in large-scale assessments. One such protocol is the day reconstruction method and event reconstruction method, which assesses how students spend their time (especially time outside of school) and which samples their experienced well-being during various activities. The proposed questions build upon a time use module proposed for PISA 2015 (Bertling and Kyllonen, 2012[64]), which was partly implemented in the 2015 main survey, as well as on methods and question formats recommended by the authors of the original day reconstruction and event reconstruction methods (Grube et al., 2008[65]; Kahneman et al., 2004[3]; Schwarz, Kahneman and Xu, 2009[66]).

Consider alternative questionnaire designs to reduce respondent burden

The large sample sizes in large-scale assessments make viable the use of matrix sampling approaches to reduce respondent burden within the constraints of overall testing time while maintaining content coverage across relevant areas. These approaches provide different respondents with different sets of items. This is already standard practice for subject-area tests in large-scale educational assessments (Comber and Keeves, 1973[67]; OECD, 2016[68]) and has more recently been proposed as a potentially viable alternative to fixed questionnaires, where all students receive the same items. A three-form matrix sampling design was applied to the student questionnaire in PISA 2012, allowing the questionnaire content to increase by 33 percent (Klieme and Kuger, 2014[69]). Similarly, a design with ten partly overlapping questionnaire booklets was implemented in the 2013 pilot of the 2015 NAEP Technology and Engineering Learning (TEL) assessment (Almonte et al., 2014[70]). NAEP now routinely uses matrix sample questionnaire designs for their large-scale pilots. New research related to PISA 2021 further compared the feasibility of different possible matrix sampling approaches for operational administration (Bertling and Weeks, 2018[71]). New analytical approaches will be required to analyse the incomplete data from these approaches.

Unfortunately, research findings to date are inconclusive regarding the risks and benefits of questionnaire matrix sampling in practical scenarios: while many researchers reported substantial increases in content coverage with a very small to negligible impact on the overall measurement model (Adams, Lietz and Berezner, 2013[72]; Almonte et al., 2014[70]; Kaplan and Wu, 2014[73]; Monseur and Bertling, 2014[74]), others have raised methodological concerns about possible biases (von Davier, 2014[75]). Application of mass imputation for all questions that were not administered to a given student might address these issues by creating full datasets (albeit with large proportions of imputed data). This approach has been explored in research contexts (Kaplan and Wu, 2014[73]) but has so far not yet been implemented in any large-scale assessment. As argued in Bertling, Borgovoni and Almonte (2016[76]), it would be beneficial to explore a matrix sampling design for survey questionnaires, and in particular, the well-being questionnaire. This could allow for the exploration of a larger number of facets of the well-being construct (e.g., a larger set of affective states in the experienced well-being questionnaire) without increasing individual student burden.

Suggested quality of life indicators

Quality of Life as a Whole

Well-being with regard to life as a whole, or overall life satisfaction, is often used as a single indicator for individual subjective well-being. Despite the importance of including overall life satisfaction as an important yardstick in any well-being instrument, it does not sufficiently capture the more specific dimensions of one’s quality of life (e.g., the quality of relationships). Unfortunately, no direct objective indicators for well-being with regard to life as a whole are available; all of the indicators below are subjective.

Life Evaluation and Life Satisfaction

Life satisfaction, an evaluation of an individual’s quality of life, is an important aspect of well-being (Diener et al., 1999[13]). Classical approaches of assessing subjective well-being rely mostly on unanchored self-report ratings: respondents are asked something similar to “Overall, how satisfied are you with life as a whole these days?” and must answer on a scale from 0 to 10 with zero indicating “not at all satisfied” and 10 indicating “completely satisfied”. The questionnaire can also ask respondents to evaluate their satisfaction with specific domains of their lives (e.g., health, personal relationships and security) (Figure 8.1). Scores in these domains can then be treated as stand-alone scores or aggregated into an overall index representing overall satisfaction across all domains.

Two alternative approaches, sometimes considered to be equivalent or interchangeable, are widely used in the well-being literature. One approach is the Cantril ladder (Cantril, 1965[77]), used in major international surveys such as the Gallup World Poll and the Gallup Student Poll. By asking respondents to indicate where, on a ladder with steps from 1 (“the worst possible life”) to 10 ("the best possible life"), they see themselves at the current point in time, the question targets the *evaluative* aspect of well-being, or how individuals perceive or evaluate their life. The alternative approach focusses on *satisfaction* instead of evaluation by asking a question similar to "How satisfied are you with your life overall these days?". This is one of the core well-being questions recommended by the OECD guidelines on measuring subjective well-being (OECD, 2013[46]).

Empirical findings are somewhat inconclusive about the comparative validity of the two distinct yet related approaches, particularly for adolescents (Casas et al., 2012[47]). Indeed, there are so far no large-scale studies that systematically compare the nuanced differences between the life evaluation and life satisfaction approaches. However, several large studies have observed that the life evaluation approach tends to create data that varies more within samples and produces average scores closer to the midpoint on the scale. On the other hand, the direct life satisfaction approach elicits skewed score distributions with a mean noticeably above the midpoint of the scale. The PISA 2015 field trial, which tested both questions in all participating countries, confirmed these findings. No clear advantages of the life evaluation approach over the life satisfaction approach were found, apart from slightly better differentiation across the scale. Moreover, the two questions correlate strongly.

In the end, the OECD has recommended using the 11-point life satisfaction scale as part of the core well-being module in its guidelines on measuring well-being (OECD, 2013[46]) for several reasons. First, the question is shorter and therefore requires less time to answer. Second, the question is easier to understand because of its lower cognitive burden and reduced reading load. Lastly, the life satisfaction question is less intrusive then the life evaluation question because it does not explicitly introduce the concept of social rank or potentially imply comparison with other individuals, which might elicit negative emotions in some respondents. The 11-point satisfaction question was introduced for the first time in the 2015 PISA student questionnaire. It is therefore recommended that the 0-10 life satisfaction question be included as a core question in the PISA 2018 well-being questionnaire. If space allows, posing the life evaluation question could further increase the robustness of the measure.

While several multi-item scales for overall subjective well-being have been proposed (Adelman, Taylor and Nelson, 1989[78]; Huebner, 2001[29]; Rees and Main, 2015[52]), their incremental value over the single-item indicators described above is unclear. Given space and cognitive burden constraints, it is therefore recommended that no multi-item scale on students' overall life evaluation or satisfaction be included in the well-being questionnaire. Rather, aggregating satisfaction ratings across multiple domains into a potential composite life satisfaction index would maintain an acceptable questionnaire length while covering all of the important facets of the well-being construct. Indeed, well-being research has also moved from studying overall subjective well-being to domain-specific subjective well-being (Elmore and Huebner, 2010[79]; Gilman and Huebner, 2000[80]; Long et al., 2012[81]; Tian, Wang and Huebner, 2015[82]). Empirical findings show that adolescents' domain-based reports on aspects of well-being and satisfaction (such as family and school) show greater validity than global life satisfaction reports (Haranin, Huebner and Suldo, 2007[83]).

Affect/Emotional Well-being

Several subjective indicators for emotional well-being are proposed, drawing on (a) both positive and negative affect as indicators of emotional well-being and (b) experienced well-being questions with regard either to very specific activities or to emotional states experienced over extended periods of time.

One way to measure affect is to ask individuals whether or to what degree they have felt specific emotions during a certain period, through questions such as "Overall, how happy did you feel yesterday?" or "Overall, how angry did you feel yesterday?". This corresponds to Watson's positive and negative affect schedule (PANAS) (Watson, Clark and Tellegen, 1988[84]), which has been used extensively in psychological research. Hedonic balance, defined as the difference between positive and negative affect, has been proposed as a measure of overall emotional well-being (Watson, Clark and Tellegen, 1988[84]; Schimmack, Diener and Oishi, 2002[85]); however, there is no agreement as to the specific emotions that need to be sampled. Laurent et al. (1999[86]) presented a version of the Positive and Negative Affect Scale for children (PANAS-C). However, there is not yet sufficient research that shows that a PANAS-type measure works well in an international comparison of students.

Instead, the KIDSCREEN-10 measure could be adapted for use in PISA (Ravens-Sieberer et al., 2014[23]). KIDSCREEN conceptualises quality of life as a multidimensional construct with physical, emotional, mental, social and behavioural components. The short ten-item measure is Rasch-scalable (Ravens-Sieberer et al., 2010[87]) and has been used by the HBSC survey since 2005 (Currie et al., 2009[88]). If space in the questionnaire allows, the WHO-5, a short 5-item affective well-being measure widely used and well-established in clinical research, might be an additional benchmark to link PISA with other surveys (see (Topp et al., 2015[22]) for a recent literature review). In direct comparison, however, the KIDSCREEN-10 should be given priority given its prior international use with students and its less clinical and diagnostic focus.

The cardinal method for measuring experienced well-being with regard to specific events or behaviours is the day reconstruction method, or the DRM (Kahneman et al., 2004[3]). In this method, respondents are asked to revisit a previous day and report in detail on their activities as well as the emotional states they experienced. The original DRM is not viable for inclusion in PISA given its time and scoring requirements. However, a PISA well-being questionnaire can ask students to report on the emotional states experienced during events of interest to PISA, such as specific classes, time spent doing homework, leisure activities with friends or time spent with parents or guardians. (This is similar to the event reconstruction method (Grube et al., 2008[65]), which itself is based on the day reconstruction method.) Affective states can then be related to specific before-school, at-school, and after-school activities. New questions might also be developed through the psychological concept of flourishing (Seligman, 2012[89]), or engagement and flow-related emotional states, because they relate more directly to academic achievement; indeed, such emotional states include the feelings of being challenged or inspired.

The specific proposed event reconstruction questions targeted at measuring experienced well-being are referenced in the following sections based on their classification under one of the well-being dimensions (self, school environment or out-of-school environment).

Self-Related Well-Being

The first broad domain of quality of life as a whole is quality of life in regards to the student as an individual, with the three sub-dimensions of health, education and skills, and psychological functioning.

Health

To capture the overall health construct, data should be collected on a set of key objective and subjective indicators,[2] some of which have already been introduced in the PISA 2015 student questionnaires or been field trialled (Bertling and Kyllonen, 2012[64]).

Objective Indicators

Quetelet's index, defined as Weight/Height^2 and better known as the *body mass index* BMI (Garrow J. S. and Webster J., 1985[90]), is a key health indicator and is widely used in international studies in both adult and youth populations. The BMI is an indicator of being either overweight or obese, growing health problems among adolescents in many countries (Lobstein, Baur and Uauy, 2004[91]; Haug et al., 2009[92]; Rokholm, Baker and Sørensen, 2010[93]; WHO, 2010[94]). Previous research has shown that being overweight is correlated with behaviours associated with health risks (such as skipping breakfast, being less physically active or watching more television), a lower overall quality of life (Haug et al., 2009[92]; Must and Tybor, 2005[95]; Williams et al., 2005[96]) and being a victim of bullying (Janssen et al., 2004[97]). Research among adolescents further suggests that dieting and unhealthy weight control behaviours are related to significant weight gain over time

(Neumark-Sztainer et al., 2012[98]). The BMI may be used as a screening tool to identify potential weight problems in individuals and to track the degree to which populations are overweight or obese. However, it should not be used as a single diagnostic tool for body fat or overall student health (Nihiser et al., 2007[99]), and interpretations of the BMI need to account for potential differences across racial or ethnic groups (James, Chen and Inoue, 2002[100]) or genders (Dupuy et al., 2011[101]). To better account for potential inaccuracies in student-provided weight and height information, the two qualifying questions "When did you last weigh yourself?" and "When did you last measure your height?" can be added. Both of these questions are currently used in the HBSC survey.

Participation in physical exercise does not only contribute positively to student health but also protects against excessive body image concerns (Gaspar et al., 2011[102]) and long-term negative physical and mental health outcomes, particularly as habits established in adolescence are likely to be carried through into adulthood (Malina, 1991[103]; Hallal et al., 2006[104]; Iannotti et al., 2009[105]; McMurray et al., 2008[106]; Sibley and Etnier, 2003[107]). Children who play sports or exercise more frequently report higher levels of subjective well-being (Abdallah et al., 2014[108]). Moreover, research indicate that physical activity may also improve cognitive performance (Martínez-Gómez et al., 2011[109]; Sibley and Etnier, 2003[107]). The World Health Organization recommends that children participate in at least 60 minutes of moderate-to-vigorous physical activity daily (Strong et al., 2005[110]). A small set of questions about students' physical exercise habits were introduced in PISA 2015, covering both moderate and vigorous exercise, participation in physical education classes, and the physical exercise performed the previous day. These questions are also valuable to the PISA well-being questionnaire. In addition, information regarding a student's typical duration of sleep and his or her behaviours associated with health risks might be collected via a brief day reconstruction checklist.

Subjective Indicators

Subjective indicators provide important information about the overall health construct beyond the objective indicators discussed above. PISA should measure such subjective indicators through instruments that have been validated in other contexts, if possible. These subjective indicators include one's perception of and satisfaction with one's body image (Rudd and Lennon, 2000[111]), satisfaction with one's sleep, perceived overall health, psychosomatic complaints and satisfaction with one's overall health.

Research indicates that girls report greater dissatisfaction with their body image than boys (Marcotte et al., 2002[112]) and being overweight increases the likelihood that adolescents engage in unhealthy weight-reduction activities and report substance abuse, risky sexual behaviour and poor mental health (Kaufman and Augustson, 2008[113]; Kvalem et al., 2011[114]; Verplanken and Velsvik, 2008[115]; Ojala et al., 2007[116]; Currie et al., 2012[117]).

The HBSC includes a short checklist of symptoms that can be used as a non-clinical measure of mental health. This checklist includes both psychological complaints (e.g., nervousness or irritability) and somatic complaints (e.g., headaches or backaches), both of which are strongly related to each other (Petersen et al., 1997[118]; Brosschot, 2002[119]) and to important facets of the overall well-being construct (Petersen et al., 1997[118]; Vingilis, Wade and Seeley, 2002[120]; Hetland, Torsheim and Aarø, 2002[121]; Ravens-Sieberer et al., 2008[122]).

Education and Skills

A student's education and skills, his or her self-perceptions of his or her ability to perform specific academic tasks, and his or her general confidence in his or her own capabilities are important aspects of the overall well-being construct.

Objective Indicators

Objective indicators for students' knowledge and skills come from the cognitive assessments in PISA and not further elaborated upon here.

Subjective Indicators

Questions about students' beliefs in their own competency or their academic self-efficacy (Bandura, 1997[123]) directly address competence, one of the three main basic psychological needs identified in self-determination theory (Ryan and Deci, 2000[124]). Research has shown that adolescents' perceptions of their school performance and their own competency are correlated with higher perceived health and well-being (Suldo, Riley and Shaffer, 2006[125]; Ravens-Sieberer, Kökönyei and Thomas, 2004[126]), higher life satisfaction (Suldo and Huebner, 2006[127]), and lower rates of bullying (Nansel et al., 2001[128]). Qualitative studies further point to positive attitudes (Edwards and Lopez, 2006[129]), personal strengths (Shikako-Thomas et al., 2009[130]), and a positive self-image (Helseth and Misvær, 2010[131]) as important determinants of student well-being. The PISA academic self-efficacy questions included in the student questionnaire might provide data on students' perceptions of their competency. Questions about students' satisfaction with their own knowledge and skills and their self-confidence are proposed as part of a question set focusing on this domain.

Psychological Functioning

Psychological functioning, also referred to as "eudaimonic well-being" or "flourishing", has been proposed as an additional component of the subjective self-related well-being construct (Seligman, 2012[89]). Psychological functioning is concerned with people's sense of meaning, purpose and engagement. It is related to "flow", defined as a gratifying experiential state that can "make life worth living" (Csikszentmihalyi, 1975[132]; Csikszentmihalyi and Csikszentmihalyi, 2006[133]), and is concerned with personal growth, self-expression and the pursuit of meaningful goals (Ryan and Deci, 2001[134]).

Some researchers consider psychological functioning to be part of the overall subjective well-being construct (Seligman, 2012[89]; Kern et al., 2015[135]), while others do not (The Children's Society, 2015[12]). Furthermore, while there is large consensus on two of the key building blocks of the subjective well-being construct (life satisfaction and affect), there is less consensus on the nature and role of psychological well-being. This might be partly due to the overlap of the psychological well-being construct with other aspects of well-being: for instance, questions targeting psychological well-being (e.g., "I like being the way I am") are very similar to questions measuring overall subjective well-being (e.g., "My life is just right") (Huebner, 1991[136]). Indeed, some authors have conceptualised psychological functioning as a higher-level construct that includes both positive and negative affect.

Four main facets of psychological functioning described in the literature are competence, autonomy, meaning/purpose and optimism (OECD, 2013[46]). This framework includes psychological functioning as part of the self-related dimension of well-being and not as a measure of overall (life as a whole) well-being, as psychological functioning focuses explicitly on the self and does not encompass environmental factors; the variables described as potential indicators of overall well-being did not have that level of specificity. Large-scale data gathered through PISA might provide empirical evidence as to whether psychological functioning variables relate more strongly to overall or to self-related well-being.

Several questionnaires for the assessment of psychological functioning have been proposed, similar to those that assess concepts such as personality, self-concept, locus of control and attribution (Huppert et al., 2009[137]; Kern et al., 2015[135]). Respondents are asked questions such as “I am always optimistic about my future” or “I am free to decide for myself how to live my life”. Ryff (1995[138]) proposed six dimensions of psychological functioning: self-acceptance, positive relations with others, personal growth, purpose in life, environmental mastery (the ability to control the environment around oneself or to create a context suitable to one’s needs) and autonomy. These dimensions overlap considerably with various components of the proposed well-being framework as well as with many of the currently-used attitudinal and self-related questions in the PISA student questionnaire. For instance, the PISA 2012 and PISA 2015 questions on perseverance and openness to problem solving overlap with the psychological well-being dimensions of personal growth and autonomy. Openness to new experiences is a particularly good predictor of psychological functioning among adolescents (Bassi et al., 2014[139]).

Psychological functioning is an important additional facet of well-being. However, it could be measured via a potential composite index instead of through a separate unidimensional psychological well-being scale based on a unique set of questions. This index could be created from questionnaire items that capture various subjective perceptions, such as perceptions of competence, knowledge and skills; autonomy, personal freedom and opportunities; meaning and purpose; and relationships.

School-Related Well-Being

Students spend a large proportion of their time at school. Their experiences and relationships at school have an important impact on their perceived quality of life; indeed, schools not only nurture academic achievement but also promote students’ health and well-being (Jourdan et al., 2008[140]). A positive school climate is associated not only with higher academic achievement but also with better self-reported student health, well-being and health behaviours (Cohen et al., 2009[141]; Jia et al., 2009[142]), lower perceived stress (Torsheim and Wold, 2001[143]) and more positive student reactions to demands at school (such as better stress management) (Huebner et al., 2004[144]). Although researchers have called for specialised measures of subjective well-being in school to account for potential differences between well-being at school and overall well-being (Huebner et al., 2005[145]), only a few studies have so far explicitly focused on examining students’ subjective well-being at school (Huebner, 2001[29]; Epstein and McPartland, 1976[146]; Karatzias, Power and Swanson, 2001[147]; Tian, Wang and Huebner, 2015[82]).

Two main sub-dimensions are proposed for school well-being: social connections and schoolwork. A few additional potential indicators are also outlined. Most proposed indicators are subjective as they concern student perceptions of their school life and their school environment rather than objective circumstances. The questionnaire should especially focus on students’ social connections and workload instead of on school infrastructure and security as other indicators might be available for this area (e.g., school records).

Social Connections at School

Social connections are students' social relationships with teachers and with other students and, more generally, general patterns of student interactions and the school climate. These factors might foster a sense of belonging to school – the feeling of being accepted, respected, included and socially supported in the school environment (Goodenow, 1993[148]) – or a sense of discrimination and loneliness. PISA has included a sense of belonging at school scale in its main student questionnaire for several assessment cycles. The sense of belonging at school correlates with measures of life satisfaction as well as experienced emotional well-being (Gilman and Anderman, 2006[149]; Millings et al., 2012[150]). Moreover, prior research has also found that student-teacher relationships and classmate support are important predictors of student adjustment and adolescent life satisfaction (Reddy, Rhodes and Mulhall, 2003[151]; Suldo et al., 2009[152]).

Findings from the HBSC show that students who perceive their school as supportive more frequently report positive health behaviours and health and well-being outcomes (Ravens-Sieberer, Kökönyei and Thomas, 2004[126]; Due et al., 2003[153]; Freeman et al., 2009[154]; Vieno et al., 2007[155]). Students who indicate that they like school are less likely to be victims of bullying (Harel-Fisch et al., 2011[156]), take fewer sexual risks (Dias, Matos and Gonçalves, 2005[157]) and less frequently report drug use (Fletcher, Bonell and Hargreaves, 2008[158]). In contrast, disliking school is related to an increased risk of dropping out (Archambault et al., 2009[159]) and a higher prevalence of health problems (Shochet et al., 2006[160]).

Bullying, defined as negative physical or verbal actions that have hostile intent, cause distress to victims, are repeated and involve a power differential between perpetrators and victims (Craig, Pepler and Atlas, 2000[161]; Mahady Wilton, Craig and Pepler, 2000[162]; Olweus, 1991[163]), has received increasing policy attention in recent years (Farrington et al., 2011[164]). Victims of physical or mental bullying, for example, are more likely to exhibit poor school performance or to drop out of the education system (Moore et al., 2008[165]; Currie et al., 2012[117]; Olweus, 1991[163]; Glew et al., 2008[166]; Olweus, 1994[167]) to experience depression, anxiety, loneliness and a range of psychosomatic symptoms (Olweus, 1991[163]; Craig, 1998[168]; Nansel et al., 2001[128]; Due et al., 2005[169]); and to abuse drugs and alcohol (Molcho, Harel and Dina, 2004[170]). Adolescents who have recently been bullied tend to report levels of subjective well-being substantially below the population average and research suggests that the effects of bullying on well-being are far stronger than the effects of other many contextual factors (The Children's Society, 2015[12]).

School-based bullying prevention programmes are very often successful (Currie et al., 2012[117]). Results from major well-being and health studies further suggest that reducing and preventing bullying could be strongly linked to improving students' well-being not only in adolescence but also in adulthood (Bond et al., 2001[171]; Clapper et al., 1995[172]; Ttofi et al., 2011[173]).

The HBSC also recommends that cyberbullying, or bullying involving modern digital communication technologies, be investigated (Ahlfors, 2010[174]). Furthermore, the perspective of the bullied can be supplemented by the perspectives of perpetrators and bystanders; questions to these groups could also be included in a well-being module (Rigby and Slee, 1991[175]; Veenstra et al., 2005[176]). Indeed, perpetrator behaviours are also associated with a range of negative health, social and academic behaviours (Glew et al., 2008[166]; Nansel et al., 2001[128]; Harel, 1999[177]; Olweus, 2011[178]; Farrington et al., 2011[164]).

Objective Indicators

Questions in the PISA survey on students' experiences with bullying, introduced in the 2015 cycle, are objective indicators of negative or dysfunctional social relationships and the lack of social integration. These questions are objective because students are asked to state in which of the listed specific, clearly described and quantifiable behaviours they have engaged. Other instruments that measure bullying have been described in the literature (Olweus, 1996[179]) and used in large-scale surveys (e.g. the HBSC).

Subjective Indicators

PISA student questionnaire items on student-student and student-teacher relationships, sense of discrimination and sense of belonging are key subjective indicators of students' connections at school. Students' sense of belonging and social connectedness at school are positively correlated to relatedness, one of the three main basic psychological needs in self-determination theory (Ryan and Deci, 2000[124]). Perceived discrimination, on the other hand, can have detrimental effects on student well-being (Schmitt et al., 2014[180]).

Schoolwork

Research among adults shows that well-being and health suffer when individuals are subjected to extreme working conditions. There are not yet any comprehensive findings for adolescents, but it is expected that extreme hours of school might have negative consequences (Karasek and Theorell, 1992[181]). Feeling pressured or stressed by schoolwork may lead to more frequent health-compromising behaviours such as smoking, drinking alcohol and drunkenness; more frequent health complaints such as headache, abdominal pain and backache; psychological problems such as feeling sad, tense or nervous (Torsheim and Wold, 2001[143]; Simetin et al., 2011[182]); and lower overall life satisfaction (Ravens-Sieberer, Kökönyei and Thomas, 2004[126]). However, students may prefer different subjects and activities, making it imperative to consider an entire day or week instead of simply one moment in time when examining well-being related to schoolwork. Students' workload and time spent at school is one part of the proposed school/life-balance composite index.

Objective Indicators

Objective indicators of student well-being related to schoolwork include the time spent on school-related activities: hours spent at school, spent on the way to and from school, and spent on homework and studying for school. The main student questionnaire already asks about some of these variables; additional questions can be asked to fill the remaining gaps.

Subjective Indicators

The subjective indicator of student well-being related to schoolwork proposed by this framework is students' self-reported emotions experienced during selected episodes associated with schoolwork.[3] For instance, students who report negative emotions in school more frequently are more likely to withdraw from school, to show antisocial behaviour, and to abuse drugs (Roeser, 2001[183]). Affective states that are especially relevant to the school environment should be prioritised. In particular, students can be asked about their emotions during mathematics, language of instruction and art/creativity classes (chosen because they represent a broad range of contents and classroom practices) and while doing homework or studying for school. For reasons of practicality, affective states are limited to a short set of both positive and negative affective states. Matrix sampling approaches would allow a larger set of events and affective states to be sampled in the future.

Other Potential Indicators

Students' perceptions of their safety at school and on their way to school, as well as their satisfaction with their safety and the general infrastructure of the school are further facets of student well-being at school. Information on these facets could come from subject-specific survey questions, school records of reported incidents and police/safety statistics of the area around the school. In addition, aggregate measures of the prevalence of bullying or other disciplinary problems in the school could be aspects of this sub-dimension.

Well-Being Outside of School

Students' experiences in their out-of-school environment constitute the third broad well-being dimension identified in this framework. Key sub-dimensions of out-of-school well-being are students' social connections outside of school (including their friendships and their relationships with parents), their material living conditions and their leisure-time activities.

Social Connections Outside of School

In addition to students' social connections at school, relationships with parents and other family members and friendships that take place outside of school are important factors for students' well-being. Research shows that having high-quality peer relationships has positive effects on adolescent health (Barker and Galambos, 2003[184]; Zambon et al., 2009[185]). On the other hand, having fewer friends in adolescence result in a lack of opportunities to learn social skills (Gifford-Smith and Brownell, 2003[186]; Sullivan, 1953[187]), potentially leading to lowered life satisfaction and more frequently experienced negative affect and bullying experiences (Larson and Richards, 1991[188]). Other findings point to the importance of family relationships and friendships as two main factors that determine self-satisfaction (Edwards and Lopez, 2006[129]; Suldo et al., 2013[189]). Indeed, research indicates that self-reported ease of communication with one's parents is associated with a range of positive health outcomes (Currie et al., 2012[117]) and that children who report talking more frequently to family members about things that matter to them also tend to report higher levels of subjective well-being (Abdallah et al., 2014[108]).

Social connections outside of school also include student's sense of and identification with their community (Davidson and Cotter, 1991[190]; Farrell, Aubry and Coulombe, 2003[191]; Prezza et al., 2001[192]; Prezza and Costantini, 1998[193]).

Objective Indicators

Time spent on activities with friends and parents may serve as objective indicators of student's social connections outside of school. This information can be collected via a short day/event reconstruction protocol focusing on selected key events, such as having dinner with one's parents and spending time with friends outside of school. These questions about a specific day can be complemented by a short set of questions from the HBSC survey. These questions inquire about the number of days per week students spend time with friends right after school and in the evenings, or the number of days they communicate via electronic media; the timespan of one week reduces the risk that a single outlier day might bias results. A final objective indicator of students' social connections outside of school is where they met these connections, whether at their current school, a previous school, in the neighbourhood or through their family.

Subjective Indicators

In order to capture students' subjective perceptions about their social connections, and their affect and satisfaction regarding these relationships, PISA can include a series of questions based on those already used to similar effect in the HBSC and KIDSCREEN-10 surveys and, as a complement to these, it can measure experienced well-being with a short set of event reconstruction questions.

Proposed questions on friendships cover the number of perceived close female and male friends (HBSC),[4] students' satisfaction with the number of friends they have, the degree to which students felt they had fun with their friends over the past week (KIDSCREEN-10), the perceived ease with which students talk to their best friend about things that bother them (HBSC), and students' experienced affect while spending time outside the home with their friends (newly developed for PISA following the event reconstruction approach). Capturing information beyond the mere number of friends is important as the quality of relationships is a stronger predictor of well-being than their quantity (The Children's Society, 2015[12]).

Proposed questions on the subjective quality of relationships with parents, guardians or other family members include the degree to which students felt they were treated fairly by their parents over the past week (KIDSCREEN-10); the degree to which students think their friends are accepted by their parents (HBSC); the perceived ease of talking to their parents, stepparents or elder siblings about things that bother them (HBSC); students' perceptions of their parents' or guardians' general behaviour and attitude towards them (HBSC); and students' experienced affect while having dinner at home with their parents (newly developed for PISA following the event reconstruction approach).

Material Living Conditions

A student's material living conditions, as measured by his or her family's socio-economic status (SES Expert Panel, 2012[194]), constitute an important determinant of overall well-being with small but robust positive associations between household income and adolescent subjective well-being (Rees, Pople and Goswami, 2011[195]). Children from highly affluent families also tend to report better health (Torsheim et al., 2004[196]; Richter et al., 2009[197]), and students' basic needs and desires are more likely to be met when they live in rich nations (Tay and Diener, 2011[198]; Diener et al., 2010[199]). Moreover, the literature indicates that poverty, and particularly perceived poverty, is a crucial limiting factor for students' well-being (Goswami, 2014[200]). Research indicates that child-reported material deprivation explained a larger proportion of the variation in children's subjective well-being than overall family socio-economic status did and that children "tend to talk about money and possessions in relative terms – e.g., having 'enough' or 'the same amount' as rather than 'more' than – others so that they fit in and are not excluded from things that others can do" (The Children's Society, 2015[12]). These findings point to the importance of subjective socio-economic status (Diemer et al., 2012[201]; Quon and McGrath, 2014[202]), which has not received as much attention as its objective counterpart.

Objective Indicators

PISA measures students' objective material living conditions through a composite *index of economic, cultural and social status* (ESCS) derived from questions about general wealth (based on several proxy variables including home possessions), parental education and parental occupation. Although no changes to the ESCS are currently envisaged, a number of additional indicators, currently used in other surveys, could add substantial value to the current ESCS indicator and could be included in the future. These include whether students receive pocket money (used in the HBSC), whether they have been on a vacation with their family (used in the HBSC and Children's Worlds), and whether they have had to go to bed hungry (used in the HBSC). The broader concept of unmet needs could further inform the measurement of poverty and deprivation as risk factors for student well-being (Diemer et al., 2012[201]).

Subjective Indicators

An indicator of subjective material living conditions would capture students' subjective perceptions of their economic standing. It would focus on perceptions of the adequacy of one's standard of living (Conger, Conger and Martin, 2010[203]; Mistry and Lowe, 2006[204]) as well as the psychological experiences of material deprivation and hardship (Iceland, 2003[205]; Mayer and Jencks, 1989[206]; Gershoff et al., 2007[207]). Research on poverty and aspirations (Dalton, Ghosal and Mani, 2015[208]; Ray, 2006[209]) suggests that poverty and the inability to aspire to change one's life for the better may lead to the underutilisation of available resources, and that subjective perceptions of poverty might play an equally important or maybe an even larger role in this than actual poverty (The Children's Society, 2015[12]). Moreover, the perception of financial constraints is strongly associated with adolescent health outcomes (Quon and McGrath, 2014[202]).

These findings underline the importance of paying attention to the subjective "experience" of poverty in addition to objective measures of socioeconomic status. Questions on how well off students believe their family to be and whether they worry about their family's financial situation could also be informative; the latter is already implemented in Children's Worlds.

Leisure Time

An individual's well-being depends on his or her ability to pursue activities that he or she enjoys and to spend time with his or her family and friends (Rees, Goswami and Bradshaw, 2010[210]; Abdallah et al., 2014[108]). This takes place during leisure time, which can be defined for students as the time awake not spent in school, on schoolwork, on the commute to school,[5] or on other obligations. Indicators of leisure time use and emotions experienced during this time are therefore important elements of overall well-being.

Objective Indicators

Both the total time available for leisure as well as how students use this time are objective indicators of students' leisure time. A proxy of the former can be derived as the difference between the hours awake minus hours spent at school, spent on the way to and from school, and spent on homework and studying for school. The main student questionnaire already asks about some of these variables; additional questions on hours awake, hours spent at school, and hours spent on the commute to school will fill the remaining gaps. A short day reconstruction protocol focusing on selected activities, such as watching television or videos, reading a book, browsing/reading on the Internet, spending time on chat/social networks/e-mail, playing video games, meeting friends, talking to parents, eating or practicing a sport, can provide information on how students use their leisure time. These activities were included in an abbreviated time-use protocol introduced to PISA 2015

(Bertling and Kyllonen, 2012[64]) and by other studies concerned with student time use (e.g. Children's Worlds and the American Time Use Survey) (Rees and Main, 2015[52]; Bureau of Labor Statistics, 2015[211]; Carson, Staiano and Katzmarzyk, 2015[212]; Larson and Verma, 1999[213]).

The use of social media should be included as a separate activity in the time use protocol, given its increasing part in the lives of adolescents. Research suggests that excessive use of social media may lead to poorer health, sleeping habits, loneliness and greater engagement in risky behaviours (Prezza, Pacilli and Dinelli, 2004[214]; Punamäki et al., 2007[215]; Koivusilta, Lintonen and Rimpelä, 2005[216]). Moreover, spending more than two hours per day on social networking sites was associated with reporting poorer mental health and higher levels of psychological distress among adolescents (Sampasa-Kanyinga and Lewis, 2015[217]).

Subjective Indicators

A combination of event reconstruction questions and a set of questions asking students to report on how they perceive and how satisfied they are with their use of time (in general) and leisure time (in particular) will provide subjective indicators of the quality of students' leisure time.

Event reconstruction questions could examine students' experienced well-being during breaks between classes at school or time spent outside of their home with friends. A short set of both positive and negative affective states would cover key emotions while keeping student burden low. Matrix sampling approaches for questionnaires would allow a larger set of events and affective states to be investigated; unfortunately, such matrix sampling will not be implemented for the 2018 PISA well-being questionnaire. Students can also be asked about their overall satisfaction with their use of time and what they do in their free time; their satisfaction with specific activities engaged in on the previous day (as part of the day reconstruction protocol); and their evaluation of the amount of time they have for themselves (already done in KIDSCREEN-10).

Other Potential Indicators

Students' perception of and satisfaction with their safety at home, safety in their neighbourhood and opportunities in their neighbourhood are also relevant to their out-of-school well-being. Unfortunately, due to space constraints, additional questions covering these themes must be prioritised. Additional information on this framework component might be drawn from other sources, such as records about the local area or geographical region a student is living in, if available.

Possible Composite Indicators

In addition to the proposed indicators representing individual cells of the framework, composite indices covering multiple cells of the framework might be of policy interest. Aggregating indicators into composite indices risks increased opaqueness as to which are the most critical areas of well-being (UNICEF, 2007[9]). However, a number of previous studies have proposed composite well-being indicators that are already widely used in applied contexts (Bradshaw, Hoelscher and Richardson, 2007[218]; Land, Lamb and Mustillo, 2001[219]; Land et al., 2007[50]; Moore et al., 2008[165]; Bradshaw et al., 2009[220]), and creating such indicators in addition to more specific indices may facilitate measuring progress over time and comparisons across sub-groups (Ben-Arieh, 2008[25]). Some potential composite indicators that are seen as especially promising for policy and practice include:

- An index of the overall quality of life;
- An index of overall subjective well-being;
- An index of overall emotional well-being, created by aggregating the subjective indicators of affective well-being across all content dimensions;
- An index of work/school-life balance, created by aggregating the well-being related to schoolwork and to leisure time;
- An index of overall social well-being, created by aggregating the well-being related to social connections at school and outside of school.

Notes

[1] The other word that the 14- and 15-year-olds discussed was "do", as in "things to do", "something to do", or "nothing to do"; this word came in as the fourth-most commonly discussed.

[2] Note, PISA will have to rely on student self-reported data for both sets of indicators. The key difference is that objective indicators are clearly quantifiable and behavioural indicators that require minimal judgment or interpretation on the part of the respondent (for example, a student does not need to provide an interpretation of his/her weight when providing his/her weight in kilograms).

[3] However, please note that this section discusses subjective indicators of well-being with reference to school in general, not schoolwork in particular. The indicators described here may also be examined in relation to just schoolwork.

[4] These questions are considered to be subjective as individuals might differ in their perception of how close "close friends" are (Keller, 2004[221]). A short qualifying statement about the definition of "close friends" should be given at the beginning of the question to maximize the comparability of the question across individuals and cultures.

[5] There is overwhelming evidence that long and difficult commutes for adults are typically perceived as unpleasant and are associated with reduced subjective well-being (Kahneman et al., 2004[3]; Office for National Statistics, 2014[222]).

References

Abdallah, S. et al. (2014), *Ways to Well-Being: Exploring the Links Between Children's Activities and Their Subjective Well-Being Research Report*, The Children's Society, London, http://eprints.whiterose.ac.uk/82855/. [108]

Adams, R., P. Lietz and A. Berezner (2013), "On the Use of Rotated Context Questionnaires in Conjunction With Multilevel Item Response Models", *Large-scale Assessments in Education*, Vol. 1/5, pp. 1-27, http://dx.doi.org/10.1186/2196-0739-1-5. [72]

Adelman, H., L. Taylor and P. Nelson (1989), "Minors' Dissatisfaction With Their Life Circumstances", *Child Psychiatry And Human Development*, Vol. 20/2, pp. 135-147, http://dx.doi.org/10.1007/BF00711660. [78]

Ahlfors (2010), "Many Sources, One Theme: Analysis of Cyberbullying Prevention and Intervention Websites", *Journal of Social Sciences*, Vol. 6/4, pp. 515-522, http://dx.doi.org/10.3844/jssp.2010.515.522. [174]

Angelini, V. et al. (2014), "Do Danes and Italians Rate Life Satisfaction in the Same Way? Using Vignettes to Correct for Individual-Specific Scale Biases", *Oxford Bulletin of Economics and Statistics*, Vol. 76/5, pp. 643-666, http://dx.doi.org/10.1111/obes.12039. [58]

Archambault, I. et al. (2009), "Adolescent Behavioral, Affective, and Cognitive Engagement in School: Relationship to Dropout", *Journal of School Health*, Vol. 79/9, pp. 408-415, http://dx.doi.org/10.1111/j.1746-1561.2009.00428.x. [159]

Bandura, A. (1997), *Self-Efficacy: The Exercise of Control*, W. H. Freeman, New York. [123]

Barker, E. and N. Galambos (2003), "Body Dissatisfaction of Adolescent Girls and Boys: Risk and Resource Factors", *The Journal of Early Adolescence*, Vol. 23/2, pp. 141-165, http://dx.doi.org/10.1177/0272431603023002002. [184]

Bassi, M. et al. (2014), "Personality and Optimal Experience in Adolescence: Implications for Well-Being and Development", *Journal of Happiness Studies*, Vol. 15/4, pp. 829–843, http://dx.doi.org/10.1007/s10902-013-9451-x. [139]

Beal, D., E. Rueda-Sabater and S. Ling Heng (2015), *Why Well-Being Should Drive Growth Strategies: The 2015 Sustainable Economic Development Assessment*, Boston Consulting Group, Boston. [21]

Ben-Arieh, A. (2008), "The Child Indicators Movement: Past, Present, and Future", *Child Indicators Research*, Vol. 1/1, pp. 3–16, http://dx.doi.org/10.1007/s12187-007-9003-1. [25]

Bertling J.P., Borgonovi F. and Almonte D.E. (2016), "Psychosocial Skills in Large-Scale Assessments: Trends, Challenges, and Policy Implications", in Lipnevich A., Preckel F. and Roberts R. (eds.), *Psychosocial Skills and School Systems in the 21st Century*, Springer, Berlin, http://dx.doi.org/10.1007/978-3-319-28606-8_14. [76]

Bertling, J. (2014), "Improving the Contextual Questionnaires for the National Assessment of Educational Progress: Plans for NAEP Core Contextual Modules", White Paper presented at the meeting of the National Assessment Governing Board, Boston, MA. [51]

Bertling, J. (ed.) (2014), *Imputation Issues Relevant to Context Questionnaire Rotation*. [73]

Bertling, J. (ed.) (2014), *Questionnaire Rotation in International Surveys: Findings From PISA*. [74]

Bertling, J. (ed.) (2014), *Spiraling of Contextual Questionnaires in the NAEP TEL Pilot Assessment*. [70]

Bertling, J. and P. Kyllonen (2012), *Domain-General Student Attitudes and Behaviors: Constructs and Items for PISA 2015 (Module 10)*, Technical Report presented at PISA 2012 Questionnaire Expert Group meeting, Heidelberg, Germany. [64]

Bertling, J. and J. Weeks (2018), *Within-Construct Questionnaire Matrix Sampling: Comparison of Different Approaches for PISA 2021*. [71]

Boarini, R. et al. (2012), "What Makes for a Better Life?: The Determinants of Subjective Well-Being in OECD Countries – Evidence from the Gallup World Poll", *OECD Statistics Working Papers*, No. 2012/3, OECD Publishing, Paris, https://dx.doi.org/10.1787/5k9b9ltjm937-en. [24]

Boarini, R., A. Kolev and A. Mcgregor (2014), "Measuring Well-being and Progress in Countries at Different Stages of Development: Towards a More Universal Conceptual Framework", *OECD Development Centre*, No. 325, OECD, Paris, http://dx.doi.org/10.1787/18151949. [5]

Bond, L. et al. (2001), "Does Bullying Cause Emotional Problems? A Prospective Study of Young Teenagers", *BMJ*, Vol. 323/7311, pp. 480-484, http://dx.doi.org/10.1136/bmj.323.7311.480. [171]

Bradshaw, J., P. Hoelscher and D. Richardson (2007), "An Index of Child Well-being in the European Union", *Social Indicators Research*, Vol. 80/1, pp. 133-177, http://dx.doi.org/10.1007/s11205-006-9024-z. [218]

Bradshaw, J. et al. (2011), "Children's subjective well-being: International comparative perspectives", *Children and Youth Services Review*, Vol. 33/4, pp. 548-556, http://dx.doi.org/10.1016/j.childyouth.2010.05.010. [30]

Bradshaw, J. et al. (2009), "A Child Well-Being Index at Small Area Level in England", *Child Indicators Research*, Vol. 2/2, pp. 201-219, http://dx.doi.org/10.1007/s12187-008-9022-6. [220]

Brosschot, J. (2002), "Cognitive-Emotional Sensitization and Somatic Health Complaints", *Scandinavian Journal of Psychology*, Vol. 43/2, pp. 113-121, http://dx.doi.org/10.1111/1467-9450.00276. [119]

Bureau of Labor Statistics (2015), *American Time Use Survey User's Guide: Understanding ATUS 2003 to 2014*, https://www.bls.gov/tus/atususersguide.pdf. [211]

Campbell, A., P. Converse and W. Rodgers (1976), *The Quality of American Life: Perceptions, Evaluations, and Satisfactions*, Russell Sage Foundation, New York, NY, US. [48]

Cantril, H. (1965), *The Patterns of Human Concern*, Rutgers University Press, New Brunswick. [77]

Carson, V., A. Staiano and P. Katzmarzyk (2015), “Physical Activity, Screen Time, and Sitting among U.S. Adolescents”, *Pediatric Exercise Science*, Vol. 27/1, pp. 151-159, http://dx.doi.org/10.1123/pes.2014-0022. [212]

Casas, F. (2011), “Subjective Social Indicators and Child and Adolescent Well-being”, *Child Indicators Research*, Vol. 4/4, pp. 555–575, http://dx.doi.org/10.1007/s12187-010-9093-z. [1]

Casas, F. et al. (2012), “Subjective Indicators of Personal Well-Being among Adolescents. Performance and Results for Different Scales in Latin-Language Speaking Countries: A Contribution to the International Debate”, *Child Indicators Research*, Vol. 5/1, pp. 1–28, http://dx.doi.org/10.1007/s12187-011-9119-1. [47]

Clapper, R. et al. (1995), “Adolescent Problem Behaviors as Predictors of Adult Alcohol Diagnoses”, *International Journal of the Addictions*, Vol. 30/5, pp. 507-523, http://dx.doi.org/10.3109/10826089509048741. [172]

Cohen, J. et al. (2009), “School Climate: Research, Policy, Practice, and Teacher Education”, *Teachers College Record*, Vol. 111/1, pp. 180-213. [141]

Comber, L. and J. Keeves (1973), *Science Education in Nineteen Countries: An Empirical Study*, Almqvist & Wiksell, Stockholm. [67]

Conger, R., K. Conger and M. Martin (2010), “Socioeconomic Status, Family Processes,and Individual Development”, *Journal of Marriage and Family*, Vol. 72/3, pp. 685-704, http://dx.doi.org/10.1111/j.1741-3737.2010.00725.x. [203]

Craig, W. (1998), “The Relationship Among Bullying, Victimization, Depression, Anxiety, and Aggression in Elementary School Children”, *Personality and Individual Differences*, Vol. 24/1, pp. 123-130, http://dx.doi.org/10.1016/s0191-8869(97)00145-1. [168]

Craig, W., D. Pepler and R. Atlas (2000), “Observations of Bullying in the Playground and in the Classroom”, *School Psychology International*, Vol. 21/1, pp. 22-36, http://dx.doi.org/10.1177/0143034300211002. [161]

Csikszentmihalyi, M. (1975), “Play and Intrinsic Rewards”, *Journal of Humanistic Psychology*, Vol. 15/3, pp. 41-63, http://dx.doi.org/10.1177/002216787501500306. [132]

Csikszentmihalyi, M. and I. Csikszentmihalyi (2006), *A Life Worth Living. Contributions to Positive Psychology*, Oxford University Press, Oxford. [133]

Cummins, R. and A. Lau (2005), *Personal Wellbeing Index-School Children (Third Edition)*, School of Psychology, Deakin University. [26]

Currie, C. et al. (2009), “The Health Behaviour in School-Aged Children: WHO Collaborative Cross-National (HBSC) Study: Origins, Concept, History and Development 1982-2008”, *International Journal of Public Health*, Vol. 54/2, pp. 131-139, http://dx.doi.org/10.1007/s00038-009-5404-x. [88]

Currie, C. et al. (2012), *Social Determinants of Health and Well-Being Among Young People. Health Behaviour in School-Aged Children*, WHO Regional Office for Europe, http://www.euro.who.int/__data/assets/pdf_file/0003/163857/Social-determinants-of-health-and-well-being-among-young-people.pdf. [117]

Dalton, P., S. Ghosal and A. Mani (2015), "Poverty and Aspirations Failure", *The Economic Journal*, Vol. 126/590, pp. 165-188, http://dx.doi.org/10.1111/ecoj.12210. [208]

Davidson, W. and P. Cotter (1991), "The Relationship Between Sense of Community and Subjective Well-Being: A First Look", *Journal of Community Psychology*, Vol. 19/3, pp. 246-253, http://dx.doi.org/10.1002/1520-6629(199107)19:3<246::AID-JCOP2290190308>3.0.CO;2-L. [190]

Dias, S., M. Matos and A. Gonçalves (2005), "Preventing HIV Transmission in Adolescents: An Analysis of the Portuguese Data From the Health Behaviour School-Aged Children Study and Focus Groups", *European Journal of Public Health*, Vol. 15/3, pp. 300-304, http://dx.doi.org/10.1093/eurpub/cki085. [157]

Diemer, M. et al. (2012), "Best Practices in Conceptualizing and Measuring Social Class in Psychological Research", *Analyses of Social Issues and Public Policy*, Vol. 13/1, pp. 77-113, http://dx.doi.org/10.1111/asap.12001. [201]

Diener, E. and M. Chan (2011), "Happy People Live Longer: Subjective Well-Being Contributes to Health and Longevity", *Applied Psychology: Health and Well-Being*, Vol. 3/1, pp. 1-43, http://dx.doi.org/10.1111/j.1758-0854.2010.01045.x. [20]

Diener, E. et al. (2010), "Income's Association with Judgments of Life versus Feelings", in Diener, E. and D. Helliwell, John F. Kahneman (eds.), *International Differences in Well-Being*, Oxford University Press, Oxford. [199]

Diener, E., S. Oishi and R. Lucas (2015), "National Accounts of Subjective Well-Being", *American Psychologist*, Vol. 70/3, pp. 234-242, http://dx.doi.org/10.1037/a0038899. [45]

Diener, E. and M. Seligman (2004), "Beyond Money Toward an Economy of Well-Being", *Psychological Science in the Public Interest*, Vol. 5/1, pp. 1-31, http://dx.doi.org/10.1111/j.0963-7214.2004.00501001.x. [2]

Diener, E. et al. (1999), "Subjective Weil-Being: Three Decades of Progress", *Psychological Bulletin*, Vol. 125/2, pp. 276-302. [13]

Dillman, D., J. Smyth and L. Christian (2014), *Internet, Phone, Mail, and Mixed Mode Surveys: The Tailored Design Method*, John Wiley & Sons Inc, Hoboken, NJ, US. [53]

Dolan, P., T. Peasgood and M. White (2008), "Do We Really Know What Makes Us Happy? A Review of the Economic Literature on the Factors Associated with Subjective Well-Being", *Journal of Economic Psychology*, Vol. 29/1, pp. 94–122, http://dx.doi.org/10.1016/j.joep.2007.09.001. [14]

Due, P. et al. (2005), "Bullying and Symptoms Among School-Aged Children: International Comparative Cross Sectional Study in 28 Countries", *European Journal of Public Health*, Vol. 15/2, pp. 128-132, http://dx.doi.org/10.1093/eurpub/cki105. [169]

Due, P. et al. (2003), "Socioeconomic Health Inequalities Among a Nationally Representative Sample of Danish Adolescents: the Role of Different Types of Social Relations", *Journal of Epidemiology and Community Health*, Vol. 57/9, pp. 692-698, http://dx.doi.org/10.1136/jech.57.9.692. [153]

Dupuy, M. et al. (2011), "Socio-Demographic and Lifestyle Factors Associated with Overweight in a Representative Sample of 11-15 Year Olds in France: Results From the WHO-Collaborative Health Behaviour in School-Aged Children (HBSC) Cross-Sectional Study", *BMC Public Health*, Vol. 11/442, pp. 1-11, http://dx.doi.org/10.1186/1471-2458-11-442. [101]

Edwards, L. and S. Lopez (2006), "Perceived Family Support, Acculturation, and Life Satisfaction in Mexican American Youth: A Mixed-Methods Exploration", *Journal of Counseling Psychology*, Vol. 53/3, pp. 279-287, http://dx.doi.org/10.1037/0022-0167.53.3.279. [129]

Elmore, G. and E. Huebner (2010), "Adolescents' Satisfaction With School Experiences: Relationships With Demographics, Attachment Relationships, and School Engagement Behavior", *Psychology in the Schools*, Vol. 47/6, pp. 525-537, http://dx.doi.org/10.1002/pits.20488. [79]

Epstein, J. and J. McPartland (1976), "The Concept and Measurement of the Quality of School Life", *American Educational Research Journal*, Vol. 13/1, pp. 15-30, http://dx.doi.org/10.3102/00028312013001015. [146]

Evans, J., I. Macrory and C. Randall (2015), *Measuring National Well-being: Life in the UK 2015*, Office for National Statistics, London. [6]

Farrell, S., T. Aubry and D. Coulombe (2003), "Neighborhoods and Neighbors: Do They Contribute to Personal Well-Being?", *Journal of Community Psychology*, Vol. 32/1, pp. 9-25, http://dx.doi.org/10.1002/jcop.10082. [191]

Federal Interagency Forum on Child and Family Statistics (2009), *America's Children: Key National Indicators of Well-Being*, US Government Printing Office, Washington, DC. [7]

Fernandes, L., A. Mendes and A. Teixeira (2012), "A Review Essay on the Measurement of Child Well-Being", *Social Indicators Research*, Vol. 106/2, pp. 239–257, http://dx.doi.org/10.1007/s11205-011-9814-9. [49]

Fletcher, A., C. Bonell and J. Hargreaves (2008), "School Effects on Young People's Drug Use: A Systematic Review of Intervention and Observational Studies", *Journal of Adolescent Health*, Vol. 42/3, pp. 209-220, http://dx.doi.org/10.1016/J.JADOHEALTH.2007.09.020. [158]

Freeman, J. et al. (2009), "The Relationship of Schools to Emotional Health and Bullying", *International Journal of Public Health*, Vol. 54/S2, pp. 251-259, http://dx.doi.org/10.1007/s00038-009-5421-9. [154]

Garrow J. S. and Webster J. (1985), “Quetelet’s Index (W/H2) as a Measure of Fatness”, *International Journal of Obesity*, Vol. 9/2, pp. 147-153. [90]

Gaspar, M. et al. (2011), “Protective Effect of Physical Activity on Dissatisfaction With Body Image in Children - A Cross-Sectional Study”, *Psychology of Sport and Exercise*, Vol. 12, pp. 563-569, http://dx.doi.org/10.1016/j.psychsport.2011.05.004. [102]

Gehlbach, H. (2015), “Seven Survey Sins”, *Journal of Early Adolescence*, Vol. 35/5-6, pp. 883–897, http://dx.doi.org/10.1177/0272431615578276. [54]

Gershoff, E. et al. (2007), “Income Is Not Enough: Incorporating Material Hardship Into Models of Income Associations With Parenting and Child Development”, *Child Development*, Vol. 78/1, pp. 70-95, http://dx.doi.org/10.1111/j.1467-8624.2007.00986.x. [207]

Gifford-Smith, M. and C. Brownell (2003), “Childhood Peer Relationships: Social Acceptance, Friendships, and Peer Networks”, *Journal of School Psychology*, Vol. 41/4, pp. 235-284, http://dx.doi.org/10.1016/s0022-4405(03)00048-7. [186]

Gilman, R. and E. Anderman (2006), “The Relationship Between Relative Levels of Motivation and Intrapersonal, Interpersonal, and Academic Functioning Among Older Adolescents”, *Journal of School Psychology*, Vol. 44/5, pp. 375-391, http://dx.doi.org/10.1016/J.JSP.2006.03.004. [149]

Gilman, R. and E. Huebner (2000), “Review of Life Satisfaction Measures for Adolescents”, *Behaviour Change*, Vol. 17/3, pp. 178-195, http://dx.doi.org/10.1375/bech.17.3.178. [80]

Gilman, R. and S. Huebner (2003), “A Review of Life Satisfaction Research With Children and Adolescents”, *School Psychology Quarterly*, Vol. 18/2, pp. 192-205, http://dx.doi.org/10.1521/scpq.18.2.192.21858. [31]

Glew, G. et al. (2008), “Bullying and School Safety”, *The Journal of Pediatrics*, Vol. 152/1, pp. 123-128.e1, http://dx.doi.org/10.1016/j.jpeds.2007.05.045. [166]

Goodenow, C. (1993), “The Psychological Sense of School Membership Among Adolescents: Scale Development and Educational Correlates”, *Psychology in the Schools*, Vol. 30/1, pp. 79-90, http://dx.doi.org/10.1002/1520-6807(199301)30:1<79::AID-PITS2310300113>3.0.CO;2-X. [148]

Goswami, H. (2014), “Children’s Subjective Well-Being: Socio-Demographic Characteristics and Personality”, *Child Indicators Research*, Vol. 7/1, pp. 119-140, http://dx.doi.org/10.1007/s12187-013-9205-7. [200]

Grube, A. et al. (2008), “The Event Reconstruction Method: An Efficient Measure of Experience-Based Job Satisfaction”, *Journal of Occupational and Organizational Psychology*, Vol. 81/4, pp. 669–689, http://dx.doi.org/10.1348/096317907X251578. [65]

Hallal, P. et al. (2006), “Adolescent Physical Activity and Health: A Systematic Review”, *Sports Medicine*, Vol. 36/12, pp. 1019–1030, http://dx.doi.org/10.2165/00007256-200636120-00003. [104]

Haranin, E., E. Huebner and S. Suldo (2007), "Predictive and Incremental Validity of Global and Domain-Based Adolescent Life Satisfaction Reports", *Journal of Psychoeducational Assessment*, Vol. 25/2, pp. 127-138, http://dx.doi.org/10.1177/0734282906295620. [83]

Hardy, C., W. Bukowski and L. Sippola (2002), "Stability and Change in Peer Relationships During the Transition to Middle-Level School", *Journal of Early Adolescence*, Vol. 22/2, pp. 117-142, http://dx.doi.org/10.1177/0272431602022002001. [40]

Harel-Fisch, Y. et al. (2011), "Negative School Perceptions and Involvement in School Bullying: A Universal Relationship Across 40 Countries", *Journal of Adolescence*, Vol. 34/4, pp. 639-652, http://dx.doi.org/10.1016/J.ADOLESCENCE.2010.09.008. [156]

Harel, Y. (1999), "A Cross-National Study of Youth Violence in Europe", *International Journal of Adolescent Medicine and Health*, Vol. 11/3-4, pp. 121-134. [177]

Haug, E. et al. (2009), "Overweight in School-Aged Children and its Relationship with Demographic and Lifestyle Factors: Results from the WHO-Collaborative Health Behaviour in School-Aged Children (HBSC) Study", *International Journal of Public Health*, Vol. 54/2, pp. 167-179, http://dx.doi.org/10.1007/s00038-009-5408-6. [92]

Helliwell, J., R. Layard and J. Sachs (2015), *World Happiness Report*, The Earth Institute, Columbia University. [17]

Helseth, S. and N. Misvær (2010), "Adolescents' Perceptions of Quality of Life: What It is and What Matters", *Journal of Clinical Nursing*, Vol. 19/9-10, pp. 1454-1461, http://dx.doi.org/10.1111/j.1365-2702.2009.03069.x. [131]

Hetland, J., T. Torsheim and L. Aarø (2002), "Subjective Health Complaints in Adolescence: Dimensional Structure and Variation Across Gender and Age", *Scandinavian Journal of Public Health*, Vol. 30/3, pp. 223-230, http://dx.doi.org/10.1177/140349480203000309. [121]

Huebner, E. (1991), "Initial Development of the Student's Life Satisfaction Scale", *School Psychology International*, Vol. 12/3, pp. 231-240, http://dx.doi.org/10.1177/0143034391123010. [136]

Huebner, E. and T. Dew (1996), "The Interrelationships of Positive Affect, Negative Affect, and Life Satisfaction in an Adolescent Sample", *Social Indicators Research*, Vol. 38/2, pp. 129-137, http://dx.doi.org/10.1007/BF00300455. [32]

Huebner, E. et al. (2004), "Life Satisfaction in Children and Youth: Empirical Foundations and Implications for School Psychologists", *Psychology in the Schools*, Vol. 41/1, pp. 81-93, http://dx.doi.org/10.1002/pits.10140. [144]

Huebner, E. et al. (2005), "Middle school students' perceptions of quality of life", *Journal of Happiness Studies*, Vol. 6/1, pp. 15-24, http://dx.doi.org/10.1007/s10902-004-1170-x. [145]

Huebner, S. (2001), *Manual for the Multidimensional Student's Life Satisfaction Scale*, University of South Carolina, Department of Psychology, Columbia, SC, 29208. [29]

Huppert, F. et al. (2009), “Measuring Well-Being Across Europe: Description of the ESS Well-Being Module and Preliminary Findings”, *Social Indicators Research*, Vol. 91/3, pp. 301-315, http://dx.doi.org/10.1007/s11205-008-9346-0. [137]

Iannotti, R. et al. (2009), “Interrelationships of Adolescent Physical Activity, Screen-Based Sedentary Behaviour, and Social and Psychological Health”, *International Journal of Public Health*, Vol. 54/Supplement 2, pp. 191-198, http://dx.doi.org/10.1007/s00038-009-5410-z. [105]

Iceland, J. (2003), “Why Poverty Remains High: The Role of Income Growth, Economic Inequality, and Changes in Family Structure, 1949-1999”, *Demography*, Vol. 40/3, pp. 499-519, http://dx.doi.org/10.1353/dem.2003.0025. [205]

James, W., C. Chen and S. Inoue (2002), “Appropriate Asian Body Mass Indices?”, *Obesity Reviews*, Vol. 3/3, pp. 139-139, http://dx.doi.org/10.1046/j.1467-789X.2002.00063.x. [100]

Janssen, I. et al. (2004), “Associations between overweight and obesity with bullying behaviors in school-aged children”, *Pediatrics*, Vol. 113/5, pp. 1187-1194, http://www.ncbi.nlm.nih.gov/pubmed/15121928. [97]

Jia, Y. et al. (2009), “The Influence of Student Perceptions of School Climate on Socioemotional and Academic Adjustment: A Comparison of Chinese and American Adolescents”, *Child Development*, Vol. 80/5, pp. 1514–1530, http://dx.doi.org/10.1111/j.1467-8624.2009.01348.x. [142]

Jourdan, D. et al. (2008), “The Future of Health Promotion in Schools Goes Through the Strengthening of Teacher Training at a Global Level”, *Promotion & Education*, Vol. 15/3, pp. 36-38, http://dx.doi.org/10.1177/1025382308095657. [140]

Kahneman, D. and A. Krueger (2006), “Developments in the Measurement of Subjective Well-Being”, *Journal of Economic Perspectives*, Vol. 20/1, pp. 3–24, http://dx.doi.org/10.1257/089533006776526030. [44]

Kahneman, D. et al. (2004), “A Survey Method for Characterizing Daily Life Experience: The Day Reconstruction Method”, *Science*, Vol. 306/5702, pp. 1776-1780, http://science.sciencemag.org/content/306/5702/1776.short. [3]

Karasek, R. and T. Theorell (1992), *Healthy Work: Stress, Productivity, and the Reconstruction of Working Life*, Basic Books. [181]

Karatzias, A., K. Power and V. Swanson (2001), “Quality of School Life: Development and Preliminary Standardisation of an Instrument Based on Performance Indicators in Scottish Secondary Schools”, *School Effectiveness and School Improvement*, Vol. 12/3, pp. 265-284, http://dx.doi.org/10.1076/sesi.12.3.265.3449. [147]

Kaufman, A. and E. Augustson (2008), “Predictors of Regular Cigarette Smoking Among Adolescent Females: Does Body Image Matter?”, *Nicotine and Tobacco Research: Official Jurnal of the Society for Research on Nicotine and Tobacco*, Vol. 10/8, pp. 1301-1309, http://dx.doi.org/10.1080/14622200802238985. [113]

Keller, M. (2004), *A Cross-Cultural Perspective on Friendship Research*. [221]

Kern, M. et al. (2015), “A Multidimensional Approach to Measuring Well-Being in Students: Application of the PERMA Framework”, *Journal of Positive Psychology*, Vol. 10/3, pp. 262–271, http://dx.doi.org/10.1080/17439760.2014.936962. [135]

Keyes, C. and J. Grzywacz (2005), “Health As a Complete State: The Added Value in Work Performance and Healthcare Costs”, *Journal of Occupational and Environmental Medicine*, Vol. 47/5, pp. 523-532, http://dx.doi.org/10.1097/01.jom.0000161737.21198.3a. [18]

Kim, H., U. Schimmack and S. Oishi (2012), “Cultural Differences in Self- and Other-Evaluations and Well-Being: A Study of European and Asian Canadians”, *Journal of Personality and Social Psychology*, Vol. 102/4, pp. 856-873, http://dx.doi.org/10.1037/a0026803. [57]

Klieme, E. and S. Kuger (2014), *PISA 2015 Draft Questionnaire Framework*, OECD, Paris, http://www.oecd.org/pisa/pisaproducts/PISA-2015-draft-questionnaire-framework.pdf. [69]

Koivusilta, L., T. Lintonen and A. Rimpelä (2005), “Intensity of Mobile Phone Use and Health Compromising Behaviours - How is Information and Communication Technology Connected to Health-Related Lifestyle in Adolescence?”, *Journal of Adolescence*, Vol. 28/1, pp. 35-47, http://dx.doi.org/10.1016/j.adolescence.2004.05.004. [216]

Kristensen, N. and E. Johansson (2008), “New Evidence on Cross-Country Differences in Job Satisfaction Using Anchoring Vignettes”, *Labour Economics*, Vol. 15, pp. 96–117, http://dx.doi.org/10.1016/j.labeco.2006.11.001. [59]

Krosnick, J. and L. Fabrigar (1997), “Designing Rating Scales for Effective Measurement in Surveys”, in Lyberg, L. et al. (eds.), *Survey Measurement and Process Quality*, John Wiley, New York. [55]

Kvalem, I. et al. (2011), “Body Evaluation and Coital Onset: A Population-Based Longitudinal Study”, *Body Image*, Vol. 8/2, pp. 110-118, http://dx.doi.org/10.1016/j.bodyim.2011.02.001. [114]

Kyllonen P. C. and Bertling J. P. (2014), *Anchoring Vignettes reduce Bias in Noncognitive Rating Scale Responses*, ETS/OECD, Princeton, NJ. [60]

Kyllonen, P. and J. Bertling (2014), “Innovative Questionnaire Assessment Methods to Increase Cross-Country Comparability”, in Rutkowski, L., M. von Davier and D. Rutkowski (eds.), *Handbook of International Large-Scale Assessment: Background, Technical Issues, and Methods of Data Analysis*, CRC Press, Boca Raton. [56]

Land, K. et al. (2007), “Measuring trends in child well-being: an evidence-based approach”, *Social Indicators Research*, Vol. 80/1, pp. 105-132, http://link.springer.com/article/10.1007/s11205-006-9023-0. [50]

Land, K., V. Lamb and S. Mustillo (2001), “Child and Youth Well-Being in the United States, 1975–1998: Some Findings from a New Index”, *Social Indicators Research*, Vol. 56/3, pp. 241-318, http://dx.doi.org/10.1023/A:1012485315266. [219]

Larson, R. and M. Richards (1991), “Daily Companionship in Late Childhood and Early Adolescence: Changing Developmental Contexts”, *Child Development*, Vol. 62/2, pp. 284-300, http://dx.doi.org/10.1111/j.1467-8624.1991.tb01531.x. [188]

Larson, R. and S. Verma (1999), “How children and adolescents spend time across the world: Work, play, and developmental opportunities.”, *Psychological Bulletin*, Vol. 125/6, pp. 701-736, http://dx.doi.org/10.1037/0033-2909.125.6.701. [213]

Laurent, J. et al. (1999), “A Measure of Positive and Negative Affect for Children: Scale Development and Preliminary Validation”, *Psychological Assessment*, Vol. 11/3, pp. 326-338, http://dx.doi.org/10.1037/1040-3590.11.3.326. [86]

Lippman, L., K. Moore and H. McIntosh (2011), “Positive Indicators of Child Well-Being: A Conceptual Framework, Measures, and Methodological Issues”, *Applied Research in Quality of Life*, Vol. 6/4, pp. 425–449, http://dx.doi.org/10.1007/s11482-011-9138-6. [27]

Lobstein, T., L. Baur and R. Uauy (2004), “Obesity in Children and Young People: A Crisis in Public Health”, *Obesity Reviews*, Vol. 5/s1, pp. 4-85, http://dx.doi.org/10.1111/j.1467-789X.2004.00133.x. [91]

Long, R. et al. (2012), “Measuring School-Related Subjective Well-Being in Adolescents”, *American Journal of Orthopsychiatry*, Vol. 82/1, pp. 50-60, http://dx.doi.org/10.1111/j.1939-0025.2011.01130.x. [81]

Mahady Wilton, M., W. Craig and D. Pepler (2000), “Emotional Regulation and Display in Classroom Victims of Bullying: Characteristic Expressions of Affect, Coping Styles and Relevant Contextual Factors”, *Social Development*, Vol. 9/2, pp. 226-245, http://dx.doi.org/10.1111/1467-9507.00121. [162]

Marcotte, D. et al. (2002), “Gender differences in depressive symptoms during adolescence: Role of gender-typed characteristics, self-esteem, body image, stressful life events, and pubertal status”, *Journal of Emotional and Behavioral Disorders*, http://dx.doi.org/10.1177/106342660201000104. [112]

Martínez-Gómez, D. et al. (2011), “Active Commuting to School and Cognitive Performance in Adolescents: The AVENA Study”, *Archives of Pediatrics and Adolescent Medicine*, Vol. 165/4, pp. 300-305, http://dx.doi.org/10.1001/archpediatrics.2010.244. [109]

Mayer, S. and C. Jencks (1989), “Growing Up in Poor Neighborhoods: How Much Does It Matter?”, *Science*, Vol. 243/4897, pp. 1441-1445, http://dx.doi.org/10.1126/science.243.4897.1441. [206]

McGill, R. et al. (2012), “Academic Adjustment Across Middle School: The Role of Public Regard and Parenting”, *Developmental Psychology*, Vol. 48/4, pp. 1003-1018, http://dx.doi.org/10.1037/a0026006. [37]

McMurray, R. et al. (2008), “Influence of Physical Activity on Change in Weight Status as Children Become Adolescents”, *International Journal of Pediatric Obesity*, Vol. 3/2, pp. 69-77. [106]

Millings, A. et al. (2012), "School Connectedness, Peer Aattachment, and Self-Esteem as Predictors of Adolescent Depression", *Journal of Adolescence*, Vol. 35/4, pp. 1061-1067, http://dx.doi.org/10.1016/J.ADOLESCENCE.2012.02.015. [150]

Mistry, R. and E. Lowe (2006), "What Earnings and Income Buy: The "Basics", Plus "A Little Extra": Implications for Family and Child Well-Being", in Yoshikawa, H., T. Weisner and E. Lowe (eds.), *Making it Work: Low-Wage Employment, Family Life, and Child Development*, Russell Sage Foundation, New York. [204]

Molcho, M., Y. Harel and L. Dina (2004), "Substance use and youth violence. A study among 6th to 10th grade Israeli school children", *International Journal of Adolescent Medicine and Health*, Vol. 16/3, http://dx.doi.org/10.1515/ijamh.2004.16.3.239. [170]

Moore, K. et al. (2008), "A Microdata Child Well-Being Index: Conceptualization, Creation, and Findings", *Child Indicators Research*, Vol. 1/1, pp. 17-50, http://dx.doi.org/10.1007/s12187-007-9000-4. [165]

Must, A. and D. Tybor (2005), "Physical Activity and Sedentary Behavior: A Review of Longitudinal Studies of Weight and Adiposity in Youth", *International Journal of Obesity*, Vol. 29/2, pp. 84-96. [95]

Nansel, T. et al. (2001), "Bullying Behaviors Among US Youth: Prevalence and Association with Psychosocial Adjustment", *Journal of the American Medical Association*, Vol. 285/16, pp. 2094-2100, http://dx.doi.org/10.1001/jama.285.16.2094. [128]

Natsuaki, M., M. Biehl and X. Ge (2009), "Trajectories of Depressed Mood From Early Adolescence to Young Adulthood: The Effects of Pubertal Timing and Adolescent Dating", *Journal of Research on Adolescence*, Vol. 19/1, pp. 47–74, http://dx.doi.org/10.1111/j.1532-7795.2009.00581.x. [38]

Neumark-Sztainer, D. et al. (2012), "Dieting and Unhealthy Weight Control Behaviors During Adolescence: Associations with 10-year Changes in Body Mass Index", *Journal of Adolescent Health*, Vol. 50/1, pp. 80-86, http://dx.doi.org/10.1016/j.jadohealth.2011.05.010. [98]

Nihiser, A. et al. (2007), "Body Mass Index Measurement in Schools*", *Journal of School Health*, Vol. 77/10, pp. 651-671, http://dx.doi.org/10.1111/j.1746-1561.2007.00249.x. [99]

OECD (2017), *PISA 2015 Assessment and Analytical Framework: Science, Reading, Mathematic, Financial Literacy and Collaborative Problem Solving*, PISA, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264281820-en. [34]

OECD (2016), *Technical Report of the Survey of Adult Skills (PIAAC) (2nd Edition, pre-publication copy)*, OECD, http://www.oecd.org/skills/piaac/PIAAC_Technical_Report_2nd_Edition_Full_Report.pdf. [68]

OECD (2015), *How's Life? 2015*, Organisation for Economic Co-operation and Development, Paris, http://www.oecd-ilibrary.org/content/book/how_life-2015-en. [36]

OECD (2013), *OECD Guidelines on Measuring Subjective Well-being*, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264191655-en. [46]

OECD. (2017), *PISA 2015 Results (Volume III) Students' Well-Being*, OECD Publishing. [35]

Office for National Statistics (2014), *Commuting and Personal Well-being*, https://webarchive.nationalarchives.gov.uk/20160107224314/http:/www.ons.gov.uk/ons/dcp171766_351954.pdf. [222]

Ojala, K. et al. (2007), "Attempts to Lose Weight Among Overweight and Non-Overweight Adolescents: A Cross-National Survey", *International Journal of Behavioral Nutrition and Physical Activity*, Vol. 4/50, http://dx.doi.org/10.1186/1479-5868-4-50. [116]

Olweus, D. (2011), "Bullying at school and later criminality: Findings from three Swedish community samples of males", *Criminal Behaviour and Mental Health*, Vol. 21/2, pp. 151-156, http://dx.doi.org/10.1002/cbm.806. [178]

Olweus, D. (1996), "Bully/Victim Problems at School: Facts and Effective Intervention", *Reclaiming Children and Youth: Journal of Emotional and Behavioral Problems*, Vol. 5/1, pp. 15-22. [179]

Olweus, D. (1994), "Bullying at School: Basic Facts and Effects of a School Based Intervention Program", *Journal of Child Psychology and Psychiatry*, Vol. 35/7, pp. 1171-1190, http://dx.doi.org/10.1111/j.1469-7610.1994.tb01229.x. [167]

Olweus, D. (1991), "Bully/Victim Problems Among School Children: Some Basic Facts and Effects of a School-Based Intervention Program", in Pepler, D. and K. Rubin (eds.), *The Development and Treatment of Childhood Aggression*, Lawrence Erlbaum Associates, Hillsdale, NJ. [163]

Parks, R. and H. Eckert (eds.) (1991), *Fitness and Performance: Adult Health and the Culture of Youth*, Human Kinetics, Champaign, USA. [103]

Petersen, A. et al. (1997), "Promoting Mental Health During the Transition Into Adolescence", in Schulenberg, J., J. Maggs and K. Hurrelmann (eds.), *Health Risks and Developmental Transitions During Adolescence*, Cambridge University Press, New York. [118]

Pollard, E. and P. Lee (2003), "Child Well-Being: A Systematic Review of the Literature", *Social Indicators Research*, Vol. 61/1, pp. 59–78, http://dx.doi.org/10.1023/A:1021284215801. [28]

Prezza, M. et al. (2001), "Sense of Community Referred to the Whole Town: Its Relations With Neighboring, Loneliness, Life Satisfaction, and Area of Residence", *Journal of Community Psychology*, Vol. 29/1, pp. 29-52, http://dx.doi.org/10.1002/1520-6629(200101)29:1<29::AID-JCOP3>3.0.CO;2-C. [192]

Prezza, M. and S. Costantini (1998), "Sense of Community and Life Satisfaction: Investigation in Three Different Territorial Contexts", *Journal of Community & Applied Social Psychology*, Vol. 8/3, pp. 181-194, http://dx.doi.org/10.1002/(SICI)1099-1298(199805/06)8:3<181::AID-CASP436>3.0.CO;2-4. [193]

Prezza, M., M. Pacilli and S. Dinelli (2004), "Loneliness and new technologies in a group of Roman adolescents", *Computers in Human Behavior*, Vol. 20/5, pp. 691-709, http://dx.doi.org/10.1016/j.chb.2003.10.008. [214]

Punamäki, R. et al. (2007), "Use of information and communication technology (ICT) and perceived health in adolescence: The role of sleeping habits and waking-time tiredness", *Journal of Adolescence*, Vol. 30/4, pp. 569-585, http://dx.doi.org/10.1016/j.adolescence.2006.07.004. [215]

Quon, E. and J. McGrath (2014), "Subjective Socioeconomic Status and Adolescent Health: A Meta-Analysis", *Health Psychology*, Vol. 33/5, pp. 433-447, http://dx.doi.org/10.1037/a0033716. [202]

Ravens-Sieberer, U. et al. (2010), "Reliability, Construct and Criterion Validity of the KIDSCREEN-10 Score: A Short Measure for Children and Adolescents' Well-Being and Health-Related Quality of Life", *Quality of Life Research*, Vol. 19/10, pp. 1487–1500, http://dx.doi.org/10.1007/s11136-010-9706-5. [87]

Ravens-Sieberer, U. et al. (2008), "An International Scoring System for Self-Reported Health Complaints in Adolescents", *European Journal of Public Health*, Vol. 18/3, pp. 294-299, http://dx.doi.org/10.1093/eurpub/ckn001. [122]

Ravens-Sieberer, U. et al. (2014), "The European KIDSCREEN Approach to Measure Quality of Life and Well-Being in Children: Development, Current Application, and Future Advances", *Quality of Life Research*, Vol. 23, pp. 791–803, http://dx.doi.org/10.1007/s11136-013-0428-3. [23]

Ravens-Sieberer, U., G. Kökönyei and C. Thomas (2004), "School and Health", in Currie, C. et al. (eds.), *Young People's Health in Context. Health Behaviour in Schoolaged Children Study: International Report from the 2001/2002 Survey (Health Policy for Children and Adolescents, No.4)*, WHO Regional Office for Europe, Copenhagen. [126]

Ray, D. (2006), "Aspirations, Poverty, and Economic Change", in *Understanding Poverty*, Oxford University Press, http://dx.doi.org/10.1093/0195305191.003.0028. [209]

Reddy, R., J. Rhodes and P. Mulhall (2003), "The Influence of Teacher Support on Student Adjustment in the Middle School Years: A Latent Growth Curve Study", *Development and Psychopathology*, Vol. 15/1, pp. 119-138, http://dx.doi.org/10.1017.S0954579403000075. [151]

Rees, G., H. Goswami and J. Bradshaw (2010), *Developing an Index of Children's Subjective Well-Being in England*, The Children's Society, London, http://www.childrenssociety.org.uk. [210]

Rees, G. and G. Main (2015), *Children's views on their lives and well-being in 15 countries: An initial report on the Children's Worlds survey, 2013-14*, http://www.isciweb.org/_Uploads/dbsAttachedFiles/10and12FullReport.pdf. [52]

Rees, G., L. Pople and H. Goswami (2011), *Understanding Children's Well-Being - Links Between Family Economic Factors and Children's Subjective Well-Being: Initial Findings From Wave 2 and Wave 3 Quarterly Surveys*, The Children's Society, https://www.childrenssociety.org.uk/sites/default/files/tcs/promoting_positive_well-being_for_children_economic_factors_march.pdf. [195]

Richter, M. et al. (2009), "The role of behavioural factors in explaining socio-economic differences in adolescent health: A multilevel study in 33 countries", *Social Science & Medicine*, Vol. 69/3, pp. 396-403, http://dx.doi.org/10.1016/j.socscimed.2009.05.023. [197]

Rigby, K. and P. Slee (1991), "Bullying among Australian School Children: Reported Behavior and Attitudes toward Victims", *The Journal of Social Psychology*, Vol. 131/5, pp. 615-627, http://dx.doi.org/10.1080/00224545.1991.9924646. [175]

Roeser, R. (2001), "To Cultivate the Positive … Introduction to the Special Issue on Schooling and Mental Health Issues", *Journal of School Psychology*, Vol. 39/2, pp. 99-110, http://dx.doi.org/10.1016/s0022-4405(01)00061-9. [183]

Rokholm, B., J. Baker and T. Sørensen (2010), "The Levelling Off of the Obesity Epidemic Since the Year 1999 - A Review of Evidence and Perspectives", *Obesity Reviews*, Vol. 11/12, pp. 835–846, http://dx.doi.org/10.1111/j.1467-789X.2010.00810.x. [93]

Rudd, N. and S. Lennon (2000), "Body Image and Appearance-Management Behaviors In College Women", *Clothing and Textiles Research Journal*, Vol. 18/3, pp. 152-162, http://dx.doi.org/10.1177/0887302X0001800304. [111]

Ryan, R. and E. Deci (2001), "On Happiness and Human Potentials: A Review of Research on Hedonic and Eudaimonic Well-Being", *Annual Review of Psychology*, Vol. 52/1, pp. 141-166, http://dx.doi.org/10.1146/annurev.psych.52.1.141. [134]

Ryan, R. and E. Deci (2000), "Self-Determination Theory and the Facilitation of Intrinsic Motivation, Social Development, and Well-Being", *American Psychologist*, Vol. 55/1, pp. 68-78, http://dx.doi.org/10.1037/0003-066X.55.1.68. [124]

Ryff, C. (1995), "Psychological Well-Being in Adult Life", *Current Directions in Psychological Science*, Vol. 4/4, pp. 99-104, http://dx.doi.org/10.1111/1467-8721.ep10772395. [138]

Sacks, D., B. Stevenson and J. Wolfers (2010), "Subjective Well-Being, Income, Economic Development and Growth", *NBER Working Paper Series*, No. 16441, National Bureau of Economic Research, Cambridge, MA, https://www.nber.org/papers/w16441.pdf. [15]

Saha, R. et al. (2010), "A Longitudinal Study of Adolescent Life Satisfaction and Parenting", *Child Indicators Research*, Vol. 3/2, pp. 149–165, http://dx.doi.org/10.1007/s12187-009-9050-x. [33]

Salomon, J., A. Tandon and C. Murray (2004), "Comparability of Self Rated Health: Cross Sectional Multi-Country Survey Using Anchoring Vignettes", *British Medical Journal*, Vol. 328, pp. 258-261, http://dx.doi.org/10.1136/bmj.37963.691632.44. [61]

Sampasa-Kanyinga, H. and R. Lewis (2015), "Frequent Use of Social Networking Sites Is Associated with Poor Psychological Functioning Among Children and Adolescents", *Cyberpsychology, Behavior, and Social Networking*, Vol. 18/7, pp. 380-385, http://dx.doi.org/10.1089/cyber.2015.0055. [217]

Schimmack, U., E. Diener and S. Oishi (2002), "Life-Satisfaction Is a Momentary Judgment and a Stable Personality Characteristic: The Use of Chronically Accessible and Stable Sources", *Journal of Personality*, Vol. 70/3, pp. 345-385, http://dx.doi.org/10.1111/1467-6494.05008. [85]

Schleicher, A. (2015), *Proposals for PISA 2018 International Questionnaire Options*, OECD, Paris. [39]

Schmitt, M. et al. (2014), "The Consequences of Perceived Discrimination for Psychological Well-Being: A Meta-Analytic Review", *Psychological Bulletin*, Vol. 140/4, pp. 921-948, http://dx.doi.org/10.1037/a0035754. [180]

Schwarz, N., D. Kahneman and J. Xu (2009), "Global and Episodic Reports of Hedonic Experience", in Belli, R., D. Alwin and F. Stafford (eds.), *Using Calendar and Diary Methods in Life Events Research*, Sage Publications, Newbury Park, CA. [66]

Seligman, M. (2012), *Flourish A Visionary New Understanding of Happiness and Well-being*, Simon and Schuster. [89]

SES Expert Panel (2012), *Improving the Measurement of Socioeconomic Status for the National Assessment of Educational Progress: A Theoretical Foundation*, U.S. Department of Education, Washington D.C., https://nces.ed.gov/nationsreportcard/pdf/researchcenter/Socioeconomic_Factors.pdf. [194]

Shikako-Thomas, K. et al. (2009), "Quality of Life From the Perspective of Adolescents With Cerebral Palsy: "I Just Think I'm a Normal Kid, I Just Happen to Have a Disability"", *Quality of Life Research*, Vol. 18/7, pp. 825-832, http://dx.doi.org/10.1007/s11136-009-9501-3. [130]

Shochet, I. et al. (2006), "School Connectedness Is an Underemphasized Parameter in Adolescent Mental Health: Results of a Community Prediction Study", *Journal of Clinical Child & Adolescent Psychology*, Vol. 35/2, pp. 170-179, http://dx.doi.org/10.1207/s15374424jccp3502_1. [160]

Sibley, B. and J. Etnier (2003), "The Relationship between Physical Activity and Cognition in Children: A Meta-Analysis", *Pediatric Exercise Science*, Vol. 15/3, pp. 243-256, http://dx.doi.org/10.1123/pes.15.3.243. [107]

Simetin, I. et al. (2011), "Inequalities in Croatian Pupils' Unhealthy Behaviours and Health Outcomes: Role of School, Peers and Family Affluence", *The European Journal of Public Health*, Vol. 21/1, pp. 122-128, http://dx.doi.org/10.1093/eurpub/ckq002. [182]

Statham, J. and E. Chase (2010), "Childhood Wellbeing: A brief overview", *Loughborough: Childhood Wellbeing Research Centre*, http://www.researchgate.net/profile/June_Statham/publication/242676811_Childhood_Wellbeing_A_brief_overview/links/549bd87c0cf2b80371372fc7.pdf. [11]

Stiglitz, J., A. Sen and J. Fitoussi (2009), *Report by the Commission on the Measurement of Economic Performance and Social Progress*, European Commission, Brussels, http://www.stiglitz-sen-fitoussi.fr. [4]

Strong, W. et al. (2005), "Evidence Based Physical Activity for School-age Youth", *The Journal of Pediatrics*, Vol. 146/6, pp. 732-737, http://dx.doi.org/10.1016/j.jpeds.2005.01.055. [110]

Suldo, S. et al. (2013), "American High School Students' Perceptions of Determinants of Life Satisfaction", *Social Indicators Research*, Vol. 118/2, pp. 485-514, http://dx.doi.org/10.1007/s11205-013-0436-2. [189]

Suldo, S. et al. (2009), "Teacher Support and Adolescents' Subjective Well-Being: A Mixed-Methods Investigation", *School Psychology Review*, Vol. 38/1, pp. 67-85. [152]

Suldo, S. and E. Huebner (2006), "Is Extremely High Life Satisfaction During Adolescence Advantageous?", *Social Indicators Research*, Vol. 78/2, pp. 179–203, http://dx.doi.org/10.1007/s11205-005-8208-2. [127]

Suldo, S., K. Riley and E. Shaffer (2006), "Academic correlates of children and adolescents' life satisfaction", *School Psychology International*, Vol. 27/5, pp. 567-582, http://spi.sagepub.com/content/27/5/567.short. [125]

Sullivan, H. (1953), *The Interpersonal Theory of Psychiatry*, Norton, New York. [187]

Tay, L. and E. Diener (2011), "Needs and Subjective Well-Being Around the World", *Journal of Personality and Social Psychology*, Vol. 101/2, pp. 354-365, http://dx.doi.org/10.1037/a0023779. [198]

The Children's Society (2015), *The Good Childhood Report 2015*, The Children's Society, https://www.childrenssociety.org.uk/sites/default/files/TheGoodChildhoodReport2015.pdf. [12]

The Foundation for Child Development (FCD) (2012), *National Child and Youth Well-Being Index (CWI)*, The Foundation for Child Development (FCD). [8]

Tian, L., D. Wang and E. Huebner (2015), "Development and Validation of the Brief Adolescents' Subjective Well-Being in School Scale (BASWBSS)", *Social Indicators Research*, Vol. 120/2, pp. 615-634, http://dx.doi.org/10.1007/s11205-014-0603-0. [82]

Topp, C. et al. (2015), "The WHO-5 Well-Being Index: A Systematic Review of the Literature", *Psychotherapy and Psychosomatics*, Vol. 84/3, pp. 167-176, http://dx.doi.org/10.1159/000376585. [22]

Torsheim, T. et al. (2004), "Material Deprivation and Self-Rated Health: A Multilevel Study of Adolescents From 22 European and North American Countries", *Social Science & Medicine*, Vol. 59/1, pp. 1-12, http://dx.doi.org/10.1016/j.socscimed.2003.09.032. [196]

Torsheim, T. and B. Wold (2001), "School-Related Stress, Support, and Subjective Health Complaints Among Early Adolescents: A Multilevel Approach", *Journal of Adolescence*, Vol. 24/6, pp. 701-713, http://dx.doi.org/10.1006/jado.2001.0440. [143]

Ttofi, M. (ed.) (2011), "Bullying perpetration and victimization as predictors of delinquency and depression in the Pittsburgh Youth Study", *Journal of Aggression, Conflict and Peace Research*, Vol. 3/2, pp. 74-81, http://dx.doi.org/10.1108/17596591111132882. [164]

Ttofi, M. (ed.) (2011), "Do the Victims of School Bullies Tend to Become Depressed Later in Life? A Systematic Review and Meta-Analysis of Longitudinal Studies", *Journal of Aggression, Conflict and Peace Research*, Vol. 3/2, pp. 63-73, http://dx.doi.org/10.1108/17596591111132873. [173]

UNICEF (2012), *Children's Well-Being From Their Own Point of View: What Affects the Children's Well-being in the First Year of Compulsory Secondary Education in Spain?*, UNICEF Spain, Madrid. [10]

UNICEF (2007), *Child Poverty in Perspective: An Overview of Child Well-Being in Rich Countries*, UNICEF Innocenti Research Centre, Florence. [9]

van Soest, A. et al. (2011), "Validating the Use of Anchoring Vignettes for the Correction of Response Scale Differences in Subjective Questions", *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, Vol. 174/3, pp. 575–595, http://dx.doi.org/10.1111/j.1467-985x.2011.00694.x. [62]

Veenstra, R. et al. (2005), "Bullying and Victimization in Elementary Schools: A Comparison of Bullies, Victims, Bully/Victims, and Uninvolved Preadolescents.", *Developmental Psychology*, Vol. 41/4, pp. 672-682, http://dx.doi.org/10.1037/0012-1649.41.4.672. [176]

Verplanken, B. and R. Velsvik (2008), "Habitual Negative Body Image Thinking as Psychological Risk Factor in Adolescents", *Body Image*, Vol. 5/2, pp. 133-140, http://dx.doi.org/10.1016/j.bodyim.2007.11.001. [115]

Vieno, A. et al. (2007), "Social Support, Sense of Community in School, and Self-Efficacy as Resources During Early Adolescence: An Integrative Model", *American Journal of Community Psychology*, Vol. 39/1-2, pp. 177-190, http://dx.doi.org/10.1007/s10464-007-9095-2. [155]

Vingilis, E., T. Wade and J. Seeley (2002), "Predictors of Adolescent Self-rated Health: Analysis of the National Population Health Survey", *Canadian Journal of Public Health / Revue Canadienne de Santé Publique*, Vol. 93/3, pp. 193-197, https://www.jstor.org/stable/41993476. [120]

von Davier, M. (2014), "Imputing Proficiency Data Under Planned Missingness in Population Models", in Rutkowski, L., M. von Davier and D. Rutkowski (eds.), *Handbook of International Large-Scale Assessment: Background, Technical Issues, and Methods of Data Analysis*, CRC Press, Boca Raton, FL. [75]

Wang, M. and J. Eccles (2012), "Adolescent Behavioral, Emotional, and Cognitive Engagement Trajectories in School and Their Differential Relations to Educational Success", *Journal of Research on Adolescence*, Vol. 22/1, pp. 31–39, http://dx.doi.org/10.1111/j.1532-7795.2011.00753.x. [42]

Watson, D., L. Clark and A. Tellegen (1988), "Development and Validation of Brief Measures of Positive and Negative Affect: The PANAS Scales", *Journal of Personality and Social Psychology*, Vol. 54/6, pp. 1063-1070, http://dx.doi.org/10.1037/0022-3514.54.6.1063. [84]

Way, N. and M. Greene (2006), "Trajectories of Perceived Friendship Quality During Adolescence: The Patterns and Contextual Predictors", *Journal of Research on Adolescence*, Vol. 16/2, pp. 293–320, http://dx.doi.org/10.1111/j.1532-7795.2006.00133.x. [41]

Way, N., R. Reddy and J. Rhodes (2007), "Students' Perceptions of School Climate During the Middle School Years: Associations with Trajectories of Psychological and Behavioral Adjustment", *American Journal of Community Psychology*, Vol. 40, pp. 194–213, http://dx.doi.org/10.1007/s10464-007-9143-y. [43]

White, J. (2007), "Wellbeing and Education: Issues of Culture and Authority", *Journal of Philosophy of Education*, Vol. 41/1, pp. 17-28, http://dx.doi.org/10.1111/j.1467-9752.2007.00540.x. [63]

WHO (2010), *Global Recommendations on Physical Activity for Health*, World Health Organization, Geneva, Switzerland, https://apps.who.int/iris/bitstream/handle/10665/44399/9789241599979_eng.pdf;jsessionid=E8D1D04FE62F915F871801A30FCC3DF6?sequence=1. [94]

Williams, J. et al. (2005), "Health-Related Quality of Life of Overweight and Obese Children", *JAMA*, Vol. 293/1, pp. 70-76, http://dx.doi.org/10.1001/jama.293.1.70. [96]

Winkelmann, L. and R. Winkelmann (1998), "Why Are the Unemployed So Unhappy? Evidence from Panel Data", *Economica*, Vol. 65/257, pp. 1-15, http://dx.doi.org/10.1111/1468-0335.00111. [16]

Wright, T., D. Bonett and D. Sweeney (1993), "Mental Health and Work Performance: Results of a Longitudinal Field Study", *Journal of Occupational and Organizational Psychology*, Vol. 66/4, pp. 277-284, http://dx.doi.org/10.1111/j.2044-8325.1993.tb00539.x. [19]

Zambon, A. et al. (2009), "The contribution of club participation to adolescent health: evidence from six countries", *Journal of Epidemiology & Community Health*, Vol. 64/01, pp. 89-95, http://dx.doi.org/10.1136/jech.2009.088443. [185]

Annex A. PISA 2018 Background questionnaires

Annex A presents a link to the background questionnaires used in the PISA 2018 survey. These are the **student questionnaire** *distributed to all participating students; the* **school questionnaire** *distributed to the principals of all participating schools; three optional questionnaires for students (the* **educational career questionnaire**, *the* **ICT familiarity questionnaire**, *and the* **well-being questionnaire**); *an optional* **questionnaire for parents**; *an optional* **questionnaire for teachers** *(both for reading teachers and for teachers of all other subjects); and the* **financial literacy questionnaire** *for students in countries that participated in the financial literacy assessment.*

PISA 2018 Background questionnaires

Please see www.oecd.org/pisa/data for a link to all of the questionnaires that were administered as part of the PISA 2018 assessment.

Annex B. PISA 2018 Expert Groups

Annex B lists the members of the expert groups who were involved in developing the PISA 2018 framework for the major domain (reading), the innovative domain (global competence) and the questionnaires. The lists of the experts involved in developing the PISA 2012 frameworks for mathematics and financial literacy and the PISA 2015 framework for science can be found in the OECD publications PISA 2012 Frameworks – Mathematics, Problem Solving and Financial Literacy *(2013) and* PISA 2015 Assessment and Analytical Framework *(2017), respectively.*

Reading expert group (REG)

Jean-François Rouet (REG Chair)

University of Poitiers

France

Paul van den Broek

Universiteit Leiden

The Netherlands

Kevin Chang

University of Hong Kong

Hong Kong, China

Sascha Schroeder

Max Planck Institute for Human Development, Berlin

Germany

Sari Sulkunen (liaison to the Global Competence expert group)

University of Jyväskylä

Finland

Dominique Lafontaine (liaison to the Questionnaire expert group)

University of Liège

Belgium

Global Competence expert group (GEG)

Responsible for the first part of the development of the Global Competence assessment

Darla Deardorff

Duke University

United States

David Kerr

University of Reading and YoungCitizens

United Kingdom

Peter Franklin

HTWG Konstanz University of Applied Sciences

Germany

Sarah Howie

University of Pretoria

South Africa

Wing On Lee

Open University of Hong Kong

Hong Kong, China

Jasmine B-Y Sim

National Institute of Education

Singapore

Sari Sulkunen (liaison to the Reading expert group)

University of Jyväskylä

Finland

Responsible for the second part of the development of the Global Competence assessment

Martyn Barrett

University of Surrey

United Kingdom

Veronica Boix Mansilla

Harvard University and Project Zero

United States

Darla Deardorff

Duke University

United States

Hye-Won Lee

Korea Institute for Curriculum and Evaluation

Korea

Questionnaire expert group (QEG)

Fons van de Vijver (QEG Chair)

Tilburg University, the North-West University and the University of Queensland

The Netherlands, South Africa and Australia

Dominique Lafontaine (liaison to the Reading expert group)

Université de Liège

Belgium

Sarah Howie

University of Pretoria

South Africa

Andrew Elliot

University of Rochester

United States

Therese Hopfenbeck

University of Oxford

United Kingdom

David Kaplan

University of Wisconsin-Madison

United States

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT

The OECD is a unique forum where governments work together to address the economic, social and environmental challenges of globalisation. The OECD is also at the forefront of efforts to understand and to help governments respond to new developments and concerns, such as corporate governance, the information economy and the challenges of an ageing population. The Organisation provides a setting where governments can compare policy experiences, seek answers to common problems, identify good practice and work to co-ordinate domestic and international policies.

The OECD member countries are: Australia, Austria, Belgium, Canada, Chile, the Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Israel, Italy, Japan, Korea, Latvia, Lithuania, Luxembourg, Mexico, the Netherlands, New Zealand, Norway, Poland, Portugal, the Slovak Republic, Slovenia, Spain, Sweden, Switzerland, Turkey, the United Kingdom and the United States. The European Union takes part in the work of the OECD.

OECD Publishing disseminates widely the results of the Organisation's statistics gathering and research on economic, social and environmental issues, as well as the conventions, guidelines and standards agreed by its members.

OECD PUBLISHING, 2, rue André-Pascal, 75775 PARIS CEDEX 16
ISBN 978-92-64-94031-4 – 2019

PISA

PISA 2018 Assessment and Analytical Framework

The OECD Programme for International Student Assessment (PISA), now in its seventh cycle, seeks to determine what is important for citizens to know and be able to do. It does this by conducting the most comprehensive and rigorous international survey of students' knowledge, skills and well-being.

The *PISA 2018 Assessment and Analytical Framework* presents the conceptual foundations of this programme. Like previous cycles, the 2018 assessment covered reading, mathematics and science, with the major focus this cycle on reading literacy. For the first time, students' global competence – their ability to understand and appreciate the perspectives and world views of others – was evaluated. Financial literacy was offered as an optional assessment.

The framework also includes the questionnaires distributed to students and school principals that were used to elicit information about students' background and the school learning environment, and the questionnaires that some countries distributed to parents and teachers to gain more information about factors associated with student outcomes. Students in some countries also completed questionnaires, included in this publication, on their educational trajectory, their familiarity with information and communications technology, and for the first time, their well-being.

Consult this publication on line at: ***https://doi.org/10.1787/b25efab8-en***

This work is published on the *OECD iLibrary,* which gathers all OECD books, periodicals and statistical databases.
Visit ***www.oecd-ilibrary.org*** for more information.

OECD*publishing*
www.oecd.org/publishing

ISBN 978-92-64-94031-4

9 789264 940314