Investigating Consistency of Questions in Primary and Middle School Science Textbooks with Objectives in Science Curriculum

Süleyman Yaman

Correspondence: Süleyman Yaman, Ondokuz Mayıs University, Samsun, Turkey.

Abstract

Due to problems related their content and use; textbooks do not achieve the expected effect in learning although they are one of the most important elements of the science curriculum. Questions in textbooks are also important criteria in determining the effect of textbooks. In this study, it was aimed to compare questions in four different science textbooks, both qualitatively and quantitatively in terms of objectives of curriculum. 2677 questions retrieved from primary school 3rd and middle school 5th and 6th grades textbooks were investigated comparatively in terms of their ratio across objectives in the curriculum, cognitive levels and types. According to the analysis, it was determined that there were differences between distributions of questions in the same and different grade level textbooks. It was also found that questions related to the objectives regarding input and processing skills were more than questions related to the objectives of output skills. Besides, it was observed that most of the questions in textbooks were aligned with traditional assessment and evaluation tools, and the ratio of the alternative type tools was found to be lower than the ratio of four different textbooks.

Keywords: textbook, question types, cognitive level of question, science curriculum

1. Introduction

There have been many efforts to improve science education and increase its quality in the world (AAAS (The American Association for the Advancement of Science), 2001; MNE (Ministry of National Education), 2004; 2013; NRC (The National Research Council), 1996). One of the most important steps of these efforts is to update or modify the curriculum. The level of success in science to get better results at both national and international exams is one of the most important reasons of this change in Turkey (Berberoglu, Kaptan and Kutlu, 2002). The main objective of the changes that are made to the curriculum is to improve the quality of education (Ayas, Çepni and Akdeniz, 1993; Ünal, Coştu and Karatas, 2004); thus, to get better results in the measurement results. For this purpose, especially since the 1990s, many reforms have been made to science education in Turkey. Changes that were made in science education took place in the science curriculum in 1992, 2000, 2004 and 2013 respectively. The focal point of such studies and changes in the curriculum is concerned with how students' progress will be ensured and how their achievement level will be improved in science education (AAAS, 2001; MNE, 2007; NRC, 1996).

It is known that achievement levels of students are low in science and this failure is also known to affect other content areas negatively in Turkey (Berberoglu, Kaptan and Kutlu, 2002). In order to overcome this failure seen in science, changes and reforms have been made periodically to the curriculum in Turkey. Since 2000, the science curriculum has been changed three times; it was named science in 2000, science and technology in 2004, and sciences in 2013. Student-centred approaches have been adopted in these three curricula, and especially the constructivist approach has come into prominence in the last two curricula. It was especially emphasized that alternative assessment came into prominence instead of traditional assessment in 2004 and 2013 curricula (MNE, 2004; 2013).

One of the reason of change in the curriculum is students' low academic performance both in national exams such as TEOG (Exam for Transition from Primary Education to Secondary Education), YGS (Higher Education Examination) and LYS (Undergraduate Placement Examination), and in international exams such as TIMSS (The Trends in International Mathematics and Science Studies) and PISA (The Programme for International Student Assessment) (Berberoglu and Kalender, 2005; MNE, 2007). However, assuming that the failure in science stems only from the curriculum and it could be overcome with changes to the curriculum is not realistic. Some of the studies revealed that many factors have effects on achievement such as students' cognitive, affective and psychomotor skills; the school's

physical environment; students' study time and habits; the education level of parents; parents' support and teacher qualifications as well (Bayraktar, 2010; Özer and Anıl, 2011). Textbooks, which are important teaching materials, are also considered to have an impact on the students' achievement (Stake and Easley, 1978; Yager, 1996). Armbruster and Ostertag (1989) state that science lessons cannot be taught without textbooks; he also claims that science cannot be performed without textbooks. Similarly, Newport (1965) also emphasizes that textbooks are very important in achieving the goals in science.

The questions in textbooks are important tools which affect the quality of textbooks. Many research have been conducted which highlighted the importance of the questions in textbooks. For example, Stern and Ahlgren (2002) state that questions in textbooks do not provide a positive contribution to the development of students, whereas Armbruster and Ostertag (1993) and Meyer, Crummey and Greer (1988), indicate that there are differences between the distributions of the questions in textbooks. According to Yılmaz, Seçken and Morgil (1998), teachers think that questions in textbooks are above the level of the students. Shepardson and Pizzini (1991), who examined questions in textbooks with respect to the Costa's classification will be discussed below, present that most questions are at input level; that is, they urge students to memorize, while the number of high-level questions guiding students to think is not enough. Furthermore, it was expressed that the questions generally taught in textbooks in science in Turkey reveals the input and processing skills in general (Kahveci, 2009). However, through studies conducted according to classification of different cognitive skills, it was discovered that the questions in textbooks solely examined lower-level skills (Brill and Yarden, 2003; Ogan-Bekiroglu, 2007).

The questions asked to students in textbooks and exams have different purposes. These questions are asked in order to reveal the students' learning levels in cognitive domains. The vast majority of questions in textbooks are related to the cognitive domains. There are many classifications of cognitive domains questions posed to students in exams or textbooks (Arı, 2013). One such classification was also made by Arthur L. Costa in 1985. Costa (1985) states that questions asked to students are important proofs in revealing the information that students learned. A consistency is found between the steps of the most widely used classification developed by Bloom, Englehart, Furst, Hill, and Krathwohl in the cognitive field and the steps of classification developed by Costa (Dávila and Talanquer, 2010).

Costa (1985) points out that questions asked to students have an important role in shaping their cognitive and thinking skills. According to Costa, there are three levels of questions: Input, processing, and output questions. Input questions revealing students' lower-level thinking skills uncover whether they remember information or not as in knowledge and comprehension levels of Bloom's Taxonomy (Dávila and Talanquer, 2010; Pizzini et el., 1992). The questions in the processing level are the questions that reveal the meaning that students give to the information they have learned. Questions corresponding with application, analysis and synthesis levels of Bloom's Taxonomy give hints about whether meaningful learning becomes reality or not (Dávila and Talanquer, 2010). The questions in output level are about to measure students' skills of using information they have interpreted in new situations, and it corresponds to the evaluation level of Bloom' taxonomy (Dávila and Talanquer, 2010; Pizzini, Shepardson and Abell., 1992). Sample questions and key concepts for the Costa's classification are presented in Appendix 1.

Since textbooks are used by the vast majority of teachers in science (Radcliffe, Caverly, Hand and Franke, 2008), the quality and quantity of questions in this main source is very important. Especially, to what extent "supplementary measurement tools" emphasized in the science curriculum (MNE, 2013) are given place in textbooks and determining what are the skill levels of the students the questions are designed to measure may be one of the criteria of whether the purpose of the curriculum comes true as required. In this sense, it is considered to be useful to compare the questions in textbooks prepared according to the science curricula established in 2013 in terms of their quantity and cognitive levels (Kahveci, 2009). Concordantly, the aim of this study is to compare the quantity of the questions which take place in 3rd, 5th and 6th grades science textbooks and cognitive levels based on Costa's classification. The purpose of this research is to answer the following question:

1. Are there any differences between the number of questions and the number of related objectives in science curriculum and weekly course hours?

2. Are there any differences between the location of questions in textbooks and their cognitive levels?

3. Are there any differences among types of questions in science textbooks?

4. Is there a difference between distributions of traditional and alternative types of questions?

2. Method

In this study, the quantitative content analysis method was used to analyse the questions in science textbooks (Berelson, 1952; cited in Rourke and Anderson, 2004). Content analysis is a process of summarizing and indicating the main components of available information and the messages they contain (Cohen, Manion and Morrison, 2007). The similar

data were combined, organized and interpreted within the framework of certain concepts and themes in order to make content analysis (Yıldırım and Şimşek, 2006).

In this study, "categorical analysis", a type of content analysis, was employed. In this analysis, at first textbooks were divided into sub-units, then these units were grouped according to predetermined criteria (Tavşancıl and Aslan, 2001). Content analysis was carried out by five steps: a) detecting units, b) selecting samples, c) segmenting, d) interpretation and e) explanation (Krippendorff, 2012). In this research, content analysis were performed by these procedures as well. First the units which will be studied (questions in science textbooks) were identified. Then, four textbooks forming the study's sample were examined. Questions in the textbooks were divided into parts according to the classification made by Costa (1985). These parts were interpreted by the help of descriptive statistics and finally they were explained regarding their consistency with curriculum and other studies.

Sample: The textbooks of the sample are those that were approved by Board of Education and distributed to state schools and used in the 2014-15 academic year. Since the examined textbooks are still in use, convenience sampling was used as the sampling method. One of the textbooks was prepared by MNE whereas the others were prepared by private publishers and the usage rights of these textbooks were transferred to MNE for five years. The examined textbooks represented a universal span since all of the textbooks used in schools were sampled as required by the science curriculum which came into use in 2013.

Compulsory education has been 12 years in Turkey since 2013. The first four years comprises primary school education, the second four years forms middle school education, and the last four years constitutes high school education. As lessons are taught three hours per week to 3rd and 4th grades in primary school, but four hours per week to 5th, 6th, 7th and 8th grades in middle school. Textbooks are distributed free of charge to students in public schools, and they are used mandatorily in all state schools after given approval by Board of Education. This study is limited to questions in four different science textbooks which were prepared for primary school 3rd and middle school 5th and 6th grades. Two textbooks taught to 5th grade, one textbook taught to 3rd grade and one textbook taught to 6th grade were included in the scope. The Board of Education approved one textbook taught to primary school 3rd grade; one textbook taught to middle school 6th grade; two different textbooks taught to primary school 5th grade and these textbooks were distributed free of charge to all students studying in state schools. The reason why textbooks in primary school 4th and middle school 7th and 8th grades were not included into the study is the fact that in these classes textbooks prepared for the according to previous curriculum (science and technology curriculum in 2004) were still in use when the study was conducted.

Data Collection and Analysis: There are two preferred methods of content analysis. The first method is to analyse a randomly selected portion of the examined content and to generalize it to the whole content. The second method is to analyse the whole examined content (Wang, 1998). As it is scientifically stronger and reflects directly the whole content without the need for generalization, for this study, the second method was preferred.

Multiple-raters analysed different parts selected from the same content in order to determine reliability of the raters. About 10% part of each textbook was examined independently by three different raters. Then the consistency among the scores of these people was determined by interrater reliability. As a result of this analysis, the analysis of the three raters was found to be consistent with the level of 0.85. The questions rated by these three different raters were re-examined and the process continued until unanimity/majority of votes was provided.

The data obtained from the textbooks was analysed with descriptive analysis according to their research problems. Due to its large variety of alternative questions types of textbooks examined and since the number of all questions types was less, the groups were combined. For this purpose, as question types such as "the self-assessment, peer assessment, diagnostic tree, banners/posters, structured grid, concept maps, puzzles, knowledge maps, concept network and so on." was included in the scope of alternative assessment, they were classified under "alternative" title, while the traditional question types were grouped according to the classifications made by Anderson (1972), Armbruster and Ostertag (1993), and Haladyna (1992, 2004).

The classification made by Costa was used for the cognitive level of questions. That in which levels, which are input, processing or output, the questions are included was determined for each textbook. Chi-square analysis was used to reveal whether there was a difference among textbooks. A sample of questions taken from each textbook for Costa's different levels was shown in Appendix 2.

3. Findings

In this section, questions in primary school 3rd and middle school 5th and 6th grades textbooks were analysed across their ratio per the objectives of curriculum, the location of the question, cognitive level, question types, and weekly course hours. And statistics on the results of this analysis were discussed. The ratio of question numbers in four different textbooks to the number of objectives in science curriculum and to the total course hours was presented in Table 1.

Table 1. The Ratio of the Questions in Science Textbooks to the Number of Objectives of Curriculum and to the Weekly Course Hours

Grade	f	no	f / no	tch	f / tch
3- (Bilim ve Kültür)	314	32	9.81	108	2.91
4- A (MNE)	803	44	18.25	144	5.58
4- B (Sevgi)	767	44	17.43	144	5.33
6- (FenBil)	793	52	15.25	144	5.51
Total	**2677**	**172**	**15.56**	**540**	**4.82**

f=Frequency, no= Number of objectives, tch=Total course hours

According to Table 1, the frequency of questions in textbooks varies significantly between primary and middle school textbooks. When total course hours in an academic year are taken into consideration (3 hours per week in primary school and totally 108 hours; 4 hours per week in middle school and totally 144 hours), the reason why primary school textbooks have less questions is an explainable situation. However, it is surprising that although there are 52 objectives in the sixth grade, the number of questions is not high at the same rate. Given the number of objectives and total class hours, Chi-square convenience test results show that there is a significant difference among the questions in textbooks according to grade levels ($\chi2$= 107.38, p < 0.05). This difference is determined to be in favour of the 5th and 6th grade. Table 2 shows the distribution of the questions located in the Table above according to their location and purpose in textbooks.

Table 2. Results of Frequency Distribution of Cognitive Levels according to Location of Questions in Science Textbooks

Books	Beginning of chapter			In-chapter			End of chapter			Total
	Input	Processing	Output	Input	Processing	Output	Input	Processing	Output	
3- (Bilim ve Kültür)	17	0	4	37	32	22	187	9	6	314
4- A (MNE)	15	19	19	152	168	90	116	118	106	803
4- B (Sevgi)	0	12	5	296	88	47	253	60	6	767
6- (FenBil)	10	13	2	209	210	93	186	47	23	793
Total	**42**	**44**	**30**	**694**	**498**	**252**	**742**	**234**	**141**	**2677**

When the distribution of the cognitive level is examined according to the location of questions in four different science textbooks in Table 2, it is observed that more than half of the questions (53.94% of all questions) take place in the text; about two-fifth (41.73%) occur at the end of the unit, and about one-twenty fifth (4.33%) is at the beginning of unit. According to these findings, it was determined that there are significant differences between cognitive levels of questions and their locations ($\chi2$=106.64, p<.05).

According to this Table, with regard to the cognitive levels of questions in science textbooks, the order from the highest to the lowest scores is as follows: input is 55.21 percent, processing is 28.99 percent, and output is 15.80 percent. When the cognitive level of questions at the beginning of the unit was examined, it is found that questions related to input and process skills appear to have a very close ratio. Given the questions in textbooks, it is determined that nearly half of questions are for input whereas one of six is for output. Considering the distribution of questions at the end of the unit, it is observed that two-thirds of questions aim to measure input skills, while 12.62 percent are for measuring output skills. The distribution of the question types located in these textbooks is presented in Table 3.

Table 3. Distributions of Questions in Science Textbooks in terms of Question Types

Grades	Question types						
	Essay	Cloze	MC	TF	Matching	Alternative	Total
3- (Bilim ve Kültür)	108	28	18	43	38	79	314
4- A (MNE)	412	77	29	14	61	210	803
4- B (Sevgi)	262	119	88	80	14	204	767
6- (FenBil)	327	177	78	90	11	110	793
Total	**1109**	**401**	**213**	**227**	**124**	**603**	**2677**

MC: Multiple-Choice, TF: True-False

According to Table 3, two-fifths (37.69%) of questions in the science textbook are "open-ended", while the ratio of question types related to supplementary assessment and evaluation tools is 22.52 percent of all the questions. According to these results, the traditional question types are dominant and their percentage is 77.48. The ratio of assessment tools (the self-assessment, peer assessment, diagnostic tree, banners/posters, structured grid, concept maps, puzzles, knowledge maps, concept network) grouped as supplementary and located in textbooks is found to be about one-fifth of all the questions. When the distribution of the questions is examined in the textbook, it is 6th grade science textbook that includes the least supplementary questions types proportionally.

Table 4. Distributions of Questions in Science Textbooks in terms of Cognitive Levels

Grades	Traditional			Supplementary			Total		
	Input	Processing	Output	Input	Processing	Output	Input	Processing	Output
3- (Bilim ve Kültür)	144	33	31	97	8	1	241	41	32
4- A (MNE)	174	234	181	109	71	34	283	305	215
4- B (Sevgi)	369	144	42	180	16	16	549	160	58
6- (FenBil)	347	235	108	58	35	10	405	270	118
Total	**1034**	**646**	**362**	**444**	**130**	**61**	**1478**	**776**	**423**

In Table 4, all of the questions acknowledged as traditional were considered as a single group and they were grouped according to supplementary measurement tools and Costa's cognitive level. Accordingly, whereas about half of the traditional questions consisted of the input level questions, it was found that about one-fifth of questions were for output skill. It was determined that while 70 percent of questions of supplementary types consisted of questions revealing input skills, only 10 percent of questions were for measuring output skills. Besides, a significant differences ($\chi 2 = 73.67$, p <0.05) was found between the distribution of the traditional type questions in the textbooks and the distribution of the supplementary type questions.

4. Discussion, Conclusion and Suggestions

This chapter presents the conclusion and suggestions based on the findings of the questions in primary school 3rd and middle school 5th and 6th grades science textbooks. The results of the analysis show that the ratio of the number of questions in science textbooks which have been used in primary and middle schools since 2013 is not consistent with the ratio of the number of objectives in curriculum. This difference was found to be especially striking between primary and middle school textbooks. Although there is a regulation which was issued by the MNE about how to prepare textbooks in Turkey, it can be seen that this regulation includes pure clarifications but it lacks in detail and quality (MNE, 2009). Because this case does not include clear information and criteria concerning which characteristics should be considered for the questions, the author(s) of the book have prepared the questions only to the extent mentioned in the curriculum. Accordingly, it means that the author/authors preparing textbooks use their preferences, knowledge, experience in design, and thoughts regarding the question in textbooks.

While quantitative differences among the textbooks can be overcome, deficiencies in their quality may cause serious obstacles that prevent education from accomplishing its goals. Haladyna (2004) states that asking a few qualified questions are more effective instead of asking too many unqualified questions. The cognitive level of questions in examined textbooks for this purpose was investigated according to the cognitive domain classification made by Costa (1985). Generally, it is observed that questions are positioned in text and at the end of the unit. In other words, the number of questions aimed at preparing students to the unit or subject is found to be quite limited. Besides, more than half of the questions are witnessed to be comprised of questions at the input level of cognitive domain. The fewest questions are those regarding measuring output skills. That is, while the majority of questions direct students to memorize information, they aim to uncover students' skills of memorization. Examining the questions in textbooks, like this study, Dávila and Talanquer (2010) claims that questions in textbooks have features of uncovering low-level cognitive skills; the number of questions that measure high-level skills are very limited. This result may mean that neither quantity nor quality of questions is taken into consideration when the questions in textbooks are prepared.

Another point worthy of attention in this study is related to the types of questions in textbooks. It was determined that an important majority of questions in examined textbooks consisted of open-ended questions. These question types have begun to be discussed recently in Turkey and it has been stated that there are efforts to use these questions in central exams (YÖK (Higher Education Council), 2015). Although the main feature of these question types aims to reveal students' high level thinking skills (output skill), it is observed that questions in examined textbooks are far from this purpose. Popham (2003) states that open-ended questions should be prepared to determine how and why events occurred. Such questions are those that explore higher level thinking skills. However, questions such as what, who, when that measure low-level skills are often used in examined textbooks.

It is determined that about one-fifth of question types belong to supplementary measurement tools which are accepted as an alternative type. This result indicates that this issue has not been mentioned enough although it is seen as a development in terms of getting closer to objectives of curriculum while comparing previous studies (MNE, 2004, 2013). Additionally, that more than two-third of supplementary question types is at Costa's classification of input level shows that the intended level for these questions has not been reached yet, as the distribution of cognitive levels of traditional question types is more balanced than those of supplementary types. This can be explained by the fact that authors of the textbooks should be more experienced in preparing questions of a traditional type. Anderson (1998) stated that it is not easy to move from traditional assessment to an alternative one. In this context, it can be articulated that a balance can be achieved in preparing future textbooks by paying more attention to the studies conducted to reach the objectives of the curriculum in creating supplementary question types.

Indeed assessment and evaluation is one of the most important elements of the education process. Millions of students, parents and hundreds of thousands of educators/administrators' concrete expectations for outcomes of the educational process can be revealed by assessment and evaluation. Curricula prepared with huge budget and the textbooks prepared for these curricula are expected to be qualified in every aspect. However, based on the findings and conclusions above, although there are some observable improvements towards objectives in the curricula in science textbooks, it is determined that these improvements remains far behind the measure which is claimed to be in the curriculum. That the number of supplementary question types increases- compared to the textbooks prepared according to the previous curricula- can be expected as an important development. However, that the increase in quality of questions is not at the same rate is seen as a significant deficiency. In addition to the quantity, the increase in the quality of the questions will crucially contribute to curricula to achieve their goals since textbooks are one of the most preferred tools in lessons by teachers (Radcliffe et al, 2008). This requires that both textbook authors and experts should to be more careful and attentive.

Comparing the results obtained in this study to the questions in science textbooks will be published in the future will increase the results of the generalizability level. It will also lead to achieving useful results by analysing the textbooks of other courses. More importantly, analysing the question of textbooks of the countries which are at the forefront of international exams or R & D (research and development) will contribute to achieving very useful data. In this study, the level of questions regarding the taxonomy of cognitive domain was examined. Besides, it will be useful to investigate to what extent the questions assess affective and psychomotor skills.

References

AAAS (American Association for the Advancement of Science). (2001). *Designs for science literacy*. Oxford, England: Oxford University Press.

Anderson, R. C. (1972). How to construct achievement tests to assess comprehension. *Review of Research in Education, 42*(2), 145-170. https://doi.org/10.3102/00346543042002145

Anderson, R. S. (1998). Why talk about different ways to grade? The shift from traditional assessment to alternative assessment. *New Directions for Teaching and Learning, 74,* 5-16. https://doi.org/10.1002/tl.7401

Armbruster, B., & Ostertag, J. (1993). Questions in elementary science and social studies textbooks. In B.K. Britton, A. Woodward, & M. Binkley (Eds.), *Learning from textbooks: Theory and practice*, (pp. 69-93), Hillsdale, NY: Lawrence Erlbaum Associates, Publ.

Ayas, A., Çepni, S., & Akdeniz, A. R. (1993). Development of the Turkish secondary science curriculum. *Science Education, 77*(4), 433-440. https://doi.org/10.1002/sce.3730770406

Bayraktar, Ş. (2010). Status of science education in turkey according to the results of trends in mathematics and science study (TIMSS 2007): Factors effecting science achievement. *Selçuk University Journal of Ahmet Keleşoğlu Education Faculty, 30,* 249-270.

Berberoglu, G., & Kalender, İ. (2005). Investigation of student achievement across years, school types and regions: The SSE and PISA analysis. *Educational Sciences & Practice, 4*(7), 21-35.

Berberoglu, G., Kaptan, F., & Kutlu, Ö. (2002). Analyzing the higher order cognitive skills of 8th gradres in Turkey. *V. National Science and Mathematics Education Congress,* September 16-18, Middle East Technical University, Ankara.

Bloom, B., Englehart, M. Furst, E., Hill, W., & Krathwohl, D. (1956). *Taxonomy of educational objectives: The classification of educational goals. Handbook I: Cognitive domain.* New York, Toronto: Longmans, Green.

Brill, G., & Yarden, A. (2003). Learning biology through research papers: A stimulus for question-asking by high-school students. *Cell Biology Education, 1*(2), 266-274. https://doi.org/10.1187/cbe.02-12-0062

Cohen, L. M., Manion, L., & Morrison, K. (2007). *Research methods in education* (6th ed). New York: Routledge.

Costa, A. L. (1985). Teacher behaviors that enable student thinking. In A.L. Costa (Ed.), *Developing minds: A resource book for teaching thinking,* (pp. 125-137), Alexandria, VA: Association for Supervision and Curriculum Development.

Dávila, K., & Talanquer, V. (2010). Classifying end-of-chapter questions and problems for selected general chemistry textbooks used in the United States. *Journal of Chemical Education, 87*(1), 97-101. https://doi.org/10.1021/ed8000232

Haladyna, T. M. (1992). The effectiveness of several multiple-choice formats. *Applied Measurement in Education, 5*(1), 73-88. https://doi.org/10.1207/s15324818ame0501_6

Haladyna, T. M. (2004). *Developing and validating multiple-choice test items (3ᵗʰ ed.).* New Jersey: Lawrence Erlbaum Associates, Publishers.

HEC (Higher Education Council). (2015). *Written exam (exam with open-ended questions).* Retrieved October 12, 2015, from'http://www.osym.gov.tr/belge/1-23308/yazili-sinav-acik-uclu-sorularla-sinav- 04022015.html

Kahveci, A. (2009). Quantitative analysis of science and chemistry textbooks for indicators of reform: A complementary perspective. *International Journal of Science Education,* 1-25.

Krippendorff, K. (2012). *Content analysis: An introduction to its methodology (3ᵗʰ ed.).* Thousand Oaks, CA: Sage Publication.

Meyer, L. A., Crummey, L., & Greer, E. A. (1988). Elementary science textbooks: Their contents, text characteristics, and comprehensibility. *Journal of Research in Science Teaching, 25,* 435-463. https://doi.org/10.1002/tea.3660250603

MNE (Ministry of National Education). (2004). *Elementary science and technology courses (4th and 5th grades) curriculum.* Retrieved February 19, 2009, from' http://ttkb.meb.gov.tr. index1024.htm.

MNE. (2007). *OBBS report: Turkish report of determining the success of the primary school students.* Ministry of National Education Research and Development Department, Ankara. Retrieved February 02, 2010, from' http://earged.meb.gov.tr

MNE. (2009). *Course books and educational materials regulations.* Retrieved October 12, 2015, from' http://mevzuat.meb.gov.tr/html/27449_0.html

MNE. (2013). *Elementary education institutions (primary and secondary schools) science courses (3, 4, 5, 6, 7 and 8th grades) curriculum.* Ankara. Retrieved January 15, 2015, from http://ttkb.meb.gov.tr/www/ogretim-programlari/icerik/72

Newport, J.F. (1965). An evaluation of selected series of elementary school science textbooks. *Science Education, 49*(5), 479-484. https://doi.org/10.1002/sce.3730490511.

NRC (National Research Council). (1996). *National science education standards.* Washington, DC: National Academy Press.

Ogan-Bekiroglu, F. (2007). To what degree do the currently used physics textbooks meet the expectations? *Journal of Science Teacher Education, 18*(4), 599-628. https://doi.org/10.1007/s10972-007-9045-8

Özer, Y., & Anıl, D. (2011). Examining the factors affecting students' science and mathematics achievement with structural equation modeling. *Hacettepe University Journal of Education Faculty, 41,* 313-324.

Pizzini, E. L., Shepardson, D. P., & Abell, S. K. (1992). The questioning level of select middle school science textbooks. *School Science and Mathematics, 92,* 74-79. https://doi.org/10.1111/j.1949-8594.1992.tb12145.x

Popham, W. J. (2003). *Test better, teach better: The instructional role of assessment.* Alexandria, Virginia: Association for Supervision and Curriculum Development.

Radcliffe, R., Caverly, D., Hand, J., & Franke, D. (2008). Improving reading in a middle school science classroom. *Journal of Adolescent and Adult Literacy, 51*(5), 398-408. https://doi.org/10.1598/JAAL.51.5.3

Rourke, L., & Anderson, T. (2004). Validity in quantitative content analysis. *Educational Technology Research and Development, 52*(1), 5-18. https://doi.org/10.1007/BF02504769

Shepardson, D. P., & Pizzini, E. L. (1991). Questioning levels of junior high school science textbooks and their implications for learning textual information. *Science Education, 75,* 673-682. https://doi.org/10.1002/sce.3730750607

Stake, R. E., & Easley, J. A. (1978). *Case studies in science education.* Center for Instructional Research and Curriculum Evaluation, Urbana: University of Illinois.

Stern, L., & Ahlgren, A. (2002). Analysis of students' assessments in middle school curriculum materials: Aiming precisely at benchmarks and standards. *Journal of Research in Science Teaching, 39*(9), 889-910. https://doi.org/10.1002/tea.10050

Tavşancıl, E., & Aslan, E. (2001). *Content analysis and application examples for oral, writing and other materials.* İstanbul: Epsilon Publishing.

Ünal, S., Coştu, B., & Karataş, F. Ö. (2004). A general look at the science curriculum development studies in Turkey. *Gazi University Journal of Gazi Education Faculty, 24*(2), 183-202.

Wang, H. A. (1998 April). *Science textbook studies reanalysis: Teachers "friendly" content analysis methods.* Paper presented the Annual Meeting of NARST, 71st, April 19-22, San Diego, CA, (ERIC Document Reproduction Service No. ED423142).

Yager, R. E. (1996). *Science/technology/society as a reform in science education.* New York: State University Press.

Yıldırım, A., & Şimşek, H. (2006). *Qualitative research methods in the social sciences (5th ed.).* Ankara: Seçkin Publishing.

Yılmaz, A., Seçken, N., & Morgil, İ. (1998). Investigate the suitability according to chemistry education of chemistry 3 textbooks at 11th grade high school. *Hacettepe University Journal of Education Faculty*, *14,* 73-83.

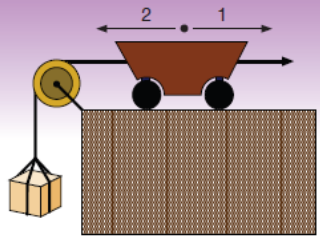**Appendix1.** Key concepts and question samples for the level of Costa's classification

Level	Key Question Words	Question Samples
Input	Identification, Completion, Listing, Selection, Observation, Narration, Summarizing, Remembering ...	- What is the best phrase describing this picture? - What are the names of the substances used in the experiment? - Summarize the events that occurred here.
Processing	Establishing a cause-effect relationship, Analysis, Synthesis, Comparison, Inference, Editing, Selection, Sorting, Classification, Creating analogies ...	- What is the reason for this event? - What is the relationship between these two events? - What distinguishes these agents from each other?
Output	Practicing a principle, Imagination, Planning, Evaluation, Forming Hypotheses, Generalization, Establishing a Model	-What is the best solution to eliminate this problem? - What changes if we take the plant from environment A to environment B? - If it continues in this way, what happens in 2025?

Appendix2. Question examples according to Costa's cognitive domain taxonomy from Science Textbooks

	Third	Fifth-A	Fifth-B	Sixth
Input	We should brush……… at least twice a day. (p. 26)	Do you know that our body consists of three-quarters water? (p. 27)	Can give you examples of the use of transparent and non-transparent substances(p. 108)	Substances that transmit electrical energy are called .... (p. 174)
Processing	Do loud noises damage our health? How? (p. 68)	What can we do to protect our urinary health? (p. 60)	Which information do fossils provide us about the past? (p. 160)	The water in teapot evaporates once heated. Which of the following is not different for evaporating water and for teapot? A) Space between the particles B) vibration, rotation and translation movements of particles C) Compressibility properties of particles D) Identification characteristics of the particles(p. 95)
Output	How do you think life would be without electricity? (p. 115)	Why doesn't water boil while there is evaporation on the surface of hot water with 78 °C?(p. 124)	"We do not inherit the earth from our ancestors; we borrow it from our children (Native American Proverb). Write down a text explaining the proverb above. (p. 183)	What kind of relationship is there between the development of microscopy and other technological tools and what we have learnt about cells? (p. 24)

2. Şekildeki düzenekte bir ucuna makara bağlanmış araba 1 yönünde kuvvet uygulanarak hareket ettiriliyor. Aşağıdakilerden hangisinin yapılması hâlinde araba 1 yönünde hızlanma hareketi yapar?

A) Yük miktarının artırılması

B) 1 yönünde uygulanan kuvvetin azaltılması

C) 2 yönünde kuvvet uygulanması

D) 1 yönünde uygulanan kuvvetin arttırılması

B Aşağıda çerçeve içinde bazı sözcük ve ifadeler verilmiştir. Bunları kullanarak eksik cümleleri uygun şekilde tamamlayalım ve defterimize yazalım.

| şekil | yavaş | tehlikeli | esnek | yön değiştirme |

| hızlı | yön | kuvvet | hareket | sallanma | dönme |

1. Kuvvet uygulanan cisimler .... ve ....değiştirebilir.

2. Bir varlığın başka bir noktaya göre yer değiştirmesi .... olarak ifade edilir.

3. Varlıkları harekete geçirmek için uyguladığımız itme ve çekmeyi .... olarak ifade ederiz.

4. Hareket eden bazı cisimleri durdurmak .... olabilir.

5. Varlıkların belli bir zaman diliminde aldıkları yolu karşılaştırarak .... ya da .... olduğunu söyleyebiliriz.

6. Kuvvetin etkisi ortadan kalkınca .... cisimler eski şekillerine döner.

7. Hareket eden bir otomobilin tekerleği .... hareketi yapar.

Figure 1. Examples of question from a textbook (Source: MNE, Grade 4)