15

Proficiency Scale Construction

INTRODUCTION

This chapter discusses the methodology used to develop the PISA reporting scales that describe a number of levels of proficiency in the different PISA literacy variables, and presents the outcomes of that development process.

For many years, the Australian Council for Educational Research (ACER) has used and progressively refined an approach to substantive interpretation of scales based on item calibration, employing a reporting mechanism generally known as “described proficiency scales”, alternatively referred to more recently as “learning metrics”, as part of its analysis and reporting of test results. The approach has its origins in work of Benjamin Wright and his collaborators at the University of Chicago from the 1960s. An early published example of a dimension laid out using Rasch-based item calibrations and illustrated with the items and their characteristics is found in Wright and Stone (1979). A similar approach has been used in a number of Australian assessment projects, dating back at least to the TORCH project that originated in Western Australia in 1982-1983 and was published later by ACER (Mossenson, Hill and Masters, 1987), the Basic Skills Testing Programme in New South Wales in 1989 (Masters et al., 1990), as well as in many more recent projects. ACER has used the approach in the reporting of PISA results from its inception: for two administrations in which reading literacy was the major test domain, two in which mathematics was the major domain, and one in which science took centre stage. The same approach was also used to report problem solving in 2003 and 2012, digital reading in 2009, and financial literacy in 2012. Reporting the results of the National Assessment of Educational Progress (NAEP) in the United States since 1990 has also used substantive descriptions of typical accomplishments at points along their reporting scales (see Bourque, 2009) for a history of NAEP reporting), using a consensus-based approach along with ‘scale anchoring’ to define levels and cut-points (Beaton and Allen, 1992).

This chapter presents the methodology, and the products of the application of the methodology, for reporting of PISA 2012 survey outcomes. PISA reports student performance not just as numerical scores, but also in terms of content, by describing what students who achieve a given level on a PISA scale typically know and can do. This chapter explains how these described proficiency scales are developed, and also how the results are reported and how they can be interpreted.

PISA has adopted an approach to reporting survey outcomes that involves the development of learning metrics, which are dimensions of educational progression. A learning metric is usually depicted as a line with numerical gradations that quantify how much of the measured variable is present. Locations along this metric can be specified by numerical ‘scores’, or can be described substantively, hence the label for these metrics used in PISA: described proficiency scales. The scales are called “proficiency scales” rather than “performance scales” because they report what students typically know and can do at given levels, rather than what the individuals who were tested actually did on a single occasion (the test administration). This is because PISA is interested in reporting general results, rather than the results of individuals. PISA uses samples of students and items to make estimates about populations: a sample of 15-year-old students is selected to represent all the 15-year-olds in a country, and a sample of test items from a large pool is administered to each student. Results are then analysed using statistical models that estimate the likely proficiency of the population, based on this sampling.

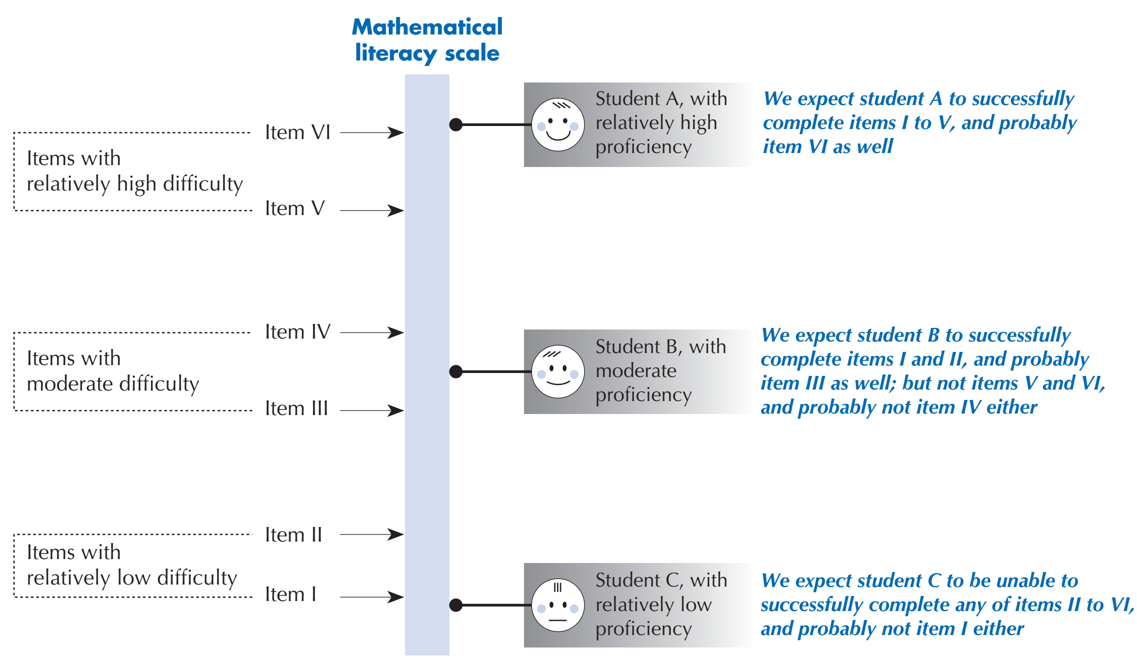
The PISA test design makes it possible to use techniques of modern item response modelling (see Chapter 9) to simultaneously estimate the ability of all students taking the PISA assessment, and the difficulty of all PISA items, locating these estimates of student ability and item difficulty on a single continuum. In this context, the single continuum is a way to represent the variable of interest – the “student ability” is determined by the extent to which a student possesses the key components of the variable, and the “item difficulty” is determined by the extent to which responding to the item requires activation of the variable.

The relative ability of students taking a particular test can be estimated by considering the proportion of test items to which they provide a correct response, and the difficulty of the items. The relative difficulty of items in a test can be estimated by considering the proportion of test takers getting each item correct, and the ability of the students. The mathematical model employed to analyse PISA data, generated from a rotated test design in which students take different but overlapping tasks, is implemented through test analysis software that uses iterative procedures to simultaneously estimate the likelihood that a particular person will respond correctly to a given test item, and the likelihood that a particular test item will be answered correctly by a given student. The result of these procedures is a set of estimates that enables a continuum (the learning metric) to be defined, which is a realisation of the variable of interest. On that continuum it is possible to estimate the location of individual students, thereby seeing how much of the variable of interest they demonstrate, and it is possible to estimate the location of individual test items, thereby seeing how much of the variable each item embodies. This continuum is referred to as the overall PISA literacy scale in the relevant test domain (such as reading, mathematics or science).

PISA assesses students, and uses the outcomes of that assessment to produce estimates of students' proficiency in relation to a number of literacy variables. These variables are defined in the relevant PISA literacy framework (OECD, 2013). For each of these literacy variables, one or more scales are defined, which stretch from very low levels of literacy through to very high levels. What such a scale means in terms of student proficiency is that students whose ability estimate places them at a certain point on the PISA literacy scale would most likely be able to successfully complete tasks at or below that location, and increasingly more likely to complete tasks located at progressively lower points on the scale, but would be less likely to be able to complete tasks above that point, and increasingly less likely to complete tasks located at progressively higher points on the scale. Figure 15.1 depicts a literacy scale, stretching from relatively low levels of literacy at the bottom of the figure, to relatively high levels towards the top. Six items of varying difficulty are placed along the scale, as are three students of varying ability. The relationship between the students and items at various levels is described.

■ Figure 15.1 ■

The relationship between items and students on a proficiency scale



It is possible to describe the scales using words that encapsulate various demonstrated competencies typical of students possessing varying amounts of the underlying literacy constructs. Each student's location on those scales is estimated, and those location estimates are then aggregated in various ways to generate and report useful information about the literacy levels of 15-year-old students within and among participating countries.

Development of the details of the method of describing proficiency in PISA reading, mathematical and scientific literacy occurred in the lead-up to the reporting of outcomes of the PISA 2000 survey and was revised in the lead-up to the PISA 2003, 2006 and 2009 surveys. Essentially, the same methodology has again been used to develop proficiency descriptions for PISA 2012. Given the volume and breadth of data that were available from the PISA 2012 assessment when mathematics was the major assessment domain, review and extension of the descriptions of mathematical literacy that had been developed from the PISA 2003 data became possible. The detailed proficiency descriptions that had been developed for the reading domain in PISA 2009 were used again, and the descriptions used for science in 2006 were used again, in both cases with the reduced data available from the 2012 administration in which those were minor assessment domains. In addition, new described proficiency scales for problem solving and for financial literacy were developed.

The Mathematics Expert Group worked with the PISA international contractor to review and revise the sets of described proficiency scale and subscales for PISA mathematics. Similarly, the international contractor worked with the Problem Solving and Financial Literacy Expert Groups to develop the described proficiency scales for these domains.

DEVELOPMENT OF THE DESCRIBED SCALES

Since PISA 2000, the development of described proficiency scales for PISA has been carried out through a process involving a number of stages. The stages are described here in a linear fashion, but in reality the development process involved some backwards and forwards movement where stages were revisited and descriptions were progressively refined.

Stage 1: Identifying possible scales

The first stage in the process involved the experts in each domain articulating possible reporting scales (dimensions) for the domain.

In the case of mathematics, a single proficiency scale was originally developed for PISA 2000. With the additional data available in the 2003 survey cycle, when mathematics was the major test domain for the first time, the possibility of reporting according to the four overarching ideas or the three competency clusters described in the PISA mathematics framework applicable at that time were both considered. Accordingly, in 2003 subscales based on the four overarching ideas – *space and shape*, *change and relationships*, *quantity* and *uncertainty* – were reported. In PISA 2006 and PISA 2009, when mathematics was again a minor domain, a single mathematics scale only was reported.

For PISA 2012, a review of the reporting structure for mathematics was carried out by the expert group as part of a comprehensive revision of the framework, in conjunction with ACER staff, and at the specific behest of the PISA Governing Board that had indicated clearly that it was interested in seeing mathematical process dimensions used as the primary basis for reporting in mathematics. As well as considering ways in which this could be done, the mathematics expert group also had to consider how the addition of an optional computer-based assessment component could be incorporated in the reporting of the PISA mathematical outcomes. The result of these considerations was firstly, that the computer-based items would be used to expand the scope of expression of the same mathematical literacy dimension that is expressed through the paper-based items; and secondly that the reporting of three process-based subscales labelled *formulating situations mathematically* (usually abbreviated to "formulate"), *employing mathematical concepts, facts, procedures and reasoning* (usually abbreviated to "employ"), and *interpreting, applying and evaluating mathematical outcomes* (with the abbreviation "interpret") would be supported. In addition, for continuity with the PISA 2003 reporting scales, the content-based scales were also reported, with the labels *space and shape*, *change and relationships*, *quantity* and *uncertainty and data* (the latter being the same dimension as the previous *uncertainty* subscale, but with a new label).

For reading in the PISA 2000 survey cycle, two main options were actively considered – scales based on the type of reading task, and scales based on the form of reading material. For the international report, the first of these was implemented, leading to the development of scales to describe the types of reading tasks, or "aspects" of reading: a subscale for *retrieving information*, a second subscale for *interpreting texts* and a third for *reflection and evaluation*. The thematic report for PISA 2000, *Reading for Change*, also reported on the development of subscales based on the form of reading material: *continuous texts* and *non-continuous texts* (OECD, 2002). Volume I of the *PISA 2009 Results* included descriptions of both of these sets of subscales as well as a combined print reading scale (OECD, 2010). The names of the aspect subscales were modified in order to better apply to digital as well as print reading tasks. The modified aspect category names are *access and retrieve* (replacing *retrieving information*), *integrate and interpret* (replacing *interpreting texts*) and *reflect and evaluate* (for *reflection and evaluation*). For digital reading, a separate, single scale was developed based on the digital reading assessment items administered in 19 countries in PISA 2009 as an international option (OECD, 2011). For PISA 2012, when reading reverted to minor domain status, a single print reading scale was reported, along with a single digital reading scale.

For science, given the small number of items in PISA 2000 and 2003, a single overall proficiency scale was developed to report results. As with mathematics in 2003, the expanded focus on science in 2006 allowed for a division into scales for reporting purposes. Two forms of scale were considered. One of these was based on definitions of scientific competencies involving the identification of scientific issues, the explanation of phenomena scientifically and the use of

scientific evidence. The other form separated scientific knowledge into “knowledge of science” involving the application of scientific concepts in the major fields of physics, chemistry, biology, earth and space science, and technology; and “knowledge about science” involving the central processes underpinning the way scientists go about obtaining and using data – in other words, understanding scientific methodology. The scales finally selected for inclusion in the PISA 2006 database were the three competency-based subscales: *identifying scientific issues*, *explaining phenomena scientifically* and *using scientific evidence* (OECD, 2007). In PISA 2009 and PISA 2012, science as a minor domain was reported as a single scale only.

Wherever subscales were under consideration, they arose clearly from the framework for the domain, they were seen to be meaningful and potentially useful for feedback and reporting purposes, and they needed to be defensible with respect to their measurement properties. Due to the longitudinal nature of the PISA project, the decision about the number and nature of reporting scales also had to take into account the fact that in some test cycles a domain will be treated as minor and in other cycles as major.

For problem solving, and for the optional assessment component of financial literacy, in both of which a rather limited volume of data were available based on a relatively small number of test items, proficiency descriptions of a single overall dimension were developed in each domain.

Stage 2: Assigning items to scales

The second stage in the process was to associate each test item used in the study with each of the subscales under consideration. Domain experts (including members of the relevant subject matter expert group, the test developers and staff of the international contractor) judged the characteristics of each test item against the relevant framework categories.

Stage 3: Skills audit

The next stage involved a detailed expert analysis of each item, and in the case of items with partial credit, for each score step within the item, in relation to the definition of the relevant subscale from the domain framework. The skills and knowledge required to achieve each score step were identified and described.

This stage involved negotiation and discussion among the experts involved, circulation of draft material, and progressive refinement of drafts on the basis of expert input and feedback. Further detail on this analysis is provided below.

Stage 4: Analysing Field Trial data

For each set of scales being considered, the Field Trial item data were analysed using item response techniques to derive difficulty estimates for each achievement threshold for each item.

Many items had a single achievement threshold (associated with students providing a correct rather than incorrect response). Where partial credit was available, more than one achievement threshold could be calculated (achieving a score of one or more rather than zero, two or more rather than one, and so on).

Within each scale, achievement thresholds were placed along a difficulty continuum linked directly to student abilities. This analysis gives an indication of the utility of each scale from a measurement perspective.

Stage 5: Defining the dimensions

The information from the domain-specific expert analysis (Stage 3) and the statistical analysis (Stage 4) were combined. For each set of scales being considered, the item score steps were ordered according to the magnitude of their associated thresholds and then linked with the descriptions of associated knowledge and skills, giving a hierarchy of knowledge and skills that defined the dimension. Clusters of skills were found using this approach, which provided a basis for understanding each dimension and describing proficiency in different regions of the scale.

Stage 6: Revising and refining with Main Survey data

When the Main Survey data became available, the information arising from the statistical analysis about the relative difficulty of item thresholds was updated. This enabled a review and revision of Stage 5. The preliminary descriptions and levels were then reviewed and revised, and the approach to defining levels and associating students with those levels that had been used in the reporting of PISA 2000, 2003, 2006 and 2009 results was applied.

DEFINING AND INTERPRETING PROFICIENCY LEVELS

How should we divide the proficiency continuum up into levels that might have some utility? And having defined levels, how should we decide on the level to which a particular student should be assigned? What does it mean to be at a level?

The relationship between the student and the items is probabilistic: that is, there is some probability that a particular student can correctly answer any particular item. If a student is located at a point above an item, the probability that the student can successfully complete that item is relatively high, and if the student is located below the item, the probability of success for that student on that item is relatively low. This leads to the question as to the precise criterion that should be used to locate a student on the same scale as that on which the items are laid out. When placing a student at a particular point on the scale, what probability of success should we deem sufficient in relation to items located at the same point on the scale? If a student were given a test comprising a large number of items each with the same specified difficulty, what proportion of those items would we expect the student to successfully complete? Or, thinking of it in another way, if a large number of students of equal ability were given a single test item having a specified item difficulty, about how many of those students would we expect to successfully complete the item?

The answer to these questions is essentially arbitrary, but in order to define and report PISA outcomes in a consistent manner, we need an approach to defining performance levels, and to associating students with those levels. This is both a technical and very practical matter of interpreting what it means to be at a level, and has very significant consequences for reporting national and international results. The methodology that was developed and used for PISA 2000, 2003, 2006 and 2009 was essentially retained for PISA 2012.

Several principles were considered for developing and establishing a useful meaning for being at a level, and therefore for determining an approach to locating cut-off points between levels and associating students with them. The overriding need to develop and promote a common understanding of the meaning of levels was recognised. First, it is important to understand that the literacy skills measured in PISA must be considered as continua: there are no natural breaking points to mark borderlines between stages along these continua. Dividing each of these continua into levels, though useful for communication about students' development, is essentially arbitrary. Like the definition of units on, for example, a scale of length, there is no fundamental difference between 1 metre and 1.5 metres – it is a matter of degree. It is useful, however, to define stages, or levels along the continua, because they enable us to communicate about the proficiency of students in terms other than numbers. The approach adopted for PISA 2000 was that it would only be useful to regard students as having attained a particular level if this would mean that we can have certain expectations about what these students are capable of in general when they are said to be at that level. It was decided that this expectation would have to mean at a minimum that students at a particular level would be more likely than not to successfully complete tasks at that level. By implication, it must be expected that they would succeed on at least half of the items on a test composed of items uniformly spread across that level. This definition of being "at a level" is useful in helping to interpret the proficiency of students at different points across the proficiency range defined at each level.

For example, students at the bottom of a level would complete at least 50% of tasks correctly on a test set at the level, while students at the middle and top of each level would be expected to achieve a higher success rate. At the top end of the bandwidth of a level would be the students who have mastered that level. These students would be likely to solve a high proportion of the tasks at that level. But, being at the top border of that level, they would also be at the bottom border of the next level up, where according to the reasoning here they should have a likelihood of at least 50% of solving any tasks defined to be at that higher level.

Further, the meaning of being at a level for a given scale should be more or less consistent for each level, indeed also for scales from the different domains. In other words, to the extent possible within the substantively based definition and description of levels, cut-off points should create levels of more or less constant breadth. Some small variation may be appropriate, but in order for interpretation and definition of cut-off points and levels to be consistent, the levels have to be about equally broad within each scale. Clearly this would not apply to the highest and lowest proficiency levels, which are unbounded.

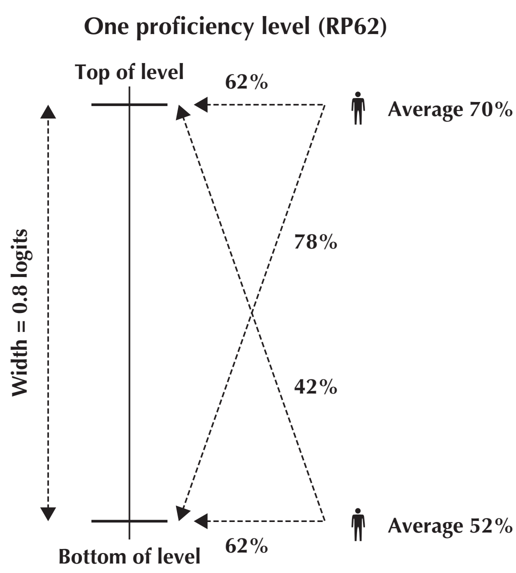
A more or less consistent approach should be taken to defining levels for the different scales. Their breadth may not be exactly the same for the proficiency scales in different domains, but the same kind of interpretation should be possible for each scale that is developed. The approach links the two variables mentioned in the preceding paragraphs, and a third related variable. The three variables can be expressed as follows:

- the expected success of a student at a particular level on a test containing items at that level (proposed to be set at a minimum that is near 50% for the student at the bottom of the level, and higher for other students in the level);
- the width of the levels in that scale (determined largely by substantive considerations of the cognitive demands of items at the level and observations of student performance on the items); and
- the probability that a student in the middle of a level would correctly answer an item of average difficulty for that level (in fact, the probability that a student at any particular level would get an item at the same level correct), sometimes referred to as the "RP-value" for the scale (where "RP" indicates "response probability").

Figure 15.2 summarises the relationship among these three mathematically linked variables under a particular scenario. The vertical line represents a segment of the proficiency scale, with marks delineating the "top of level" and "bottom of level" for any level one might want to consider, with a width of 0.8 logits between the boundaries of the level (but note that this width can vary somewhat for different domains). The RP62 indicates that any person will be located on the scale at a point that gives him or her a 62% chance of getting an item at that same level correct. The person represented near the top of the level shown has a 62% chance of getting an item correct that is located at the top of the level, and similarly the person represented at the bottom of the level has the same chance of correctly answering a question at the bottom of the level. A person at the bottom of the level will have an average score of about 52% correct on a set of items spread uniformly across the level. Of course that person will have a higher likelihood (62%) of getting an item at the bottom of the level correct, and a lower likelihood (about 42%) of getting an item at the top of the level correct. A person at the top of the level will have an average score of about 70% correct on a set of items spread uniformly across the level. Of course that person will have a higher likelihood (about 78%) of getting an item at the bottom of the level correct, and a lower likelihood (62%) of getting an item at the top of the level correct.

■ Figure 15.2 ■

Calculating the RP value used to define PISA proficiency level



PISA 2000 implemented the following solution that was then used in all subsequent survey administrations: start with the range of described abilities for each bounded level in each scale (the desired band breadth); then determine the highest possible RP value that will be common across domains potentially having bands of slightly differing breadth that would give effect to the broad interpretation of the meaning of being at a level (an expectation of correctly responding to a minimum of 50% of the items in a test comprising items spread uniformly across that level). The value RP=0.62 is a probability value that satisfies the logistic equations through which the scaling model is defined, subject to the two constraints mentioned earlier (a width per level of about 0.8 logits and the expectation that a student would get at

least half of the items correct on a hypothetical test composed of items spread evenly across the level). In fact RP=0.62 satisfies the requirements for any scales having band widths up to about 0.97 logits.

With the constraint of a minimum 50% mentioned above, which is central to the definition of the PISA proficiency levels, the RP value required for scales composed of bands of other widths is given by the equation in Figure 15.2, where *x* is the width of the bands.

$$RP \geq \frac{\exp\left(x/2\right)}{1+\exp\left(x/2\right)}$$

The highest and lowest levels are unbounded. For a certain high point on the scale and below a certain low point, the proficiency descriptions could, arguably, cease to be applicable. At the high end of the scale, this is not such a problem since extremely proficient students could reasonably be assumed to be capable of at least the achievements described for the highest level. At the other end of the scale, however, the same argument does not hold. A lower limit therefore needs to be determined for the lowest described level, below which no meaningful description of proficiency is possible. It was proposed that the floor of the lowest described level be set so that it was the same breadth as the other bounded levels. Student performance below this level is lower than that which PISA can reliably assess and, more importantly, describe.

REPORTING THE RESULTS FOR PISA MATHEMATICS

In this section, the way in which levels of mathematical literacy are defined, described and reported will be discussed. They will be exemplified using a number of items from the PISA 2012 assessment. The mathematics scale and content subscales were developed from the corresponding scale and subscales established in PISA 2003 (OECD, 2004), whereas the process subscales were created as a completely new measure.

Building an item map for mathematics

The data from the PISA mathematics assessment were processed to generate a set of item difficulty measures initially for the 290 paper-based and computer-based items used in the Field Trial that took place in 2011, and reviewed using the 150 items included in the Main Survey. In fact, when the difficulty measures that were estimated for each of the partial credit steps of the polytomous items are also taken into account, 168 item difficulty estimates were generated from the Main Survey items.

The 6-step analysis of items described earlier was carried out as the mathematics items were developed. This analysis included judgements about the elements of the PISA mathematics framework that were relevant to each item. For example, each item was analysed to determine which of the newly defined process categories was most significantly involved in a successful response.

Following data analysis and the resultant generation of difficulty estimates for each of the 168 item steps (and the additional item steps from the Field Trial items), the items and item steps were associated with their difficulty estimates, with their framework classifications, and with their brief qualitative descriptions. Figure 15.3 shows a map of some of this information from a sample of items from the PISA 2012 test, the items that made up two complete clusters in the test that were released publicly following the release of PISA 2012 results. Each row in Figure 15.3 represents an individual item or item step. The selected items and item steps have been ordered according to their difficulty, with the most difficult of these steps at the top, and the least difficult at the bottom. The difficulty estimate for each item and step is given in PISA scale units, along with the associated classifications and descriptions.

When a map such as this is prepared using all available items, it becomes possible to look for factors that are associated with item difficulty. This can be done by referring to the ways in which mathematical literacy is associated with questions located at different points ranging from the bottom to the top of the scale. For example, the item map in Figure 15.3 shows that the easiest items tend to involve identifying mathematical information presented in a table or graph and linking that information to some element of the problem context. The most difficult items, by contrast, are based on knowledge of particular mathematical content or procedures, and they involve several steps that require some creativity or strategic control in linking the context to the mathematical representation of aspects of the context, and often substantial mathematical processing or calculation to devise a solution.

■ Figure 15.3 ■

A map for selected mathematics items

Item Code	Item Name	Item difficultiy on PISA scale	Description of item demand	Process: Formulate	Process: Employ	Process: Interpret	Content: Change and relationships	Content: Quantity	Content: Space and shape	Content: Uncertainty and data
PM995Q02	Revolving Door Q2	840.3	Apply knowledge of circle geometry and reasoning to interpret a given geometric model and to formulate it mathematically enabling a solution	●					●	
PM923Q04	Sailing Ships Q4	702.1	Devise and implement a multi-step strategy involving significant modelling and extended calculation to formulate then solve a complex real world problem involving fuel costs and volume, equipment costs	●			●			
PM957Q03	Helen the Cyclist (E) Q3	696.6	Interpret information about distance and speed, devise a representation to help formulate a model for average speed, calculate average speed including converting units		●		●			
PM991Q02.2	Garage Q2.2	687.3	Interpret task demand from text and diagrams, formulate area calculation process from given measurements and specification (correct working and justification)		●				●	
PM991Q02.1	Garage Q2.1	663.2	Interpret task demand from text and diagrams, formulate area calculation process from given measurements and specification (partially correct result)		●				●	
PM903Q01.2	Drip Rate Q1.2	657.7	Interpret text and equation linking four variables, provide explanation of effect of specified change to one variable on a second variable if all other variables remain unchanged		●		●			
PM942Q02	Climbing Mount Fuji Q2	641.6	Follow multi-step strategy to interpret information, formulate and use a model that connects given time, speeds, and distance, and implement a time calculation	●			●			
PM903Q03	Drip Rate Q3	631.7	Interpret formula linking three variables in medical context, check consistency of units, substitute two values into given equation, transpose equation and solve		●		●			
PM903Q01.1	Drip Rate Q1.1	610.5	Interpret text and equation linking four variables, provide partial explanation of effect of specified change to one variable on a second variable if all other variables remain unchanged		●		●			
PM942Q03.2	Climbing Mount Fuji Q3.2	610.0	Identify and mathematise the defined task goal; use the model to calculate an average from given data in context, in specified units		●			●		
PM934Q01	London Eye Q1	592.3	Interpret text and diagram to form a strategy: identify, extract and use data from geometric sketch to formulate a model, apply it to calculate a length		●				●	
PM942Q03.1	Climbing Mount Fuji Q3.1	591.3	Identify and mathematise the defined task goal; use the model to calculate an average from given data in context, in specified units (answer correct but expressed in wrong units)		●			●		
PM00FQ01	Apartment Purchase Q1	576.2	Interpret graphic representation, use geometric reasoning to identify relevant dimensions needed to carry out specified area calculation with several components	●					●	
PM995Q03	Revolving Door Q3	561.3	Use reasoning to formulate and apply a proportional model involving several steps	●				●		
PM985Q03	Which Car? Q3	552.6	Interpret information on tax rate for a purchase to formulate a simple model, locate and extract data from table, and calculate a percentage		●			●		
PM923Q03	Sailing Ships Q3	538.5	Use geometry knowledge (trigonometry, or Pythagoras) to form a simple model to solve a right-angled triangle in context, evaluate and select answer from given options		●				●	
PM923Q01	Sailing Ships Q1	511.7	Interpret text and quantitative information; use reasoning and calculation to implement a percentage increase, and select from given options		●			●		
PM957Q02	Helen the Cyclist Q2	510.6	Interpret information about distance and speed, devise a simple proportional model to calculate a time corresponding to given distance and average speed		●		●			
PM985Q02	Which Car? Q2	490.9	Identify smallest of four decimal numbers from data table, use place value in context		●			●		
PM924Q02	Sauce Q2	489.1	Follow a multi-step strategy to devise and apply a suitable proportional model and perform the resultant percent calculation	●				●		
PM934Q02	London Eye Q2	481.0	Interpret text to understand task, extract and use data from graphic to formulate simple model, involving reasoning about fractions of a circle	●					●	
PM942Q01	Climbing Mount Fuji Q1	464.0	Interpret text to understand task; formulate strategy - define a time period in required unit (days), and combine information to devise a method to calculate a daily average; perform the calculation	●				●		
PM957Q01	Helen the Cyclist Q1	440.5	Interpret information about the distance travelled in two time periods to verify a given conclusion about the corresponding average speeds		●		●			
PM918Q05	Charts Q5	428.2	Identify and extract relevant data from a bar graph, model trend and use it to interpolate		●					●
PM991Q01	Garage Q1	419.6	Use spatial reasoning: devise a comparison strategy to identify correct representational model from given options			●			●	
PM918Q02	Charts Q2	415.0	Interpret bar graph; identify and extract data value defined by comparative condition to answer a question about the context			●				●
PM918Q01	Charts Q1	347.7	Interpret bar graph, identify and extract data value to answer a question about the context			●				●
PM985Q01	Which Car? Q1	327.8	Identify data in a table meeting specifications of simple mathematical relationships			●				●

More generally, the difficulty of mathematics questions in PISA 2012 is associated with a number of item characteristics that can be seen as calling forth varying levels of activation by students of each member of the set of fundamental mathematical capabilities described in the mathematics framework. That set of capabilities has been useful in exposing the ways in which cognitive demand varies among different items, and has provided a rich means of describing different levels of proficiency.

- Mathematical *communication* involves understanding the stated task objectives and the mathematical language used, recognising what information is relevant and what is the nature of the response needed; and also may involve the active steps including some or all of presenting the response, solution steps, description of the reasoning used and justification of the answer provided. Demand for this capability increases according to the complexity of material to be interpreted in understanding the task, the need to link multiple information sources or to move repeatedly among information elements; and with the need to provide a detailed written solution or explanation.
- Item complexity and difficulty is also affected by the nature and extent of *strategic thinking* that is required to progress towards a problem solution. In the simplest problems, the solution path is specified or it is obvious, and involves perhaps just a single processing step, while in other problems a solution strategy may involve drawing on several elements of mathematical knowledge, linking them in a particular sequence of related steps, and exercising quite a degree of control to keep sight of the objective and the way the stages of a solution will lead to meeting essential sub-goals that will fit together in achieving the overall problem objective.
- PISA problems very frequently are set in some context of the kind individuals may encounter in their school, work or daily life. Contextualised problems may require the student to impose a transformation of information into a suitable mathematical form. This process of *mathematisation* lies at the heart of the mathematical process referred to as *formulating*. In the most difficult problems it can involve making simplifying assumptions, identifying relevant variables and devising a suitable way to express them mathematically, and understanding the relationships between the contextual elements and their mathematical expression. It can also involve forging links between mathematical results or mathematical information and the situation that information is intended to describe. Translating or *interpreting* mathematical results in relation to specific elements of the problem context, and validating the adequacy of the solution with respect to the context are also part of this mathematical capability.
- A widely recognised element of much mathematical work is the myriad ways in which mathematical information, relationships and processes can be expressed. Mathematical *representations* can take the form of equations, graphs, charts, tables, formulae and so on. These vary in familiarity to students, and in their complexity, and this variation can directly affect the difficulty of tasks that involve the use or construction of mathematical representations. Students may be presented with mathematical representations they must use or process in some way. Or they may be required to create or devise a representation of data, information or relationships in order to solve a problem. Representations can be simple, or more complex. Multiple representations may be involved or required in order to solve a problem, and tasks that involve linking two or more different representations tend to be more difficult.
- One of the most important drivers of item difficulty lies in the particular mathematical content knowledge that must be activated to solve problems, such as the number and nature of definitions, facts, rules, algorithms and procedures, especially the need to understand and manipulate symbolic expressions, formulae, functional relations or other algebraic expressions, but also the need to perform arithmetic calculations and to understand the formal rules that govern them. A problem that requires counting or adding small integers clearly imposes a different level of cognitive demand compared to an item that requires manipulating and solving an equation, or applying the Pythagoras theorem.
- Finally, the nature of the *reasoning* involved in solving a mathematical problem, and the degree to which mathematical *argumentation* must be understood or applied as part of the solution process contribute in important ways to item difficulty. The nature, number or complexity of elements that need to be brought together in making inferences, and the length and complexity of the chain of inferences needed are significant contributors to increased demand for activation of the *reasoning and argument* competency.

Levels of mathematical literacy

The approach to reporting used by the OECD has been defined in previous cycles of PISA and is based on the definition and description of a number of levels of literacy proficiency. Descriptions were developed to characterise typical student performance at each level. The levels were used to summarise the performance of students, to compare performances across subgroups of students, and to compare average performances among groups of students, in particular among the students from different participating countries. A similar approach has been used here to analyse and report PISA 2012 outcomes for mathematics.

For mathematics in PISA 2003, when the fully articulated PISA mathematics scale was first developed, student scores were transformed to the PISA scale, with a mean of 500 and a standard deviation of 100, and six levels of proficiency were defined and described. For PISA 2012, the new items together with link items from previous PISA survey administrations that were administered again in PISA 2012 were calibrated independently as a set and then equated with the PISA 2003 scale.

The mathematics level definitions on the PISA scale are given in Figure 15.4. The same definitions apply to the overall mathematical proficiency scales, and to each of the process-based and content-based subscales.

■ Figure 15.4 ■

Mathematical literacy performance band definitions on the PISA scale

Level	Score points on the PISA scale
6	Above 669.3
5	From 607.0 to less than 669.3
4	From 544.7 to less than 607.0
3	From 482.4 to less than 544.7
2	From 420.1 to less than 482.4
1	From 357.8 to less than 420.1
Below level 1	Below 357.8

The information about the items in each band is used to develop summary descriptions of the kinds of mathematical knowledge and understanding associated with different levels of proficiency. These summary descriptions can then be used to encapsulate typical mathematical proficiency of students associated with each level. As a set, they describe development in mathematical literacy.

The PISA 2003 proficiency descriptions have been revised and enriched using information from the new items developed for PISA 2012 including those delivered via computer, and the revised descriptions are presented in Figure 15.5. They are further described and illustrated in the first volume of the *PISA 2012 Results* (OECD, 2014a).

■ Figure 15.5 ■

Summary descriptions of the six proficiency levels on the mathematical literacy scale

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

Figures 15.6, 15.7 and 15.8 provide the summary descriptions of skills and knowledge and understanding required to complete tasks located within the defined bands for the process subscales: *Formulating situations mathematically*; *Employing mathematical concepts, facts, procedures and reasoning;* and *Interpreting, applying and evaluating mathematical outcomes* respectively.

■ Figure 15.6 ■

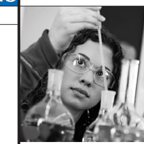
Summary descriptions of the six proficiency levels on the mathematical process subscale *Formulating situations mathematically*

Level	What students can typically do
6	Students at or above Level 6 can typically apply a wide variety of mathematical content knowledge to transform and represent contextual information or data, geometric patterns or objects into a mathematical form amenable to investigation. At this level, students can devise and follow a multi-step strategy involving significant modelling steps and extended calculation to formulate and solve complex real world problems in a range of settings, for example involving material and cost calculations in a variety of contexts, or to find the area of an irregular region on a map; can identify what information is relevant (and what is not) from contextual information about travel times, distances and speed to formulate appropriate relationships among them; can apply reasoning across several linked variables to devise an appropriate way to present data in order to facilitate pertinent comparisons; can devise algebraic formulations that represent a given contextual situation.
5	At this level, students show an ability to use their understanding in a range of mathematical areas to transform information or data from a problem context into mathematical form. They can typically transform information from different representations involving several variables, into a form suitable for mathematical treatment. They can typically formulate and modify algebraic expressions of relationships among variables; can use proportional reasoning effectively to devise computations; they are typically able to draw together information from different sources to formulate and solve problems involving geometric objects, features and properties, or analyse geometric patterns or relationships and express them in standard mathematical terms; they can transform a given model according to changed contextual circumstances; can formulate a sequential calculation process based on text descriptions; can activate statistical concepts such as randomness, or sample, and apply probability, to formulate a model.
4	At Level 4, students show an ability to link information and data from related representations (for example, a table and a map, or a spreadsheet and a graphing tool) and apply a sequence of reasoning steps in order to formulate a mathematical expression needed to carry out a calculation or otherwise to solve a contextual problem. At this level, students can typically formulate a linear equation from a text description of a process, for example in a sales context, and can formulate and apply cost comparisons to compare prices of sale items; can identify which of given graphical representations corresponds to a given description of a physical process; they can specify a sequential calculation process in mathematical terms; they can identify geometrical features of a situation and use their geometric knowledge and reasoning to analyse a problem, for example to estimate areas or to link a contextual geometric situation involving similarity to the corresponding proportional reasoning; they can typically combine multiple decision rules needed to understand or implement a calculation where different constraints apply; and they can formulate algebraic expressions when the contextual information is reasonably straight-forward, for example to connect distance and speed information in time calculations.
3	At this level, students show an ability to identify and extract information and data from text, tables, graphs, maps or other representations, and make use of them to express a relationship mathematically, including interpreting or adapting simple algebraic expressions related to an applied context. Students at this level can transform a textual description of a simple functional relationship into a mathematical form, for example with unit costs or payment rates; can form a strategy involving two or more steps to link problem elements or to explore mathematical characteristics of the elements; can apply reasoning with geometric concepts and skills to analyse patterns or to identify properties of shapes or a specified map location, or to identify information needed to carry out some pertinent calculations, including calculations involving the use of simple proportional models and reasoning, where the relevant data and information is immediately accessible; and can typically understand and link probabilistic statements to formulate probability calculations in contexts such as in a manufacturing process, or a medical test.
2	At this level, students can understand written instructions and information about simple processes and tasks in order to express them in a mathematical form. They can typically use data presented in text or in a table (for example giving information about cost of some product or service) to formulate a computation required, such as, to identify the length of a time period, or to present a cost comparison, or to calculate an average; can analyse a simple pattern, for example by formulating a counting rule or identifying and extending a numeric sequence; can work effectively with different two- and three-dimensional standard representations of objects or situations, for example devising a strategy to match one representation with another, or to compare different scenarios, or identify random experiment outcomes mathematically using standard conventions.
1	At this level students can recognise or modify and use an explicit simple model of a contextual situation. Students can choose between several such models to match the situation. For example, choose between and additive and a multiplicative model in a shopping context; choose among given two-dimensional objects to represent a familiar three-dimensional object; select one of several given graphs to represent growth of a population.

■ Figure 15.7 [Part 1/2] ■

Summary descriptions of the six proficiency levels on the mathematical process subscale *Employing mathematical concepts, facts, procedures and reasoning*

Level	What students can typically do
6	Students at or above Level 6 are typically able to employ a strong repertoire of knowledge and procedural skills in a wide range of mathematical areas. They can form and follow a multi-step strategy to solve a problem involving several stages; can apply reasoning in a connected way across several problem elements; can set up and solve an algebraic equation with more than one variable; can generate relevant data and information to explore problems, for example using a spreadsheet to sort and analyse data; are able to justify their results mathematically and to explain their conclusions and support them with well-formed mathematical arguments. At Level 6 students' work is consistently precise and accurate.
5	Students at Level 5 typically are able to employ a range of knowledge and skills to solve problems. They can sensibly link information in graphical and diagrammatic form to textual information. They can apply spatial and numeric reasoning skills to express and work with simple models in reasonably well-defined situations and where the constraints are clear. They usually work systematically, for example to explore combinatorial outcomes, and can typically sustain accuracy in their reasoning across a small number of steps and processes. They are generally able to work competently with expressions and can work with formulae and can use proportional reasoning; and are able to work with and transform data presented in a variety of forms.



■ Figure 15.7 **[Part 2/2]** ■

Summary descriptions of the six proficiency levels on the mathematical process subscale *Employing mathematical concepts, facts, procedures and reasoning*

Level	What students can typically do
4	At Level 4, students can typically identify relevant data and information from contextual material and use it to perform such tasks as calculating distances, and using proportional reasoning to apply a scale factor, convert different units to a common scale, or to relate different graph scales to each other. They are able to work flexibly with distance-time-speed relationships, and can carry out calculations in contexts that require a sequence of arithmetic calculations. They show some ability to use algebraic formulations, to follow a straightforward strategy and describe it.
3	Students at Level 3 frequently have sound spatial reasoning skills enabling them, for example, to use the symmetry properties of a figure, or to recognise patterns presented in graphical form, or to use angle facts, to solve a geometric problem. Students at this level can connect two different mathematical representations, such as data in a table and in a graph, or an algebraic expression with its graphical representation, enabling them for example to understand the effect of changing data in one representation on the other. They typically show some ability to handle percentages, fractions and decimal numbers and to work with proportional relationships.
2	Students at Level 2 are able to apply small reasoning steps to make direct use of given information to solve a problem, for example to implement a simple calculation model, or to identify a calculation error, or to analyse a distance-time relationship, or to analyse a simple spatial pattern; at this level students show an understanding of place value in decimal numbers and can use that understanding to compare numbers presented in a familiar context; can correctly substitute values into a simple formula; can recognise which of a set of given graphs correctly represents a set of percentages and can apply reasoning skills to understand and explore different kinds of graphical representations of data; and they typically show some insight into simple probability concepts.
1	Students at Level 1 can identify simple data relating to a real-world context, for example presented in a structured table or in an advertisement where the text and data labels match directly; can perform practical tasks such as decomposing money amounts into lower denominations; use direct reasoning from textual information that points to an obvious strategy to solve a given problem, particularly where the mathematical procedural knowledge required would be limited, for example, to arithmetic operations with whole numbers, or to ordering and comparing whole numbers; they demonstrate a partial understanding of graphing techniques and conventions; and can make use of symmetry properties to explore characteristics of a figure such as comparing side lengths and angles.

■ Figure 15.8 ■

Summary descriptions of the six proficiency levels on the mathematical process subscale *Interpreting, applying and evaluating mathematical outcomes*

Level	What students can typically do
6	At Level 6, students are able to link multiple complex mathematical representations in an analytic way to identify and extract data and information that enables contextual questions to be answered, and are able to effectively present their interpretations and conclusions in written form. For example they may interpret two time-series graphs in relation to different contextual conditions; or link a relationship expressed both in a graph and in numeric form (such as in a price calculator) or in a spreadsheet and graph, to present an argument or conclusion about contextual conditions. Students at this level are also typically able to apply mathematical reasoning to data or information presented in order to generate a chain of linked steps to support a conclusion (for example, analysis of a map using scale information; or analysis of a complex algebraic formula in relation to the variables represented; or translating data into a new time-frame; or performing a three-way currency conversion; or systematic use of a data generation tool to find the information needed to answer a question). Students at this level are able to bring together analysis and data and their interpretation across several different problem elements or across different questions about a context, showing a depth of insight and a capacity for sustained reasoning.
5	At Level 5, students are able to combine several processes in order to formulate conclusions based on interpretation of mathematical information with respect to context, such as formulating or modifying a model, solving an equation or carrying out computations, and using several reasoning steps to make the links to the identified context elements. At this level, students are able to make links between context and mathematics involving spatial or geometric concepts and complex statistical and algebraic concepts. They can easily interpret and evaluate a set of plausible mathematical representations, such as several graphs, to identify which one best reflects the contextual elements under analysis. Students at this level have begun to develop the ability to communicate conclusions and interpretations in written form.
4	At Level 4 students are typically able to apply appropriate reasoning steps, possibly multiple steps, to extract information from a complex mathematical situation, and to interpret complicated mathematical objects, including algebraic expressions. They can interpret complex graphical representations to identify data or information that answers a question; can perform a calculation or data manipulation (for example in a spreadsheet) to generate additional data needed to decide whether a constraint (such as a measurement condition, or a size comparison) is met; they can interpret simple statistical or probabilistic statements in such contexts as public transport, or health and medical test interpretation to link the meaning of the statements to the underlying contextual issues; they can conceptualise a change needed to a calculation procedure in response to a changed constraint; they can analyse two data samples, for example relating to a manufacturing process, to make comparisons and draw and express conclusions.
3	Students at Level 3 begin to show the ability to use reasoning, including spatial reasoning, to support their interpretations of mathematical information in order to make inferences about features of the context. They combine reasoning steps systematically to make various connections between mathematical and contextual material or when required to focus on different aspects of a context, for example where a graph shows two data series or a table contains data on two variables that must be actively related to each other to support a conclusion. They are able to test and explore alternative scenarios, using reasoning to interpret the possible effects of changing some of the variables under observation. They can use appropriate calculation steps to assist their analysis of data and to support the formation of conclusions and interpretations, including calculations involving proportions and proportional reasoning, and in situations where systematic analysis across several related cases is needed. At this level students can interpret and analyse relatively unfamiliar data presentations to support their conclusions.
2	At Level 2, students link contextual elements of the problem to the mathematics, for example by performing appropriate calculations or reading tables. Students at this level can typically make comparisons repeatedly across several similar cases; for example they can interpret a bar graph to identify and extract data to apply in a comparative condition where some insight is required. They can apply basic spatial skills to make connections between a situation presented visually and its mathematical elements; they can identify and carry out necessary calculations to support such comparisons as costs across several contexts; and they may be able to interpret a simple algebraic expression as it relates to a given context.
1	At Level 1, students are able to interpret data or information expressed in a direct way in order to answer questions about the context described. They can interpret given data to answer questions about simple quantitative relational ideas (such as 'larger', 'shorter time', 'in between') in a familiar context, for example by evaluating measurements of an object against given criterion values, or by comparing average journey times for two methods of transport, or comparing specified characteristics of a small number of similar objects. Similarly, they can make simple interpretations of data in a timetable or schedule to identify times or events. Students at this level may show rudimentary understanding of such concepts as randomness and data interpretation, for example by identifying the plausibility of a statement about chance outcomes of a lottery, or by understanding numeric and relational information in a well-labelled graph, and by understanding basic contextual implications of links between related graphs.

Figures 15.9, 15.10, 15.11 and 15.12 provide the summary descriptions of skills, knowledge and understanding required to complete tasks located within the defined bands for the mathematical content subscales: *Change and relationships, Space and Shape, Quantity* and *Uncertainty and data* respectively.

■ Figure 15.9 ■

Summary descriptions of the six proficiency levels on the mathematical content subscale *Change and relationships*

Level	What students can typically do
6	At Level 6, students use significant insight, abstract reasoning and argumentation skills and technical knowledge and conventions to solve problems involving relationships among variables and to generalise mathematical solutions to complex real-world problems. They are able to create and use an algebraic model of a functional relationship incorporating multiple quantities. They apply deep geometrical insight to work with complex patterns. And they are typically able to use complex proportional reasoning, and complex calculations with percentage to explore quantitative relationships and change.
5	At Level 5, students solve problems by using algebraic and other formal mathematical models, including in scientific contexts. They are typically able to use complex and multi-step problem-solving skills, and to reflect on and communicate reasoning and arguments, for example in evaluating and using a formula to predict the quantitative effect of change in one variable on another. They are able to use complex proportional reasoning, for example to work with rates, and they are generally able to work competently with formulae and with expressions including inequalities.
4	Students at Level 4 are typically able to understand and work with multiple representations, including algebraic models of real-world situations. They can reason about simple functional relationships between variables, going beyond individual data points to identifying simple underlying patterns. They typically employ some flexibility in interpretation and reasoning about functional relationships (for example in exploring distance-time-speed relationships) and are able to modify a functional model or graph to fit a specified change to the situation; and they are able to communicate the resulting explanations and arguments.
3	At Level 3, students can typically solve problems that involve working with information from two related representations (text, graph, table, formulae), requiring some interpretation, and using reasoning in familiar contexts. They show some ability to communicate their arguments. Students at this level can typically make a straightforward modification to a given functional model to fit a new situation; and they use a range of calculation procedures to solve problems, including ordering data, time difference calculations, substitution of values into a formula, or linear interpolation.
2	Students at Level 2 are typically able to locate relevant information on a relationship from data provided in a table or graph and make direct comparisons, for example to match given graphs to a specified change process. They can reason about the basic meaning of simple relationships expressed in text or numeric form by linking text with a single representation of a relationship (graph, table, simple formula), and can correctly substitute numbers into simple formulae, sometimes expressed in words. At this level, student can use interpretation and reasoning skills in a straightforward context involving linked quantities.
1	Students at Level 1 are typically able to evaluate single given statements about a relationship expressed clearly and directly in a formula, or in a graph. Their ability to reason about relationships, and to change in those relationships, is limited to simple expressions and to those located in familiar situations. They may apply simple calculations needed to solve problems related to clearly expressed relationships.

■ Figure 15.10 ■

Summary descriptions of the six proficiency levels on the mathematical content subscale *Space and shape*

Level	What students can typically do
6	At Level 6, students are able to solve complex problems involving multiple representations or calculations; identify, extract, and link relevant information, for example by extracting relevant dimensions from a diagram or map and using scale to calculate an area or distance; they use spatial reasoning, significant insight and reflection, for example by interpreting text and related contextual material to formulate a useful geometric model and applying it taking into account contextual constraints; they are able to recall and apply relevant procedural knowledge from their mathematical knowledge base such as in circle geometry, trigonometry, Pythagoras's rule, or area and volume formulae to solve problems; and they are typically able to generalise results and findings, communicate solutions and provide justifications and argumentation.
5	At Level 5, students are typically able to solve problems that require appropriate assumptions to be made, or that involve reasoning from assumptions provided and taking into account explicitly stated constraints, for example in exploring and analysing the layout of a room and the furniture it contains. They solve problems using theorems or procedural knowledge such as symmetry properties, or similar triangle properties or formulas including those for calculating area, perimeter or volume of familiar shapes; they use well-developed spatial reasoning, argument and insight to infer relevant conclusions and to interpret and link different representations, for example to identify a direction or location on a map from textual information.
4	Students at Level 4 typically solve problems by using basic mathematical knowledge such as angle and side-length relationships in triangles, and doing so in a way that involves multistep, visual and spatial reasoning, and argumentation in unfamiliar contexts; they are able to link and integrate different representations, for example to analyse the structure of a three dimensional object based on two different perspectives of it; and typically they can compare objects using geometric properties.
3	At Level 3, students are able to solve problems that involve elementary visual and spatial reasoning in familiar contexts, such as calculating a distance or a direction from a map or a GPS device; they are typically able to link different representations of familiar objects or to appreciate properties of objects under some simple specified transformation; and at this level students can devise simple strategies and apply basic properties of triangles and circles, and can use appropriate supporting calculation techniques such as scale conversions needed to analyse distances on a map.
2	At Level 2, students are typically able to solve problems involving a single familiar geometric representation (for example, a diagram or other graphic) by comprehending and drawing conclusions in relation to clearly presented basic geometric properties and associated constraints. They can also evaluate and compare spatial characteristics of familiar objects in a situation where given constraints apply (such as comparing the height or circumference of two cylinders having the same surface area; or deciding whether a given shape can be dissected to produce another specified shape).
1	Students at Level 1 can typically recognise and solve simple problems in a familiar context using pictures or drawings of familiar geometric objects and applying basic spatial skills such as recognising elementary symmetry properties, or comparing lengths or angle sizes, or using procedures such as dissection of shapes.

■ Figure 15.11 ■

Summary descriptions of the six proficiency levels on the mathematical content subscale *Quantity*

Level	What students can typically do
6	At Level 6 and above, students conceptualise and work with models of complex quantitative processes and relationships; devise strategies for solving problems; formulate conclusions, arguments and precise explanations; interpret and understand complex information, and link multiple complex information sources; interpret graphical information and apply reasoning to identify, model and apply a numeric pattern. They are able to analyse and evaluate interpretive statements based on data provided; work with formal and symbolic expressions; plan and implement sequential calculations in complex and unfamiliar contexts, including working with large numbers, for example to perform a sequence of currency conversions, entering values correctly and rounding results. Students at this level work accurately with decimal fractions; they use advanced reasoning concerning proportions, geometric representations of quantities, combinatorics and integer number relationships; and they interpret and understand formal expressions of relationships among numbers, including in a scientific context.
5	At Level 5, students are able to formulate comparison models and compare outcomes to determine best price; interpret complex information about real-world situations (including graphs, drawings and complex tables, for example two graphs using different scales); they are able to generate data for two variables and evaluate propositions about the relationship between them. Students are able to communicate reasoning and argument; recognise the significance of numbers to draw inferences; provide a written argument evaluating a proposition based on data provided. They can make an estimation using daily life knowledge; calculate relative and/or absolute change; calculate an average; calculate relative and/or absolute difference, including percentage difference, given raw difference data; and they can convert units (for example calculations involving areas in different units).
4	At Level 4, students are typically able to interpret complex instructions and situations; relate text-based numerical information to a graphic representation; identify and use quantitative information from multiple sources; deduce system rules from unfamiliar representations; formulate a simple numeric model; set up comparison models; and explain their results. They are typically able to carry out accurate and more complex or repeated calculations, such as adding 13 given times in hour/minute format; carry out time calculations using given data on distance and speed of a journey; perform simple division of large multiples in context; carry out calculations involving a sequence of steps and accurately apply a given numeric algorithm involving a number of steps. Students at this level can perform calculations involving proportional reasoning, divisibility or percentages in simple models of complex situations.
3	At Level 3, students typically use basic problem-solving processes, including devising a simple strategy to test scenarios, understand and work with given constraints, use trial and error, and use simple reasoning in familiar contexts. At this level students typically can interpret a text description of a sequential calculation process, and correctly implement the process; identify and extract data presented directly in textual explanations of unfamiliar data; interpret text and diagrams describing a simple pattern; perform calculations including working with large numbers, calculations with speed and time, conversion of units (for example from an annual rate to a daily rate). They understand place value involving mixed 2- and 3-decimal values and including working with prices; and are typically able to order a small series of (4) decimal values; calculate percentages of up to 3-digit numbers; and apply calculation rules given in natural language.
2	At Level 2, students can typically interpret simple tables to identify and extract relevant quantitative information; interpret a simple quantitative model (such as a proportional relationship) and apply it using basic arithmetic calculations. They are able to identify the links between relevant textual information and tabular data to solve word problems; interpret and apply simple models involving quantitative relationships; identify the simple calculation required to solve a straight-forward problem; carry out simple calculations involving the basic arithmetic operations, as well as ordering 2- and 3-digit whole numbers and decimal numbers with one or two decimal places, and calculate percentages.
1	At Level 1, students are typically able to solve basic problems in which relevant information is explicitly presented; the situation is straightforward and very limited in scope. Students at this level are able to handle situations where the required computational activity is obvious and the mathematical task is basic, such as a one-step simple arithmetic operation, or to total the columns of a simple table and compare the results; they can typically read and interpret a simple table of numbers; they can extract data and perform simple calculations; use a calculator to generate relevant data, extrapolate from the data generated, using reasoning and calculation with a simple linear model.

■ Figure 15.12 ■

Summary descriptions of the six proficiency levels on the mathematical content subscale *Uncertainty and data*

Level	What students can typically do
6	At Level 6, students are able to interpret, evaluate and critically reflect on a range of complex statistical or probabilistic data, information and situations to analyse problems. Students at this level bring insight and sustained reasoning across several problem elements; they understand the connections between data and the situations they represent and are able to make use of those connections to explore problem situations fully; they bring appropriate calculation techniques to bear to explore data or to solve probability problems; and they can produce and communicate conclusions, reasoning and explanations.
5	At Level 5, students are typically able to interpret and analyse a range of statistical or probabilistic data, information and situations to solve problems in complex contexts that require linking of different problem components. They can use proportional reasoning effectively to link sample data to the population they represent, can appropriately interpret data series over time and are systematic in their use and exploration of data. Students at this level can use statistical and probabilistic concepts and knowledge to reflect, draw inferences and produce and communicate results.
4	Students at Level 4 are typically able to activate and employ a range of data representations and statistical or probabilistic processes to interpret data, information and situations to solve problems. They can work effectively with constraints, such as statistical conditions that might apply in a sampling experiment, and they can interpret and actively translate between two related data representations (such as a graph and a data table). Students at this level can perform statistical and probabilistic reasoning to make contextual conclusions.
3	At Level 3, students are typically able to interpret and work with data and statistical information from a single representation that may include multiple data sources, such as a graph representing several variables, or from two simple related data representations such as a simple data table and graph. They are able to work with and interpret descriptive statistical, probabilistic concepts and conventions in contexts such as coin tossing or lotteries and make conclusions from data, such as calculating or using simple measures of centre and spread. Students at this level can perform basic statistical and probabilistic reasoning in simple contexts.
2	Students at Level 2 are typically able to identify, extract and comprehend statistical data presented in a simple and familiar form such as a simple table, a bar graph or pie chart; they can identify, understand and use basic descriptive statistical and probabilistic concepts in familiar contexts, such as tossing coins or rolling dice. At this level students can interpret data in simple representations, and apply suitable calculation procedures that connect given data to the problem context represented.
1	At Level 1, students can typically identify and read information presented in a small table or simple well-labelled graph to locate and extract specific data values while ignoring distracting information, and to recognise how these relate to the context. Students at this level can recognise and use basic concepts of randomness to identify misconceptions in familiar experimental contexts such as lottery outcomes.

Levels of proficiency in problem solving

The computer-based assessment of problem solving was the major innovative component of the PISA 2012 survey. Six proficiency levels were defined and described, and these are presented in Figure 15.13. The scale is further illustrated in Volume V of the *PISA 2012 Results* (OECD, 2014b). For a discussion of factors influencing item difficulty in these problem solving items, see Philpot et al. (forthcoming).

■ Figure 15.13 ■

Summary descriptions of the six proficiency levels on the problem solving scale

Level	Score range	What students can typically do
6	Equal to or higher than 683.1 points	At Level 6, students can develop complete, coherent mental models of diverse problem scenarios, enabling them to solve complex problems efficiently. They can explore a scenario in a highly strategic manner to understand all information pertaining to the problem. The information may be presented in different formats, requiring interpretation and integration of related parts. When confronted with very complex devices, such as home appliances that work in an unusual or unexpected manner, they quickly learn how to control the devices to achieve a goal in an optimal way. Level 6 problem-solvers can set up general hypotheses about a system and thoroughly test them. They can follow a premise through to a logical conclusion or recognise when there is not enough information available to reach one. In order to reach a solution, these highly proficient problem-solvers can create complex, flexible, multi-step plans that they continually monitor during execution. Where necessary, they modify their strategies, taking all constraints into account, both explicit and implicit.
5	618.2 to less than 683.1 points	At Level 5, students can systematically explore a complex problem scenario to gain an understanding of how relevant information is structured. When faced with unfamiliar, moderately complex devices, such as vending machines or home appliances, they respond quickly to feedback in order to control the device. In order to reach a solution, Level 5 problem-solvers think ahead to find the best strategy that addresses all the given constraints. They can immediately adjust their plans or backtrack when they detect unexpected difficulties or when they make mistakes that take them off course.
4	553.3 to less than 618.2 points	At Level 4, students can explore a moderately complex problem scenario in a focused way. They grasp the links among the components of the scenario that are required to solve the problem. They can control moderately complex digital devices, such as unfamiliar vending machines or home appliances, but they don't always do so efficiently. These students can plan a few steps ahead and monitor the progress of their plans. They are usually able to adjust these plans or reformulate a goal in light of feedback. They can systematically try out different possibilities and check whether multiple conditions have been satisfied. They can form an hypothesis about why a system is malfunctioning, and describe how to test it.
3	488.4 to less than 553.3 points	At Level 3, students can handle information presented in several different formats. They can explore a problem scenario and infer simple relationships among its components. They can control simple digital devices, but have trouble with more complex devices. Problem-solvers at Level 3 can fully deal with one condition, for example, by generating several solutions and checking to see whether these satisfy the condition. When there are multiple conditions or inter-related features, they can hold one variable constant to see the effect of change on the other variables. They can devise and execute tests to confirm or refute a given hypothesis. They understand the need to plan ahead and monitor progress, and are able to try a different option if necessary.
2	423.4 to less than 488.4 points	At Level 2, students can explore an unfamiliar problem scenario and understand a small part of it. They try, but only partially succeed, to understand and control digital devices with unfamiliar controls, such as home appliances and vending machines. Level 2 problem-solvers can test a simple hypothesis that is given to them and can solve a problem that has a single, specific constraint. They can plan and carry out one step at a time to achieve a sub-goal, and have some capacity to monitor overall progress towards a solution.
1	358.5 to less than 423.4 points	At Level 1, students can explore a problem scenario only in a limited way, but tend to do so only when they have encountered very similar situations before. Based on their observations of familiar scenarios, these students are able only to partially describe the behaviour of a simple, everyday device. In general, students at Level 1 can solve straightforward problems provided there is only a simple condition to be satisfied and there are only one or two steps to be performed to reach the goal. Level 1 students tend not to be able to plan ahead or set sub-goals.

Levels of financial literacy

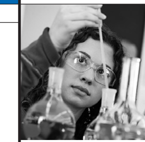
For the optional PISA 2012 assessment of financial literacy, five proficiency levels were defined and described. The factors identified to explain the variance in item difficulty included familiarity of experience with (financial) products, life stage relevance, understanding and use of financial terms, understanding and application of financial products, reading demands, conceptual understanding of numeracy, application of numeracy skills, and capacity to make effective (financial) decisions.

The proficiency descriptions are presented in Figure 15.14, and these are further explained and illustrated in Volume VI of the *PISA 2012 Results* (OECD, 2014c).

■ Figure 15.14 [Part 1/2] ■

Summary descriptions of the five proficiency levels on the financial literacy scale

Level	Score range	What students can typically do
5	Equal to or higher than 624.6 points	Students apply their understanding of a wide range of financial terms and concepts to contexts that may only become relevant to their lives in the long term. They analyse complex financial products. They take into account features of financial documents that are significant but unstated or not immediately evident, such as transaction costs. They work with a high level of accuracy and solve non-routine financial problems. They describe the potential outcomes of financial decisions, showing an understanding of the wider financial landscape, such as income tax.
4	549.9 to less than 624.6 points	Students apply their understanding of less common financial concepts and terms to contexts that will be relevant to them as they move towards adulthood, such as bank account management and compound interest in saving products. They interpret and evaluate a range of detailed financial documents such as bank statements, and explain the functions of less commonly used financial products. They make financial decisions taking into account longer-term consequences such as the impact of loan repayment on cost. They solve routine problems in less common financial contexts.



■ Figure 15.14 [Part 2/2] ■
Summary descriptions of the five proficiency levels on the financial literacy scale

Level	Score range	What students can typically do
3	475.1 to less than 549.9 points	Students apply their understanding of commonly used financial concepts, terms and products to situations that are relevant to them. They begin to consider the consequences of financial decisions and they make simple financial plans in familiar contexts. They make straightforward interpretations of a range of financial documents. They apply a range of basic numerical operations, including calculating percentages. They choose the numerical operations needed to solve routine problems in relatively common financial literacy contexts, such as budget calculations.
2	400.3 to less than 475.1 points	Students begin to apply their knowledge of common financial products and commonly used financial terms and concepts. They use given information to make financial decisions in contexts that are immediately relevant to them. They recognise the value of a simple budget. They interpret prominent features of everyday financial documents. They apply single basic numerical operations, including division, to answer financial questions. They show an understanding of the relationships between different financial elements, such as the amount of use and the costs incurred.
1	Less than 400.3 points	Students identify common financial products and terms, and interpret information relating to basic financial concepts. They recognise the difference between needs and wants and they make simple decisions on everyday spending. They recognise the purpose of everyday financial documents and apply single and basic numerical operations (addition, subtraction or multiplication) in financial contexts that they are likely to have experienced personally.

References

Bourque, M. L. (2009), *A History of NAEP Achievement Levels: Issues, Implementation, and Impact 1989-2009*, National Assessment Governing Board, Washington D.C.

Beaton, A. E. and **N. L. Allen** (1992), "Interpreting scales through scale anchoring", *Journal of Educational and Behavioural Statistics,* No. 17(2), pp. 191-204.

Masters, G., J. Lokan, B. Doig, S.T. Khoo, J. Lindsey, L. Robinson and **S. Zammit** (1990), *Profiles of Learning: The Basic Skills Testing Program in New South Wales 1989,* Australian Council for Educational Research, Hawthorn, Victoria.

Mossenson, L., P. Hill, and **G. Masters** (1987), *Tests of Reading Comprehension [TORCH],* Australian Council for Educational Research, Hawthorn, Victoria.

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

OECD (2014b), *PISA 2012 Results: Creative Problem Solving (Volume V): Students' Skills in Tackling Real-Life Problems*, PISA, OECD Publishing, Paris.
http://dx.doi.org/10.1787/9789264208070-en

OECD (2014c), *PISA 2012 Results: Students and Money: Financial Literacy Skills for the 21st Century (Volume VI),* PISA, OECD Publishing, Paris.
http://dx.doi.org/10.1787/9789264208094-en

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

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

OECD (2010), *PISA 2009 Results: What Students Know and Can Do – Student Performance in Reading, Mathematics and Science (Volume 1)*, PISA, OECD Publishing, Paris.
http://dx.doi.org/ 10.1787/9789264091450-en

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

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

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

Philpot, R., D. Ramalingam, B. McCrae, and **J. Dossey** (forthcoming), "Factors that influence the difficulty of problem solving items", in Csapó, B. and J. Funke (eds.), *The Nature of Problem Solving,* OECD Publishing, Paris.

Wright, B. and **M. Stone,** (1979), *Best Test Design*, MESA Press, Chicago.