



How Fifth Graders Are Assessed Through Central Exams in Turkey: A Comparison with TIMSS 2019 Assessment Framework*†

Research Article

Ahmet DELIL¹

¹ Celal Bayar University, Faculty of Education, Department of Mathematics Education, Manisa, Turkey, ORCID: 0000-0003-0234-2816

To cite this article: Delil, A. (2019). How Fifth Graders Are Assessed Through Central Exams in Turkey: A Comparison with TIMSS 2019 Assessment Framework, *International Online Journal of Educational Sciences*, 11 (3), 222-234.

ARTICLE INFO

Article History:

Received: 22.03.2019

Available online:
19.07.2019

ABSTRACT

In this study, Turkish fifth-grade bursary examination mathematics items, all of which are of multiple choice, from years 1999 to 2018 are analyzed based on the cognitive and content domains of TIMSS 2019 assessment framework. According to document analysis of the data collected through official website of Ministry of National Education, the distribution of the cognitive demands of all items indicates that there is an emphasis on applying or reasoning domain with about 74%. This result leads to the conclusion that Turkish fifth-grade mathematics items require upper cognitive demands comparing with TIMSS target percentages. Also, it is observed that there are dramatic changes in cognitive demands of the items from year to year which may mean that there is no framework followed for the bursary examinations. Examining the content domains of the bursary examination, findings reveal that the majority of the items are in numbers (60%), and it is followed by measurement and geometry (34%), and data (6%), which are similar with the target percentages of TIMSS 2019 that are 50, 35, and 15 respectively. Finally, suggestions are given in the light of findings and educational reforms in Turkey.

© 2019 IOJES. All rights reserved

Keywords:

TIMSS, cognitive domains, content domains, fifth grade mathematics, central examinations.

Introduction

The main high stakes exams conducted in Turkey are placement exams at eighth grade to middle schools, at twelfth grade to universities, and at the end of higher education to governmental jobs. There are also central exams from fifth to eleventh grades for boarding and scholarship, or bursary examinations in short. So, the term central exam, or sometimes called external exam, refers to regional or nationwide exam for some purposes (Woessmann, 2002, p.2). In Turkey, these purposes can be placement for further studies,

* This work was supported by Scientific Research Project Office of Manisa Celal Bayar University. Project Number: 2018-202

† This study was partially presented in International Conference on Science, Mathematics, Entrepreneurship and Technology Education, April 12-14, 2019, IZMIR.

‡ Corresponding author's address: Manisa Celal Bayar University

Telephone: +905325045446

e-mail: delil@hotmail.com

DOI: <https://doi.org/10.15345/iojes.2019.03.015>

boarding and scholarship, comparison of student achievements, comparison of country achievements, or summative evaluations. As Woessmann (2003) states, central exams are like “currencies” that function as unit of value in the economic system, and they are a measure of value for shareholders in education systems.

The name for boarding and scholarship examinations in Turkey is the Elementary and Middle-school Bursary Examinations (EMBE), and starting from 1999, Turkish Ministry of National Education (MONE) has been conducting the bursary examination on fifth to eleventh grade students that have annual family income below 13480 ₺ (Turkish Lira), which is now about 2200 € (Euro) and the amount of bursary is 260 ₺ per month which is continuously paid until the end of the middle school (EMBE, 2019). Thus, the earliest central examination in terms of grades conducted in Turkey is the fifth grade bursary examination, and it is the closest exam to TIMSS (The Trends in International Mathematics and Science Study) fourth grade exam in terms of its content and grade level.

First conducted in 1995 only on fourth, and repeated in 1999 with fourth and eighth grade students, TIMSS reports the trend data of students’ mathematics and science achievement worldwide for every four years. While there are cognitive domains classified as *knowing*, *applying* and *reasoning* with the target percentages 40, 40 and 20 respectively; there are content domains as *number*, *measurement and geometry*, and *data* with the target percentages 50, 30 and 20 respectively, and the item types are *multiple-choice* and *constructed-response* with about the same proportions in TIMSS 2019 fourth grade mathematics assessments framework. TIMSS also collects data through context questionnaire from students about their experiences, instruction, and attitudes toward learning, from teachers and school principals about school and classroom resources, and from parents about students’ home contexts for learning (Mullis and Martin, 2017).

While Turkish fourth grade students participated in TIMSS in 2011 and 2015 only, eighth graders participated in 1999, 2007, 2011 and 2015. The final TIMSS 2015 results show that the performance of Turkish fourth graders is below the average which is similar to the previous TIMSS 2011 results. One of the consequences of the poor performances of Turkish students in past years TIMSS and the other international examinations were the reforms in primary school curricula in 2005, 2013 and 2017. After these changes in mathematics curriculum, unsatisfactory small progress in TIMSS mathematics examinations has been observed (MONE, 2015, p.21). So, linking TIMSS assessments and national assessments it may be useful to compare the similarities and differences between them. It may also be worthwhile to examine the reflections of reforms on central exams.

Efforts to understand the components of human learning have led to the conclusion that every learning incident is in terms of three types of domains: cognitive (knowledge based domain), affective (attitudinal based domain) and psychomotor (skills based domain) and the best known classification of cognitive domain is so called Bloom’s Taxonomy which classifies cognitive domain under six categories. Bloom and his colleagues used the term cognitive “to include activities such as remembering and recalling knowledge, thinking, problem solving, creating” (Bloom, Engelhart, Furst, Hill and Krathwohl, 1956, p.2). Since then many other classifications of cognitive domain have been made (e.g. Anderson and Krathwohl, 2001; Marzano and Kendal, 2008). Nitko (2004) gives some suggestions for choosing the most suitable taxonomy: integrity, perspective, reform, simplicity and reporting. Bearing these suggestions in mind, TIMSS 2019 assessment framework is one of the most recent and simplest which classifies cognitive domain in three categories as knowing, applying and reasoning (Mullis and Martin (Eds.), 2017). Unlike Bloom’s taxonomy, TIMSS documents do not state that these domains are in a hierarchical order. However, many other frameworks including Marzano and Kendal’s one are in a hierarchical structure (Marzano and Kendal, 2008, p.16). Generally, cognitive skills under knowing are assumed as in the lowest level, and cognitive skills under reasoning (Table 1) are assumed as in the highest level (Bloom et al., 1956; Anderson and Krathwohl, 2001; Marzano and Kendal, 2008). According to Wilson (2000, p.5), ‘thinking skills’ are ambiguous as there is no

consensus among the taxonomies of different researchers as to what should be included in this category. But, most taxonomies include “collecting information, sorting information, analyzing information, drawing conclusions from the information, brainstorming new ideas, problem solving, determining cause and effect, evaluating options, planning and setting goals, monitoring progress, decision making, and reflecting on one’s own progress” and these are also mostly included in TIMSS 2019 assessment framework (Mullis and Martin (Eds.), 2017, p.22-24). Mastery of basics in school such as reading, writing, mathematics, science etc. “are not sufficient to meet the demands of the labor market and active citizenship” and ‘higher order’ thinking skills are needed (Wilson, 2000, p.6). Similarly, Mullis and Martin (Eds.) (2017, p.22) state that responding correctly to TIMSS items requires students to know the mathematics content being assessed and to make use of a range of cognitive skills.

There are some studies on the analysis of teacher made test items, central examination questions, curricular learning outcomes and book exercises based on item types, cognitive domains or content domains. For instance, Incikabi (2012) examined the distribution of eighth graders’ high stakes mathematics examination questions in terms of cognitive levels, content domains and item types based on TIMSS 2007 assessment framework. He reported that multiple choice item format is the only item type that is included in eighth grade high stakes exams contrary to TIMSS assessments which also includes constructed response items. He concluded that there is a resemblance considering content domain, but there are differences in cognitive domain distribution. Delil, Özcan and Işlak (2018) examined learning outcomes of 2017 Turkish primary school mathematics curriculum in terms of TIMSS 2015 assessment framework, and reported that while 50% were in number, 46% were in geometric shapes and measures, 4% in data display content domains; 58% were in knowing, 32% were in applying and the rest 10% were in reasoning cognitive domains. They concluded that for primary school students it is affirmative to have most of the learning outcomes to be at knowing cognitive domain but it is suggested to increase cognitive levels as grade level increases. Delil and Yolcu Tetik (2015) analyzed eighth grade high stakes mathematics examination items in terms of their cognitive demands, and they reported that 87% of the items are either in knowing or in applying cognitive domain. Taştekinöğlü and Aydın (2014) analyzed fourth grade teacher made mathematics tests based on TIMSS 2011 framework, and they found that, concerning content domains, of all 249 questions 80% were in numbers, 16% in geometrical shapes and measures, and 4% were in data display. But concerning cognitive domains, 67% were in knowing, 18% in applying and 15% were in reasoning. They suggested that fourth grade mathematics school exam panel should cooperate in preparing questions that require upper cognitive levels.

Being the closest exam to TIMSS fourth grade mathematics examination in terms of its grade level and content, analyzing fifth graders bursary examination mathematics items in terms of their cognitive levels, item types and content domains may shed light on why Turkish students have difficulty in doing mathematics in comparative studies like TIMSS. So, in this study the research problem is that how are the item types, and the content and cognitive domain requirements of the Turkish fifth grade bursary examinations’ mathematics items from year 1999 to year 2018 according to TIMSS 2019 assessment framework?

By this effort, it is aimed to observe the similarities and differences between the bursary examination mathematics items after and before the reforms in Turkish mathematics curriculum in terms of their cognitive and content requirements. Besides, there will be opportunity to compare content and cognitive demands of bursary examination items with those of TIMSS which may shed light on reasons why Turkish students show a performance below the average. There is no study comparing bursary examinations with TIMSS assessments framework in the literature, and besides the large amount of studies on alignment between Bloom’s cognitive standards and central exams, there are limited studies on alignment between TIMSS cognitive standards and central exams, and it is hoped that this study may contribute to close the gap in the literature as another significance of this study.

Method

As a qualitative research this is a descriptive study. Descriptive studies involve four consecutive steps (Yıldırım and Şimşek, 2016, p. 240): (1) forming a framework for the descriptive analysis, (2) processing the data in the thematic framework, (3) describing the findings, and (4) interpreting the findings. This study is defined as a descriptive one since the data were arranged and interpreted according to TIMSS 2019 assessment framework.

Selection of Items

In this study 520 Turkish fifth grade bursary examination items from 1999 to 2018 are analyzed. The items were obtained from the official website of Turkish Ministry of National Education. From years 1999 to 2002 there are 30 mathematics items for each year, but from 2003 to 2018 there are 25 mathematics items for each year in the examinations.

The Coding Scheme and Classification of the Items

Since all of the item types of the bursary examination are of multiple choice, and since the content domains of the items are also unambiguous, the only problem is to work out the cognitive demands of the items reliably. According to Neuendorf (2002, p. 142), there is an essential need for intercoder reliability, since it will be called expert analysis otherwise. So, two or more coders must be “calibrated” against one another. The coding scheme used in this study is derived from TIMSS 2019 assessment framework. The characteristics of cognitive processes in the cognitive domains are given in Table 1 (Mullis and Martin (Eds.), 2017, p.23-24).

To carry out the coding procedure, the coding panel (one the author and the other a mathematics education expert) working independently, coded the 25 mathematics items of an exam by using coding scheme of the study (see, Table 1). At first, the coders read *Mathematics Cognitive Domains* part of TIMSS 2019 Mathematics Framework (Mullis and Martin (Eds.), 2017) separately. Then, the coders came together and read the framework. They trained on some released fourth grade mathematics samples of TIMSS 2007 and TIMSS 2011. The disagreements were discussed until an agreement was reached. Then the items of the selected exam were coded separately by the coders. After the first coding, the intercoder reliability was found as 0.84 and this score is found large enough (Neuendorf, 2002). Each classification for which the coders did not agree was then discussed until an agreement was reached on how the classification would be made. Then the researchers, independently, coded the rest of 520 items. For each year’s exam items, some disagreements were occurred and disagreements were dissolved through a discussion based on the TIMSS 2019 assessment framework.

For evaluative purposes, it is assumed that cognitive skills under *knowing*, *applying* and *reasoning* are of lower, middle and higher-order level, respectively. In case if cognitive demands of an item belong to more than one cognitive domain, for this item, cognitive domain of the later in hierarchical order is chosen. Examples of classification of items are presented in Appendices.

Table 1: The coding scheme of the study

(1) KNOWING	(2) APPLYING	(3) REASONING
1.1 Recall Recall definitions, terminology, number properties, units of measurement, geometric properties, and notation (e.g., $a \times b = ab$, $a + a + a = 3a$).	2.1 Determine Determine efficient/appropriate operations, strategies, and tools for solving problems for which there are commonly used methods of solution.	3.1 Analyze Determine, describe, or use relationships among numbers, expressions, quantities, and shapes.
1.2 Recognize Recognize numbers, expressions, quantities, and shapes. Recognize entities that are mathematically		3.2 Integrate/Synthesize Link different elements of knowledge, related representations, and procedures to solve problems.

equivalent (e.g., equivalent familiar fractions, decimals, and percents; different orientations of simple geometric figures).

1.3 Classify/Order Classify numbers, expressions, quantities, and shapes by common properties.

1.4 Compute Carry out algorithmic procedures for $+$, $-$, \times , \div , or a combination of these with whole numbers, fractions, decimals, and integers. Carry out straightforward algebraic procedures.

1.5 Retrieve Retrieve information from graphs, tables, texts, or other sources.

1.6 Measure Use measuring instruments; and choose appropriate units of measurement.

2.2 Represent/Model Display data in tables or graphs; create equations, inequalities, geometric figures, or diagrams that model problem situations; and generate equivalent representations for a given mathematical entity or relationship.

2.3 Implement Implement strategies and operations to solve problems involving familiar mathematical concepts and procedures.

3.3 Evaluate Evaluate alternative problem solving strategies and solutions.

3.4 Draw Conclusions Make valid inferences on the basis of information and evidence.

3.5 Generalize Make statements that represent relationships in more general and more widely applicable terms.

3.6 Justify Provide mathematical arguments to support a strategy or solution.

SOURCE: IEA's Trends in International Mathematics and Science Study – TIMSS 2019 Copyright © 2019 International Association for the Evaluation of Educational Achievement (IEA). Publisher: TIMSS and PIRLS International Study Center, Lynch School of Education, Boston College.

Findings

The bursary examination mathematics items and TIMSS assessments are compared in terms of item format and presented in Table 2 below. While all of the bursary examination questions are of multiple choice, TIMSS 2019 target percentages are both multiple choice and constructed response with about the same proportions (Mullis and Martin, 2017).

Table 2: Percentages of item formats of bursary examination and TIMSS targets

Item Format	Bursary Examination	TIMSS 2019 Target
Multiple Choice	100	50
Constructed Response	0	50

The content domains of TIMSS 2019 assessment framework are given as *number, measurement and geometry*, and *data* with target percentages 50, 30 and 20 respectively. The content domains of bursary examinations are presented in Table 3. It is seen that of all bursary examination items, about 60% are in number, 33% percent are in measurement and geometry and 6% are in data content domains. In this table, since the first educational reform was in 2005, the second was in 2013 and the final one was in 2017, the years 1999-2005 are called pre reform years, 2006-2013 are called 1st reform years, and so on.

Table 3: Content domains of bursary examination mathematics items

Year	Number of Items in Content Domains			
	Number	Measurement and geometry	Data	Total
1999	20	8	2	30
2000	21	7	2	30
2001	22	7	1	30
2002	24	5	1	30
2003	17	7	1	25
2004	14	10	1	25
2005	15	9	1	25
1999-2005	133	53	9	195

(Pre reform %)	(68.2%)	(27.2%)	(4.6%)	
2006	13	10	2	25
2007	12	11	2	25
2008	12	10	3	25
2009	14	9	2	25
2010	11	11	3	25
2011	11	12	2	25
2012	16	8	1	25
2013	16	7	2	25
2006-2013	105	78	17	200
(1 st reform %)	(52.5 %)	(39 %)	(8.5 %)	
2014	15	9	1	25
2015	14	10	1	25
2016	15	8	2	25
2017	14	9	2	25
2014-2017	58	36	6	100
(2 nd reform %)	(58 %)	(36 %)	(6 %)	
2018	17	7	1	25
2018-2018	17	7	1	25
(3 rd reform %)	(68 %)	(28 %)	(4 %)	
1999-2018	313	174	33	520
(Total %)	(60.19%)	(33.46%)	(6.35%)	

Figure 1 illustrates the comparison of the fifth grade bursary examination items and TIMSS 2019 targets in terms of content domains. A *Chi-square test of goodness-of-fit* was performed to determine whether the content domains of all bursary examination items' percentages and TIMSS target percentages are consistent with each other or not. Evaluated value $X^2_{(df=2, N=200)} = 8.698$, $p < 0.05$ suggests that there is no consistency between TIMSS target percentages and bursary examination items in terms of content domains.

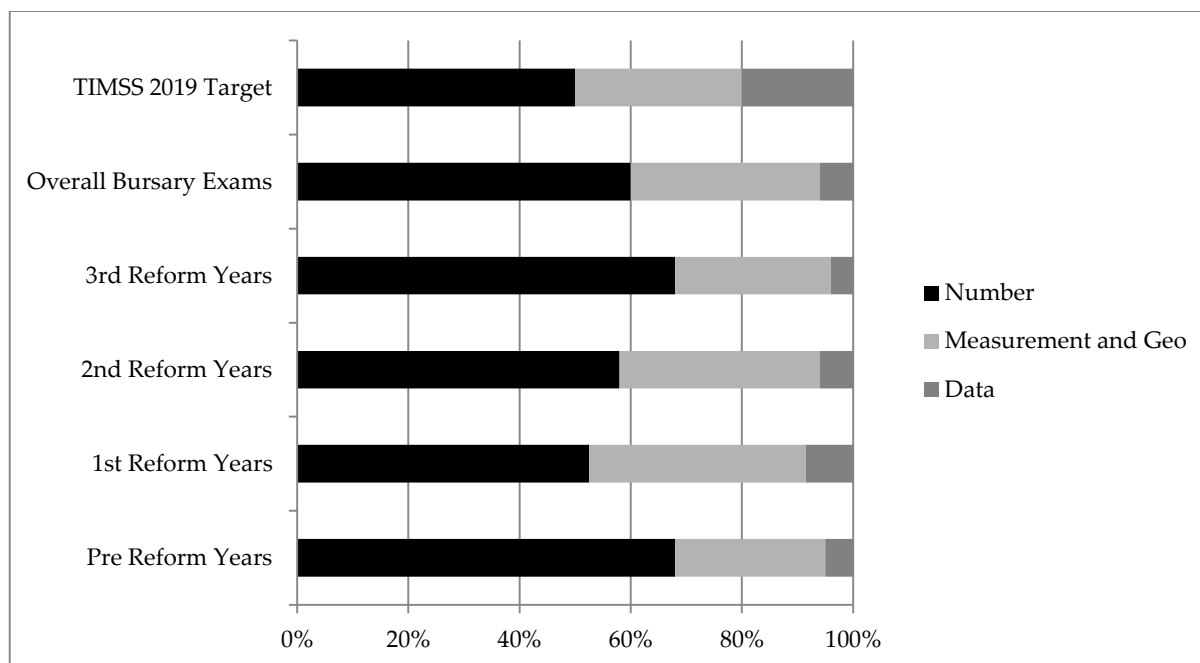


Figure 1: Percentages of content domains of TIMSS and bursary examination items

The cognitive requirements of all 520 items of fifth grade Turkish bursary examinations from 1999 to 2018 are summarized in Table 4 below. It is seen that the cognitive requirements of the mathematics items of the years 1999-2018 are in knowing, applying, and reasoning cognitive domains with percentages about 26, 38, and 36 respectively.

Table 4: Cognitive demands of bursary examination mathematics items

Year	Number of Items in Cognitive Domains			Total
	Knowing	Applying	Reasoning	
1999	5	14	11	30
2000	8	9	13	30
2001	6	14	10	30
2002	6	13	11	30
2003	3	11	11	25
2004	4	5	16	25
2005	8	5	12	25
1999-2005 (Pre Reform)	40 (20.5%)	71 (36.4%)	84 (43.1%)	195
2006	2	5	18	25
2007	1	14	10	25
2008	7	12	6	25
2009	4	12	9	25
2010	4	11	10	25
2011	3	15	7	25
2012	7	9	9	25
2013	8	13	4	25
2006-2013 (1 st Reform)	36 (18 %)	91 (45.5 %)	73 (36.5 %)	200
2014	6	14	5	25
2015	11	4	10	25
2016	14	8	3	25
2017	8	8	9	25
2014-2017 (2 nd Reform)	39 (39 %)	34 (34 %)	27 (27 %)	100
2018	18	4	3	25
2018-2018 (3 rd Reform)	18 (72 %)	4 (16 %)	3 (12 %)	25
1999-2018 (TOTAL %)	133 (25.58 %)	200 (38.46 %)	187 (35.96 %)	520 (100%)

The target percentages given by TIMSS 2019 framework for knowing, applying, and reasoning domains are 40, 40 and 20, respectively (Mullis and Martin, 2017, p.14). In the following, Figure 2 gives an overall summary of the results. By this figure, target percentages of TIMSS 2019, pre-reform, 1st reform, 2nd reform, 3rd reform and all mathematics items can be compared. Also, it is seen that in terms of distribution of cognitive requirements, the most that resembles with TIMSS target percentages is the 2nd reform period items with 39%, 34% and 27% in knowing, applying, and reasoning domains respectively.

A *Chi-square test of goodness-of-fit* was performed to determine whether the cognitive requirements of all bursary examination items' percentages and TIMSS 2019 target percentages are consistent with each other or not. Evaluated value $X^2_{(df=2, N=200)} = 7.592$, $p < 0.05$ suggests that there is no consistency between TIMSS

target percentages and bursary examination items in terms of cognitive demands. The only period that seems to resemble with TIMSS targets is the second reform years.

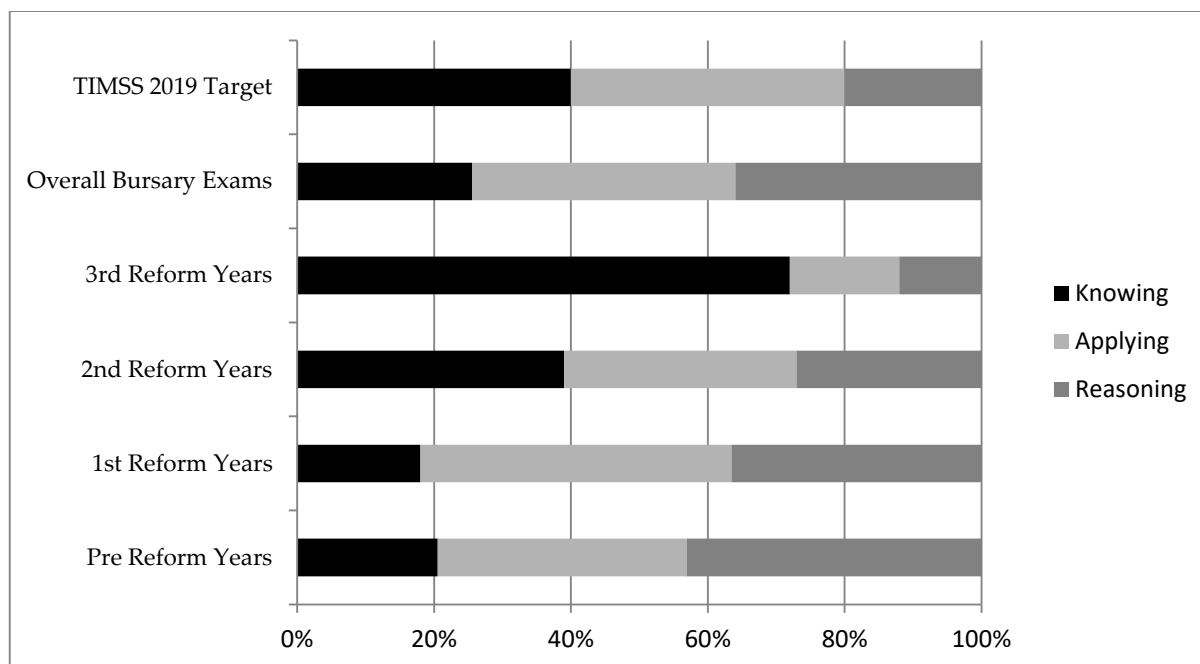


Figure 2. Percentages of cognitive domains of TIMSS and bursary examination items

Discussion and Recommendations

This study aims to give a broad picture of the fifth grade central examinations conducted in Turkey and to make a comparison with TIMSS assessments. Concerning item formats, findings reveal that the only item type that is included in bursary examination is *multiple choice* in contrast to TIMSS 2019 that also includes constructed response item type. National or international central exams exclusively use multiple choice item types due to its validity and reliability advantages. Nevertheless, being familiar with different kinds of item formats such as multiple choice, constructed response, essays and performance assessment tasks may have a big effect on student performance (Ben-Simon and Cohen, 2004). So, one reason of the low performance of Turkish students in international surveys like TIMSS may be the students' unfamiliarity with item types conducted, and such a result was also shared previously by Incikabi (2012, p.310).

When bursary examinations are compared with TIMSS in terms of content domains it is seen that of all fifth grade bursary examination mathematics items, 60% are in *number*, 33% are in *measurement and geometry* and 6% are in *data* content domains. Comparing with TIMSS content domain target percentages that are 50, 30 and 20 in number, measurement and geometry, and data respectively, while there is a resemblance in number and measurement and geometry, there is a small gap in data content domain.

The distribution of cognitive demands of overall Turkish fifth grade bursary examination mathematics items between the years 1999 and 2018 indicates that there is an emphasis on *applying* or *reasoning* cognitive domains, with a total percentage of 74. This result is inconsistent with the target percentages of TIMSS 2019 mathematics framework which shows that 60 percentages of TIMSS mathematics items require *applying* or *reasoning* in total, and it means that Turkish fifth grade bursary examination items require upper cognitive demands than those in TIMSS.

Starting from 2005, three major reforms took place in Turkish mathematics curriculum. One may ask if these reforms has brought any changes to the central examinations or not. Table 4 shows that, the pre reform and 1st reform examinations are similar in distribution of cognitive requirements of items: while about 20

percent in both require knowing the rest 80 percent require applying or reasoning. This result may suggest that the first reform did not bring much change considering cognitive requirements. But, while the 2nd reform is very much similar to TIMSS targets, after the final reform the cognitive demands of items that is allocated to knowing is raised up to 72 percent, leaving only 28 percent for applying and reasoning both. Actually, Figure 2 demonstrates a decrease in reasoning cognitive domain requirements gradually, and it has not been reported by bursary examination test developers whether this is a systematical approach or not. Literature reveal that there is not a coherence between TIMSS assessments and Turkish curricular outcomes, teacher made tests or high stakes examinations cognitive domain distributions, and this could be the reason behind the low performance of Turkish students in TIMSS assessments. So, it may be suggested that more emphasis should be given to applying and reasoning cognitive domain in the next bursary examinations in order to have a balance in cognitive domains that is closer to TIMSS targets.

A deeper analysis indicates that there are dramatic changes from year to year (see Table 4). For example, from 1999 to 2003 the numbers of items that require *applying* or *reasoning* are similar but from 2004 to 2006 it is seen that there is a strong emphasis on *reasoning* domain. Also, for years 2008, 2011, 2013 and 2014, it is seen that the item developers focused heavily on *applying* domain, but in 2018 they focused on *knowing* domain. These are all suggesting that, the test developers have not followed a cognitive framework for the central examinations that they have developed in Turkey. This result is consistent with the results of Delil and Yolcu Tetik (2015), in which they concluded that eight grade high stakes examinations had no cognitive framework followed. Test items need to reflect cognitive domains, as well as content domains. Also, the weight of cognitive domains should not change so dramatically from year to year. In this way, the trend of student achievement can be traced reliably. In fact, Turkish central examination or high-stakes test developers have not published any framework to show content or cognitive domains of the test items, but providing such information would very much help test developers, schools and examinees to plan their work more efficiently.

Acknowledgements

The author wishes to extend his gratitude to H. Delil for his illuminative discussions during analysis and cognitive domain classifications of the items.

Appendices

Examples for classifications of Turkish fifth grade bursary examination items (Appendix A) and the fourth grade TIMSS 2007 released items (Appendix B) are given below. While the classifications of the bursary examination items were done by the coding panel of this research, the classifications of the TIMSS 2007 released items were done by TIMSS experts.

Appendix A

Example 1

Problem #10 in Bursary Examination 1999:

What is the result of the operations $\left(\frac{3}{8} : \frac{1}{8}\right) - \left(2 - \frac{1}{4}\right)$?

- (A) $\frac{1}{2}$ (B) $\frac{5}{4}$ (C) $\frac{32}{25}$ (D) $\frac{11}{8}$

Coding: *knowing*

Example 2

Problem #2 in Bursary Examination 1999:

How many elements have a set of having a total of 16 subsets?

- (A) 1 (B) 2 (C) 3 (D) 4

Coding: *applying*

Example 3:

Problem #3 in Bursary Examination 1999:

A load is planned to be carried equally by 6 people. However, since 2 persons do not attend, each of the rest has to carry 15 kilograms more. How many kilograms is the load?

- (A) 90 (B) 120 (C) 180 (D) 240

Coding: *reasoning*

Appendix B

Example 1:

TIMSS-2007 Released Items (Page 24, ID: M041014)

In which of the following are the numbers arranged from LARGEST to SMALLEST?

- (A) 36, 43, 66, 87 (B) 66, 43, 36, 87 (C) 87, 66, 36, 43 (D) 87, 66, 43, 36

Cognitive Domain: *knowing* (TIMSS 2007 Released Items, 2009).

Example 2:

TIMSS-2007 Released Items (Page: 17, ID: M031301)

Al wanted to find how much his cat weighed. He weighed himself and noted that the scale read 57 kg. He then stepped on the scale holding his cat and found that it read 62 kg.

What was the weight of the cat in kilograms?

Answer: ----- kilograms

Cognitive Domain: *applying* (TIMSS 2007 Released Items, 2009).

Example 3:

TIMSS-2007 Released Items (Page: 46, ID: M031235)

A group of 8 children have 74 sweets altogether. How many more sweets are needed for the children to be able to share them equally?

Answer: -----

Cognitive Domain: *reasoning* (TIMSS 2007 Released Items, 2009).

REFERENCES

- Anderson, L.W. (Ed.), Krathwohl, D.R. (Ed.), Airasian, P.W., Cruikshank, K.A., Mayer, R.E., Pintrich, P.R., Raths, J. and Wittrock, M.C. (2001). *A taxonomy for learning, teaching, and assessing: A Revision of Bloom's Taxonomy of Educational Objectives*. (Complete edition). New York: Longman.
- Ben-Simon, A. and Cohen, Y. (2004). International assessments: Merits and pitfalls. Proceedings of the 30th Annual Conference of the International Association for Educational Assessment (IAEA). Philadelphia, June 2004.
- Bloom, B.S. (Ed.), Engelhart, M.D., Furst, E.J., Hill, W.H. and Krathwohl, D.R. (1956). *Taxonomy of Educational Objectives, the classification of educational goals – Handbook I: Cognitive Domain*. New York: David McKay.
- Çelik, S., Kul, Ü. and Çalık Uzun, S. (2018). *Using Bloom's revised taxonomy to analyse learning outcomes in mathematics curriculum*. Abant İzzet Baysal Üniversitesi Eğitim Fakültesi Dergisi, 18 (2), 775-795.
- Delil, A. and Yolcu Tetik, B. (2015). *An analysis of Turkish eighth grade high stakes mathematics examination questions based on TIMSS-2015 framework*. MCBU Sosyal Bilimler Dergisi, Cilt:13, S.4, Doi Number: 10.18026/cbusos.87313.
- Delil, A., Özcan, B.N. and Işlak, O. (2018). Yeni İlkokul Matematik Dersi Öğretim Programı Kazanımlarının TIMSS-2015 Çerçevesine Göre Değerlendirilmesi. (An Analysis of the New Turkish Mathematics Curriculum Acquisitions in Terms of TIMSS-2015 Assessment Frameworks). VIII. International Congress on Research in Education. 9-11 May 2018, Manisa.
- Delil, H. (2006). *An Analysis of Geometry Problems in 6-8 Grades Turkish Mathematics Textbooks*. (Master thesis). Middle East Technical University, Ankara.
- EMBE (2019). MEB İOKBS Başvuru ve Uygulama e-Kılavuzu (Ministry of National Education Elementary and Middle-school Bursary Examinations E-guide). Retrieved on 15.03.2019 from http://www.meb.gov.tr/meb_iys_dosyalar/2019_03/11110605_bursluluk_ekilavuz_2019_.pdf
- Incikabi, L. (2012). *After the reform in Turkey: A content analysis of SBS and TIMSS assessment in terms of mathematics content, cognitive domains, and item types*. Education as Change, 16:2, 301-312.
- Marzano, R.J. and Kendall, J.S. (2008). *Designing and Assessing Educational Objectives: Applying the New Taxonomy*. Corwin Press.
- MONE (2016). TIMSS 2015 Ulusal Matematik ve Fen Ön Raporu: 4. ve 8. Sınıflar (TIMSS 2015 National Mathematics and Science Report: 4th and 8th graders). Ministry of National Education, Ankara.
- Mullis, I.V.S., and Martin, M.O. (Eds.) (2017). TIMSS 2019 Assessment Frameworks. Retrieved from Boston College, TIMSS and PIRLS International Study Center website: <http://timssandpirls.bc.edu/timss2019/frameworks/>
- Neuendorf, K.A. (2002). *The content analysis guidebook*. Sage Publications.
- Nitko, A. J. (2004). *Educational assessment of students*. Pearson Merrill-Hill, NJ: Upper Saddle Rive.
- Taştekinöğlü, E. and Aydın, G. (2014). 4. Sınıf Matematik Sınav Sorularının TIMSS 2011 Bilişsel Alanları ve Öğretim Programlarıyla Karşılaştırılması. (A Comparison of Fourth Grade Mathematics Examinations with TIMSS 2011 Cognitive Domains and Mathematics Curricula). III. Türkiye Lisansüstü Çalışmalar Kongresi Bildiriler Kitabı - I (Sosyoloji - İlahiyat - Eğitim), 15-18 Mayıs 2014, Sakarya. Retrieved on 07.02.2019 from http://tlck.org.tr/wp-content/uploads/2018/09/III_TLCK_1_kitap_BASKI.pdf
- TIMSS (2009). TIMSS-2007 Released Items. Retrieved on 19.05.2012 from <http://timss.bc.edu/timss2007/items.html>
- TIMSS (2013). TIMSS-2011 Released Items. Retrieved on 19.02.2019 from <https://timssandpirls.bc.edu/timss2011/international-released-items.html>
- Wilson, V. (2000). *Can thinking skills be taught?* Retrieved on 19.08.2013 from http://www.peopledev.co.za/library/can_think_skill_be_taught.pdf%E2%80%8E

- Woessmann, L. (2002). How Central Exams Affect Educational Achievement: International Evidence from TIMSS and TIMSS-Repeat. A paper prepared for the conference Taking Account of Accountability: Assessing Politics and Policy. John F. Kennedy School of Government Harvard University (June 10 – 11, 2002).
- Woessmann, L. (2003). Central Exams as the "Currency" of School Systems: International Evidence on the Complementarity of School Autonomy and Central Exams, CESifo DICE Report, ISSN 1613-6373, ifo Institut für Wirtschaftsforschung an der Universität München, München, Vol. 01, Iss. 4, pp. 46-56.