



**HAL**  
open science

# The inconsistency between curricular expectations and assessment results

Hülya Kılıç, Oğuzhan Doğan, Ayşegül Kılıç

► **To cite this version:**

Hülya Kılıç, Oğuzhan Doğan, Ayşegül Kılıç. The inconsistency between curricular expectations and assessment results. Thirteenth Congress of the European Society for Research in Mathematics Education (CERME13), Alfréd Rényi Institute of Mathematics; Eötvös Loránd University of Budapest, Jul 2023, Budapest, Hungary. hal-04413530

**HAL Id: hal-04413530**

**<https://hal.science/hal-04413530>**

Submitted on 23 Jan 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# The inconsistency between curricular expectations and assessment results

Hülya Kılıç, Oğuzhan Doğan and Ayşegül Kılıç

Yeditepe University, Türkiye; [hulya.kilic@yeditepe.edu.tr](mailto:hulya.kilic@yeditepe.edu.tr)

*To pursue high-ranked lycées in Türkiye, many eight grade students study for High School Entrance Exam which takes place at the end of the middle school. The exam consists of 20 mathematics items aligned with middle school math curriculum. In recent years, the exam has aimed to assess students' mathematical thinking skills besides their mathematical knowledge. However, the mean scores obtained on the test are below the curricular expectations. Therefore, we wonder how enacted curriculum support eight grade students' studies for the exam. We prepared a math exam consist of 8 items and handed it to 127 students from different schools. We analyzed 68 students' papers who got minimum score 4 out of 40. The analysis revealed that as aligned with national and international assessment results, students could answer procedural questions. Still, they fail to answer questions about reasoning, connections and problem solving skills.*

*Keywords: Mathematical thinking skills, middle school, curriculum, national and international assessment.*

## Introduction

Curriculum reforms are triggered by changes in the philosophy of current education system, needs of society, development of science and technology and other local or global issues that impact other systems and people's lives such as pandemic, wars and economic crises. As having four components, namely, objectives, content, method and assessment, any inconsistency between the components of a curriculum causes a failure in achieving desired learning outcomes (Özmantar et al., 2018). Students' performances in national and international standardized tests can be considered as a tool to check whether curricular expectations are met or not for the particular level of education or subject matter (Breakspear, 2012). In Türkiye, the first official middle school curriculum was launched in 1924 and it was revised 14 times till 2018 (Özmantar et al., 2018). More rapid changes took place after 2000s such that the mathematics curriculum was revised 7 times since then. Although technological changes and social needs are the major factors for such reforms, Turkish students' poor mathematics performance in both national and international tests is one of the factors behind the changes in recent years (MoNE, 2019).

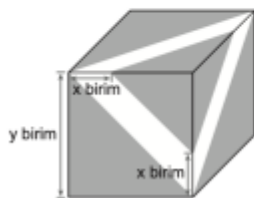
The latest curriculum reform in all grade levels was made in 2018 in Türkiye. In the same year, the format of High School Entrance (HSE) Exam was changed. Previously, the math items were designed to assess students' understanding and application of simple procedures but nowadays, students' mathematical thinking skills (reasoning, connections, problem solving, etc.) are aimed to be assessed via contextualized problems (Kedikli & Katrancı, 2022). Although fostering students' mathematical thinking skills was one of the major goals of mathematics curriculum since 1920s, such skills were not specifically addressed in standardized multiple-choice exams till 2018. As noted in 2023 *Education Vision* report of the MoNE, scores obtained from recent PISA and TIMSS results forced the national authorities to make changes in the standardized tests format (MoNE, 2019).

In both PISA and TIMSS not only students' procedural or algorithmic knowledge is measured but also their reasoning, connections and problem solving skills are assessed. Therefore, national authorities decided to modify math questions of HSE exam to address both curricular goals and assessment criteria of TIMSS and PISA by allocating more items in the exam to assess students' mathematical thinking skills (MoNE, 2019). As an initial step, modified versions of some released PISA or TIMSS items are beginning to be used in the latest exams. For instance, *Transit System* problem which was asked in 2003 PISA (OECD, 2003) was modified and asked in HSE exam in 2020.

In this paper, we identify procedural knowledge as applying algorithmic procedures to solve given simple problems or doing calculations and mathematical thinking skills as ability to explore, justify and make reasoning about given situations, understand, use and relate multiple representations, understand and solve non-routine or complex problems (Stein et al., 2007). Moreover, mathematical thinking skills are defined as a composition of problem solving, reasoning, connections and communication skills in Turkish math curriculum (MoNE, 2018). For instance, in 2018 HSE exam, the first question illustrated below entails procedural knowledge while the second question entails use of connection skills. In the first question the student will rewrite given number in scientific notation but in the second question student will use knowledge of area of a square, algebraic expressions and binomial expansion.

**Question 1:** If  $0.00013 \cdot 10^a$  is greater than 1000, determine the minimum integer value of  $a$ .

**Question 2:** (Question is shortened) If white-color strips are stacked to all faces of the cube as follows and the remaining parts are painted, determine the area of painted region.



**Figure 1: Sample questions from 2018 HSE exam**

Although HSE exam was reformed in 2018 by illustrating items to assess students' thinking skills, students still performed poorly in mathematics as in the previous years. In years between 2018 and 2022, students' average in math test out of 20 questions were as follows: 6.99, 5.09, 9.36, 7.56 and 4.74 (e.g., MoNE, 2022). It seemed that students performed well in 2020 but it was because of pandemic such that students were asked questions related to topics covered during the first semester of the 8<sup>th</sup> grade. The lower scores in 2018 might not be surprising since items aimed to measure students' mathematical thinking were new for both students and teachers. Even though some sample items have been released by MoNE in monthly base via official website ([odsgm.meb.gov.tr](http://odsgm.meb.gov.tr)) for each academic year since then, the studies revealed that teachers were not supported enough how to implement new curriculum nor how to guide students for HSE exam when the new curriculum was launched (Erden, 2000). Because very limited and short-term professional development programs were provided for teachers, they did not feel confident about how to modify their teaching practices to support students' mathematical thinking skills (Kertil et al., 2021). Teachers were informed about

what has changed in the curriculum in the form of a meeting organized by local authorities and then they were suggested to attend professional development programs according to their needs provided on-site or online. In other words, teachers were not given a compulsory workshop about how to support students' mathematical thinking skills, the volunteers participated in such workshops. In Türkiye, although teachers are able to choose a professional development program or participate in working groups provided by MoNE-Teacher Network Platform (oba.gov.tr), such programs are still short-term programs. Even some new textbooks were written in line with new curriculum and HSE exam context, teachers found that compulsory textbooks used in the schools did not cover enough samples, therefore they had to recommend some other additional textbooks for practice for students (Kertil et al., 2021). Despite of having limited resources for adapting new curriculum into lessons, we still expect both teachers and students have learned about the new format of HSE exam and perform better in the exam in each successive year. However, it was not the case such that students obtained the lowest score in 2022 HSE exam.

In this paper, we want to explore how 8<sup>th</sup> grade students perform on problems entails both procedural knowledge and mathematical thinking skills via HSE-exam like questions. We want to use our findings as a basis for a large-scale study in which we investigate how to foster middle grade students' mathematical thinking skills and their performance in national and international exams.

## Methodology

### Participants

In this study we used convenient sampling such that we asked our collaboration schools and former graduates of our program whether they became volunteer to implement the math exam prepared by us (the researchers) to their eighth graders. Totally, we applied the exam to 127 students (67 female, 60 male) from 6 different schools. Four of the schools were public schools and two of the schools were private schools.

### Data collection and analysis

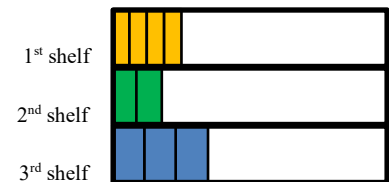
We prepared a math exam in line with middle school math objectives that 8<sup>th</sup> grade students had already covered. The test consisted of 8 items and graded out of 40 points. The points assigned for questions varied between 4 points to 6 points. In terms of TIMSS item-type classification, Question 1 can be classified as knowing, Question 6 and 7 as reasoning and the rest can be thought as applying type of questions. More specifically, Questions 2, 3 and 4 entail use of connections skills besides procedural knowledge while Questions 5 and 8 require problem solving. Additionally, pictures and tables are illustrated in questions 2, 4, 5 and 6 such that students need to use the given table in question 2 and the picture in question 6 to solve the problem. Moreover, the type of the problems in the test aligned with the types in the latest HSE exams (Kedikli & Katrancı, 2022) except our questions were not in multiple-choice form but students need to show their calculations and explain their answers. Some sample questions with their rubrics for scoring and explanation of related skills are given below.

**Question 4.** To keep documents in an office, different size (width) and coloured folders are used. The folders can be arranged without leaving spaces on a shelf. By using the following information determine the size of each folder. (5p)

- The length of a shelf is 90 cm.
- When 15 folders are put on the 2<sup>nd</sup> shelf, they occupy the half of the shelf.

- Total width of three folders put on the 2<sup>nd</sup> shelf equals to total width of 2 folders put on the 3<sup>rd</sup> shelf.
- When all shelves are full then the number of folders in the first shelf is 5 more than twice of the folders in the third shelf.

**Rubric:** 1) Finding the width of the folder on the 2<sup>nd</sup> shelf (2p). Recognizing that there are  $15 \cdot 2 = 30$  folders in total. Then size of one folder is  $90 \div 30 = 3$  cm (*Procedural knowledge*). 2) Finding the width of the folder on the 3<sup>rd</sup> shelf (1p). Using given information that  $3 \cdot 3 = 2 \cdot x$  then  $x = 4.5$  cm (*Procedural knowledge and Connection skill: Converting verbal expressions into mathematical operations*) 3) Finding the width of the folders on the 1<sup>st</sup> shelf (2p). Using given information to find number of folders in the 1<sup>st</sup> row and then determine its size.  $90 \div 4.5 = 20$  folders on the 3<sup>rd</sup> shelf (*Procedural knowledge*). Then the number of folders on the 1<sup>st</sup> shelf is  $20 \cdot 2 + 5 = 45$ . Finally, the width of the folder is  $90 \div 45 = 2$  cm. (*Connection skills: Setting up relationship between the number of folders, the length of the shelf and size of the folder*)



**Question 7.** In an amusement park, you can get either single coin to use a single vehicle or decide how many vehicles you will use and purchase a pre-paid Fun-card for using the vehicles. The prices for each option are as follows:

Each coin: 5 TL

Fun-card: 10 TL for the card and additional 3 TL for each vehicle to be used

If Ali decides to go to the amusement park during the weekend, which offer should he choose? Why? (5p)

**Rubric:** Try different number of vehicles by showing calculations (2p) (*Problem solving and reasoning: thinking about different cases to make decision*) and note that if Ali uses fewer than 5 vehicles then he should buy coins, if he decides to use 5 vehicles, then he can choose either of them, if he wants to use more than 5 vehicles then he should buy the card (1p for each). If students just estimate a number and make their decision accordingly then they deserve 2 points (*problem solving: using trial and error strategy*). An answer without explanation or justification no points.

The split-half reliability of the test was calculated as .87. Two researchers graded all the tests in line with the rubric. We achieved full agreement on grading the students' work. Although we applied the test 127 students, we decided to analyze 68 students' (39 female, 29 male) paper who obtained at least 4 points in the test.

## Results

We initially analyzed all data collected from 127 students to determine the cut off score for further analysis since there were many students who received no credit from the test. Since the median of the scores was 4, we decided to analyze papers of the students who received 4 points at minimum. The mean and standard deviation of each item for 68 students is shown in Table 1.

**Table 1: Descriptive analysis of achievement test**

	Mean	Std. dev.	Max.	Min.
<b>Q1</b>	3.15	1.39	4	0
<b>Q2</b>	1.06	1.91	5	0
<b>Q3</b>	3.15	1.69	5	0
<b>Q4</b>	2.60	2.17	5	0
<b>Q5</b>	2.44	2.77	6	0
<b>Q6</b>	0.34	1.02	4	0
<b>Q7</b>	1.99	1.73	5	0
<b>Q8</b>	0.43	1.34	6	0
<b>Total</b>	15.15	8.56	34	4

To understand in which questions students performed well we looked the frequencies of the items that students received full credit, partial credit or no credit as shown in Table 2.

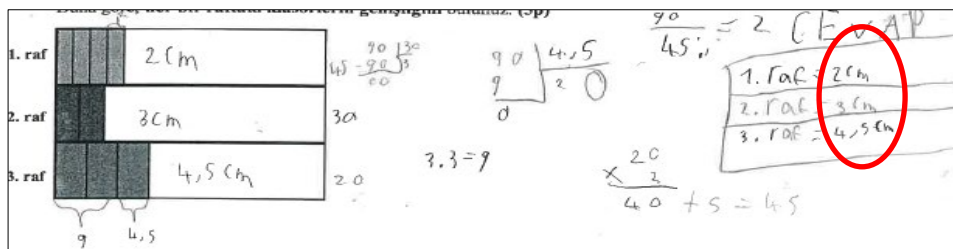
**Table 2: Frequencies of distribution of scores per item in the test**

	Full answer	Partial answer	No or wrong answer
Q1	43 (63%)	17 (25%)	8 (12%)
Q2	11 (16%)	8 (12%)	49 (72%)
Q3	9 (13%)	47 (69%)	12 (18%)
Q4	23 (34%)	22 (32%)	23 (34%)
Q5	22 (32%)	12 (18%)	34 (50%)
Q6	4 (6%)	4 (6%)	60 (88%)
Q7	10 (15%)	38 (56%)	20 (29%)
Q8	3 (4%)	7 (10%)	58 (85%)

*Note: Due to rounding up numbers, total may not be 100% in some of the cells.*

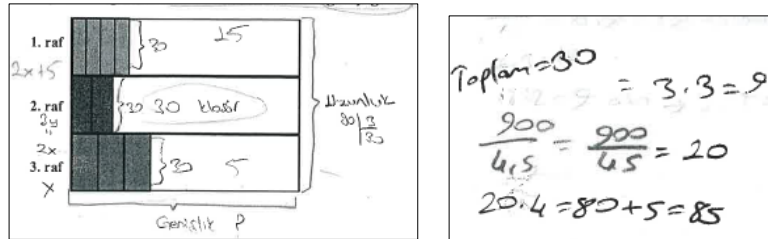
As seen in Table 2, the students' performance on some of the questions highly varied. Therefore, we decided to analyze whether students' performance differed in terms of question types. We found significant differences when we compared students' performances in procedural questions (Q2, Q3 and Q4) versus problem solving (Q5 and Q8) ( $t_{67,.05}=6.902, p=.000$ ) and procedural questions versus reasoning (Q6 and Q7) ( $t_{67,.05}=7.660, p=.000$ ). Obviously, students' performance in knowledge question (Q1) was significantly higher than the rest of the test ( $t_{67,.05}=9.854, p=.000$ ). However, students' performance in problem solving and reasoning questions was poor and did not vary significantly ( $t_{67,.05}=.199, p=.843$ ). We want to illustrate some student work that students partially or fully answered in the test.

As seen in Table 2, approximately, one third of the students could not answer Question 4 while one third answered it partially. As shown in figure 2, one of the students found the sizes of the folders correctly as 2 cm, 3 cm and 4.5 cm, respectively by applying appropriate procedures, doing necessary connections in terms of length of a shelf and sizes of folders and utilizing correct calculations.



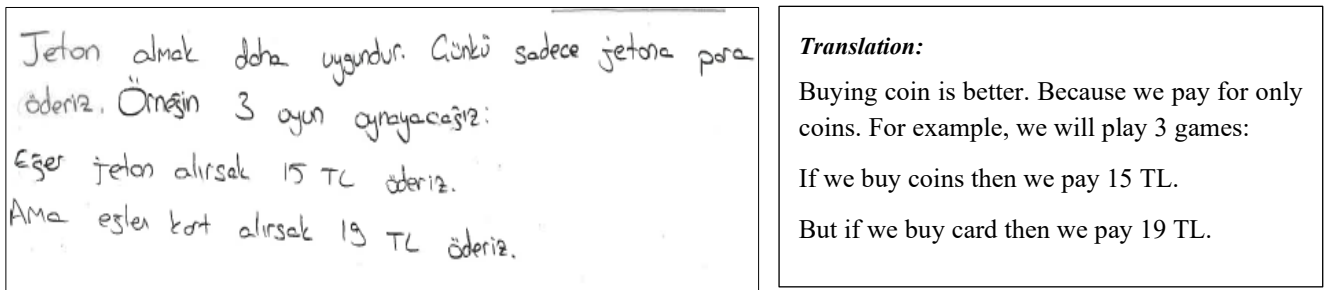
**Figure 2: A student's full answer for Question 4**

However, some of the students misinterpreted the given information and set up wrong relationships. As shown in figure 3, one of the students (on the left) confused the terms “length” and “width” such that he thought the vertical axis as the length of the row while another student (on the right) mistranslated given statement into mathematics. Therefore, they both could not finalize their solution.

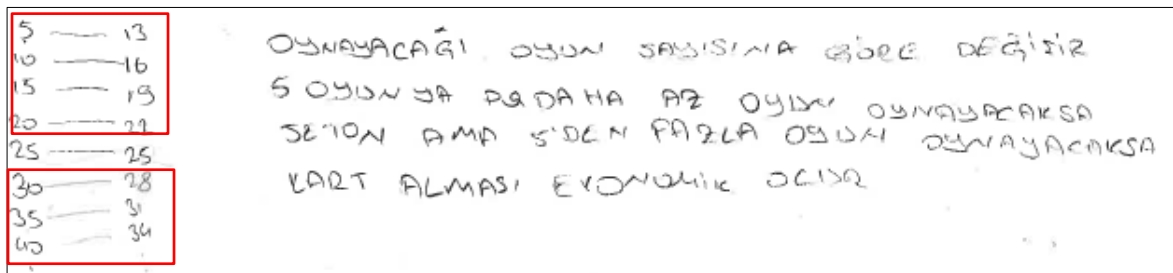


**Figure 3: Some students’ partial answer for Question 4**

As explained in previous section, in Question 7 we want students to think about all possibilities for given situation and make a decision accordingly. However, as shown in figure 4 students just thought about a few cases to decide which offer will be better. The student used trial and error strategy to find an appropriate answer, which was thought as an indicator of immature problem solving skill. A few students thought all possible cases and gave full answer as shown in figure 5.



**Figure 4: A student’s partial answer for Question 7**



**Figure 5: A student’s full answer for Question 7**

As seen in the students’ papers, some students attempted to utilize mathematical operations to solve given problems but they failed to finalize their solutions since they failed to understand given problem statements or interpret given tables or charts correctly. That is, they had some procedural knowledge about mathematics but they lacked of making connections between multiple representations, reasoning and problem solving skills.

## Discussion

In this study, we tried to explore how students performed on a mathematics test which consisted of items aiming to assess students' knowledge as well as their mathematical thinking skills. Although our sample size was limited, some of our findings were compatible with findings of large-scale national and international studies.

The students in our sample were better in answering algorithmic questions but they failed to answer questions entails problem solving and reasoning. The analysis of 8<sup>th</sup> grade students' performance in previous HSE and TIMSS exams support this finding (e.g., MoNE 2022; Mullis et al., 2021). TIMSS 2019 report revealed that Turkish students' mean score in problem solving questions was even lower than their overall mean (Mullis et al, 2021). Even though improving students' mathematical knowledge and thinking skills is emphasized in the curriculum, students' poor performance in national and international tests raise a question about the difference between official and enacted curriculum. Indeed, some teachers complained about having no or a few supports about how to improve students' mathematical thinking skills in the lesson (e.g. Erden, 2020). They noted that they need a long-term professional development about how to enact the new curriculum in the school otherwise, they continue in teaching their own ways (e.g. Kertil et al., 2021). Moreover, they stated that teaching resources in terms of textbooks, teacher guides and materials should be available to pursue and attain the learning goals of the curriculum (e.g. Kertil et al., 2021). Whenever the gap between official and enacted curriculum gets larger, it will not be surprising to get similar student performance in the exams (Stein et al, 2007) as we observed in our small-scale study. Therefore, whenever major revisions are made in the curriculum of a specific subject or grade level, all responsible teachers should be given a long-term and sustainable support to implement and adapt new curriculum into their teaching practices.

Even though students performed better in algorithmic problems, their performance was still below our expectations in relation to curricular goals. As noted above, some of the students partially answered or failed to answer some procedural questions, e.g. question 4, because they failed to understand the problem statement or illustrations (charts, pictures, etc.). Since reading and understanding mathematical language and representations is related to mathematical literacy as well as mathematical connection and communication skills, as seen in the results, students' mathematical literacy skills were also poor as noted in recent PISA reports (OECD, 2019).

Briefly, the main purpose of curriculum reform is to attain and improve the quality in education. However, desired outcomes may not be achieved unless the reforms are addressed during the enactment. Although we can achieve a perfect alignment between the written objectives and the assessment items in standardized tests, we may still obtain low scores in the tests unless we do required changes in our teaching and learning methods or resources. This seems to be one of the reasons behind Turkish students' failure in items require mathematical thinking skills in standardized tests. They do not know or are not used to discuss how mathematical concepts are related to each other or other disciplines or how mathematical fact is represented in different ways or what the problem solving strategies other than calculating or solving equations are (Erden, 2020). Therefore, mathematics teachers should be provided with enough professional support and resources about how to improve students' mathematical thinking skills in every stage of middle school and students should



be given more opportunities to use mathematics in their daily lives both in schools and at their homes to achieve an increase in students' performance in national and international standardized tests.

## References

- Breakspear, S. (2012). *The policy impact of PISA: An exploration of the normative effects of international benchmarking in school system performance*. OECD Education Working Papers, No. 71, OECD Publishing. <https://doi.org/10.1787/5k9fdfqffr28-en>
- Erden, B. (2020). Teachers' views related to skill-based questions in Turkish, mathematics and science lessons. *Academia Eğitim Araştırmaları Dergisi*, 5(2), 270–292.
- Kedikli, D., & Katrancı, Y. (2022). Analysis of skills-based middle school mathematics questions. *Journal of Education for Life*, 36(3), 673-696. <https://doi.org/10.33308/26674874.2022363446>
- Kertil, M., Gulbagci-Dede, H., & Ulusoy, E. G. (2021). Skill-based mathematics questions: What do middle school mathematics teachers think about and how do they implement them? *Turkish Journal of Computer and Mathematics Education*, 12(1), 151–186. <https://doi.org/10.16949/turkbilmat.774651>
- Ministry of National Education [MoNE] (2018). *Ortaokul Matematik Öğretim Programı* [Middle School Mathematics Curriculum]. MEB.
- Ministry of National Education [MoNE] (2019). *2023 Eğitim Vizyonu* [2023 Education Vision]. MEB.
- Ministry of National Education [MoNE] (2022). *2022 High School Entrance Exam*. MEB. [https://cdn.eba.gov.tr/icerik/2022/06/2022\\_LGS\\_rapor.pdf](https://cdn.eba.gov.tr/icerik/2022/06/2022_LGS_rapor.pdf)
- Mullis, I. V. S., Martin, M. O., Fishbein, B., Foy, P., & Moncaleano, S. (2021). *Findings from the TIMSS 2019 Problem solving and inquiry tasks*. Retrieved from Boston College, TIMSS & PIRLS International Study Center. <https://timssandpirls.bc.edu/timss2019/psi/>
- OECD (2003). *PISA Problem solving items and scoring guides*. Retrieved from U.S. Department of Education. [https://nces.ed.gov/surveys/pisa/pdf/items2\\_solving.pdf](https://nces.ed.gov/surveys/pisa/pdf/items2_solving.pdf)
- OECD (2019). *PISA 2018 results (Volume I): What students know and can do*, PISA, OECD Publishing. <https://doi.org/10.1787/5f07c754-en>
- Özmantar, M. F., Akkoç, H., Kuşdemir-Kayıran, B., & Özyurt, M. (Eds.). (2018). *Ortaokul matematik öğretim programları: Tarihsel bir inceleme* [History of middle school mathematics curricula]. Pegem Akademi.
- Stein, M. K., Remillard, J., & Smith, M. S. (2007). How curriculum influences student learning. In F. K. Lester, Jr. (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 319–369). IAP.