European Journal of Education Studies
ISSN: 2501 - 1111
ISSN-L: 2501 - 1111
Available on-line at: www.oapub.org/edu

doi: 10.5281/zenodo.3653287
Volume 6 | Issue 11 | 2020

CLASSIFICATION OF ASSESSMENT AND EVALUATION ACTIVITIES IN AN EIGHTH-GRADE TURKISH SCIENCE TEXTBOOK ACCORDING TO PISA SCIENCE LITERACY PROFICIENCY LEVELS[i]

Merve Nur Genç[ii],
Mustafa Sami Topçu
Department of Mathematics
and Science Education,
Yıldız Technical University,
Istanbul, Turkey

Abstract:
This qualitative study used a document analysis method to examine the assessment and evaluation activities at the end of each chapter and unit of an eighth-grade science course textbook according to the Program for International Student Assessment (PISA) science literacy proficiency scale. The textbook was approved by the Turkish Ministry of Education as the only textbook to be used for a period of five years starting from the 2018–2019 academic year. Descriptive analysis of data obtained from the PISA science literacy proficiency scale was undertaken, expert evaluation was used as a measure of the proficiency level corresponding to each item. Chi-Square goodness of fit test was applied to determine the significance of these values. Intraclass correlation values were calculated to reliably determine the consistency among the opinions of the experts included in this study. The study found that the textbook included activities at proficiency levels 1, 2, 3, 4, and 5; however, no assessment and evaluation activities were found to be included at level 6. It was also found that 62.23% of the activities were at level 1, 18.02% were at level 2, 11.15% were at level 3, 6.43% were at level 4, and 2.14% were at level 5. Additionally, adequacy levels of the activities were found to differ according to the corresponding units. It is recommended that the content and activities of the textbooks be reviewed again to develop high-level skills according to the PISA science literacy proficiency scale, and that activities used to measure high-level skills be included in the science textbooks.

[i] This study is a part of the master thesis of Merve Nur Genç under the supervision of Prof. Dr. Mustafa Sami Topçu.
[ii] Correspondence: email mervgnc@gmail.com

Keywords: PISA, science literacy, textbook, science education, assessment and evaluation activities

1. Introduction

The Program for International Student Assessment (PISA) is a test used to evaluate the ability of 15-year-old students to transfer the knowledge and skills they have received in educational environments to real-life situations. The test is applied to these students from countries in the Organization for Economic Cooperation and Development (OECD), and to students from other participating countries (OECD, 2017). The aim of the PISA program is to test students' knowledge and skills in the fields of science, reading literacy, interdisciplinary problem-solving, and mathematics (Dohn, 2007). In light of the data collected from the PISA assessments and questionnaires, which have been implemented every three years since 2000, it is possible to make predictions about the quality of education systems around the world owing to the knowledge and skills outcome of these fields (Bybee et al., 2009). Every three years, the PISA selects one of these domains (reading literacy, science, and mathematics) as the main focus of the PISA assessment. For instance, in 2000 and 2009 reading literacy was determined as the area of focus, while 2003 and 2012 focused on mathematics, and 2006 and 2015 focused on science. The main focus of the 2018 PISA implementation was reading literacy (OECD, 2019). According to the PISA, science literacy can be defined as "*the ability to understand nature, make comments, and draw inferences about it (nature), be able to identify scientific problems using scientific concepts, be able to use scientific process skills to solve them, and be able to willingly engage with ideas and professions related to science*" (OECD, 2019). A review of the contents of the PISA tests shows that the aim of the test is to research and determine the extent to which students have mastered these skills in daily life rather than determine whether they have acquired scientific knowledge. In this context, the PISA science literacy scale is used to assess the extent to which students have mastered certain identified critical thinking skills (OECD, 2017).

2. PISA Science Literacy Scale

Seven levels of proficiency in science literacy are described in the PISA. The skills and procedures that students can perform are described according to these proficiency levels. The explanations regarding the proficiency levels are given in the table below (OECD, 2019).

Table 1: PISA science literacy proficiency scale (OECD, 2019, p. 113)

Level	The Lowest Score	Percentage of Students at this Level (OECD Mean)	What students can do at this level (Task Description)
6	708	0.8 %	Students can draw on a range of interrelated scientific ideas and concepts from the physical, life, and earth and space sciences and use content, procedural and epistemic knowledge in order to offer explanatory hypotheses of novel scientific phenomena, events and processes or to make predictions. In interpreting data and evidence, they are able to discriminate between relevant and irrelevant information and can draw on knowledge external to the normal school curriculum. They can distinguish between arguments that are based on scientific evidence and theory and those based on other considerations. Level 6 students can evaluate competing designs of complex experiments, field studies or simulations and justify their choices.
5	633	6.8 %	Students can use abstract scientific ideas or concepts to explain unfamiliar and more complex phenomena, events and processes involving multiple causal links. They are able to apply more sophisticated epistemic knowledge to evaluate alternative experimental designs and justify their choices, and use theoretical knowledge to interpret information or make predictions. Level 5 students can evaluate ways of exploring a given question scientifically and identify limitations in interpretations of data sets, including sources and the effects of uncertainty in scientific data.
4	559	24.9 %	Students can use more complex or more abstract content knowledge, which is either provided or recalled, to construct explanations of more complex or less familiar events and processes. They can conduct experiments involving two or more independent variables in a constrained context. They are able to justify an experimental design by drawing on elements of procedural and epistemic knowledge. Level 4 students can interpret data drawn from a moderately complex data set or less familiar context, draw appropriate conclusions that go beyond the data and provide justifications for their choices.
3	484	52.3 %	Students can draw upon moderately complex content knowledge to identify or construct explanations of familiar phenomena. In less familiar or more complex situations, they can construct explanations with relevant cueing or support. They can draw on elements of procedural or epistemic knowledge to carry out a simple experiment in a constrained context. Level 3 students are able to distinguish between scientific and non-scientific issues and identify the evidence supporting a scientific claim.
2	410	78.0 %	Students are able to draw on everyday content knowledge and basic procedural knowledge to identify an appropriate scientific explanation, interpret data and identify the question being addressed in a simple experimental design. They can use basic or everyday scientific knowledge to identify a valid conclusion from

				a simple data set. Level 2 students demonstrate basic epistemic knowledge by being able to identify questions that can be investigated scientifically.
1	1a	335	94.1 %	Students are able to use basic or everyday content and procedural knowledge to recognise or identify explanations of simple scientific phenomena. With support, they can undertake structured scientific enquiries with no more than two variables. They are able to identify simple causal or correlational relationships and interpret graphical and visual data that require a low level of cognitive demand. Level 1a students can select the best scientific explanation for given data in familiar personal, local and global contexts.
	1b	261	99.3 %	Students can use basic or everyday scientific knowledge to recognise aspects of familiar or simple phenomena. They are able to identify simple patterns in data, recognise basic scientific terms and follow explicit instructions to carry out a scientific procedure.

Comments about students' proficiency across the country can be made based those data obtained from this scale (OECD, 2007). In the scale, starting from level 1b up to level 6, proficiency levels below level 2 are defined as "*low proficiency levels*", while the proficiency levels of 5 and 6 are introduced as "*high proficiency levels*" (Ministry of National Education (MoNE), 2016). Students should be able to put together the complex elements of a problem to solve challenging questions posed by the PISA, and should be able to make appropriate explanations by using their ideas and creativity to solve those problems (Aygogdu Iskenderoglu & Bakı, 2011). A total of 13.6% of students from OECD-member countries reached levels 5 and 6 for the 2018 PISA. Concerning the performance of Turkish students, only 2.4% reached the top two proficiency levels. In addition, 0.3% of Turkish students were found to be below level 1, while 24.8% were at level 1, 32.8% were at level 2, 27.3% were at level 3, 12.3% were at level 4, 2.3% were at level 5, and 0.1% at level 6; Turkish students scored below the mean scores of students from other OECD countries. Assessment and evaluation activities in textbooks can also affect this result, because textbooks play a vital role in science teaching and learning (Abd-El-Khalick et al., 2008). Textbooks are tools that transfer 99% of all of the knowledge provided by teachers and other teaching materials to students (Alkan, 1996). Furthermore, textbooks are frequently used by teachers among those teaching materials most (Karna et al., 2012).

In the relevant literature, no study could be found in which the assessment and evaluation activities of secondary school science-textbooks had been examined according to the PISA science literacy assessment scale. Within such a context, and considering the PISA science literacy proficiency scale, the levels to which the assessment and evaluation activities in the eighth-grade science course textbook correspond to the PISA science literacy proficiency scale are determined. The textbook in question was accepted by the MoNE in Turkey as the only textbook to be used for a period of five years following the 2018–2019 academic year.

3. Material and Methods

This research was designed as a case study and uses a qualitative research method. The document analysis method was used to analyze the study data. Documents have an important place in qualitative studies (Creswell, 2012), and document analysis involves the analysis of written materials related to the case or cases that concern the aim of a research investigation (Yıldırım & Şimşek, 2014). An eighth-grade science textbook, which is to be valid for a five-year period starting with the 2018–2019 academic year, was examined according to the Science Literacy Proficiency Scale of the 2018 PISA. PISA science-literacy proficiency levels of assessment, and evaluation activities in the book were analyzed by both the researcher as well as experts in the scientific field. The opinions of 10 experts were used to develop this study. Mode values, which have descriptive statistical values, were analyzed to determine those levels of assessment and evaluation at which the question items at the end of each chapter and unit were concentrated.

To accept the mode values determined as the agreed level of measurement and evaluation for each item, Chi-square goodness of fit analysis was performed on those data collected from the expert group. The values expected as a result of the analysis do not indicate fitness, but significance. After the level of science literacy competence for all assessment and evaluation items were determined, the intraclass correlation value—a fitness statistic—was analyzed using an analysis program to determine the degree of compatibility of the experts' responses. It is recommended that, in the process of calculating this reliability, at least three people conduct this assessment; however, using more people in this regard will make the study more reliable (Bademci, 1991). Accordingly, the evaluations of 10 experts were used to make this study more reliable. Results of these analyses showed which assessment and evaluation activities corresponded to which levels in the PISA science literacy proficiency scale. After the adequacy levels of the questions had been determined frequency and percentage tables, which show the levels of questions for each unit, were then formed and examined.

4. Results

Data obtained from Seasons and Climate, DNA and Genetic Code, Pressure, Matter and Industry, Simple Machines, Energy Transformations and Environmental Science, Electric Charges, and Electric Energy units in the textbook were investigated by drawing up tables. Table 2 shows the distribution of end-of-chapter and end-of-unit assessment, as well as the evaluation activities of these units individually.

Table 2: The distribution of assessment and evaluation activities in the eighth-grade science textbook

Units	End-of-Chapter Assessments	End-of-Unit Assessments
	N	N
Seasons and Climate	6	32
DNA and Genetical Code	11	19
Pressure	2	26
Matter and Industry	12	29
Simple Machines	3	30
Energy Transformations and Environmental Science	10	19
Electric Charges and Electric Energy	8	26
Numbers of the Activities	52	181
Total:	233	

The PISA science literacy proficiency levels of 233 assessment and evaluation activities in the textbook were reported based on experts' opinions. The question posed to students for Matter and Industry unit's end-of-unit evaluation activity was:

"44g of substance A and 66g of substance B are put in a container, and a reaction occurs between these substances (substances A and B completely disappear). This reaction makes substances C and D. Since 40 g of substance D is formed as a result of the reaction, how many g of substance C have been produced?" (Aytac et al., 2018, p. 109).

This was coded as a proficiency level 2 question. This question was also classified at proficiency level 2 by the experts because, based on the knowledge that the mass is preserved during chemical reactions, students were expected to be able to comment on the masses of the products produced when a basic reaction occurs. The following evaluation questions are posed at the end of the unit Electric Charges and Electric Energy:

"Tools mounted on top of tall buildings to protect against lightning are called..............."; If two objects charged with electricity repel each other, they are charged with…………....."; The objects in which positive and negative charges are equal are called………..… objects." (Aytac et al., 2018, p. 235)

These was coded at proficiency level 1. The students make use of the keywords given, and use easy procedures and within a narrow field of knowledge information while they solve the questions posed. They can define basic concepts in the questions to some small degree and match these definitions directly by placing them in situations. These examples are included in the PISA science literacy proficiency level 1. A question posed to students for Simple Machines unit's end-of-unit evaluation activity was:

"A farmer wants to move a big rock from his field. He uses a large wooden plank and a small rock. By how placing the plank and rock can the farmer move the big rock by applying

little effort force? Draw and show the mechanism. How does this mechanism help the farmer? Draw and show the directions of the forces applied."(Aytac et al., 2018, p. 165).

This question was coded at proficiency level 5. This activity was shown as an example of proficiency level 5, as students were expected to draw a prototype diagram and explain the causal relationships related to the problem by using their high-level reasoning and flexible thinking skills.

Table 3: Results of assessment and evaluation activity items of the Simple Machines unit according to experts' opinions

Activity Items	Mode	Chi-Square Value
M138	5	*3.6
M139	1	*6.4
M140	1	*6.4
M141	1	*3.6
M142	1	*6.4
M143	1	*6.4
M144	1	*6.4
M145	1	*6.4
M146	1	*6.4
M147	1	*3.6
M148	1	*6.4
M149	1	*6.4
M150	1	*6.4
M151	1	*6.4
M152	1	*6.4
M153	1	*6.4
M154	1	*6.4
M155	1	*6.4
M156	1	*3.6
M157	1	*6.4
M158	1	*6.4
M159	1	*6.4
M160	1	*3.6
M161	1	*6.4
M162	1	*6.4
M163	1	*3.6
M164	1	*6.4
M165	3	*3.6
M166	1	*3.6
M167	1	*3.6
M168	2	*3.6
M169	3	*3.6
M170	3	*3.6

*Significant at $p<0.05$

In the unit Simple Machines, 28 items at level 1, one item at level 2, three items at level 3, and one item at level 5 were determined; however, no item corresponding to levels 4 and 6 was found. The intraclass correlation values that indicate the consistency between the experts' opinions are given in the table below.

Table 4: The intraclass correlation values between the experts

Experts	Intraclass Correlation Value
For All Experts	0.852
For One Expert	0.982

As can be seen in Table 4, the intraclass correlation coefficient, which is determined as the measure of agreement among the experts, is 0.98 for one expert. This value shows that the consistency among expert opinions was perfectly reliable. The resultant reliability of all expert opinions was calculated as 0.85. This value shows that there was a high level of consistency among those experts who examined the textbook. It is seen, when the eighth-grade science textbook is examined as a whole, that there were assessment and evaluation activities at the levels 1, 2, 3, 4, and 5, but that no activity was included at level 6. It was found that 62.23% of the activities were at proficiency level 1, 18.02% were at proficiency level 2, 11.15% were at proficiency level 3, 6.43% were at proficiency level, and 2.14% were at proficiency level 5.

Table 5: The proficiency levels of assessment and evaluation activities in the eighth-grade science textbook

Proficiency Levels	N	%
1	145	62.23
2	42	18.02
3	26	11.15
4	15	6.43
5	5	2.14
6	0	0
Total	233	100

It was found that the proficiency levels of the activities differed according to the unit concerned. The table showing the levels of activities by units is given below:

Table 6: PISA proficiency levels of science literacy of units in the eighth-grade science textbook

Book Name	Units	PISA Science Literacy Proficiency Levels													
		Level 1		Level 2		Level 3		Level 4		Level 5		Level 6		Total	
		N	%	N	%	N	%	N	%	N	%	N	%	N	%
Eighth-Grade Secondary School Science Textbook	**Seasons and Climate**	31	81.57	5	13.15	2	5.26	0	0	0	0	0	0	38	100
	DNA and Genetical Code	13	43.33	5	16.66	5	16.66	4	13.33	3	10	0	0	30	100
	Pressure	17	60.71	6	21.42	4	14.28	1	3.57	0	0	0	0	28	100
	Matter and Industry	24	8.53	11	26.82	3	7.31	3	7.31	0	0	0	0	41	100
	Simple Machines	28	84.84	1	3.03	3	9.09	0	0	1	3.03	0	0	33	100
	Energy Transformations And Environmental Science	13	44.82	8	27.58	5	17.24	3	10.34	0	0	0	0	29	100
	Electric Charges and Electric Energy	19	55.88	6	17.64	4	11.76	4	11.76	1	2.94	0	0	34	100

In the textbook, 81.57% of questions included in the unit titled Seasons and Climate were determined to be level 1 questions, 13.15% were determined to be level 2 questions, and 5.26% were level 3 questions . The unit DNA and Genetic Code included 43.33% level 1 questions, 16.66% level 2 questions, 16.66% level 3 questions, 13.33% level 4 questions, and 10% level 5 questions. The unit Pressure included 60.71% level 1 questions, 21.42% level 2 questions, 14.28% level 3 questions, and 3.57% level 4 questions. The unit Matter and Industry included 8.53% level 1 questions, 26.82% level 2 questions, 7.31% level 3 questions, and 7.31% level 4 questions. The unit Energy Transformations and Environmental Science included 44.82% level 1 questions, 27.58% level 2 questions, 17.24% level 3 questions, and 10.34% level 4 questions. The unit Simple Machines included 84.84% level 1 questions, 3.03% level 2, 9.09% level 3, and 3.03% level 5 questions. Comparatively, the unit Electric Charges and Electric Energy included 55.88% level 1, 17.64% level 2, 11.76% level 3, 11.76% level 4, and 2.94% level 5 questions.

5. Discussion and Conclusion

This study aimed to examine the assessment and evaluation activities at the end of each chapter and unit of an eighth-grade science course textbook according to the PISA science literacy proficiency scale; the book was accepted by the Ministry of Education as the only textbook to be used for five years starting from the 2018–2019 academic year. In examining the questions in the book, most activities comprise level 1 and level 2 questions that represent/correspond to low proficiency levels. The results of PISA 2018 indicate that 24.8% of Turkish students were at level 1, and that 32.8% of them were at level 2 (OECD, 2019). In addition, 80% of questions in the book were determined as being at proficiency levels 1 and 2, so it is not surprising that the researchers encountered such a result. Accordingly, the number of the assessment and evaluation activities, especially those aiming at high-level thinking, should be increased in the textbooks. According to Savran (2004), all of the questions in the PISA measure the success of individuals using critical-thinking skills such as creative thinking, interpreting and evaluating the knowledge, problem solving, analyzing, and drawing conclusions. However, the lack of inclusion of adequate of questions representing all proficiency levels in the textbook can be considered to be a deficiency.

Karamustafaoğlu et al. (2016) reported that the activities and studies in science textbooks were inadequate and had deficiencies concerning the evaluation of the learning process. A doctoral dissertation examining the assessment and evaluation activities in physics textbooks proposed by the MoNE, which accorded with PISA science literacy proficiency levels (Türk, 2018), reported that the activities in the books were mostly at levels 1, 2, 3, and 4. It is also stated that activities corresponding to levels 5 and 6 were almost excluded. Aydoğdu İskenderoğlu and Baki (2011) examined eighth-grade mathematics textbooks according to PISA mathematics literacy proficiency levels, and noted that the activities in the books were at levels 1, 2, 3 and 4. These results are consistent with the PISA 2018 assessment results issued by the OECD (OECD, 2019). According to the results, the majority of students who participated in the PISA assessments from Turkey were at proficiency levels 1, 2, 3, and 4. Concerning the examination of the results of the current study—in which the evaluation activities in the eighth-grade science textbook are investigated—current end-of chapter and end-of-unit assessment and evaluation activities in the textbook consist of questions that do not allow students to go beyond certain situations by commenting on the directly given situations. This does not comply with PISA's point of view regarding transferring knowledge and skills to real life situations. It is therefore recommended that the assessment and evaluation activities in the textbooks are reviewed so that high-level skills can be developed for the PISA science literacy proficiency scale. Excepting this current research, no study exists in the literature in which secondary school science textbooks are examined according to PISA science literacy proficiency criteria. In this context, it is recommended that the levels of textbook science questions used in other secondary school levels according to PISA science literacy proficiency criteria be investigated for future studies. It is expected that these studies will be useful to practitioners in the process of writing new books in order to increase the PISA science literacy success level.

References

Abd-El-Khalick, F., Waters, M., & An-Phong, L. (2008). Representations of Nature Of Science in High School Chemistry Textbooks Over The Past Four Decades. *Journal of Research in Science Teaching, 45*(7), 835-855.

Alkan, C. (1996). *Eğitim Teknolojsi.* Ankara: Atilla Kitabevi.

Aygogdu Iskenderoglu, T., & Bakı, A. (2011). Classification of the Questions in an 8th Grade Mathematics Textbook with Respect to the Competency Levels of PISA. *Education and Science, 36*(161), 287-301.

Aytac, A., Türker, S., Bozkaya, T., & Üçüncü, Z. (2018). *Ortaokul ve İmam Hatip Ortaokulu Fen Bilimleri 8.Sınıf Ders Kitabı.* Ankara: Tutku Yayıncılık.

Bademci, V. (1991). *Varyans analiziyle güvenirlik hesaplanması.* Ankara, Ankara: Hacettepe Üniversitesi Yayınları.

Bybee, R., McCrae, B., & Laurie, R. (2009). PISA 2006: An assessment of scientific literacy. *Journal of Research in Science Teaching, 46*(8), 865-883.

Creswell, J. W. (2012). *Educational research: Planning, conducting, and evaluating quantitative.* Boston: Pearson Education.

Dohn, N. B. (2007). Knowledge and Skills for PISA Assessing the Assessment. *Journal of Philosophy of Education, 41*(1).

Karamustafaoğlu, S., Salar, U., & Celep, A. (2015). Ortaokul 5. Sınıf Fen Bilimleri Ders Kitabına Yönelik Öğretmen Görüşleri. *Gazi Eğitim Bilimleri Dergisi, 1*(2), 93-117.

Karna, P., Hakonen, R., & Kuusela, J. (2012). *Luonnontieteellinen osaaminen perusopetuksen 9. luokalla 2011.* Helsinki.

Lavonen, J. (2015). Foreword One. *Diversities and Interculturality in Textbooks: Finland as an Example* (s. 4-7). içinde Cambridge Scholars Publishing.

MEB. (2016). *PISA 2015 Ulusal Raporu.* T.C. Milli Eğitim Bakanlığı, Eğitimi Araştırma ve Geliştirme Dairesi Başkanlığı, Ankara.

OECD. (2007). *PISA 2006 Science Competencıes For Tomorrow's World Volume 1: Analysıs .* OECD. Paris: OECD Publishing.

OECD. (2017). *PISA 2015 Assessment and Analytical Framework: Science, Reading, Mathematic, Financial Literacy and Collaborative Problem Solving.* Paris: OECD Publishing.

OECD. (2019). *PISA 2018 Assessment and Analytical Framework.* PISA. Paris: OECD Publishing.

OECD. (2019). What can students do in science? *PISA 2018 Results (Volume I): What Students Know and Can Do* (s. 111-117). içinde Paris: OECD Publishing. https://doi.org/10.1787/5f07c754-en. adresinden alındı

Savran, N. Z. (2004). PISA-Projesi'nin Türk Eğitim Sistemi Açısından Değerlendirilmesi. *Gazi Üniversitesi Türk Eğitim Bilimleri Dergisi, 2*(4), s. 379-414.

Türk, O. (2018). *Ortaöğretim Fizik Ders Kitaplarındaki Ölçme ve Değerlendirme Etkinliklerinin PISA Sınavı ile Karşılaştırılması.* Doktora Tezi, Gazi Üniversitesi, Matematik ve Fen Bilimleri Eğitimi Ana Bilim Dalı, Ankara.

Yıldırım, A., & Şimşek, H. (2014). *Sosyal Bilimlerde Nitel Araştırma Yöntemleri.* Ankara: Seçkin Yayıncılık.