This chapter defines “reading literacy” as assessed in the Programme for International Student Assessment (PISA) in 2015 and the competencies required for reading literacy. It describes the cognitive processes (aspects) involved in reading that are assessed, the types of texts and response formats used in the assessment, and how student performance in reading is measured and reported.
In PISA 2015, reading literacy is assessed as a minor domain, providing an opportunity to make comparisons in student performance over time. This framework uses the same description and illustrations of the PISA reading assessment as included in the 2009 framework, when reading was re-examined and updated for use as the major domain in that cycle. The framework does not, however, cover digital reading (also referred to as electronic reading in 2009). This is because the 2009 report provided separate scales for print reading and digital reading. Since reading is a minor domain in PISA 2015, and since digital reading was not assessed in all participating countries in 2009 or in 2012, there are no separate data on digital reading, nor was digital reading included as part of the overall concept of reading literacy.

For PISA 2015, the computer is the primary mode of delivery for all domains, including reading literacy. However, paper-based assessment instruments are provided for countries that choose not to test their students by computer. The reading literacy component of both the computer-based and paper-based instruments is composed of the same clusters of reading trend items. The number of trend items in the minor domains are increased compared with previous PISA assessments, thereby 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.

With the move to computer-based delivery for 2015, the 2012 text classification “medium: print and digital” is a potential source of confusion. For 2015, the terminology has been updated to “fixed text” and “dynamic text” to distinguish between delivery mode and the space in which the text is displayed (hereafter referred to as “text display space”), regardless of whether it is printed or on screen. It is important to note, however, that the constructs of the 2009 framework remain unchanged.

| Mode: refers only to the delivery channel. The following distinctions are made: |
| Paper-based: items delivered on paper |
| Computer-based: items delivered on computer |
| Text display space: In 2009, a broad classification, “medium”, was used to describe the features of print and digital texts. For 2015, the classification remains, but is renamed “text display space”. |
| Fixed text: what was previously called “print-medium text”. As this type of text is presented on a screen in PISA 2015, the term “print” no longer applies. |
| Dynamic text: what was previously called “digital-medium text”. As “print-medium” texts are also presented on a screen in PISA 2015, the term “digital” applies to both text display spaces. |
| Digital reading: The term “digital reading assessment” is retained for historical purposes to refer specifically to the 2009/2012 optional domain. |

Note: This new terminology is intended to be provisional, for use only in 2015 when items previously delivered on paper and classified as “print” are delivered on a screen. The purpose is to make a clearer distinction between the mode of delivery and the features of the classification previously known as “medium”. In 2018, when reading literacy will once again become the major domain, both the framework and these terms will be revisited and updated.

In 2015, only fixed-text items are used in the assessment, and these are delivered primarily in a computer-based mode. This is shown in Table 3.1 below.

Table 3.1 Relationship between mode and text display space for 2015

<table>
<thead>
<tr>
<th>Mode/Text display space</th>
<th>Fixed text</th>
<th>Dynamic text</th>
</tr>
</thead>
<tbody>
<tr>
<td>Paper-based mode</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Computer-based mode</td>
<td>✓</td>
<td>✓ (but not assessed in 2015)</td>
</tr>
</tbody>
</table>

Reading literacy was the major domain assessed in 2000, for the first PISA cycle and in 2009, for the fourth PISA cycle. For the sixth PISA cycle (PISA 2015), reading is a minor domain and its framework has not changed from the PISA 2009 cycle (OECD, 2010). There were two major modifications to the PISA 2009 version of the reading framework: the incorporation of an assessment of digital reading and the elaboration of the constructs of reading engagement and metacognition. However, reading is a minor domain in PISA 2015. The reading of digital texts is not included and no data on engagement or metacognition in reading are collected.
The PISA framework for assessing the reading literacy of students towards the end of compulsory education, therefore, must focus on skills that include finding, selecting, interpreting and evaluating information from a full range of texts, including those encountered both inside and outside the classroom.

**DEFINING READING LITERACY**

Definitions of reading and reading literacy have changed over time in parallel with changes in society, economy and culture. The concept of learning, particularly the concept of lifelong learning, has expanded the perception of reading literacy. Literacy is no longer considered to be 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.

Cognitive-based theories of reading literacy emphasise the interactive nature of reading and the constructive nature of comprehension, in the print medium (Binkley and Linnakylä, 1997; Bruner, 1990; Dole et al., 1991) and to an even greater extent in the digital medium (Fastrez, 2001; Legros and Crinon, 2002; Leu, 2007; Reinking, 1994). 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. While constructing meaning, the reader uses various processes, skills and strategies to foster, monitor and maintain understanding. These processes and strategies are expected to vary with context and purpose as readers interact with a variety of continuous and non-continuous texts in the print medium and (typically) with multiple texts in the digital medium.

The PISA 2015 definition of reading literacy, the same as used in PISA 2009, as shown in Box 3.1:

**Box 3.1 The 2015 definition of reading literacy**

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

**Reading literacy…**

The term “reading literacy” is preferred to “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, or even reading aloud, whereas the intention of this survey is to measure something broader and deeper. Reading literacy includes a wide range of cognitive competencies, from basic decoding, to knowledge of words, grammar and larger linguistic and textual structures and features, to knowledge about the world.

In this study, “reading literacy” is intended to express the active, purposeful and functional application of reading in a range of situations and for various purposes. According to Holloway (1999), reading skills are essential to the academic achievement of middle- and high school students. PISA assesses a wide range of students. Some will go on to university; some will pursue further studies in preparation for joining the labour force; some will enter the workforce directly after completing compulsory education. Achievement in reading literacy is not only a foundation for achievement in other subject areas within the education system, but also a prerequisite for successful participation in most areas of adult life (Cunningham and Stanovich, 1998; Smith et al., 2000). Indeed, regardless of their academic or labour-force aspirations, students’ reading literacy is important for their active participation in their community and economic and personal life.

Reading literacy skills matter not just for individuals, but for economies as a whole. Policy makers and others are coming to recognise that in modern societies, human capital – the sum of what the individuals in an economy know and can do – may be the most important form of capital. Economists have for many years developed models showing generally that a country’s education levels are a predictor of its economic growth potential (Coulombe et al., 2004).

**…is understanding, using, reflecting on…**

The word “understanding” is readily connected with “reading comprehension”, a well-accepted element of reading. The word “using” refers to the notions of application and function – doing something with what we read. “Reflecting on” is added to “understanding” and “using” to emphasise the notion that reading is interactive: readers draw on their own thoughts and experiences when engaging with a text. Of course, every act of reading requires some reflection, drawing on information from outside the text. Even at the earliest stages, readers draw on symbolic knowledge to decode a text.
and require a knowledge of vocabulary to construct meaning. As readers develop their stores of information, experience and beliefs, they constantly, often unconsciously, test what they read against outside knowledge, thereby continually reviewing and revising their sense of the text.

...and engaging with...

A reading literate person 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 in 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.

...written texts...

The term “written texts” is meant to include all those coherent texts in which language is used in its graphic form, whether printed or digital. Instead of the word “information”, which is used in some other definitions of reading, the term “texts” was chosen because of its association with written language and because it more readily connotes literary as well as information-focused reading.

These texts do not include aural language artefacts, such as voice recordings; nor do they include film, TV, animated visuals or pictures without words. They do include visual displays, such as diagrams, pictures, maps, tables, graphs and comic strips that include some written language (for example, captions). These visual texts can exist either independently or they can be embedded in larger texts. Digital texts are distinguished from printed texts in a number of respects, including physical readability; the amount of text visible to the reader at any one time; the way different parts of a text and different texts are connected with one another through hypertext links; and, given these text characteristics, the way that readers typically engage with digital texts. To a much greater extent than with printed or hand-written texts, readers need to construct their own pathways to complete any reading activity associated with a digital text.

...in order to achieve one's goals, develop one's knowledge and potential, and 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” spells out the 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 personal life and lifelong education. 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.

ORGANISING THE DOMAIN OF READING

This section describes how the domain is represented, a vital issue because the organisation and representation of the domain determines the test design and, ultimately, the evidence about student proficiencies that can be collected and reported.

Reading is a multidimensional domain. While many elements are part of the construct, not all can be taken into account in building the PISA assessment. Only those considered most important were selected.

The PISA reading literacy assessment is built on three major task characteristics to ensure a broad coverage of the domain:

- **situation**, which refers to the range of broad contexts or purposes for which reading takes place
- **text**, which refers to the range of material that is read
- **aspect**, which refers to the cognitive approach that determines how readers engage with a text.

In PISA, features of the text and aspect variables (but not of the situation variable) are also manipulated to influence the difficulty of a task.

Reading is a complex activity. The elements of reading do not exist independently of one another in neat compartments. The assignment of texts and tasks to framework categories does not imply that the categories are strictly partitioned or that the materials exist in atomised cells determined by a theoretical structure. The framework scheme is provided to ensure coverage, to guide the development of the assessment and to set parameters for reporting, based on what are considered the marked features of each task.
Examples of reading items are available in the *PISA 2012 Assessment and Analytical Framework* (OECD, 2013) and on the PISA website ([www.oecd.org/pisa/](http://www.oecd.org/pisa/)).

**Situation**

The PISA situation variables were adapted from the Common European Framework of Reference (CEFR) developed for the Council of Europe (Council of Europe, 1996). The four situation variables – personal, public, educational and occupational – are described in the following paragraphs.

The *personal* situation relates to texts that are intended to satisfy an individual’s personal interests, both practical and intellectual. This category also includes texts that are intended to maintain or develop personal connections with other people. It includes personal letters, fiction, biography, and informational texts that are intended to be read to satisfy curiosity, as a part of leisure or recreational activities. In the digital medium it includes personal e-mails, instant messages and diary-style blogs.

The *public* category describes the reading of texts that relate to activities and concerns of the larger society. The category includes official documents and information about public events. In general, the texts associated with this category assume a more or less anonymous contact with others; they also therefore include forum-style blogs, news websites and public notices that are encountered both on line and in print.

The content of *educational* texts is usually designed specifically for the purpose of instruction. Printed text books 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 instead assigned by an instructor. The model tasks are those usually identified as “reading to learn” (Sticht, 1975; Stiggins, 1982).

Many 15-year-olds will move from school into the labour force within one to two years. A typical *occupational* reading task is one that involves the accomplishment of some immediate task. It might include searching for a job, either in a print newspaper’s classified advertisement section, or on line; or following workplace directions. The model tasks of this type are often referred to as “reading to do” (Sticht, 1975; Stiggins, 1982).

*Situation* is used in PISA reading literacy to define texts and their associated tasks, and refers to the contexts and uses for which the author constructed the text. The manner in which the situation variable is specified is therefore about supposed audience and purpose, and is not simply based on the place where the reading activity is carried out. 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 *personal* in PISA. As Hubbard (1989) 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 unofficially at school as well. These texts are classified as *public* in PISA. Conversely, textbooks are read both in schools and in homes, and the process and purpose probably differ little from one setting to another. Such texts are classified as *educational* in PISA.

The four categories overlap. In practice, for example, a text may be intended both to delight and to instruct (personal and educational); or to provide professional advice that is also general information (occupational and public). While content is not a variable that is specifically manipulated in this study, by sampling texts across a variety of situations the intent is to maximise the diversity of content that is included in the PISA reading literacy survey.

Table 3.2 shows the desired distribution of items by situation for reading tasks.

<table>
<thead>
<tr>
<th>Situation</th>
<th>Percentage of total items</th>
</tr>
</thead>
<tbody>
<tr>
<td>Personal</td>
<td>30</td>
</tr>
<tr>
<td>Educational</td>
<td>25</td>
</tr>
<tr>
<td>Occupational</td>
<td>15</td>
</tr>
<tr>
<td>Public</td>
<td>30</td>
</tr>
<tr>
<td>Total</td>
<td>100</td>
</tr>
</tbody>
</table>
Text

Reading requires material for the reader to read. In an assessment, that material – a text (or a set of texts) related to a particular task – must be coherent within itself. That is, the text must be able to stand alone without requiring additional material to make sense to the proficient reader. While it is obvious that there are many different kinds of texts and that any assessment should include a broad range, it is not so obvious that there is an ideal categorisation of kinds of texts.

PISA 2009 and PISA 2012

In PISA 2009 and PISA 2012, the addition of digital reading to the framework made this issue still more complex. There were four main text classifications, because of the print and digital reading assessments proposed in these surveys:

- medium: print and digital
- environment: authored, message-based and mixed (only applicable to digital medium)
- text format: continuous, non-continuous, mixed and multiple
- text type: description, narration, exposition, argumentation, instruction and transaction.

PISA 2015

As explained above, in PISA 2015 only the items used previously for the “print reading assessment” are delivered on computer or paper, and there are only two text classifications:

- text format
- text type.

Text display space is a third classification of text with two categories, fixed texts and dynamic text. It is not used in PISA 2015 but will be integrated into the PISA 2018 survey.

In PISA 2015, the term “text display space” is used to describe the features of the space – fixed or dynamic – and not the mode in which the text is presented.

Fixed texts usually appear on paper in forms such as single sheets, brochures, magazines and books, but tend to appear more and more on a screen as PDFs and on e-readers. This development results in further blurring the distinction between what was labelled “print reading” and “digital reading” in the PISA 2009 framework. As PISA 2015 uses only what was labelled “print reading” in 2009 there are no conceptual change in this aspect for PISA 2015. The physical status of the fixed text encourages (though it may not compel) the reader to approach the content of the text in a particular sequence. In essence, such texts have a fixed or static existence. In real life and in the assessment context, the extent or amount of the text is immediately visible to the reader.

When moving the fixed-text “print” reading trend items from paper to computer-based delivery in the 2015 assessment, care needed to be taken to use navigation tools typical of dynamic texts sparingly and only the most obvious among them. Effects of presenting the original paper-based items on the computer were examined during the mode-effect study in the field trial.

Dynamic texts only appear on a screen. Dynamic text is synonymous with hypertext: a text or texts with navigation tools and features that make possible and indeed even require non-sequential reading. Each reader constructs a “customised” text from the information encountered at the links he or she follows. In essence, such texts have an unfixed, dynamic existence. In dynamic texts, typically only a fraction of the available text can be seen at any one time, and often the extent of text available is unknown. No dynamic texts are included in PISA 2015.

The environment classification was a new variable for the PISA 2009 reading framework. Since it applies only to dynamic texts it is not discussed in the 2015 PISA framework.

Text format

An important classification of texts is the distinction between continuous and non-continuous texts.

Texts in continuous and non-continuous format appear in both the print and digital media. Mixed and multiple format texts are also prevalent in both media, particularly so in the digital medium. Each of these four formats is elaborated as follow:

Continuous texts are formed by sentences organised into paragraphs. These may fit into even larger structures, such as sections, chapters, and books (e.g. newspaper reports, essays, novels, short stories, reviews and letters including on e-book readers).
Non-continuous are most frequently organised in matrix format, composed of a number of lists (Kirsch and Mosenthal, 1990) (e.g. lists, tables, graphs, diagrams, advertisements, schedules, catalogues, indexes and forms). They thus require a different approach to reading than continuous texts do.

Many texts are single, coherent artefacts consisting of a set of elements in both a continuous and non-continuous format. In well-constructed mixed texts, the constituents (e.g. a prose explanation, along with a graph or table) are mutually supportive, with coherence and cohesion links throughout. Mixed text in the print medium is a common format in magazines, reference books and reports. In the digital medium, 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.

Multiple texts are defined as those that have been generated independently, and make sense independently; they are juxtaposed for a particular occasion or may be loosely linked together for the purposes of the assessment. The relationship between the texts may not be obvious; they may be complementary or may contradict one another. For example, a set of websites from different companies providing travel advice may or may not provide similar directions to tourists. Multiple texts may have a single “pure” format (for example, continuous), or may include both continuous and non-continuous texts.

Table 3.3 shows the desired distribution of items by text format.

<table>
<thead>
<tr>
<th>Text format</th>
<th>Percentage of total items</th>
</tr>
</thead>
<tbody>
<tr>
<td>Continuous</td>
<td>60</td>
</tr>
<tr>
<td>Non-continuous</td>
<td>30</td>
</tr>
<tr>
<td>Mixed</td>
<td>5</td>
</tr>
<tr>
<td>Multiple</td>
<td>5</td>
</tr>
<tr>
<td>Total</td>
<td>100</td>
</tr>
</tbody>
</table>

Text type

A different categorisation of text is by text type: description, narration, exposition, argumentation, instruction and transaction.

Texts, as they are found in the world, typically resist categorisation; they are usually not written with rules in mind, and tend to cut across categories. In order to ensure that the reading instrument represents different types of reading, PISA categorises texts based on their predominant characteristics.

The following classification of texts used in PISA is adapted from the work of Werlich (1976).

Description is the type of text in which the information refers to properties of objects in space. The typical questions that descriptive texts answer are “what” questions (e.g. a depiction of a particular place in a travelogue or diary, a catalogue, a geographical map, an online flight schedule or a description of a feature, function or process in a technical manual).

Narration is the type of text in which the information refers to properties of objects in time. Narration typically answers questions relating to “when”, or “in what sequence”. “Why characters in stories behave as they do” is another question that narration typically answers (e.g. a novel, a short story, a play, a biography, a comic strip, fictional texts and a newspaper report of an event).

Exposition is the type of text in which the information is presented as composite concepts or mental constructs, or those elements into which concepts or mental constructs can be analysed. The text provides an explanation of how the different elements interrelate in a meaningful whole, and often answers questions about “how” (e.g. a scholarly essay, a diagram showing a model of memory, a graph of population trends, a concept map and an entry in an online encyclopaedia).

Argumentation is the type of text that presents the relationship among concepts or propositions. Argument texts often answer “why” questions. An important sub-classification of argument texts is persuasive and opinionative texts, referring to opinions and points of view. Examples of text in the text type category argumentation are a letter to the editor, a poster advertisement, the posts in an online forum and a web-based review of a book or film.
Instruction is the type of text that provides directions on what to do. The text presents directions for certain behaviours in order to complete a task (e.g. a recipe, a series of diagrams showing a procedure for giving first aid, and guidelines for operating digital software).

Transaction is the kind of text that aims to achieve a specific purpose outlined in the text, such as requesting that something is done, organising a meeting or making a social engagement with a friend. Before the spread of digital communication, this kind of text was a significant component of some kinds of letters and, as an oral exchange, the principal purpose of many phone calls. This text type was not included in Werlich’s (1976) categorisation. It was used for the first time in the PISA 2009 framework because of its prevalence in the digital medium (e.g. everyday e-mail and text message exchanges between colleagues or friends that request and confirm arrangements).

Aspect
Whereas navigation tools and features are the visible or physical features that allow readers to negotiate their way into, around and between texts, aspects are the mental strategies, approaches or purposes that readers use to negotiate their way into, around and between texts.

Five aspects guide the development of the reading literacy assessment tasks:
- retrieving information
- forming a broad understanding
- developing an interpretation
- reflecting on and evaluating the content of a text
- reflecting on and evaluating the form of a text.

As it is not possible to include sufficient items in the PISA assessment to report on each of the five aspects as a separate subscale, these five aspects are organised into three broad aspect categories for reporting on reading literacy:
- access and retrieve
- integrate and interpret
- reflect and evaluate.

Retrieving information tasks, which focus the reader on separate pieces of information within the text, are assigned to the access and retrieve scale.

Forming a broad understanding and developing an interpretation tasks focus the reader on relationships within a text. Tasks that focus on the whole text require readers to form a broad understanding; tasks that focus on relationships between parts of the text require developing an interpretation. The two are grouped together under integrate and interpret.

Tasks addressing the last two aspects, reflecting on and evaluating the content of a text and reflecting on and evaluating the form of a text, are grouped together into a single reflect and evaluate aspect category. Both require the reader to draw primarily on knowledge outside the text and relate it to what is being read. Reflecting on and evaluating content tasks are concerned with the notional substance of a text; reflecting on and evaluating form tasks are concerned with its structure or formal features.
Figure 3.1 shows the relationship between the five aspects targeted in the test development and the three broad reporting aspects.

An elaboration of the three broad aspect categories is given below.

**Access and retrieve**

Accessing and retrieving involves going to the information space provided and navigating in that space to locate and retrieve one or more distinct pieces of information. Access and retrieve tasks can range from locating the details required by an employer from a job advertisement, to finding a telephone number with several prefix codes, to finding a particular fact to support or disprove a claim someone has made.

While retrieving describes the process of selecting the required information, accessing describes the process of getting to the place, the information space, where the required information is located. Some items may require retrieving information only, especially in fixed texts where the information is immediately visible and where the reader only has to select what is appropriate in a clearly specified information space. On the other hand, some items in the dynamic space require little more than accessing (for example, clicking to select an item in a list of search results). However, only the former processes are involved in the access and retrieve tasks in PISA 2015 as the digital reading assessment is not offered. Such access and retrieve items in the fixed-text display space might require readers to use navigation features, such as headings or captions, to find their way to the appropriate section of the text before locating the relevant information. The process of accessing and retrieving information involves skills associated with selecting, collecting and retrieving information.

**Integrate and interpret**

Integrating and interpreting involves processing what is read to make internal sense of a text.

Integrating focuses on demonstrating an understanding of the coherence of the text. Integrating involves connecting various pieces of information to make meaning, whether it be identifying similarities and differences, making comparisons of degree, or understanding cause-and-effect relationships.

Interpreting refers to the process of making meaning from something that is not stated. When interpreting, a reader is identifying the underlying assumptions or implications of part or all of the text.

Both integrating and interpreting are required to form a broad understanding. A reader must consider the text as a whole or in a broad perspective. Students may demonstrate initial understanding by identifying the main topic or message or by identifying the general purpose or use of the text.

Both integrating and interpreting are also involved in developing an interpretation, which requires readers to extend their initial broad impressions so that they develop a deeper, more specific or more complete understanding of what they have read. Integrating tasks include identifying and listing supporting evidence, and comparing and contrasting information in which the requirement is to draw together two or more pieces of information from the text. In order to process either explicit or implicit information from one or more sources in such tasks, the reader must often infer an intended relationship or category. Interpreting tasks may involve drawing an inference from a local context, for example, interpreting the meaning of a word or phrase that gives a particular nuance to the text. This process of comprehension is also assessed in tasks that require the student to make inferences about the author's intention, and to identify the evidence used to infer that intention.

The relationship between the processes of integration and interpretation may therefore be seen as intimate and interactive. Integrating involves first inferring a relationship within the text (a kind of interpretation), and then bringing pieces of information together, therefore allowing an interpretation to be made that forms a new integrated whole.

**Reflect and evaluate**

Reflecting and evaluating involves drawing upon knowledge, ideas 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.

Reflect items may be thought of as those that require readers to consult their own experience or knowledge to compare, contrast or hypothesise. Evaluate items are those that ask readers to make a judgement drawing on standards beyond the text.

Reflecting on and evaluating the content of a text requires the reader to connect information in a text to knowledge from outside sources. Readers must also assess the claims made in the text against their own knowledge of the world.
Often readers are asked to articulate and defend their own points of view. To do so, readers must be able to develop an understanding of what is said and intended in a text. They must then test that mental representation against what they know and believe on the basis of either prior information, or information found in other texts. Readers must call on supporting evidence from within the text and contrast it with other sources of information, using both general and specific knowledge as well as the ability to reason abstractly.

Reflecting on and evaluating the form of a text requires readers to stand apart from the text, to consider it objectively and to evaluate its quality and appropriateness. Implicit knowledge of text structure, the style typical of different kinds of texts and register play an important role in these tasks. Evaluating how successful an author is in portraying some characteristic or persuading a reader depends not only on substantive knowledge but also on the ability to detect subtleties in language.

Some examples of assessment tasks characteristic of reflecting on and evaluating the form of a text include determining the usefulness of a particular text for a specified purpose and evaluating an author’s use of particular textual features in accomplishing a particular goal. The student may also be called upon to describe or comment on the author’s use of style and to identify the author’s purpose and attitude. To some extent, every critical judgement requires the reader to consult his or her own experience; some kinds of reflection, on the other hand, do not require evaluation (for example, comparing personal experience with something described in a text). Thus evaluation might be seen as a subset of reflection.

Inter-relation and interdependence of the three aspects

The three broad aspects defined for PISA reading literacy are not conceived of as entirely separate and independent, but rather as interrelated and interdependent. Indeed, from a cognitive-processing perspective, they can be considered semi-hierarchical: it is not possible to interpret or integrate information without having first retrieved it; and it is not possible to reflect on or evaluate information without having made some sort of interpretation. In PISA, however, the framework description of reading aspects distinguishes approaches to reading that are demanded for different contexts and purposes; these are then reflected in assessment tasks that emphasise one or other aspect. Table 3.4 shows the desired distribution of items by aspect.

<table>
<thead>
<tr>
<th>Aspect</th>
<th>Percentage of total items</th>
</tr>
</thead>
<tbody>
<tr>
<td>Access and retrieve</td>
<td>25</td>
</tr>
<tr>
<td>Integrate and interpret</td>
<td>50</td>
</tr>
<tr>
<td>Reflect and evaluate</td>
<td>25</td>
</tr>
<tr>
<td>Total</td>
<td>100</td>
</tr>
</tbody>
</table>

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 collect evidence of students’ proficiency in reading literacy.

The distribution of tasks across the major framework variables of situation, text and aspect was discussed in the previous section. In this section some of the other major issues in constructing and operationalising the assessment are considered: factors affecting item difficulty, and how difficulty can be manipulated; the choice of response formats; and some issues around coding and scoring. Considerations of moving the fixed-text “print-medium” trend items to computer-based delivery in 2015 are also discussed further in this section.

Factors affecting item difficulty

The difficulty of any reading literacy task depends on an interaction among several variables. Drawing on Kirsch and Mosenthal’s work (see, for example, Kirsch, 2001; Kirsch and Mosenthal, 1990), we can manipulate the difficulty of items by applying knowledge of the following aspect and text format variables.

In access and retrieve tasks, difficulty depends on the number of pieces of information that the reader needs to locate, the amount of inference required, the amount and prominence of competing information, and the length and complexity of the text.

In integrate and interpret tasks, difficulty is affected by 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; the degree and prominence of competing information in the text; and by the nature of the text: the less familiar and the more abstract the content, and the longer and more complex the text, the more difficult the task is likely to be.
In reflect and evaluate tasks, difficulty is affected by the type of reflection or evaluation required (from least to most difficult, the types of reflection are: connecting; explaining and comparing; hypothesising and evaluating); the nature of the knowledge that the reader needs to bring to the text (a task is more difficult if the reader needs to draw on narrow, specialised knowledge rather than broad and common knowledge); the relative abstraction and length of the text; and by the depth of understanding of the text required to complete the task.

In tasks relating to continuous texts, difficulty is influenced by the length of the text, the explicitness and transparency of its structure, how clearly the parts are related to the general theme, and whether there are text features, such as paragraphs or headings, and discourse markers, such as sequencing words.

In tasks relating to non-continuous texts, difficulty is influenced by the amount of information in the text; the list structure (simple lists are easier to negotiate than more complex lists); whether the components are ordered and explicitly organised, for example with labels or special formatting; and whether the information required is in the body of the text or in a separate part, such as a footnote.

Response formats
The form in which the evidence is collected – the response format – varies according to what is considered appropriate given the kind of evidence that is being collected, and also according to the pragmatic constraints of a large-scale assessment. As in any large-scale assessments the range of feasible item formats is limited, with multiple-choice (simple and complex) and constructed response items (where students write their own answer) being the most manageable formats.

Students in different countries are more or less familiar with various response formats. Including items in a variety of formats is likely to provide some balance between more and less familiar formats for all students, regardless of nationality.

To ensure proper coverage of the ability ranges in different countries, to ensure fairness given the inter-country and gender differences observed, and to ensure a valid assessment of the reflect and evaluate aspect, 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 in print reading might also impact on the measurement of trends.

Table 3.5 shows target coding requirements for PISA reading tasks. The distribution is shown in relation to the three aspects of reading literacy assessment. Items that require expert judgement consist of open constructed responses. Items that do not require coder judgement consist of simple multiple-choice, complex-multiple choice and closed constructed response items. The closed constructed response items are those that require the student to generate a response, but require minimal judgement on the part of a coder. For example, a task in which a student is asked to copy a single word from the text, where only one word is acceptable, would be classified as a closed constructed response item. Such items impose a minor cost burden in operational terms and therefore from a pragmatic perspective, these closed constructed response items can be grouped with multiple choice items.

<table>
<thead>
<tr>
<th>Aspect</th>
<th>% of tasks requiring expert judgement in coding</th>
<th>% of tasks not requiring expert judgement in coding</th>
<th>% of test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Access and retrieve</td>
<td>11</td>
<td>14</td>
<td>25</td>
</tr>
<tr>
<td>Integrate and interpret</td>
<td>14</td>
<td>36</td>
<td>50</td>
</tr>
<tr>
<td>Reflect and evaluate</td>
<td>18</td>
<td>7</td>
<td>25</td>
</tr>
<tr>
<td>Total</td>
<td>43</td>
<td>57</td>
<td>100</td>
</tr>
</tbody>
</table>

Table 3.5 indicates that while there is some distribution of items that require coder judgement and those that do not across the aspects, they are not distributed evenly. The reflection and evaluation aspect tasks are assessed through a larger percentage of constructed response items, which require expert coder judgement.

Given that the delivery of the 2015 assessment is computer-based, it may be possible to use computer coding for some responses not requiring expert judgement without affecting the construct or attributes of the items.

Coding and scoring
Codes are applied to test items, either by a more or less automated process of capturing the alternative chosen by the student for a multiple-choice answer, or by a human judge (expert coder) selecting a code that best captures the kind of response given by a student to an item that requires a constructed response. The code is then converted to a score for the item.
For multiple-choice or closed-response format items, the student has either chosen the designated correct answer or not, so the item is scored as 1 (full credit) or 0 (no credit), respectively. For more complex scoring of constructed response items, some answers, even though incomplete, indicate a higher level of reading literacy than inaccurate or incorrect answers, and thus receive partial credit.

Transition from paper-based to computer-based delivery

The main mode of delivery for the previous PISA assessments was paper. In moving to computer-based delivery in 2015, care must be taken to maintain comparability between the assessments. Some of the factors considered when transposing items from paper to computer mode are discussed below.

- **Item types**: The computer provides a range of opportunities for designers of test items, including new item formats (e.g. drag-and-drop, hotspots). Since the purpose of the 2015 assessment is to study trends, there is less opportunity to exploit innovative item types. The majority of response formats remains unchanged in 2015, although some drop-down or hotspot items may be used to enable computer coding of items that were previously scored by experts, but only where no expert judgement is required and the item construct is not affected.

- **Stimulus presentation**: A feature of fixed texts defined in the construct is that “the extent or amount of the text is immediately visible to the reader”. Clearly, it is impossible, both on paper and on a screen, to have long texts displayed on a single page or screen. To allow for this and still satisfy the construct of fixed texts, pagination is used for texts rather than scrolling. Texts that cover more than one page are presented in their entirety before the student sees the first question.

- **IT skills**: Just as paper-based assessments rely on a set of fundamental skills for working with printed materials, so computer-based assessments rely on a set of fundamental 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.

There is research evidence that a computer-based testing environment can influence students’ performance in reading. Some early studies indicated that reading speed was slower in a computer-based environment (Dillon, 1994) and less accurate (Muter et al., 1982), although these studies were conducted on proofreading tasks, not in an assessment situation.

There is a large body of more recent literature on paper- and computer-based tests’ equivalency (see e.g. Macedo-Rouet et al., 2009; Paek, 2005); however these still reveal conflicting findings. A meta-analysis of studies looking at K-12 students’ mathematics and reading achievement (Wang et al, 2008) indicated that, overall, administration mode has no statistically significant effect on scores.

A mode-effects study was conducted as part of the OECD Programme for the International Assessment of Adult Competencies (PIAAC) field trial. 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 was 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 showing that responses could be measured along the same literacy and numeracy scales. This study, along with the results, was written up as part of the *Technical Report of the Survey of Adult Skills* (OECD, 2014). Given this evidence, it was hypothesised that 2009 reading items could 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 Box 1.2.)

**Reporting proficiency in reading**

PISA reports results in terms of proficiency scales that are interpretable for the purposes of policy. In PISA 2015, reading is a minor domain, and fewer reading items are administered to participating students. A single reading literacy scale is reported based upon the overall combined scale for reading.

To capture the progression of complexity and difficulty in PISA 2015, this reading literacy scale is based on the PISA 2009 combined print reading literacy scale and is divided into seven levels. Figure 3.2 describes these seven levels of reading proficiency. Level 6 is the highest described level of proficiency (Level 5 was the highest level before PISA 2009). The bottom level of measured proficiency is Level 1b (for the PISA 2009 and all subsequent PISA reading assessments, Level 1 was re-labelled as Level 1a and a new level was added, Level 1b, that describes students who would previously have been rated as “below Level 1”). These different levels of proficiency allow countries to know more about the kinds of tasks students with very high and very low reading proficiency are capable of performing. Levels 2, 3, 4 and 5 remain the same in PISA 2015 as in PISA 2000.
### Figure 3.2  ● **Summary description of the seven levels of reading proficiency in PISA 2015**

<table>
<thead>
<tr>
<th>Level</th>
<th>Lower score limit</th>
<th>Characteristics of tasks</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>698</td>
<td>Tasks at this level typically require the reader to make multiple inferences, comparisons and contrasts that are both detailed and precise. They require demonstration of a full and detailed understanding of one or more texts and may involve integrating 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. Reflect and evaluate tasks may require the reader to 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 access and retrieve tasks at this level is precision of analysis and fine attention to detail that is inconspicuous in the texts.</td>
</tr>
<tr>
<td>5</td>
<td>626</td>
<td>Tasks at this level that involve retrieving information require the reader to locate and organise several pieces of deeply embedded information, inferring which information in the text is relevant. Reflective tasks require critical evaluation or hypothesis, drawing on specialised knowledge. Both interpretative and reflective 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.</td>
</tr>
<tr>
<td>4</td>
<td>553</td>
<td>Tasks at this level that involve retrieving information require the reader to locate and organise several pieces of embedded information. Some tasks at this level require interpreting the meaning of nuances of language in a section of text by taking into account the text as a whole. Other interpretative tasks require understanding and applying categories in an unfamiliar context. Reflective tasks at this level require readers to 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.</td>
</tr>
<tr>
<td>3</td>
<td>480</td>
<td>Tasks at this level require the reader to locate, and in some cases recognise the relationship between, several pieces of information that must meet multiple conditions. Interpretative tasks at this level require the reader to 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 expectation or negatively worded. Reflective tasks at this level may require connections, comparisons and explanations, or they may require the reader to evaluate a feature of the text. Some reflective 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.</td>
</tr>
<tr>
<td>2</td>
<td>407</td>
<td>Some tasks at this level require the reader to locate one or more pieces of information, which may need to be inferred and may need to meet several conditions. Others require recognising the main idea in a text, understanding relationships, or construing 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 reflective 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.</td>
</tr>
<tr>
<td>1a</td>
<td>335</td>
<td>Tasks at this level require the reader to locate one or more independent pieces of explicitly stated information; to recognise the main theme or author's purpose in a text about a familiar topic, or to 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 reader is explicitly directed to consider relevant factors in the task and in the text.</td>
</tr>
<tr>
<td>1b</td>
<td>262</td>
<td>Tasks at this level require the reader to 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. The text typically provides support to the reader, such as repetition of information, pictures or familiar symbols. There is minimal competing information. In tasks requiring interpretation the reader may need to make simple connections between adjacent pieces of information.</td>
</tr>
</tbody>
</table>
Note

1. This does not preclude the use of several texts in a single task, but each of the texts should be coherent in itself.

References


