Reading Literacy

DEFINITION OF THE DOMAIN

Definitions of reading and reading literacy have changed over time in parallel with changes in society, the economy and culture. The concept of learning, and particularly the concept of lifelong learning, has expanded perceptions of reading literacy and the demands made on it. Literacy is no longer considered an ability only acquired in childhood during the early years of schooling. Instead, it is viewed as an expanding set of knowledge, skills and strategies which individuals build on throughout life in various situations, and through interaction with their peers and with the larger communities in which they participate.

Through a consensus-building process involving the reading experts selected by the participating countries and the OECD/PISA advisory groups, the following definition of reading literacy was adopted for the survey:

"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."

This definition goes beyond the notion of reading literacy as decoding and literal comprehension: it implies that reading literacy involves understanding, using and reflecting on written information for a variety of purposes. It thus takes into account the active and interactive role of the reader in gaining meaning from written texts. The definition also recognises the full scope of situations in which reading literacy plays a role for young adults, from private to public, from school to work, from active citizenship to lifelong learning. It spells out the idea that literacy enables the fulfilment of individual aspirations – from defined aspirations such as gaining an educational qualification or obtaining a job, to those less immediate goals which enrich and extend one's personal life. Literacy also provides the reader with a set of linguistic tools that are increasingly important for meeting the demands of modern societies with their formal institutions, large bureaucracies and complex legal systems.

Readers respond to a given text in a variety of ways as they seek to use and understand what they are reading. This dynamic process involves many factors, some of which can be manipulated in large-scale assessments such as OECD/PISA. These include the reading situation, the structure of the text itself and the characteristics of the questions that are asked about the text (the test rubric). All of these factors are regarded as important components of the reading process and were manipulated in the creation of the items used in the assessment.

In order to use text format, characteristics of the items and situations in constructing the assessment tasks, and later in interpreting the results, the range for each of these factors had to be specified. This allowed for the categorisation of each task so that the weighting of each component could be taken into account in the final assembly of the survey.

TEXT FORMAT

At the heart of the OECD/PISA assessment is a distinction 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. The primary classification of continuous texts is by rhetorical purpose, or text type.

- *Non-continuous texts* (or documents, as they are known in some approaches) can be categorised in two ways. One is the formal structure approach used in the work of Kirsch and Mosenthal (1989-1991). Their work classifies texts by the way underlying lists are put together to construct the various non-continuous text types. This approach is useful for understanding the similarities and differences between types of non-continuous texts. The other method of classification is by everyday descriptions of the formats of these texts. This second approach is used in classifying non-continuous texts in OECD/PISA.

Continuous texts

Text types are standard ways of organising continuous texts by content and author's purpose.

- *Narration* is the type of text in which the information refers to properties of objects in time. Narrative texts typically provide answers to "when", or "in what sequence" questions.

- *Exposition* is the type of text in which the information is presented as composite concepts or mental constructs, or elements into which concepts or mental constructs can be analysed. The text provides an explanation of how the component elements interrelate in a meaningful whole and often answers "how" questions.

- *Description* is the type of text in which the information refers to properties of objects in space. Descriptive texts typically provide an answer to "what" questions.

- *Argumentation* is the type of text that presents propositions as to the relationship between concepts, or other propositions. Argumentative texts often answer "why" questions. Another important sub-classification of argumentative texts is persuasive texts.

- *Instruction* (sometimes called injunction) is the type of text that provides directions on what to do and includes procedures, rules, regulations and statutes specifying certain behaviours.

- A *document or record* is a text that is designed to standardise and conserve information. It can be characterised by highly formalised textual and formatting features.

- *Hypertext* is a set of text slots linked together in such a way that the units can be read in different sequences, allowing readers to follow various routes to the information.

Non-continuous texts

Non-continuous texts are organised differently from continuous texts and so require different kinds of reading approaches. The reader should refer to the work of Kirsch and Mosenthal (1989-1991) for a discussion of the structural approach. According to their work, lists are the most elementary non-continuous texts. They consist of a number of entries that share some property(ies). This shared property may be used as a label or title for the list. Lists may have their entries ordered (*e.g.*, the names of students in a class arranged alphabetically) or unordered (*e.g.*, a list of supplies to be bought at a shop).

Classifying non-continuous texts by their format, as shown below, provides a familiar means of discussing what types of non-continuous texts may be included in the assessment.

- *Charts and graphs* are iconic representations of data. They are used for the purposes of scientific argumentation, and also in journals and newspapers to display numerical and tabular public information in a visual format.

- *Tables and matrices*. Tables are row and column matrices. Typically, all the entries in each column and each row share properties and thus the column and row labels are part of the information structure of the text. Common tables include schedules, spreadsheets, order forms and indexes.

- *Diagrams* often accompany technical descriptions (*e.g.*, demonstrating parts of a household appliance), expository texts and instructive texts (*e.g.*, illustrating how to assemble a household appliance). It is often useful to distinguish procedural (how to) from process (how something works) diagrams.

- *Maps* are non-continuous texts that indicate the geographical relationships between places. There is a variety of types of maps. Road maps mark the distance and routes between identified places. Thematic maps indicate the relationships between locations and social or physical features.

- *Forms* are structured and formatted texts which request the reader to respond to specific questions in specified ways. Forms are used by many organisations to collect data. They often contain structured or pre-coded answer formats. Typical examples are tax forms, immigration forms, visa forms, application forms, statistical questionnaires, etc.

- *Information sheets* differ from forms in that they provide, rather than request, information. They summarise information in a structured way and in such a format that the reader can easily and quickly locate specific pieces of information. Information sheets may contain various text forms as well as lists, tables, figures and sophisticated text-based graphics (headings,

fonts, indentation, borders, etc.) to summarise and highlight information. Timetables, price lists, catalogues and programmes are examples of this type of non-continuous text.

- *Calls and advertisements* are documents designed to invite the reader to do something, *e.g.*, to buy goods or services, attend gatherings or meetings, elect a person to a public office, etc. The purpose of these documents is to persuade the reader. They offer something and request both attention and action. Advertisements, invitations, summonses, warnings and notices are examples of this document format.

- *Vouchers* testify that their owner is entitled to certain services. The information that they contain must be sufficient to show whether the voucher is valid or not. Typical examples are tickets, invoices, etc.

- *Certificates* are written acknowledgements of the validity of an agreement or a contract. They are formalised in content rather than format. They require the signature of one or more persons authorised and competent to bear testimony of the truth of the given statement. Warranties, school certificates, diplomas, contracts, etc. are documents that have these properties.

Figure 2.1 ■ **Distribution of reading literacy tasks, by text format and type**

Reading as a major domain (PISA 2000)

Reading as a minor domain (PISA 2003)

Text format and type	Percentage of tasks by text format and type (%)		Percentage of tasks by text format and type, based on the whole test (%)	
■ ***Continuous***				
Narrative	21	17	14	11
Expository	36	67	24	43
Descriptive	14	17	9	11
Argumentative and persuasive	20	-	13	-
Injunctive	10	-	7	-
TOTAL[1]	**100**	**100**	**68**	**64**
■ ***Non-Continuous***				
Charts and graphs	37	20	12	7
Tables	29	40	9	14
Diagrams	12	-	4	-
Maps	10	10	3	4
Forms	10	30	3	11
Advertisements	2	-	1	-
TOTAL[1]	**100**	**100**	**32**	**36**

1. Data may not always add up to the totals indicated because of roundings.

The distribution and variety of texts that students are asked to read for OECD/PISA are important characteristics of the assessment. Figure 2.1 shows the distributions of tasks for continuous and non-continuous texts in PISA 2000 (reading as major domain) and in PISA 2003 (reading as minor domain). It can be readily seen that in both 2000 and 2003 continuous texts represent two-thirds of the tasks or items contained in the assessment. Within this category, in both cycles, the largest percentage comes from expository texts.

CHARACTERISTICS OF THE ITEMS

Three sets of variables are used to describe the characteristics of the items: the processes (aspects), which set out the task for the examinee; item types, which set out the ways in which examinees are asked to demonstrate their proficiency at the task; and rules for marking, which specify how examinees' answers are to be evaluated. Each of these will be discussed in turn, though the first requires considerably more attention.

Five processes (aspects)

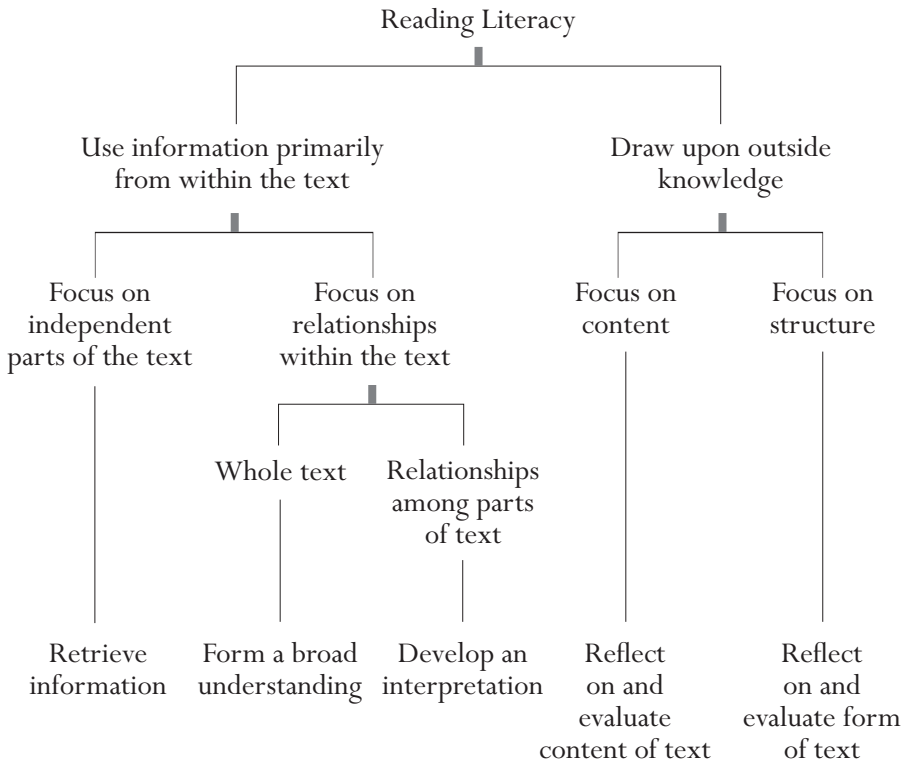
In an effort to simulate authentic reading situations, the OECD/PISA reading assessment measures the following five processes associated with achieving a full understanding of a text, whether the text is continuous or non-continuous. Examinees are expected to demonstrate their proficiency in all of these processes:

- retrieving information,
- forming a broad general understanding,
- developing an interpretation,
- reflecting on and evaluating the content of a text, and
- reflecting on and evaluating the form of a text.

The full understanding of texts involves all of these processes. It is expected that all readers, irrespective of their overall proficiency, will be able to demonstrate some level of competency in each of them (Langer, 1995). While there is an interrelationship between the five aspects – each may require many of the same underlying skills – successfully accomplishing one may not ensure successful completion of any other. Some view them as being in the repertoire of each reader at every developmental level rather than forming a sequential hierarchy or set of skills.

Figure 2.2 identifies the key distinguishing characteristics of the five processes of reading measured in OECD/PISA. While this figure necessarily oversimplifies each process, it provides a useful scheme for organising and remembering the relationships between them. As depicted in this figure, the five processes can be distinguished in terms of four characteristics. The first deals with the extent to which the reader is expected to use information primarily from within the text or to draw also upon outside knowledge.

Figure 2.2 ■ **Characteristics distinguishing the five processes (aspects) of reading literacy**



A second characteristic involves the extent to which the reader is asked to focus on independent parts of the text or on the relationships within the information contained in the text. Sometimes readers are expected to retrieve independent pieces of information while at other times they are asked to demonstrate their understanding of the relationships between parts of the text. Focusing on either the whole text or on relationships between parts of the text is the third distinguishing characteristic. The fourth characteristic relates to whether the reader is asked to deal with the content or substance of the text rather than its form or structure. The five processes of reading are represented in the last line of Figure 2.2 at the ends of the various branches. By starting at the top of the figure and following each branch one can see which characteristics are associated with each process.

The following discussion attempts to define each process operationally and to associate it with particular kinds of items. Although each process is discussed in terms of a single text, each can also apply to multiple texts when these are presented together as a unit within the test. The description of each process has two parts. The first provides a general overview of the process, while the second describes particular ways in which the process might be assessed.

Retrieving information

In the course of daily life, readers often need a particular piece of information: a telephone number or the departure time for a bus or train. They may want to find a particular fact to support or refute a claim someone has made. In situations such as these, readers are interested in retrieving isolated pieces of information. To do so, readers must scan, search for, locate and select relevant information. The processing involved is most frequently at the sentence level, though in some cases the information may be in two or more sentences or in different paragraphs.

In assessment tasks that call for retrieving information, examinees must match information given in the question with either identically worded or synonymous information in the text and use this to find the new information called for. In these tasks, retrieving information is based on the text itself and on explicit information included in it. Retrieving tasks require the examinee to find information based on requirements or features specified in questions. The examinee has to detect or identify one or more essential elements of a question: characters, place/time, setting, etc. and then to search for a match that may be literal or synonymous.

Retrieving tasks can involve various degrees of ambiguity. For example, the examinee may be required to select explicit information, such as an indication of time or place in a text or table. A more difficult version of this same type of task might involve finding synonymous information. This sometimes involves categorisation skills, or it may require discriminating between two similar pieces of information. The different levels of proficiency associated with this process of comprehension can be measured by systematically varying the elements that contribute to the difficulty of the task.

Forming a broad general understanding

To form a broad general understanding of what has been read, a reader must consider the text as a whole or in a broad perspective. There are various assessment tasks in which readers are asked to form a broad general understanding. Examinees may demonstrate initial understanding by identifying the main topic or message or by identifying the general purpose or use of the text. Examples include tasks that require the reader to select or create a title or thesis for the text, to explain the order of simple instructions, or to identify the main dimensions of a graph or a table. Others include tasks that require the examinee to describe the main character, setting or milieu of a story, to identify a theme or message of a literary text, or to explain the purpose or use of a map or a figure.

Within this process some tasks might require the examinee to match a particular piece of text to the question. For example, this would happen when a theme or main idea is explicitly stated in the text. Other tasks may require the examinee to focus on more than one specific reference in the text – for instance, if the reader had to deduce the theme from the repetition of a particular category

of information. Selecting the main idea implies establishing a hierarchy among ideas and choosing the most general and overarching. Such a task indicates whether the examinee can distinguish between key ideas and minor details, or can recognise the summary of the main theme in a sentence or title.

Developing an interpretation

Developing an interpretation requires readers to extend their initial impressions so that they develop a more specific or complete understanding of what they have read. Tasks in this category call for logical understanding; readers must process the organisation of information in the text. To do so, readers must demonstrate their understanding of cohesion even if they cannot explicitly state what cohesion is. In some instances, developing an interpretation may require the reader to process a sequence of just two sentences relying on local cohesion, which might even be facilitated by the presence of cohesive markers, such as the use of "first" and "second" to indicate a sequence. In more difficult instances (*e.g.*, to indicate relations of cause and effect), there might not be any explicit markings.

Examples of tasks that might be used to assess this process include comparing and contrasting information, drawing inferences, and identifying and listing supporting evidence. "Compare and contrast" tasks require the examinee 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. This process of comprehension is also assessed in tasks that require the examinee to make inferences about the author's intention, and to identify the evidence used to infer that intention.

Reflecting on and evaluating the content of a text

Reflecting on and evaluating the content of a text requires the reader to connect information in a text to knowledge from other 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 that with other sources of information, using both general and specific knowledge as well as the ability to reason abstractly.

Assessment tasks representative of this category of processing include providing evidence or arguments from outside the text, assessing the relevance of particular pieces of information or evidence, or drawing comparisons with moral or aesthetic rules (standards). The examinee might be asked to offer or identify alternative pieces of information that might strengthen an author's argument, or to evaluate the sufficiency of the evidence or information provided in the text.

The outside knowledge to which textual information is to be connected may come from the examinee's own knowledge, from other texts provided in the assessment, or from ideas explicitly provided in the question.

Reflecting on and evaluating the form of a text

Tasks in this category require readers to stand apart from the text, consider it objectively and evaluate its quality and appropriateness. Knowledge of such things as text structure, genre and register play an important role in these tasks. These features, which form the basis of an author's craft, figure strongly in understanding standards inherent in tasks of this nature. 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 nuances in language – for example, understanding when the choice of an adjective might colour interpretation.

Some examples of assessment tasks characteristic of reflecting on the form of a text include determining the utility of a particular text for a specified purpose and evaluating an author's use of particular textual features in accomplishing a particular goal. The examinee 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.

Figure 2.3 shows the distribution of reading literacy tasks by each of the three subscales generated from the five reading processes (aspects) defined above. The largest category of tasks, which accounts for approximately 50 per cent of the test, is represented by the two branches of Figure 2.2 that ask students to focus on relationships within a text. These tasks require students either to form a broad understanding or to develop an interpretation. They have been grouped together for reporting purposes into a single process called interpreting texts. In PISA 2000 and 2003, the next largest category was made up of the 29 per cent of the tasks that require students to demonstrate their skill at retrieving isolated pieces of information. Each of these processes – forming a broad understanding, retrieving information and developing an interpretation – focuses on the degree to which the reader can understand and use information contained primarily

Figure 2.3 ■ **Distribution of reading literacy tasks, by reading process (aspect)**

Reading as a major domain (PISA 2000)

Reading as a minor domain (PISA 2003 and 2006)

Reading process (aspect)	Percentage of tasks (%)	
Retrieving information	29	29
Interpreting texts	49	50
Reflection and evaluation	22	21
TOTAL	**100**	**100**

within the text. The remaining of the tasks approximately 20 per cent required students to reflect on either the content or information provided in the text or on the structure and form of the text itself.

Item types

Figure 2.4 indicates that in PISA 2000 and 2003, around 43 per cent of the reading literacy tasks in the OECD/PISA assessment were open constructed-response items which required judgement on the part of the marker. The remaining tasks consist of closed constructed-response items that require little judgement on the part of the marker, as well as simple multiple-choice items, for which students choose one of several alternative answers, and complex multiple-choice items, for which students choose more than one response.

This table also reveals that while multiple-choice and open constructed-response items are represented across the processes, they are not distributed evenly. A larger percentage of multiple-choice items are associated with the two processes dealing with interpreting relationships within a text. This is shown in the second row of Figure 2.4. In contrast, while reflection and evaluation tasks account for around 20 per cent in PISA 2000 and 2003, only 2 per cent in 2000 are multiple-choice. Of the reflection and evaluation tasks, around 20 per cent are open constructed-response items that require judgement on the part of the marker.

Figure 2.4 ■ **Distribution of reading literacy tasks, by reading process (aspect) and item type**

Reading as a major domain (PISA 2000)

Reading as a minor domain (PISA 2003)

Process (aspect)	Item types: Percentage of multiple-choice items		Percentage of complex multiple-choice items		Percentage of closed constructed-response items		Percentage of open constructed-response items[1]		TOTAL[2]	
Retrieving information	8		2	4	6	14	13	11	29	29
Interpreting texts	32	29	2	4	2	7	13	11	49	50
Reflection and evaluation	2		2				18	21	22	21
TOTAL[2]	**42**	**29**	**6**	**7**	**9**	**21**	**44**	**43**	**100**	**100**

1. This category includes short-response items.

2. Data may not always add up to the totals indicated because of rounding.

Marking

Marking is relatively simple with dichotomously scored multiple-choice items: the examinee has either chosen the designated answer or not. Partial-credit models allow for more complex marking of items. Here, because some wrong answers are more complete than others, examinees who provide an "almost right" answer receive partial credit. Psychometric models for such polytomous scoring are well-established and in some ways are preferable to dichotomous scoring, as they utilise more of the information in the responses. Interpretation of polytomous marking is more complex, however, as each task has several locations on the difficulty scale: one for the full-credit answer and others for each of the partial-credit answers. Partial-credit marking is used for some of the more complex constructed-response items in OECD/PISA.

SITUATIONS

The manner in which situation was defined was borrowed from the Council of Europe's (2001) work on language. Four situation variables were identified: reading for private use, reading for public use, reading for work and reading for education. While the intention of the OECD/PISA reading literacy assessment was to measure the kinds of reading that occur both within and outside classrooms, the manner in which situation was defined could not be based simply on where the reading activity is carried out. For example, textbooks are read both in schools and in homes, and the process and purpose of reading these texts differ little from one setting to another. Moreover, reading also involves the author's intended use, different types of content and the fact that others (*e.g.,* teachers and employers) sometimes decide what should be read and for what purpose.

Thus, for the purpose of this assessment, situation can be understood as a general categorisation of texts based on the author's intended use, on the relationship with other persons implicitly or explicitly associated with the text, and on the general content. The sample texts were drawn from a variety of situations to maximise the diversity of content included in the reading literacy survey. Close attention was also paid to the origin of texts selected for inclusion in this survey. The goal was to reach a balance between the broad definition of reading literacy used in OECD/PISA and the linguistic and cultural diversity of participating countries. This diversity helped to ensure that no one group would be either advantaged or disadvantaged by the assessment content.

The four situation variables taken from the work of the Council of Europe can be described as follows:

- *Reading for private use (personal)*. This type of reading is carried out to satisfy an individual's own interests, both practical and intellectual. It also includes reading to maintain or develop personal connections to other people. Contents typically include personal letters, fiction, biography and informational texts read for curiosity, as a part of leisure or recreational activities.

- *Reading for public use*. This type of reading is carried out to participate in the activities of the wider society. It includes the use of official documents as well

as information about public events. In general, these tasks are associated with more or less anonymous contact with others.

- *Reading for work (occupational)*. While not all 15-year-olds will actually have to read at work, it is important to assess their readiness to move into the world of work since, in most countries, over 50% of them will be in the labour force within one to two years. The prototypical tasks of this type are often referred to as "reading to do" (Sticht, 1975; Stiggins, 1982) in that they are tied to the accomplishment of some immediate task.
- *Reading for education*. This type of reading is normally involved with acquiring information as part of a larger learning task. The materials are often not chosen by the reader, but assigned by a teacher. The content is usually designed specifically for the purpose of instruction. The prototypical tasks are those usually identified as "reading to learn" (Sticht, 1975; Stiggins, 1982).

Figure 2.5 shows the distribution of reading literacy tasks in the assessment across all four situations for two scenarios: when reading was a major domain (PISA 2000) and when it is a minor domain (PISA 2003). A more even distribution of tasks across situations is achieved in 2003.

Figure 2.5 ■ **Distribution of reading literacy tasks, by situation**

Reading as a major domain (PISA 2000)
Reading as a minor domain (PISA 2003)

Situation	Percentage of tasks	
Personal	20	21
Public	38	25
Occupational	14	25
Educational	28	29
TOTAL	**100**	**100**

REPORTING OUTCOMES

Scaling the reading literacy tasks

The reading literacy tasks are constructed and administered to nationally representative samples of 15-year-olds in participating countries to ensure that the assessment provides the broadest possible coverage of reading literacy as defined here. However, no individual student can be expected to respond to the entire set of tasks. Accordingly, the survey is designed to give each student participating in the study a subset of the total pool of tasks, while at the same time ensuring that each of the tasks is administered to nationally representative samples of students. Summarising the performance of students across this entire pool of tasks thus poses a challenge.

One may imagine the reading literacy tasks arranged along a continuum in terms of difficulty for students and the level of skill required to answer each item correctly. The procedure used in OECD/PISA to capture this continuum of difficulty and ability is Item Response Theory (IRT). IRT is a mathematical model used for estimating the probability that a particular person will respond correctly to a given task from a specified pool of tasks. This probability is modelled along a continuum which 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. This continuum of difficulty and proficiency is referred to as a "scale".

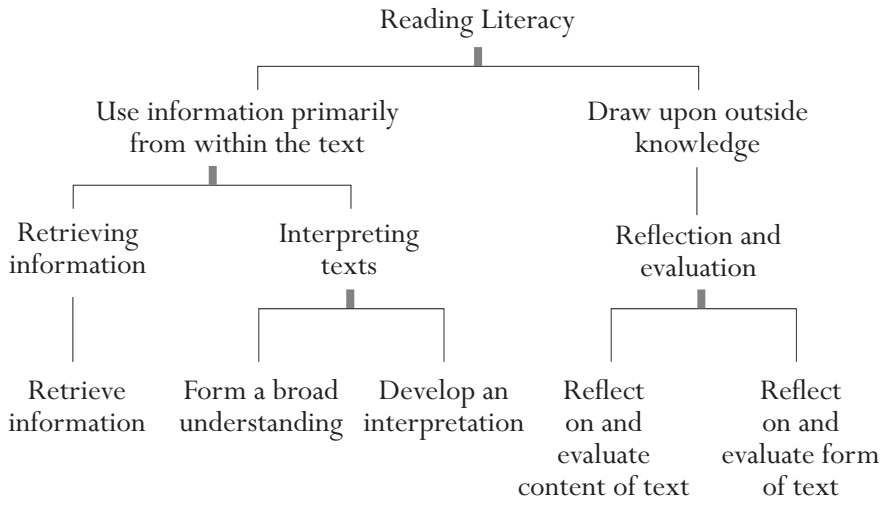
Reporting

PISA 2003 will follow the reporting scheme used in PISA 2000, which reported outcomes in terms of a proficiency scale based on theory and interpretable in policy terms. The results of the reading literacy assessment were first summarised on a single composite reading literacy scale having a mean of 500 and a standard deviation of 100. In addition, student performance was also represented on five subscales: three process (aspect) subscales (retrieving information, interpreting texts, and reflection and evaluation; OECD, 2001a) and two text format subscales (continuous and non-continuous text; OECD, 2002b). These five subscales make it possible to compare mean scores and distributions among subgroups and countries by various components of the reading literacy construct. Although there is a high correlation between these subscales, reporting results on each subscale may reveal interesting interactions among the participating countries. Where such features occur, they can be examined and linked to the curriculum and teaching methodology used. In some countries, the important question may be how to teach the current curriculum better. In others, the question may not only be *how* to teach but also *what* to teach.

The reading process (aspect) subscales

Figure 2.6 summarises the reading literacy tasks in terms of three processes. There are two reasons for reducing the number of process subscales from five to three for reporting purposes. The first is pragmatic. In 2003 and 2006 reading, as a minor domain, will be restricted to about 30 items instead of the 141 that were used in 2000 when reading was a major domain. The amount of information, therefore, will be insufficient to report trends over five process subscales. The second reason is conceptual. The three process subscales are based on the set of five processes shown in Figure 2.2. *Forming a broad understanding* and *developing an interpretation* have been grouped together in an "interpreting texts" subscale because, in both, the reader processes information in the text: in the case of *forming a broad understanding*, the whole text and in the case of *developing an interpretation*, one part of the text in relation to another. *Reflecting on and evaluating the content of a text* and *reflecting on and evaluating the form of a text* have been collapsed into a single "reflection and evaluation" subscale because the distinction between reflecting on and evaluating content and reflecting on and evaluating form, in practice, was found to be somewhat arbitrary.

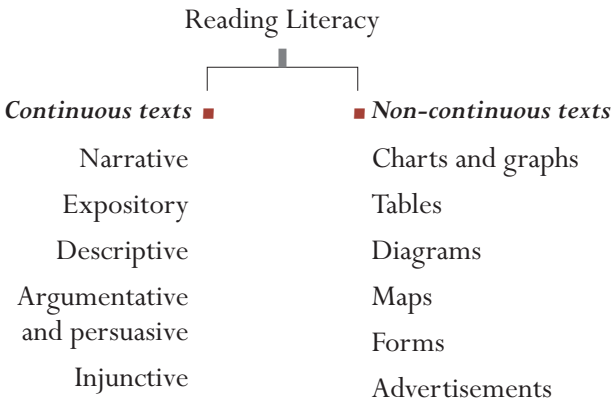
Figure 2.6 ■ **Relationship between the reading literacy framework and the process (aspect) subscales**



The text format subscales

PISA 2003 will also offer the possibility of providing results based on text format subscales, as reported in *Reading for change: Performance and engagement across countries* (OECD, 2002b). Figure 2.7 summarises the various text formats and the associated tasks along the two text format subscales. Organising the data in this way provides the opportunity to examine to what extent countries differ with respect to ability to deal with texts in different formats. In reporting results for 2000, two-thirds of the tasks were used to create the continuous text subscale while the remaining one-third of the tasks were used to create the non-continuous text subscale. There is a similar distribution of tasks between the two text formats in 2003.

Figure 2.7 ■ **Relationship between the reading literacy framework and the text format subscales**



The scores on the composite scale as well as on each of the five subscales represent varying degrees of proficiency. A low score indicates that a student has very limited knowledge and skills, while a high score indicates that a student has quite advanced knowledge and skills. Use of Item Response Theory makes it possible not only to summarise results for various subpopulations of students, but also to determine the relative difficulty of the reading literacy tasks included in the survey. In other words, just as individuals receive a specific value on a scale according to their performance in the assessment tasks, each task receives a specific value on a scale according to its difficulty, as determined by the performance of students across the various countries that participate in the assessment.

Building an item map

The complete set of reading literacy tasks used in OECD/PISA varies widely in text format, situation and task requirements, and hence also in difficulty. This range is captured through what is known as an item map. The item map provides a visual representation of the reading literacy skills demonstrated by students along the scales. The map should contain a brief description of a selected number of released assessment tasks along with their scale values. These descriptions take into consideration the specific skills the item is designed to assess and, in the case of open-ended tasks, the criteria used for judging the item correct. An examination of the descriptions provides some insight into the range of processes required of students and the proficiencies they need to demonstrate at various points along the reading literacy scales.

Figure 2.8 shows an example of an item map from PISA 2000. An explanation of how to interpret it may be useful. The score assigned to each item is based on the theory that someone at a given point on the scale is equally proficient in all tasks at that point on the scale. It was decided that, for the purposes of OECD/PISA, "proficiency" should mean that students at a particular point on the reading literacy scale would have a 62 per cent chance of responding correctly to items at that point. For example, in Figure 2.8 an item appears at 421 on the composite scale. This means that students scoring 421 on the composite reading literacy scale will have a 62 per cent chance of correctly answering items graded 421 on the scale. This does not mean that students receiving scores below 421 will always answer incorrectly. Rather, students scoring below 421 will be expected to answer correctly an item of that level of difficulty less than 62 per cent of the time. Conversely, students having scores above 421 will have a greater than 62 per cent chance of responding correctly. It should be noted that the item will also appear on a process subscale and on a format subscale as well as on the combined reading literacy scale. In this example, the item at 421 on the composite scale requires students to identify the purpose that two short texts have in common by comparing the main ideas in each of them. It is an interpretation item and thus appears on the interpreting texts scale as well as on the continuous texts scale.

Figure 2.8 ■ **An example of a PISA 2000 item map**

○ Types of Process (Aspect) ■ Text Format **Composite item map**	**Types of Process (Aspect)**			**Text Format**	
	Retrieving Information	**Interpreting**	**Reflecting and evaluating**	**Continuous**	**Non- continuous**
822: **HYPOTHESISE** about an unexpected phenomenon by taking account of outside knowledge along with all relevant information in a **COMPLEX TABLE** on a relatively unfamiliar topic. (score 2)			○		■
727: **ANALYSE** several described cases and **MATCH** to categories given in a **TREE DIAGRAM**, where some of the relevant information is in footnotes. (score 2)		○			■
705: **HYPOTHESISE** about an unexpected phenomenon by taking account of outside knowledge along with some relevant information in a **COMPLEX TABLE** on a relatively unfamiliar topic. (score 1)			○		■
652: **EVALUATE** the ending of a **LONG NARRATIVE** in relation to its implicit theme or mood (score 2)			○	■	
645: **RELATE NUANCES OF LANGUAGE** in a **LONG NARRATIVE** to the main theme, in the presence of conflicting ideas. (score 2)		○		■	
631: **LOCATE** information in a **TREE DIAGRAM** using information in a footnote. (score 2)	○				■
603: **CONSTRUE** the meaning of a sentence by relating it to broad context in a **LONG NARRATIVE**.		○		■	
600: **HYPOTHESISE** about an authorial decision by relating evidence in a graph to the inferred main theme of **MULTIPLE GRAPHIC DISPLAYS**.			○		■
581: **COMPARE AND EVALUATE** the style of two open **LETTERS**.			○	■	
567: **EVALUATE** the ending of a **LONG NARRATIVE** in relation to the plot. (score 1)			○	■	
542: **INFER AN ANALOGICAL RELATIONSHIP** between two phenomena discussed in an open **LETTER**.		○		■	
540: **IDENTIFY** the implied starting date of a **GRAPH**.	○				■
539: **CONSTRUE THE MEANING** of short quotations from a **LONG NARRATIVE** in relation to atmosphere or immediate situation. (score 1)		○		■	
537: **CONNECT** evidence from a **LONG NARRATIVE** to personal concepts in order to justify opposing points of view. (score 2)			○	■	…

Figure 2.8 *(continued)* ■ **An example of a PISA 2000 item map**

○ Types of Process (Aspect)

■ Text Format

Composite item map	Types of Process (Aspect)			Text Format	
	Retrieving Information	Interpreting	Reflecting and evaluating	Continuous	Non- continuous
529: **EXPLAIN** a character's motivation by linking events in a **LONG NARRATIVE**.		○		■	
508: **INFER THE RELATIONSHIP** between **TWO GRAPHIC DISPLAYS** with different conventions.		○			■
486: **EVALUATE** the suitability of a **TREE DIAGRAM** for particular purposes.			○		■
485: **LOCATE** numerical information in a **TREE DIAGRAM**. (score 1)	○				■
480: **CONNECT** evidence from a **LONG NARRATIVE** to personal concepts in order to justify a single point of view. (score 1)			○	■	
478: **LOCATE AND COMBINE** information in a **LINE GRAPH** and its introduction to infer a missing value.	○				■
477: **UNDERSTAND** the structure of a **TREE DIAGRAM**.		○			■
473: **MATCH** categories given in a **TREE DIAGRAM** to described cases, when some of the relevant information is in footnotes. (score 1)		○			■
447: **INTERPRET** information in a single paragraph to understand the setting of a **NARRATIVE**.		○		■	
445: Distinguish between variables and **STRUCTURAL FEATURES** of a **TREE DIAGRAM**.			○		■
421: **IDENTIFY** the common **PURPOSE** of **TWO SHORT TEXTS**.		○		■	
405: **LOCATE** pieces of explicit information in a **TEXT** containing strong organizers.	○			■	
397: Infer the **MAIN IDEA** of a simple **BAR GRAPH** from its title.		○			■
392: **LOCATE** a literal piece of information in a **TEXT** with clear text structure.	○			■	
367: **LOCATE** explicit information in a short, specified section of a **NARRATIVE**.	○			■	
363: **LOCATE** an explicitly stated piece of information in a **TEXT** with headings.	○			■	
356: **RECOGNISE THEME** of an article having a clear subheading and considerable redundancy.		○		■	

Levels of reading literacy proficiency

Just as students within each country are sampled to represent the national population of 15-year-old students, each reading literacy task represents a class of tasks from the reading literacy domain. Hence, it represents: proficiency in a type of processing and in dealing with a type of text that 15-year-old students should have acquired. One obvious question is, what distinguishes tasks at the lower end of the scale from those in the middle and upper ranges of the scale? Also, do tasks that fall around the same place on the scale share some characteristics that result in their having similar levels of difficulty? Even a cursory review of the item map reveals that tasks at the lower end of each scale differ from those at the higher end. A more careful analysis of the range of tasks along each scale provides indications of an ordered set of information-processing skills and strategies. Members of the reading expert group examined each task to identify a set of variables that seemed to influence its difficulty. They found that difficulty is in part determined by the length, structure and complexity of the text itself. However, they also noted that in most reading units (a unit being a text and a set of questions), the questions range across the reading literacy scale. This means that while the structure of a text contributes to the difficulty of an item, what the reader has to do with that text, as defined by the question or directive, interacts with the text and affects the overall difficulty.

The members of the reading expert group and test developers identified a number of variables that can influence the difficulty of any reading literacy task. One salient factor is the process involved in retrieving information, developing an interpretation or reflecting on what has been read. Processes range in complexity and sophistication from making simple connections between pieces of information, to categorising ideas according to given criteria, and to critically evaluating and hypothesising about a section of text. In addition to the process called for, the difficulty of retrieving information tasks varies with the number of pieces of information to be included in the response, the number of criteria which the information must satisfy, and whether or not what is retrieved needs to be sequenced in a particular way. In the case of interpretative and reflective tasks, the amount of a text that needs to be assimilated is an important factor affecting difficulty. In items that require reflection on the reader's part, difficulty is also conditioned by the familiarity or specificity of the knowledge that must be drawn on from outside the text. In all processes of reading, the difficulty of the task depends on how prominent the required information is, how much competing information is present, and whether or not the reader is explicitly directed to the ideas or information required to complete the task.

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 five levels:

Level	Score points on the PISA scale
1	335 to 407
2	408 to 480
3	481 to 552
4	553 to 625
5	More than 625

Expert panels judged that the tasks within each level of reading literacy shared many of the same task features and requirements, and differed in systematic ways from tasks at higher or lower levels. As a result, these levels appear to be a useful way to explore the progression of reading literacy demands within each scale. This progression is summarised in Figure 2.9. This process will be repeated for the major domains for each cycle.

Figure 2.9 ■ **Reading literacy levels map**

	Retrieving information	Interpreting texts	Reflecting and evaluating
5	Locate and possibly sequence or combine multiple pieces of deeply embedded information, some of which may be outside the main body of the text. Infer which information in the text is relevant to the task. Deal with highly plausible and/or extensive competing information.	Either construe the meaning of nuanced language or demonstrate a full and detailed understanding of a text.	Critically evaluate or hypothesise, drawing on specialised knowledge. Deal with concepts that are contrary to expectations and draw on a deep understanding of long or complex texts.
	Continuous texts: Negotiate texts whose discourse structure is not obvious or clearly marked, in order to discern the relationship of specific parts of the text to its implicit theme or intention. ***Non-continuous texts***: Identify patterns among many pieces of information presented in a display which may be long and detailed, sometimes by referring to information external to the display. The reader may need to realise independently that a full understanding of the section of text requires reference to a separate part of the same document, such as a footnote.		
4	Locate and possibly sequence or combine multiple pieces of embedded information, each of which may need to meet multiple criteria, in a text with familiar context or form. Infer which information in the text is relevant to the task.	Use a high level of text-based inference to understand and apply categories in an unfamiliar context, and to construe the meaning of a section of text by taking into account the text as a whole. Deal with ambiguities, ideas that are contrary to expectation and ideas that are negatively worded.	Use formal or public knowledge to hypothesise about or critically evaluate a text. Show accurate understanding of long or complex texts.
	Continuous texts: Follow linguistic or thematic links over several paragraphs, often in the absence of clear discourse markers, in order to locate, interpret or evaluate embedded information or to infer psychological or metaphysical meaning. ***Non-continuous texts***: Scan a long, detailed text in order to find relevant information, often with little or no assistance from organisers such as labels or special formatting, to locate several pieces of information to be compared or combined.		

Figure 2.9 *(continued)* ■ **Reading literacy levels map**

	Retrieving information	Interpreting texts	Reflecting and evaluating
3	Locate, and in some cases recognise the relationship between pieces of information, each of which may need to meet multiple criteria. Deal with prominent competing information.	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. Compare, contrast or categorise taking many criteria into account. Deal with competing information.	Make connections or comparisons, give explanations, or evaluate a feature of text. Demonstrate a detailed understanding of the text in relation to familiar, everyday knowledge, or draw on less common knowledge.
	Continuous texts: Use conventions of text organisation, where present, and follow implicit or explicit logical links such as cause and effect relationships across sentences or paragraphs in order to locate, interpret or evaluate information. ***Non-continuous texts***: Consider one display in the light of a second, separate document or display, possibly in a different format, or combine several pieces of spatial, verbal and numeric information in a graph or map to draw conclusions about the information represented.		
2	Locate one or more pieces of information, each of which may be required to meet multiple criteria. Deal with competing information.	Identify the main idea in a text, understand relationships, form or apply simple categories, or construe meaning within a limited part of the text when the information is not prominent and low-level inferences are required.	Make a comparison or connections between the text and outside knowledge, or explain a feature of the text by drawing on personal experience and attitudes.
	Continuous texts: Follow logical and linguistic connections within a paragraph in order to locate or interpret information; or synthesise information across texts or parts of a text in order to infer the author's purpose. ***Non-continuous texts***: Demonstrate a grasp of the underlying structure of a visual display such as a simple tree diagram or table, or combine two pieces of information from a graph or table.		
1	Locate one or more independent pieces of explicitly stated information, typically meeting a single criterion, with little or no competing information in the text.	Recognise the main theme or author's purpose in a text about a familiar topic, when the required information in the text is not prominent.	Make a simple connection between information in the text and common, everyday knowledge.
	Continuous texts: Use redundancy, paragraph headings or common print conventions to form an impression of the main idea of the text, or to locate information stated explicitly within a short section of text. ***Non-continuous texts***: Focus on discrete pieces of information, usually within a single display such as a simple map, a line graph or a bar graph that presents only a small amount of information in a straightforward way, and in which most of the verbal text is limited to a small number of words or phrases.		

Interpreting the reading literacy levels

Not only does each level represent a range of tasks and associated knowledge and skills, it also represents a range of proficiencies demonstrated by students. As mentioned previously, the reading literacy levels were initially set by the members of the reading expert group to represent a set of tasks with shared characteristics. These levels also have shared statistical properties. The average student within each level can be expected to successfully perform the average task within that level 62 per cent of the time. In addition, the width of each level is in part determined by the expectation that a student at the lower end of any level will score 50 per cent on any hypothetical test made up of items randomly selected from that level.

Since each reading literacy scale represents a progression of knowledge and skills, students at a particular level not only demonstrate the knowledge and skills associated with that particular level but the proficiencies associated with the lower levels as well. Thus the knowledge and skills assumed at each level build on and encompass the proficiencies laid down in the next lower level. This means that a student who is judged to be at Level 3 on a reading literacy scale is proficient not only in Level 3 tasks but also in Level 1 and 2 tasks. This also means that students who are at Levels 1 and 2 will be expected to get the average Level 3 item correct less than 50 per cent of the time. Put another way, they will be expected to score less than 50 per cent on a test made up of items drawn from Level 3.

Figure 2.10 shows the probability that individuals performing at selected points along the combined reading literacy scale will give a correct response to tasks of varying difficulty. One is a Level 1 task, one is a Level 3 task, and the third task receives two score points: one at Level 4 and the other at Level 5. It is readily seen here that a student with a score of 298, who is estimated to be below Level 1, has only a 43 per cent chance of responding correctly to the Level 1 task that is at 367 on the reading literacy scale. This person has only a 14 per cent chance of responding to the item from Level 3 and almost no chance of responding correctly to the item from Level 5. Someone with a proficiency of 371, in the middle of Level 1, has a 63 per cent chance of responding to the item at 367, but only slightly more than one chance in four of responding correctly to the task at 508, and only a seven per cent chance of responding correctly to the task selected from Level 5. In contrast, someone at Level 3 would be expected to respond correctly 89 per cent of the time to tasks at 367 on the reading literacy scale, and 64 per cent of the time to tasks at 508, near the middle of Level 3. However, he or she would only have just over one chance in four (27 per cent) of correctly responding to items from the middle of Level 5. Finally, a student at Level 5 is expected to respond correctly most of the time to most of the tasks. As shown in Figure 2.10, a student having a score of 662 on the combined reading literacy scale has a 98 per cent chance of answering the task at 367 correctly, a 90 per cent chance of answering the item at Level 3 (508) correctly and a 65 per cent of responding correctly to the task selected from near the centre of Level 5 (652).

Figure 2.10 ■ **Probability (as a percentage) of responding correctly to selected tasks of varying difficulty for students with varying levels of proficiency**

	Level 1 item at 367 points	Level 3 item at 508 points	Level 4 item at 567 points	Level 5 item at 652 points
Below Level 1 *(Proficiency of 298 points)*	43	14	8	3
Level 1 *(Proficiency of 371 points)*	63	27	16	7
Level 2 *(Proficiency of 444 points)*	79	45	30	14
Level 3 *(Proficiency of 517 points)*	89	64	48	27
Level 4 *(Proficiency of 589 points)*	95	80	68	45
Level 5 *(Proficiency of 662 points)*	98	90	82	65

Figure 2.10 also implicitly raises questions concerning the highest and lowest designated levels. Even though the top of the reading literacy scale is unbounded, it can be stated with some certainty that students of extremely high proficiency are capable of performing tasks characterised by the highest level of proficiency. There is more of an issue for students who are at the bottom end of the reading literacy scale. Level 1 begins at 335, yet a certain percentage of students in each country is estimated to be below this point on the scale. While there are no reading literacy tasks with a scale value below 335, it is not correct to say that these students are without any reading literacy skills or are "totally illiterate". However, on the basis of their performance in the set of tasks used in this assessment, they would be expected to score less than 50 per cent on a set of tasks selected from Level 1. They are classified, therefore, as performing below Level 1.

Since comparatively few young adults in our societies have no literacy skills, the framework does not call for a measure of whether or not 15-year-old students can read in a technical sense. That is, OECD/PISA does not measure the extent to which 15-year-old students are fluent readers or how competent they are at word recognition tasks or spelling. It does, however, reflect the contemporary view that students should, upon completing compulsory education, be able to construct, extend and reflect on the meaning of what they have read across a wide range of continuous and non-continuous texts commonly associated with a variety of situations both within and outside school. While it was not possible to say what knowledge and skills students performing below Level 1 may possess with regard to reading literacy, their level of proficiency indicates that these students are unlikely to be able to use reading independently as a tool to assist them in acquiring knowledge and skills in other areas. ⌟