Assessment Background: What PISA Measures and How

Luisa Araújo, Patrícia Costa, and Nuno Crato

Abstract This chapter provides a short description of what the Programme for International Student Assessment (PISA) measures and how it measures it. First, it details the concepts associated with the measurement of student performance and the concepts associated with capturing student and school characteristics and explains how they compare with some other International Large-Scale Assessments (ILSA). Second, it provides information on the assessment of reading, the main domain in PISA 2018. Third, it provides information on the technical aspects of the measurements in PISA. Lastly, it offers specific examples of PISA 2018 cognitive items, corresponding domains (mathematics, science, and reading), and related performance levels.

1 Introduction

PISA seeks to capture a common dimension of cognitive skills across countries. These skills are thought to be a good indication of the knowledge and skills that are essential for full participation in contemporary societies (OECD 2019a), and the attained level of these cognitive skills is viewed as an important determinant of economic growth (Heckman and Jacobs 2009). More specifically, PISA reinforces the idea that "…direct measures of cognitive skills offer a superior approach to understanding how

The author was partially supported by the Project CEMAPRE/REM - UIDB/05069/2020 - financed by FCT/MCTES through national funds.

L. Araújo (✉)
Instituto Superior de Educação e Ciências, ISEC, Lisbon, Portugal
e-mail: luisa.araujo@iseclisboa.pt

P. Costa
European Commission Joint Research Centre, Ispra, Italy
e-mail: patricia.costa@lese-conference.org

N. Crato
Cemapre/REM, ISEG, University of Lisbon, Lisbon, Portugal
e-mail: ncrato@iseg.ulisboa.pt

N. Crato (ed.), *Improving a Country's Education*,
https://doi.org/10.1007/978-3-030-59031-4_12

human capital affects the economic fortunes of nations", as expressed by Hanushek and Woessmann (2015, p.28). That is, as it is nowadays widely recognized, the quality of one's education is a better indicator of life outcomes than the quantity of education, as measured in years of schooling or similar indicators (Heckman and Jacobs 2009).

PISA results are complemented by other ILSA studies, and it is reassuring that high correlations across studies have been found. In particular, consider the Third International Mathematics and Science Study (TIMSS), a curriculum-sensitive ILSA conducted by the International Association for the Evaluation of Educational Achievement (IEA). PISA and TIMSS assess similar mathematics and science knowledge and skills at approximately the same time during schooling and a comparison between the two reveals that "… the correlation between the TIMSS 2003 tests of 8th graders and the PISA 2003 tests of 15-year-olds across the 19 countries participating in both is as high as 0.87 in mathematics and 0.97 in science. It is also 0.86 in both mathematics and science across the 21 countries participating both in the TIMSS 1999 tests and the PISA 2000–02 tests" (OECD 2010, p. 38).

A corresponding comparison of PISA with IEA's Program for International Reading Literacy Study (PIRLS) is not possible since this ILSA is designed to assess the reading skills of 4th graders, when most students are between 9 and 10 years of age. Still, a close look at both the PIRLS 2016 and the PISA 2018 assessment frameworks shows a very similar definition of reading. In PIRLS 2016 "Reading literacy is the ability to understand and use those written language forms required by society and/or valued by the individual. Readers can construct meaning from texts in a variety of forms. They read to learn, to participate in communities of readers in school and everyday life, and for enjoyment (Mullis et al. 2015, p.12). In PISA 2018, "reading literacy is understanding, using, evaluating, reflecting on and engaging with texts in order to achieve one's goals, to develop one's knowledge and potential and to participate in society" (OECD 2019c, p.28).

PISA, as the other ILSA such as PIRLS and TIMSS, also collects contextual information on students' socio-demographic and dispositional characteristics, students' home environment and teaching and schools' learning contexts (Lenkeit et al. 2015). This is done through the application of several questionnaires.

PISA results attract public attention mainly because of the country rankings they present in a comparative perspective and of the results' policy implications suggested by the OECD (Araújo et al. 2017). Educational implications can be drawn from statistical associations between cognitive performance and the information collected in the various questionnaires. In PISA 2018, such associations between cognitive performance and learning variables are discussed at length through several OECD volumes; main findings appear in the *Combined Executive Summaries* (OECD 2019b). For example, two findings with clear educational implications are: (1) students who perceived greater support from teachers scored higher in reading and (2) students whose parents discuss their progress on the initiative of the teacher had higher achievement in reading.

2 How Cognitive Skills Are Measured

All the ILSA here discussed use multistage sampling, unequal sampling probabilities, and stratification, but there are some differences.

PISA adopts a two-stage stratified sample design in which the primary sampling unit consists of at least 150 schools having 15-year-old students. Schools are sampled systematically from the school sampling frame, with probabilities proportional to a measure of the school size, which is a function of the estimated number of PISA-eligible 15-year-old students enrolled in the school. The second sampling unit includes *students* (around 5000 students) within the sampled schools.

TIMSS and PIRLS also employ a two-stage random sample design. In the first stage a sample of schools is drawn, but in the second stage one or more *complete classes of students* are selected from each of the sampled schools.

In PISA, TIMSS, and PIRLS, students' test scores are computed according to Item Response Theory (IRT) and standardised with a mean of around 500 and standard deviation of around 100. Even though the methodology is quite similar, the scores in these three ILSA are not directly comparable.

From the students' score points, proficiency levels are identified based on the PISA main domain scales. In this sense, PISA results can also be reported in terms of percentages of the student population at each of the predefined level. To define the proficiency levels and their cut off scores, IRT techniques are used to estimate simultaneously the difficulty and the ability of all students participating in PISA. Higher proficiency levels characterize the knowledge, skills, and capabilities needed to perform tasks of increasing complexity.

In PISA, TIMSS, and PIRLS, each student completes one booklet containing a subset of all the material. The booklets are created by combining different blocks of items in order to match to the framework characteristics. For the cognitive assessment of PISA 2018, the total testing time was 2 h and for TIMMS 2015 (8th grade), 1.5 h. PISA reading questions include a variety of items, including the conventional multiple-choice format and a complex multiple-choice format. TIMSS cognitive assessments primarily use multiple choice and constructed response items.

In all these surveys, national estimates are generated from the sample with different weights. To increase accuracy, these ILSA use plausible values (multiple imputations) drawn from a posteriori distribution which is constructed by combining the IRT scaling of the test items with a latent regression model with information from the student context questionnaire within a population model. For each student, 10 plausible values are computed in PISA (since 2015) and 5 plausible values are computed in all cycles of TIMSS and PIRLS.

All these ILSA studies allow for cross-country comparisons and for trend monitoring over time. In order to guarantee the comparability across countries, along years and delivery modes (paper and computer), linking procedures are used by considering a large number of common items in which the parameters are fixed to the same values. These items serve as *anchors* of the reporting scales and support the validity of cross-country and trend comparisons (OECD 2019c).

3 The Measurement of Student Performance in PISA

In PISA 2018, reading was the major domain of assessment, as it was in 2000 and 2009. The texts and items were selected based on a conceptual framework (OECD 2019a), which included five subscales. Three of the PISA 2018 assessment subscales have already been used in 2000 and 2009: "locating information", "understanding" and "evaluating and reflecting", (OECD, 2009). Two assessment subscales were newly created to describe students' literacy with single-source and with multiple source texts. Additionally, PISA 2018 included for the first time a measure of reading fluency in order to assess the reading skills of students in the lower proficiency levels. Reading fluency is defined as "the ease and efficiency with which one can read and understand a piece of text" (OECD 2019c, p. 270).

This was an important addition. As recognized in the PISA assessment framework, research shows that many students have difficulties with reading comprehension because they have not developed effortless decoding or the automaticity in word recognition that enables readers to focus on comprehension processes (OECD 2019a). Numerous research studies on reading processes have confirmed this (Adams 1990, 2009; Perfetti et al. 2005). Although comprehension can be developed throughout schooling and reading comprehension skills can be improved (Catts 2009; Elbro and Buch-Iverson 2013), it is fundamental that students acquire the basic reading skills that will allow them to read fluently, which implies reading words and text fast and accurately (Perfetti et al. 2005).

In order to simplify the interpretation of results, PISA scale is categorized into six ordinal proficiency levels. Each proficiency level requires a certain set of competencies, knowledge, and understanding items to be successfully completed. The minimum level is 1, although students can still score below the lower threshold of level 1. The maximum level is 6, with no ceiling. Mean scores are included in level 3. Table 1 reproduces the score limits for reading for PISA 2018.

Table 1 PISA 2018 reading scores levels of proficiency

Level 6	Above 698.32 score points
Level 5	From 625.61 to less than 698.32 score points
Level 4	From 552.89 to less than 625.61 score points
Level 3	From 480.18 to less than 552.89 score points
Level 2	From 407.47 to less than 480.18 score points
Level 1a	From 334.75 to less than 407.47 score points
Level 1b	From 262.04 to less than 334.75 score points
Level 1c	From 189.33 to less than 262.04 score points
Below level 1c	Less than 189.33 score points

Students scoring below level 2 are considered low-performers
Students scoring above level 4 are considered high-performers
Source OECD, PISA 2018 Database, Table I.B1.4; Figure I.4.1

Students scoring below level 2 are considered *low-performers* and those scoring above level 4 are considered *high-performers*. In 2015, recognizing the worrisome number of low performers and the need to better discriminate those students, PISA has subdivided level 1 in 1a and 1b. In 2018, PISA introduced an additional lower level, 1c.

Reading comprehension in PISA is assessed by asking students to locate information in a text, to retrieve literal information, to generate inferences and to evaluate and reflect on the content and form of texts. Evaluating a text is a more complex skill than simply identifying the requested information, and the six difficulty levels that PISA establishes are related to the tasks students need to perform. Locating explicit information in a text is a very basic reading task typical of level 1, whereas reflecting on the content of a text is a complex skill that characterizes questions at level 6. The difficulty level of the test items correspond to what the OECD refers to as *aspect* and reflect the cognitive processes involved in the task: "the *access and retrieve* aspect assessing the lowest benchmark proficiency levels (1 & 2), followed by the *Integrate and interpret* level (3 & 4) and with the *Reflect and evaluate* levels at the highest text processing level (5 & 6)" (OECD 2019a).

Level 2 marks the point at which students have acquired the basic skills to read and can use reading for learning. "At a minimum, these students [scoring at least level 2] are able to identify the main idea in a text of moderate length, find information based on explicit criteria, and reflect on the purpose and form of texts when explicitly directed to do so." Low performers are not able to attain this basic level.

Students who attained the highest proficiency levels 5 or 6 in reading, "are able to comprehend lengthy texts, deal with concepts that are abstract or counterintuitive, and establish distinctions between fact and opinion, based on implicit cues pertaining to the content or source of the information". (OECD 2019c).

The test items used to assess these text processing abilities are a mixture of multiple-choice questions and questions requiring students to construct their own responses. Such question and formats appear for a wide range of texts types; narrative, expository, descriptive and argumentative texts. Text types are presented as both continuous texts, organized in paragraphs and non-continuous, matrix-like formats, or with the appearance of a list. Since the purpose of assessing reading performance in PISA is to obtain a measure of reading comprehension, even the questions that require the students to construct a written response do not ask for extensive responses (OECD 2019a).

4 Questionnaire Data

PISA includes compulsory questionnaires and optional questionnaires. Compulsory questionnaires are the student background questionnaire (distributed to all participating students) and the school questionnaire (distributed to the principals of all participating schools). The student questionnaire, which takes about 35 minutes to complete, includes socio-demographic information about the students, such as age, gender, type of educational program the student is completing, immigrant background and parental occupation, a proxy for socio-economic status https://www.oecd.org/pisa/pisaproducts/PISA-2018-INTEGRATED-DESIGN.pdf. The school questionnaire that principals complete covers school learning experiences, school management, assessment, and school climate. For example, student truancy and bullying, cooperation among teachers and among students, and teacher enthusiasm and encouragement of reading are measures of school climate, a construct that includes social and academic dimensions believed to predict academic achievement and social skills (Costa and Araújo 2018; Chirkina and Khavenson 2018).

In 2018, the optional PISA questionnaires included three questionnaires for students (the educational career questionnaire, the ICT familiarity questionnaire, and the well-being questionnaire); one questionnaire for parents; one questionnaire for teachers (both for reading teachers and for all other subjects teachers); and one financial literacy questionnaire for students in countries that participated in the financial literacy assessment.

PIRLS and TIMSS usually include the following questionnaires: student, home (for 4th grade students and distributed to the parents of the students participating in the survey), teachers, schools, and curricular background data.

Teacher questionnaires in PISA are answered by the teachers of the sampled schools, while the PIRLS and TIMSS questionnaires are answered by the teachers of the assessed classes.

5 Examples of Cognitive Items in PISA 2018 and Other ILSA—What Questions Look Like

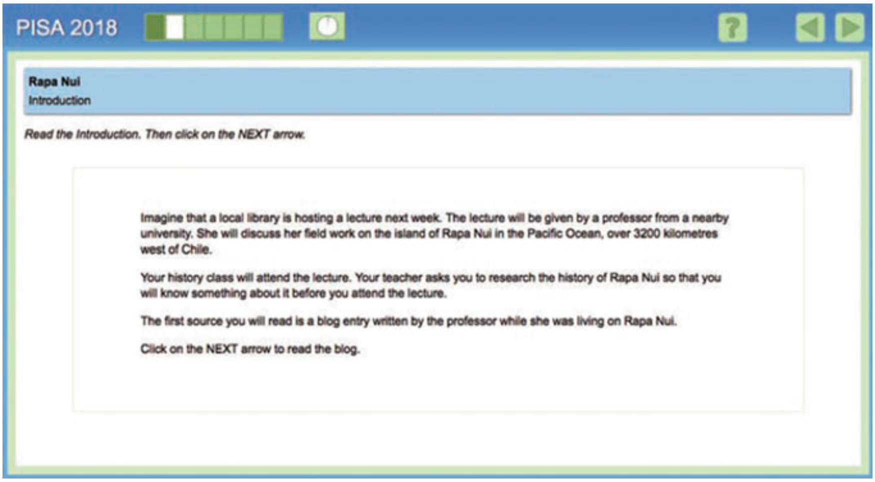
In the next pages we show examples of PISA reading items, followed by examples of some science and mathematics items, both from PISA and from TIMSS. Firstly, we will focus on the Rapa Nui Unit,[1] which is a scenario-based example. In this kind of unit, the student is given both a context and a purpose that helps to shape the way he/she searches for, comprehends, and integrates information. Rapa Nui refers to an island; the student is preparing to attend a lecture about a professor's field work, which was conducted on this island. This unit begins with a fictional scenario and is a multiple-source unit. It consists of three texts: a webpage from the professor's blog, a book review, and a news article from an online science magazine. The blog post is multiple-source text given that the comments section represents different authors. Both the book review and the news article are classified as single text, static, continuous, and argumentative. The Rapa Nui scenario prompts the student to integrate information in questions that are related to one text and then to demonstrate the ability to handle information from multiple texts. This design allows students

[1]Example of a PISA 2018 reading scenario. "Released items from the PISA 2018 computer-based reading assessment", in OECD (2019c).

with varying levels of ability to demonstrate proficiency on at least some questions of the unit. Specifically, this unit is intended to be of moderate to high difficulty.

5.1 Example 1: Rapa Nui—Scenario

1. Introduction



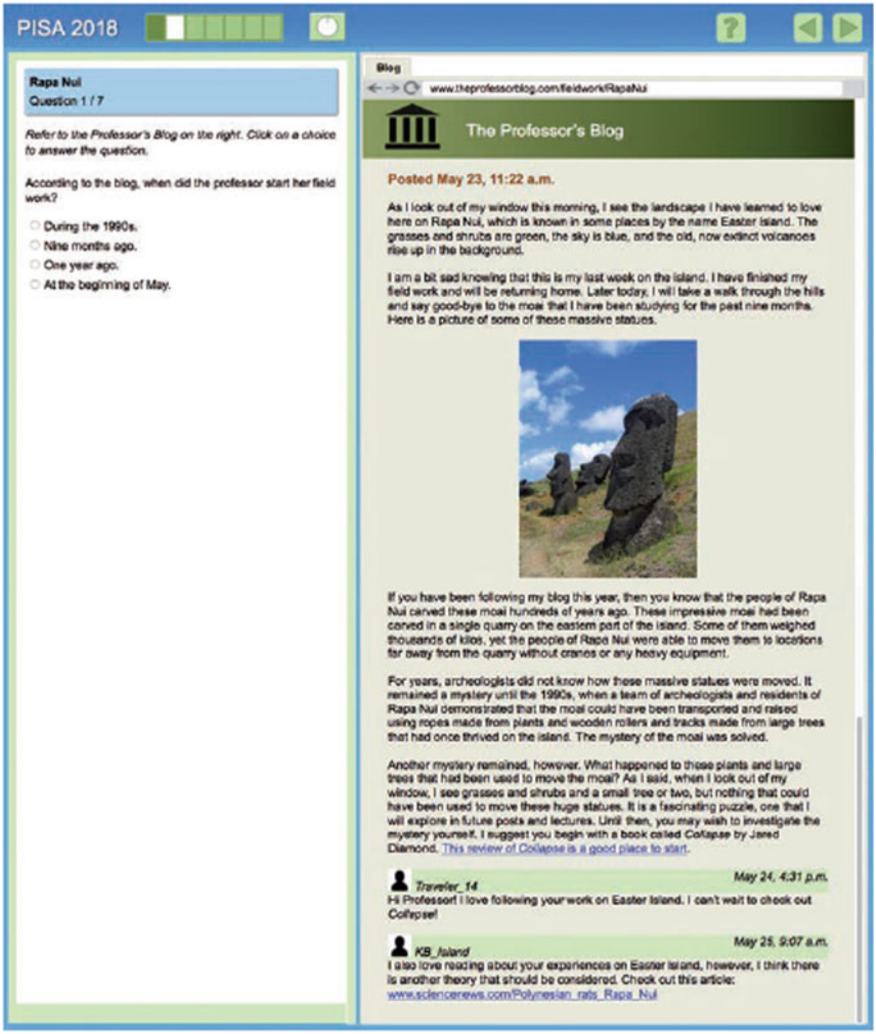
Item #1 is a single source item and the student must find the correct information within the blog post. The cognitive process required to engage in this task is that of assessing and retrieving information within a piece of text and its difficulty level is 4.

Item #2 is an open response (human coded) item[2] where the student must understand the second mystery mentioned in the Blog Post. It involves the cognitive process of representing literal meaning and its difficulty level is 3.

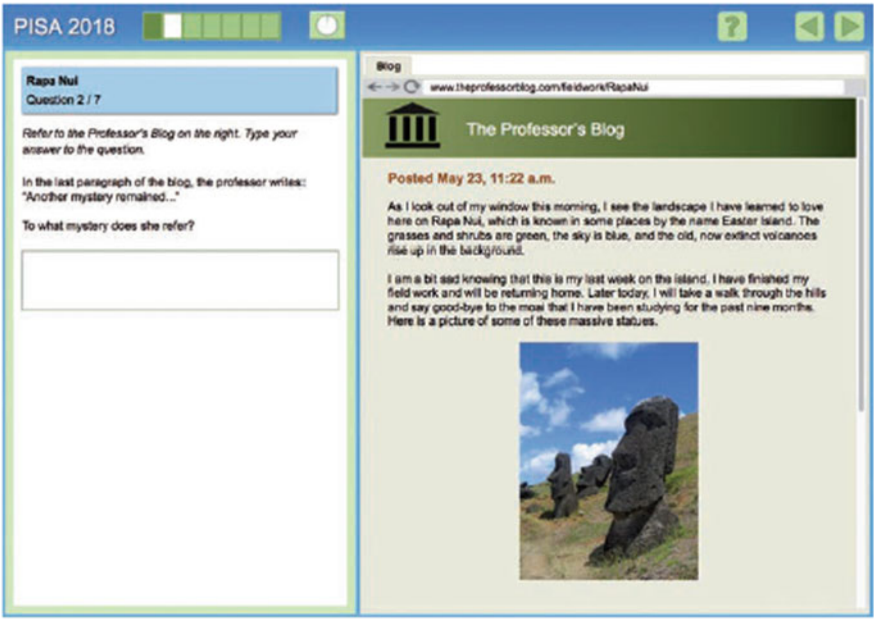
Item #6 asks students to integrate information across the texts with respect to the differing theories put forward by several scientists. This item involves integrating and generating inferences across multiple sources and is a complex multiple-choice item with a complexity level of 5.

[2]More information and the coding guide used can be found at "Released items from the PISA 2018 computer-based reading assessment", in OECD (2019c).

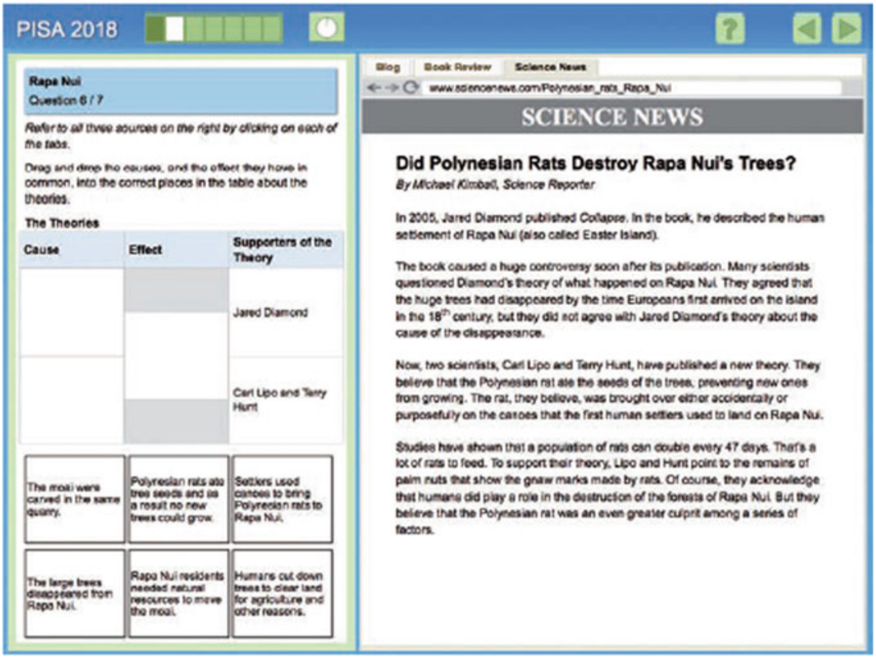
2. Released Item #1. The Professor's Blog - (Item number CR551Q01)



3. Released Item #2. The Professor's Blog (Item number CR551Q05)



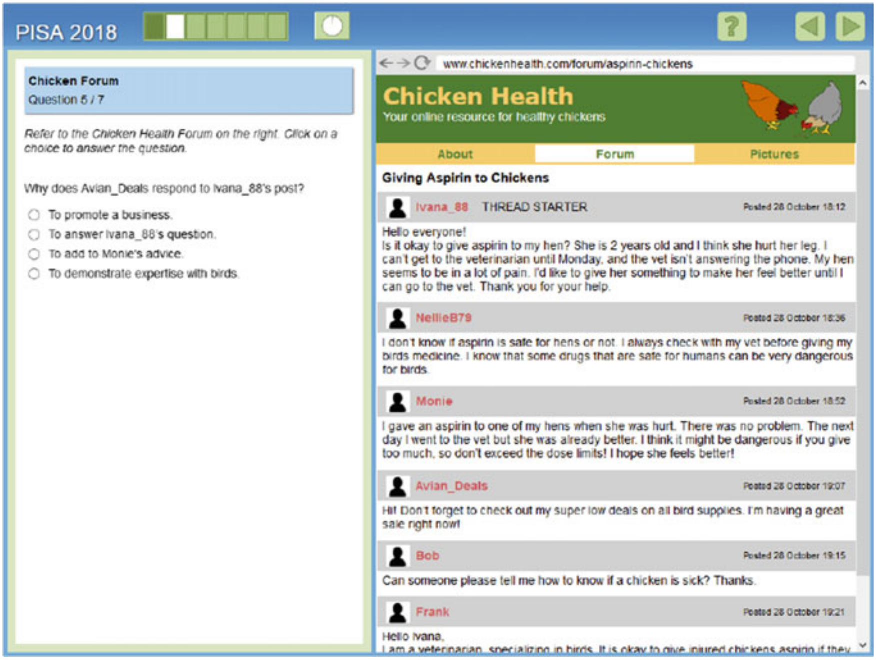
4. Released item # 6. Science News (Item number CR551Q10)



Next, we present an example of a reading proficiency level 1 task in PISA 2018. The item is part of the Chicken Forum Scenario[3] and describes a person who is seeking information about how to help an injured chicken. In this particular item it is expected that the student makes an inference from the information provided in a post. The item is classified as a single multiple choice one and it involves integrating and generating inferences as a cognitive process.

5.2 Example 2: Chicken Forum (Item Number CR548Q05)[4]

1. Released Item #5



Example 3 presents Science items from PISA and from TIMSS (8th grade). The PISA item is a multiple choice item classified as level 4 and it is an item "that requires students to be able to relate the rotation of the earth on its axis to the phenomenon of day and night and to distinguish this from the phenomenon of the seasons, which

[3]The units Chicken Forum was administered in the PISA 2018 Field Trial but was not selected for the Main Survey.

[4]More information can be found https://www.oecd.org/pisa/test/ and in the document (OECD, 2019f).

arises from the tilt of the axis of the earth as it revolves around the sun. All four alternatives given are scientifically correct" (OECD 2004, p. 289).

5.3 *Example 3: Science Items—PISA and TIMSS*

1. PISA 2003 item: DAYLIGHT[5]

Read the following information and answer the questions that follow.

DAYLIGHT ON 22 JUNE 2002

Today, as the Northern Hemisphere celebrates its longest day, Australians will experience their shortest.

In Melbourne*, Australia, the Sun will rise at 7:36 am and set at 5:08 pm, giving nine hours and 32 minutes of daylight.

Compare today to the year's longest day in the Southern Hemisphere, expected on 22 December, when the Sun will rise at 5:55 am and set at 8:42 pm, giving 14 hours and 47 minutes of daylight.

The President of the Astronomical Society, Mr Perry Vlahos, said the existence of changing seasons in the Northern and Southern Hemispheres was linked to the Earth's 23-degree tilt.

*Melbourne is a city in Australia at a latitude of about 38 degrees South of the equator.

Question 1: DAYLIGHT *S129Q01*

Which statement explains why daylight and darkness occur on Earth?

A The Earth rotates on its axis.
B The Sun rotates on its axis.
C The Earth's axis is tilted.
D The Earth revolves around the Sun.

2. TIMSS 2011 item: Recognizes the major cause of tides[6]

[5] We cannot help noticing the scientifically incorrect statement of the third paragraph: There is no such thing as the longest day in the Southern Hemisphere with the sun rising and setting at specific times; the length of the day and the specific times depend on the latitude.

[6] SOURCE: TIMSS 2011 Assessment. Copyright © 2013 International Association for the Evaluation of Educational Achievement (IEA). Publisher: TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College, Chestnut Hill, MA and International Association for the Evaluation of Educational Achievement (IEA), IEA Secretariat, Amsterdam, the Netherlands.

Which of the following is the major cause of tides?

A. heating of the oceans by the Sun
B. gravitational pull of the Moon
C. earthquakes on the ocean floor
D. changes in wind direction

Example 4 shows Mathematics items from PISA and from TIMSS (8th grade). Both items are open-ended items.

5.4 Example 4: Mathematics—PISA and TIMSS

1. PISA 2012 item: DRIP RATE[7]

Infusions (or intravenous drips) are used to deliver fluids and drugs to patients.

Nurses need to calculate the drip rate, *D*, in drops per minute for infusions.

They use the formula $D = \frac{dv}{60n}$ where

d is the drop factor measured in drops per millilitre (mL)

v is the volume in mL of the infusion

n is the number of hours the infusion is required to run.

Question 1: DRIP RATE PM903Q01 – 0 1 2 9

A nurse wants to double the time an infusion runs for.

Describe precisely how *D* changes if *n* is **doubled** but *d* and *v* do not change.

....

....

....

[7] More information can be found at https://www.oecd.org/pisa/test/ - PISA 2012, Mathematics items.

2. TIMSS 2011 item: Ann and Jenny divide 560 zeds[8]

> Ann and Jenny divide 560 zeds between them. If Jenny gets $\frac{3}{8}$ of the money, how many zeds will Ann get?
>
> Answer: ______________

6 Conclusion

This chapter offers a short description of what PISA measures and how it measures it. As such, it provides basic information about PISA's assessment framework and technical specifications related to sampling and statistical procedures and analyses. For more detailed information, readers can access OECD documents, namely the PISA assessment framework reports and the technical reports published by OECD for every assessment cycle. The PISA questionnaires can be accessed through the OECD/PISA database webpage (https://www.oecd.org/pisa/data/2018database/). More examples of released items can be found in https://www.oecd.org/pisa/test/PISA2018_Released_REA_Items_12112019.pdf. In order to have a good insight about PISA student results it is important to get acquainted with a few testing items. We hope this concluding assessment background chapter provides information to better understand PISA analyses.

References

Adams, M. J. (1990). *Beginning to read: Thinking and learning about print*. Cambridge, MA: MIT Press.

Adams, M. J. (2009). The challenge of advanced texts: The interdependence of reading and learning. In E. Hiebert (Ed.), *Reading more, reading better* (pp. 163–189). New York, NY: Guilford Press.

Araújo, L., Saltelli, A., & Schnepf, S.V. (2017). Do PISA data justify PISA-based education policy? *International Journal of Comparative Education and Development, 19*(1), 20–34. https://doi.org/10.1108/IJCED-12-2016-0023.

Catts, H. W. (2009). The narrow view of reading promotes a broad view of comprehension. *Language, Speech, and Hearing Services in the Schools, 40,* 178–183.

[8]SOURCE: TIMSS 2011 Assessment. Copyright © 2013 International Association for the Evaluation of Educational Achievement (IEA). Publisher: TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College, Chestnut Hill, MA and International Association for the Evaluation of Educational Achievement (IEA), IEA Secretariat, Amsterdam, the Netherlands.

Chirkina, T. A., & Khavenso, T. E. (2018). School climate: A history of the concept and approaches to defining and measuring it on PISA questionnaires. *Russian Education & Society, 60*(2), 133–160. https://doi.org/10.1080/10609393.2018.1451189.
Costa, P., & Araújo, L. (2018). Skilled students and effective schools: Reading achievement in Denmark. *Sweden, and France, Scandinavian Journal of Educational Research, 62,* 850–864. https://doi.org/10.1080/00313831.2017.1307274.
Elbro, C., & Buch-Iverson, I. (2013). Activation of prior knowledge for inferences making: Effects on reading comprehension. *Scientific Studies of Reading 17*, 435–452.
Hanushek, E., & Woessmann, L. (2015). *The knowledge capital of nations: Education and the economics of growth.* Massachusetts: Massachusetts Institute of Technology. MIT Press.
Heckman, J., & Jacobs, B. (2009). *Policies to create and destroy human capital in Europe.* IZA DP No. 4680.
Lenkeit, J., Chan, J., Hopfenbeck, T. N., & Baird, J. (2015). A review of the representation of PIRLS related research in scientific journals. *Educational Research Review, 16*, 102–115. https://doi.org/10.1016/j.edurev.2015.10.002 [Crossref], [Web of Science ®], [Google Scholar].
Mullis, I. V. S., & Martin, M. O. (Eds.). (2015). *PIRLS 2016 assessment framework* (2nd ed.). Retrieved from Boston College, TIMSS & PIRLS International Study Center website: https://timssandpirls.bc.edu/pirls2016/framework.html.
OECD. (2004). *Learning for tomorrow's world. First results from PISA 2003*. Paris: OECD.
OECD. (2009). *PISA 2009 assessment framework. Key competencies in reading, mathematics and science.* https://www.oecd.org/dataoecd/11/40/44455820.pdf.
OECD. (2010). *The high cost of low educational performance: The long-run economic impact of improving PISA outcomes, PISA*. OECD Publishing. https://doi.org/10.1787/9789264077485-en.
OECD. (2019a). *PISA 2018 assessment and analytical framework*. Paris: PISA, OECD Publishing. https://doi.org/10.1787/b25efab8-en.
OECD. (2019b). *PISA 2018 results—Combined executive summaries, Volume I, II & III,* https://www.oecd.org/pisa/Combined_Executive_Summaries_PISA_2018.pdf. Accessed February 10, 2020.
OECD. (2019c). *PISA 2018 Results (Volume I): What students know and can do, PISA*. Paris: OECD Publishing. 10.1787/5f07c754-en.
OECD. (2019d). Released items from the PISA 2018 computer-based reading assessment. In *PISA 2018 results (Volume I): what students know and can do*. Paris: OECD Publishing. https://doi.org/10.1787/098bab1a-en.
OECD. (2019e). *PISA 2018 technical report*. https://www.oecd.org/pisa/data/pisa2018technicalreport/. Accessed June 10, 2020.
OECD. (2019f). *PISA 2018 released field trial and main survey new reading items. Version: October 2019*. https://www.oecd.org/pisa/test/PISA2018_Released_REA_Items_12112019.pdf. Accessed June 10, 2020.
Perfetti, C. A., Landi, N., & Oakhill, J. (2005). The acquisition of reading comprehension skills. In M. J. Snowling & C. Hulme (Eds.), *The science of reading: A handbook* (pp. 227–274). Oxford, UK: Blackwell.

Luísa Araújo holds a Ph.D. in Curriculum and Instruction from the University of Delaware with a concentration in literacy studies and bilingual education. Her research interests include literacy and early reading development and second/foreign language learning. Her teaching activities include teaching courses on educational evaluation, research methods, and reading development. She has served as a director of several bachelor and master degrees in education and worked as a researcher for the European Commission in the area of Education and Training at the Joint Research Center. She is currently full professor in the School of Education in the Institute of Higher Education and Science (ISEC) in Lisbon, Portugal.

Patrícia Costa has worked at the Joint Research Centre of the European Commission since June 2011. She holds a PhD in the area of Pshycometrics/Statistics and received a master and bachelor's degrees in mathematics. Patrícia has been working as a statistician and psychometrician and contributing to research and policy support in the areas of Educational Assessment and Comparative Education Statistics. Her recent research is focused in secondary data analysis of international large-scale surveys and validation of instruments.

Nuno Crato is Professor of Mathematics and Statistics at the Lisbon School of Economics & Management of the University of Lisbon. Visiting senior research scientist at the European Commission Joint Research Center, his academic research focus on stochastic models, time series applications on financial and social problems, and statistical data-based evaluation of policy measures, namely education. He has published extensively on econometrics and statistics and worked on education policy analysis.

From 2011 to 2015, he was the Education and Science Minister of Portugal. During his tenure, mandatory schooling was raised from nine to 12 grade years, English was introduced as a mandatory subject starting at third grade, the dropout rate was reduced from c. 25% to 13.7%, retention rates improved, and Portuguese students achieved the best results ever in international surveys, namely PISA and TIMSS.

He was President of the Portuguese Mathematical Society (2004–2010) and director of the International Institute of Forecasters (2016–2020).

A prolific science writer, some of his books have been translated and published in various countries, including the U.S., U.K., Italy, and Brazil, namely his Figuring It Out (Springer 2010). His recently co-organized Data-Driven Policy Impact Evaluation (Springer 2019) is becoming an international reference in the field.

He has been a vocal voice in educational debates, publishing critical articles and books on education, advocating a structured curriculum, external evaluation of students, and a better content-knowledge training of teachers. He is founding organizer of the Lisbon Economics and Statistics of Education conference series and directs Iniciativa Educação, a program for fostering education.

For his work, he has received prizes from the European Mathematical Society (2003) and the European Union (2007). He was awarded Commander (2008) and Great-Cross (2015) of Prince Henry Order of the Portuguese Republic.