

Linking Mathematics TIMSS Achievement to National Examination Scores and School Marks: Unexpected Gender Differences in Slovenia

Barbara Japelj Pavešić

Educational Research Institute, Ljubljana, Slovenia

Gašper Cankar

National Examinations Centre, Ljubljana, Slovenia

Abstract: In the article, we present the results of the Slovene national study of three different assessments of mathematics for students in Grade 8 and Grade 13: the independent TIMSS or TIMSS Advanced outcomes, the national external examination scores and internal teacher's marks. Grade 8 students who participated in TIMSS also took the national assessment (NA) one year after TIMSS assessment; TIMSS Advanced math students took the Matura examination from mathematics two months after the TIMSS Advanced assessments. The data on school marks from mathematics were collected with the nationally added questions to the international TIMSS and TIMSS Advanced questionnaires for students, together with the series of questions about the effort put into solving the test. One year later, the outcomes from TIMSS assessment, national examinations (Grade 9 and Grade 13) and school grades for each student were linked and the differences between boys and girls, attitudes toward mathematics and plans for future education were analysed. It was found that gender differences at scores from national exams as well as in school marks differ from gender differences in TIMSS and TIMSS Advanced achievement (i.e. Grade 13 students' Matura results are slightly in favour of girls while TIMSS Advanced show better achievement for boys). Comparison of three outcomes reveal some characteristics of both national examinations and teachers' marking not evident otherwise. Matura gives to the most able students proportionally less opportunity to demonstrate the highest cognitive level of knowledge. Boys who demonstrated the same knowledge in TIMSS as girls get lower national marks as girls, in exams and by teachers. Girls put less effort than boys in solving the TIMSS test which could help to explain the changing gender gap from TIMSS to the national examinations. In Grade 8, the marks and TIMSS scores also show inconsistencies on student level. They are differently associated with attitudes toward mathematics which can provide some ideas for improvement of low motivation for learning mathematics in Slovenia.

Keywords: TIMSS; mathematics; school grades; national examination; gender difference

In Slovenia, the central database of the national examination results and participation in the international assessments of mathematics enabled the national study of links between international and national measurements of knowledge of mathematics aimed for better understanding of different gaps in achievement between specific groups of students in both measurements. TIMSS assessed the representative sample of all students at the penultimate grade of elementary school, at Grade 8, and TIMSS Advanced, at the end of general upper secondary school, at Grade 13 from

<https://doi.org/10.14712/23363177.2018.294>

www.orbisscholae.cz

© 2018 The Authors. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>).

78 mathematics and collected also the rich set of background data about learning and teaching mathematics, every four years from 1995 to 2015.

All Slovene students in Grade 9 take the national external assessment (NA) from mathematics, Slovene language and one of other school subjects that varies between years and schools. NA scores are used as secondary criteria for admission to the upper secondary schools with limited number of accepted candidates, which are usually highly demanding general gymnasia. All students at the end of Grade 13 in general upper secondary schools, gymnasia, take compulsory national external mathematics examination, Matura. Passing the examination from mathematics and additional 4 subjects is required for admission into any academic university study and scores are used as criteria for the most elite university studies, such as medicine, biological sciences and law with limited numbers of places for new students. Besides the assessments, all students are given marks for mathematics knowledge internally by their school teachers. Teachers' marks from elementary school are used as the main criteria for student's admission into upper secondary schools and marks from those schools are used as the secondary criteria for students' admission to some elite university studies. There is no central database of teachers' internal marks and no specific national study of marking.

The main reason for the study were unexplained gender differences in mathematics achievement that are not consistent across all measurements of knowledge. In Grade 13, the independently measured mathematics achievement from TIMSS Advanced 2008 and 2015 assessments is higher for boys while Matura scores are traditionally higher for girls. In Grade 8, TIMSS achievement didn't differ among boys and girls in last 20 years and same is true for results from the NA, but school marks are still higher for girls. Comparisons within elementary schools (Grades 1–9) also show unexplained trends of increased frequency of the highest marks. In view of the importance of assessment results for individual students, especially their influence on admission into demanding academic study courses or upper secondary schools, our research questions were following: Do the highest achieving students in all assessments form the same groups in Grade 13 and Grade 8? How are scores from assessments related to teacher's marks? Are girls in Slovene schools in mathematics marked higher than boys because an assessment includes oral part? Does the background data suggest any explanation why their marks are higher? What are possible differences in scope and content of TIMSS Advanced and National tests that produce observed differences?

1 Gender differences in mathematics

Gender differences in mathematics are well researched area. In general, many articles address the gender differences in achievement, characteristics of differences in externally measured achievement or teacher's marks, across time, countries and different student samples. Research also often focuses on personal or sociological

factors leading to higher achievement of boys or girls. Our research is focused on the differences between externally measured achievement and internal school marks. TIMSS Advanced study found the notable gender differences in pre-university mathematics achievement in most participating countries, all in favor of boys. Even more, there was no country where girls outperformed boys, even if majority of advanced mathematics students in some of the participating populations were girls, as it is the case of Slovenia and Portugal (Mullis, Martin, Foy, & Hooper, 2016a). International examination of excellence gap trends from TIMSS data up to 2012 found shrinking sex-based gaps in Grade 8 mathematics (Rutkowski, Rutkowski, & Plucker, 2012). These findings are in line with other research findings of increasing gender difference over years of schooling, from an elementary to a high school. The increasing trend of gender differences in favor of boys was found in solving complex math problems although so small that results of the study still supports gender similarities hypothesis (Hyde, 2005). From the meta-analysis of 242 studies of gender differences in mathematics published from 1990–2007 it is evident that young boys and girls are similar in mathematics, but small difference exists in high school in favor of boys when solving complex problems (Lindberg, Hyde, Petersen & Linn, 2010). In international comparisons, across different countries or cultures, gender differences are not consistent. The general gender gap could be different in different cultures, samples or age of students present many articles and were reviewed by Ceci, Williams, and Barnett (2009).

In general, gender gap in academic achievement requires complex studies on all aspects, starting from early schooling (Halpern, 2014). Baye and Monseur (2016) studied gender differences of students aged 14 to 15 years from TIMSS and PISA, from 1995 to 2015 in an international perspective and confirmed that gender differences vary by content area, students' educational levels, and students' proficiency levels and that males are more frequently among the highest performing students in mathematics and science. Opposite, the meta-analysis of more than three hundred studies of differences in school marks in mathematics, mostly from North America, confirms that small difference exists in favor of girls. Authors found that differences were smaller in the elementary school than in college, larger in North America than in few other countries, but not linked to the year of study or the scale of marks (Voyer & Voyer, 2014).

With the development of statistical methods for meta-analysis, researchers often observe the gender gap of mean outcomes and the gap of outcomes at the tails of distributions which could help to explain small gender differences in means but the higher number of most successful male students in mathematics and science internationally (Ceci, Williams, & Barnett, 2009). Bergold, Wendt, Kasper, and Steinmayr (2017) found that already in Grade 4, across many countries participated in TIMSS and PIRLS 2011, boys were more likely than girls to perform at the top level on general academic performance tests, not only in mathematics. Reilly, Neumannn and Andrews (2015) in their meta-analysis confirmed the small mean differences but greater male variability in mathematics achievement of Grade 12 USA students

80 in NAEP studies from 1990–2011. Many studies tried to find explanations for gender differences in achievement. For Van Houtte, different study culture of boys and girls can be linked to lower achievement of boys (2004). Self-discipline and higher ability self-concept of German girls are linked to their higher math performance according to Steinmayr and Spinath (2008). Study with focus on relation of gender differences to attitudes of girls and women in different cultures showed small gender differences in mathematics achievement of students in Grade 8 or aged 15 years, from TIMSS and PISA 2003 in the meta study of 69 countries, but found more positive math attitudes and affect among boys (Else-Quest, Shibley Hyde, & Linn, 2010). However, the study of Ceci et al. (2009) found that factor connected to the underrepresentation of women in science could be “that mathematics-capable women disproportionately choose non-mathematics fields and that such preferences are apparent among math-competent girls during adolescence”. The study of Implicit Association Tests in 34 countries revealed that people implicitly associate science with males more than with females which has consequences for gender differences in Grade 8 science and mathematics achievement on national level (Nosek et al., 2009). Analyses of external achievement and internal school marks are rare in the literature compared to studies of gender gaps in one measurement. However, the study of school marks and national assessment outcomes in Sweden identified some factors related to the variation in assigning school marks in schools beside the individual student’s achievement, such as education of parents (Klapp Lekholm & Cliffordson, 2008). The study of differences in school marks and scores from national exams among Grade 8 students in Croatia found that there is slight gender difference in achievement between both but no differences in marks given by female and male teachers although authors did not observe mathematics but other school subjects (Burusic, Babarovic, & Seric, 2012). Girls are receiving higher marks for their demonstrated knowledge because they have better verbal intelligence, higher agreeableness, stronger self-discipline, as well as certain aspects of the motivation by Spinath, Eckert, and Steinmayr (2014).

2 Slovene assessments of mathematics

In Slovenia, the target population for TIMSS Advanced 2015 was the population of general upper secondary school students who took the Matura Examination two months after TIMSS Advanced assessment. Slovenia differs from other countries in TIMSS with very large mathematics coverage index (34%), which is the percentage of students taking the most advanced mathematics pre-university course as part of the whole appropriate age cohort in Slovenia. Slovene school system requires from all future students of academic university studies to finish the same, most advanced, general 4-year upper secondary school programs for all school subjects and pass the final National Examination (called Matura) at the end. The final examination consists of exams from three compulsory and two chosen subjects. Mathematics is compul-

sory and can be taken at basic level (BL) or advanced level (AL) by student. All students follow the same mathematics curriculum and they finally decide individually about their level of mathematics exam only few weeks before the examination. With passing the advanced level of math exam, students can reach higher maximum math score (up to 8 on AL instead of up to 5 on BL). A sum of examination scores from all five subjects is used as a student's qualification for the admission to the university. Therefore, mathematics examination results are important motivation for students who compete to enter the most elite university studies, regardless of whether they are going to study mathematics further or not.

Elementary schools in Slovenia include nine grades for students aged 6 to 15 years. Mathematics is one of the most important subjects, it is taught at least 4 lessons per week in every grade and from Grade 6 by specialist teachers with finished university degree from educational mathematics. Students in elementary and upper secondary schools are given marks from 1 (unsatisfactory) to 5 (excellent) by their teachers, for written tests and oral questioning, according to their demonstrations of reaching the national standards of knowledge prescribed in curriculum. At the end of school year, for each student, a teacher summarises all marks given that year in one final mark for the official student's report. In elementary school, the large majority of students receive marks 4 and 5 even from mathematics. At the end of Grades 6 and 9 students take national examinations from mathematics. The student's result and teachers' marks for every school subject for Grades 7 to 9 are used for student's placement in the upper secondary school. The procedure is centralised and include also the private upper secondary schools. While majority of students is placed in the chosen schools, there is usually a strong competition for few of the most demanding general schools (gymnasias) leading to most prestigious university studies. Three school marks for mathematics from Grades 7 to 9 and one from NA are therefore important especially for the best students when competing to reach the desired future academic education.

From the report on gender differences in mathematics achievement found in data from PISA 2012 (OECD, 2015), Slovenia was among countries with notable gender gap between high achievers in favour of boys while no gap was found among low achieving 15-year-old students. Also, the difference in achievement among highest achieving students was likely to change into being in favour of girls after accounting for gender differences in mathematics self-beliefs.

In international comparative assessment of mathematics, TIMSS 2015, mean mathematics achievement of Slovene Grade 8 students is above the international average but not extremely high. The achievement from science of the same students scored higher on the international scale than mathematics, and both increased over all 5 measurements in the last 20 years (Mullis, Martin, Foy, & Hooper, 2016b). TIMSS Advanced results are average compared to other nations in general but very high for Slovene subgroup of students taking the advanced level of national examination. This subgroup is in fact more comparable to populations from other countries than the extremely large whole population of all future university students in Slovenia.

- 82 TIMSS and TIMSS Advanced assessments covered almost the same mathematics contents as the national mathematics examinations and national curriculum, as can be seen from the analysis of contents covered by teacher reports in TIMSS and TIMSS Advanced (Mullis et al., 2016a; 2016b). Therefore, the data from these international assessments enables the direct comparison of independent measurement of mathematics achievement with the mathematics knowledge demonstrated through examinations and school marks on student level.

3 Data and methods

Data sources for our study were databases of TIMSS (IEA, 2017b) and TIMSS Advanced 2015 (IEA, 2017a) student background data with mathematics achievement for Slovenia, extended with data about school marks, effort put into TIMSS or TIMSS Advanced assessments, and relevant national examination scores from mathematics. The questions about student final summarised school mark from mathematics and some other school subjects and series of questions about effort students put into solving the test were added as national options in TIMSS and TIMSS Advanced background questionnaires for students. Upper secondary school students were also asked about their chosen level of the national math exam (basic or advanced).

There were some differences between TIMSS and TIMSS Advanced. TIMSS sampled Grade 8 students from the whole national population while TIMSS Advanced target population were students enrolled into most advanced mathematics programmes in the country and could be in different grades in each country. In all participating countries, target populations for TIMSS Advanced covered less than 20% of age cohort in the country, except Slovenia, where all future university students (34% of age cohort) follow the same most advanced mathematics in their gymnasias. The students sampled were given the test with math items only. Items were mostly multiple choice or extended open ended questions. Students were given the list of formulae and were allowed to use calculators, although items were designed to be independent of calculator use. In case of using calculator, student was required to describe his procedure and write all partial results to get full credit. They were not allowed or needed to use geometry tools. Students taking the physics test were sampled separately and were in general not the same students as in the math sample. Assessments, coding of the answers and preparations of the dataset were done similarly in TIMSS and TIMSS Advanced.

Grade 8 students who participated in TIMSS in Spring 2015, all took NA one year later; all TIMSS Advanced math students took the Matura examination from mathematics two months after the TIMSS Advanced assessments in Spring 2015. A year and half after the TIMSS assessments, the Grade 8 student background data were merged with their scores from NA math examinations from the end of their Grade 9. Data for upper secondary school students were merged with scores from their Matura mathematics examination. Although all personal information was removed from

resulting datasets, they were still considered sensitive and all statistical analyses were done according to the policies for highly sensitive individual examination data inside the safe room at the National Examination Center who holds databases on national examinations in Slovenia. We used IDB-Analyzer and SPSS statistical package as recommended for all the large-scale assessments data which report learning outcomes in the form of plausible values, to calculate descriptive statistics, correlations, regressions and means inside benchmark levels of the achievement. Some of the results are also presented via ordinal dominance graphs (Cankar, 2016). Ordinal dominance (OD) graphs are a unique way to represent information which is exact, easily understandable and allows comparison of ordinal data. They are associated with Mann Whitney's *U*-test and the area under curve can be meaningfully interpreted and compared with other OD graphs (Bamber, 1975). The area under curve that is sometimes called ordinal mean effect (OME(X)) can be literally interpreted in the following way (Jewett, 1983): If we took randomly one subject from each group compared, OME(X) would be the probability that the subject from Group A (that's on the X axis) has higher score or equal to the subject from Group B (on the Y axis).

4 Results

4.1 Comparisons of gender differences in international assessments of TIMSS

National analysis of TIMSS Advanced data for Slovene upper secondary school students revealed high differences in achievement between the students who intended to take advanced or basic level of national mathematics exam. In international comparisons, the group of Slovene students taking advanced national mathematics exam reached TIMSS Advanced mathematics scores similar to the highest ranked advanced Russian student subpopulation. The mathematics coverage index for Slovene students was higher, 8.2% of age cohort, showing that almost a quarter of the whole future university student population (which represents 34.4% of age cohort) in Slovenia takes advanced level of the national exam. The mean score of students taking basic level of national exam was similar to the achievement of Swedish and Italian students as seen from the Table 1 (source: original TIMSS Advanced 2015 Exhibits M1.2 and M1.6 with added Slovene subpopulations with regard to the chosen level of national exam from mathematics (NE) data).

TIMSS Advanced mathematics achievement of boys was higher than achievement of girls in 6 countries; three other countries showed no gender difference. Slovene boys reached higher scores than girls overall and in two subpopulations. Because proportions of boys and girls are unequal in basic and advanced exam group in Slovenia, the overall difference for both groups together is even larger (27 points) than differences in each group (23 and 22 points respectively) – a situation known in statistics as Simpson's paradox (Cankar, 2010). The percentage of boys and girls

84 Table 1 Distribution of TIMSS Advanced 2015 mathematics achievement by gender, Grade 13.

Country	Coverage index	Total		Girls		Boys		Difference (boys-girls)	
		Mean	SE	Mean	SE	Mean	SE		SE
Slovenia, advanced level of NE	8.2%	549	(3.4)	538	(3.6)	562	(6.0)	-23*	(3.3)
Russian Federation 6 hours+	1.9%	540	(7.8)	530	(9.0)	549	(4.5)	-20*	(5.2)
Lebanon	3.9%	532	(3.1)	533	(4.8)	531	(3.9)	-2	(6.1)
USA	11.4%	485	(5.2)	470	(5.3)	500	(6.4)	-30*	(5.8)
Russian Federation	10.1%	485	(5.7)	480	(6.0)	489	(6.2)	-9*	(4.3)
Portugal	28.5%	482	(2.5)	481	(3.0)	483	(3.1)	-2	(3.6)
France	21.5%	463	(3.1)	449	(3.1)	475	(3.4)	-26*	(2.8)
Slovenia	34.4%	460	(3.4)	449	(3.5)	476	(4.9)	-27*	(4.7)
Norway	10.6%	459	(4.6)	453	(5.1)	463	(5.2)	-10*	(4.8)
Slovenia, basic level of NE	26.2%	433	(3.3)	425	(4.0)	447	(4.5)	-22*	(3.2)
Sweden	14.1%	431	(4.0)	424	(5.1)	436	(4.6)	-13*	(5.3)
Italy	24.5%	422	(5.3)	427	(6.1)	419	(6.6)	-8	(7.5)

Notes: Coverage index is fraction of student population as part of whole age cohort in a country; * Significant difference at the 0.05 level; Russian Federation 6 hours+ is subpopulation of students having 6 hours of math per week.

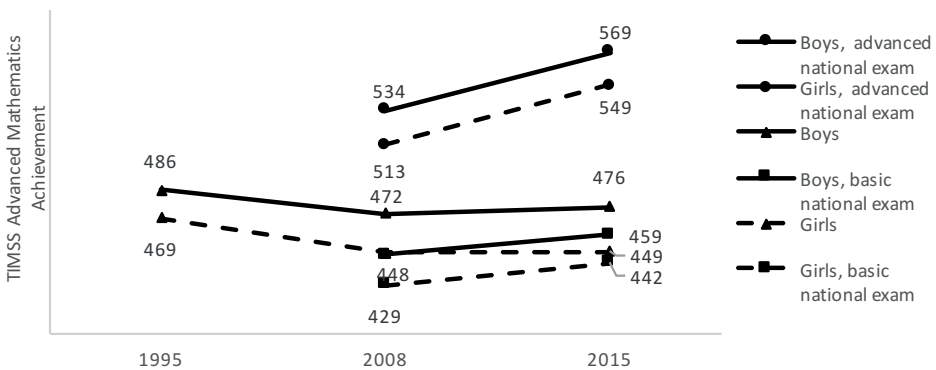


Figure 1 Trends in TIMSS Advanced 2015 mathematics achievement in Slovenia by gender and sub-populations, Grade 13.

from the whole age cohort reaching the top benchmark was similar (boys 8% and girls 7%), although there was share of 60% of girls among all TIMSS advance population and therefore, there are more males in the upper tail of distribution of achievement. Gender differences in TIMSS Advanced mathematics achievement in Slovenia stayed unchanged over the last 20 years (Figure 1).

For Grade 8 students, the analysis of trends of mathematics achievement gap between girls and boys shows no gender differences in the last 20 years. However, in 2015 boys outscored girls in one content domain (numbers) and girls slightly outscored boys in one other content domain (algebra), but no differences were seen in scores achieved in different cognitive domains of mathematics knowledge (Mullis et al., 2016b). Ratio between variance of achievement of boys over the variance of girls was 1.02, not significant according to the published literature.

4.2 Comparisons of gender differences in national marks from mathematics

Upper secondary school students can get a maximum of 5 points if they choose the basic level of the national math exam and maximum of 8 points at the advanced level. Scores from the national mathematics examination of upper secondary school students are higher for girls. As shown in the Table 2, in general, girls scored about 0.5 point (out of maximum 8 points) higher than boys. However, girls outperformed boys only among students who chose basic level of the national exam. There is no difference between boys and girls in their mean scores at the national exam at the end of elementary school at Grade 9 from mathematics, but the difference exists in scores from Slovene language, in favour of girls.

Table 2 School marks and results from national test in mathematics for Grade 13.

Grade 13	School marks				Difference (girls–boys)	National test in mathematics				Difference (girls–boys)
	Boys	SE	Girls	SE		Boys	SE	Girls	SE	
Mathematics										
Total	3.25	0.04	3.31	0.05	0.06	4.08	1.72	4.64	1.47	0.56*
Advanced exam	4.49	0.07	4.55	0.03	0.06*	6.02	1.60	6.11	1.39	0.09
Basic exam	2.82	0.04	2.97	0.05	0.15*	3.28	0.98	3.84	0.68	0.56*
Slovene language										
	3.66	0.05	3.90	0.05	0.24*					

* Difference is statistically significant at the 0.05 level.

Grade 13 girls are marked higher from Slovene language and from mathematics, when we compare two subgroups of students by the level of intended mathematics examination separately.

Although girls and boys reach similar marks at the national examination at the end of elementary school, their school marks given by teachers in classes differ. In mathematics and all science subjects, Grade 8 girls are marked significantly higher than boys (Table 3). The ratio of variances of marks of boys over girls was 1.1, slightly higher than in case of achievement, but still small.

86 Table 3 School marks and results from national test in mathematics for Grade 9¹.

Grade 8	School marks					National test in mathematics				
	Boys	SE	Girls	SE	Difference (girls–boys)	Boys	SE	Girls	SE	Difference (girls–boys)
Mathematics	3.46	0.03	3.63	0.03	0.17*	51.09	0.76	52.07	0.78	0.98
Biology	3.67	0.04	4.09	0.04	0.42*					
Geography	3.72	0.03	4.00	0.04	0.28*					
Chemistry	3.64	0.04	3.94	0.03	0.30*					
Physics	3.59	0.03	3.72	0.04	0.13*					
Slovene lang.	3.47	0.03	4.02	0.03	0.55*					

¹ Grade 9 was the TIMSS Grade 8 population one year later.

* Difference is statistically significant at the 0.05 level.

Different pattern of gender gap in assessments and in school marks for both populations was not expected. Students who reach each TIMSS Advanced benchmark level and took basic level of national exam scored lower than student who took advanced level, although the same math knowledge should be valued similarly (Table 4). School marks of girls who scored below high benchmark are higher than marks of similarly successful boys. The gap in marks disappears in groups of student who scored above the high TIMSS Advanced benchmark, the most able students. However, percentages of boys and girls over benchmarks vary, more boys than girls reached higher benchmark levels in both subpopulations.

Table 4 Scores from the national math exam and school marks by reached TIMSS Advanced benchmark levels of mathematics knowledge, Grade 13.

TIMSS Advanced benchmarks	% of students	Mean national exam score (SE)		% of girls	Mean school mark of girls (SE)		% of boys	Mean school mark of boys (SE)	
		exam score	(SE)		school mark	(SE)		school mark	(SE)
Basic level of the national exam									
Below intermediate	69	2.92	(0.05)	76	2.77	(0.05)*	64	2.55	(0.05)
Intermediate to High	27	3.78	(0.05)	22	3.53	(0.07)*	30	3.23	(0.07)
High to Advanced	4	4.25	(0.12)	2	4.16	(0.28)	6	3.64	(0.18)
Above advanced	0	-	-						
Advanced level of the national exam									
Below intermediate	10	4.03	(0.28)	14	4.02	(0.17)	7	3.67	(0.26)
Intermediate to High	39	5.42	(0.10)	43	4.49	(0.06)*	35	4.24	(0.11)
High to Advanced	40	6.42	(0.10)	36	4.77	(0.05)	43	4.69	(0.09)
Above advanced	11	7.21	(0.16)	7	4.80	(0.12)	15	4.92	(0.06)

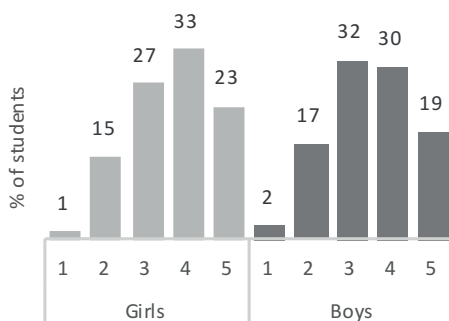
* The mean is significantly higher than the mean of the opposite sex (at the 0.05 level).

Having the mark satisfactory (2) reported 26% students, good (3) reported 29% students, very good (4) reported 23% students and excellent (5) reported 19% of students. However, students are distributed to the reached benchmark levels not consistently with their school mark. Almost half of student with excellent mark (48%) from mathematics reach at most the high knowledge in TIMSS Advanced, while 43% of girls but 64% of boys scored above the high benchmark. From this data, we can't confirm that high achievers from TIMSS are at the same the high achievers recognised by Matura examination.

Grade 8 students show no gender gap in mathematics achievement, but evidently higher marks for girls. Graphs of distributions of marks differs from almost equal distributions of TIMSS scores for both genders. Therefore, we examined if the similar pattern of lower marks for high achieving boys than for high achieving girls exists already in Grade 8. We assigned students to groups by percentiles determined through marks: the TIMSS achievement was calculated at percentiles defined by the distribution of students by school marks (Table 5) and scores used as limits of intervals for levels of knowledge. 21% of students were marked excellent and top 21% scored 572 points or more on TIMSS scale. In we estimate students who are marked differently than expected if marks and TIMSS scores would be aligned. Lower marked students are in majority boys and there are more girls among students with the highest school mark.

Table 5 Distribution of students by school marks and TIMSS achievement, Grade 8.

School marks	% of students	Percentiles of TIMSS math scores according to given school marks	% of girls	% of boys	Difference	T
Unsatisfactory (1)	2	below 372	35.47	64.53	29.06	2.73
Satisfactory (2)	16	from 372 to below 452	44.97	55.03	10.06	2.47
Good (3)	30	from 452 to below 515	45.03	54.97	9.95	3.20
Very good (4)	31	from 515 to below 572	51.04	48.96	-2.08	-0.81
Excellent (5)	21	at or above 572	53.84	46.16	-7.69	-2.43



Distributions of marks girls and boys

If we look at the intervals defined by TIMSS scores from Table 5 as the expected marks for students, the analysis of difference reveals that only 49% of students are marked as expected from TIMSS score (Figure 2). However, girls more often have higher given mark than it would be expected (29% vs 21%) and boys have more often lower mark than it would be expected (29% vs 22%) from their TIMSS score.

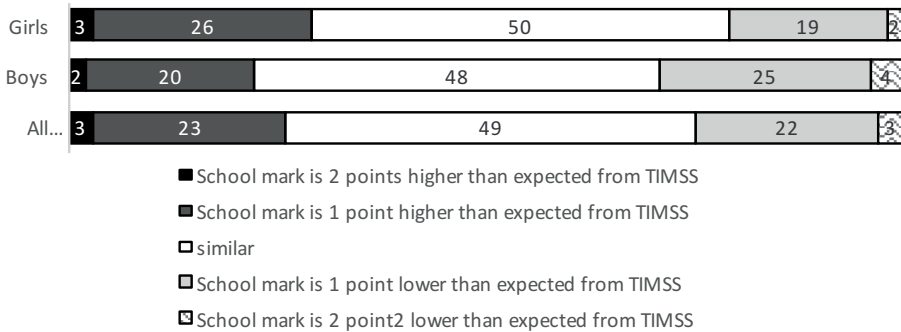


Figure 2 Distribution of percentages of girls and boys by differences between given school marks from mathematics and expected marks from TIMSS score, grade 8.

Precise data show that girls get higher marks than boys and that differences increase with marks. Out of 21% highest achievers in TIMSS, the excellent mark was given to only 56% of boys but almost 70% of girls. 9% of boys but 2% of girls received middle mark (3, good) although they were top achievers in TIMSS. We may conclude that school marks base on some additional skills and knowledge not assessed by TIMSS but giving priority to girls. This is the opposite as expected as in Slovene system, by curriculum requirements, school marks should strictly reflect the achieved standards of knowledge and not subjective teacher estimate of student effort in school.

4.3 The frameworks and structure of tests

The gender differences are higher in upper secondary schools, therefore the first step in searching for explanation for gender differences was the analysis of the content and cognitive domains covered by TIMSS Advanced and the national mathematics exam. The comparison of the *TIMSS Advanced Framework* (IEA, 2013) and the *Standards for the National Mathematics Examination* (Benko et al., 2015) shows that the national tests covered more contents than TIMSS Advanced in the area of logic, sets, probability and statistics as well as in algebra (Table 6) with trigonometry divided among functions (algebra) and shapes (geometry).

Table 6 TIMSS Advanced frameworks and national mathematics exam contents, Grade 13.

Content domains		TIMSS Advanced	National exam			
Algebra		Expressions and operations; Equations and inequalities; Functions	Expressions and operations; Equations and inequalities; Functions; Conic sections; Sequences and series; Differential of functions			
Calculus		Limits; Derivatives; Integrals	Limits; Derivatives; Integrals			
Geometry		Non-coordinate and coordinate geometry; Trigonometry	Geometry on plane and in space; shapes; Vectors on plane and in space			
Logic & sets			Basics of logic; Sets, Number sets			
Probability & Statistics			Combinatorics, Probability, Statistics			
		National Exam				
Cognitive domains	TIMSS Advanced	Written part 1 (Basic and Advanced level)	Written part 2 (Advanced level)	Oral exam (Basic level)	Oral exam (Advanced level)	
Knowing	35	at least 30	at least 10	at least 30	at least 10	
Applying	35	30–50	40–60	30–50	40–60	
Reasoning	30	up to 30	up to 40	up to 30	up to 40	

The attention to each cognitive domain in TIMSS Advanced tests is given in percentages of time for solving test items. The Matura mathematics exam covers the same cognitive domains with different attentions at the basic (BL) and advanced level (AL). The exam consists of two parts, written and oral. The written part is further divided into the first test which is the same for all students and second additional test for students who take AL. Regarding cognitive domains, the first written and oral parts for BL are similar to TIMSS Advanced test. But the second written part and oral part for AL give twice less attention to knowing and half more attention to applying and reasoning than TIMSS test.

The gender differences in percentages of student solving each TIMSS Advanced test item were mostly in favour of boys. In Slovenia, only few items out of 114 were solved better by girls. Average points of gender differences were larger for items from reasoning domain (9.21 points) and smaller for items from knowing and applying domains (5.2 and 7.1 points). Therefore, the gender difference in TIMSS Advanced is increasing with the cognitive expectation of items.

4.4 Relations between TIMSS achievement, national scores and school marks

The final score of the national exam for upper secondary school students is the sum of scores for each part. Score given at the oral part of examination contributes only with 20% to the final score. The correlation between TIMSS scores and grades from the Matura mathematics exam were found to be high for the written part and low for the oral parts of the exam. That is most certainly due to ceiling effect since most students perform excellent on the oral part and variability is rather low. The oral questioning of the exam is assessed by the teachers from the student's school and is regarded by students as not fully external although the questions are prepared nationally in advance and randomly blindly drawn from the collection by the student at the exam. The final grades of the Matura exam are also highly correlated with TIMSS Advanced scores, higher at AL than at BL (Table 7).

Table 7 Correlations between TIMSS and national scores by gender.

Correlated scores	Girls	Difference	Boys
Grade 13			
Scores for Matura, written part, BL & TIMSS Adv.	0.56	=	0.57
Scores for Matura, written part, AL & TIMSS Adv.	0.59	>	0.56
Scores for Matura, oral part, BL & TIMSS Advanced scores	0.27	>	0.24
Scores for Matura, oral part, AL & TIMSS Advanced scores	0.23	<	0.26
Matura final marks, BL (1–5) & TIMSS Advanced scores	0.56	=	0.56
Matura final marks, AL (1–8) & TIMSS Advanced scores	0.58	>	0.55
School marks (1–5) & TIMSS Advanced scores	0.61	<	0.65
School marks (1–5) & Matura final marks	0.66	<	0.76
Grade 9			
Scores for NA & TIMSS scores from Grade 8	0.77	=	0.77
School marks (1–5) & TIMSS scores from Grade 8	0.52	>	0.48

In upper secondary schools, correlations between school marks and grades from the Matura Exam are the highest. Unexpectedly, the TIMSS Advanced scores have the highest correlation with school marks although these marks are not external and come from written and oral assessments. In elementary schools, the correlation of TIMSS scores with external national exam scores is higher than with the internal school marks.

In order to describe the gender gaps, we use ordinal dominance graphs, graphical comparison of two distributions that quickly shows which group dominates and in which part of distribution (Figure 3). Results along diagonal line would indicate equal groups without any dominance. The ordinal dominance graphs by gender for the different test scores in Slovenia show equality of girls in boys on TIMSS in Grade 8, slight dominance of girls when observing results from NA in Grade 9 and larger dominance of girls when observing school marks for same population. Looking at Grade 13

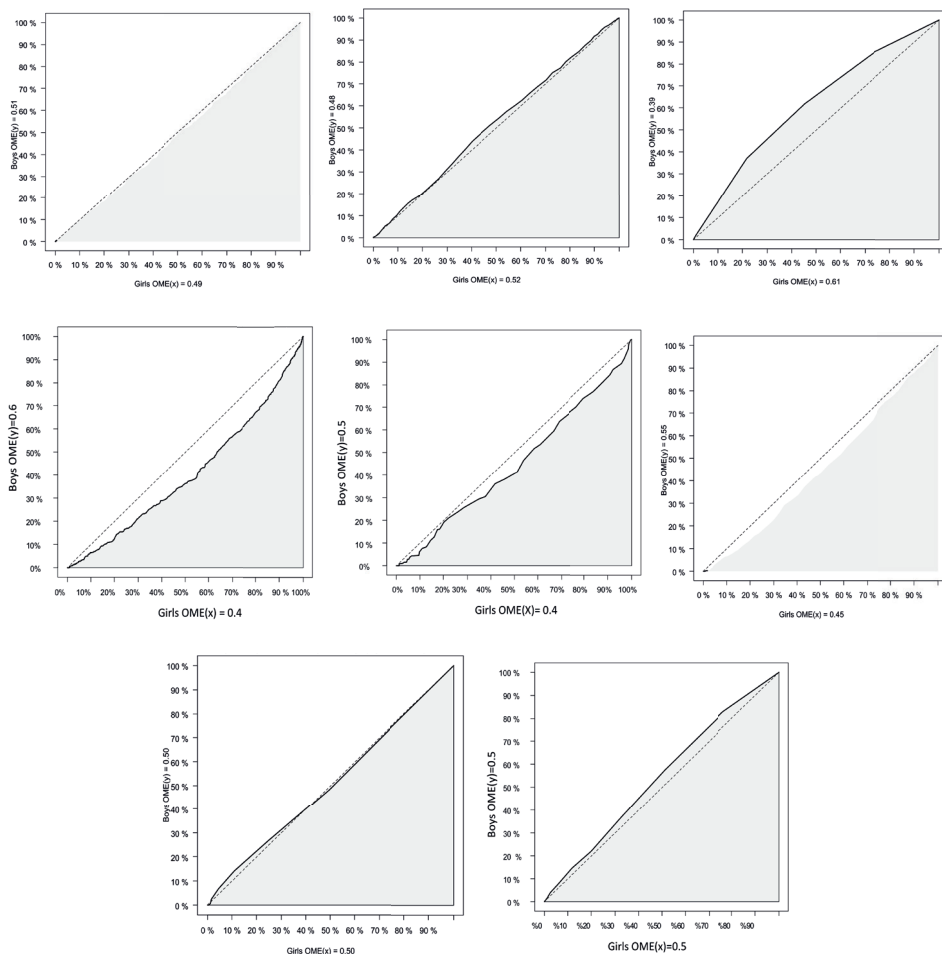


Figure 3 Ordinal dominance graphs for TIMSS and TIMSS Advanced test, national mathematics examination and school marks by gender, Grade 13.

population, results of boys dominate on TIMSS Advanced and written part of Matura at AL, while girls slightly dominate on oral part at BL.

Grade 8 girls are essentially given higher grades from teachers at mathematics lessons in schools than boys but reached similar score as boys at NA exam from mathematics and from TIMSS test. In upper secondary school, the highest gender differences in TIMSS Advanced scores occur around the middle part of the scale. Boys scored higher at the upper half of the scale for written part of the Matura exam at AL and BL, but differences are smaller than for TIMSS Advanced. Girls who reached the upper half of the scale for the oral part of the Matura Exam at BL received higher grades than boys. There was no gender difference in oral part at the AL of the Matura mathematics exam.

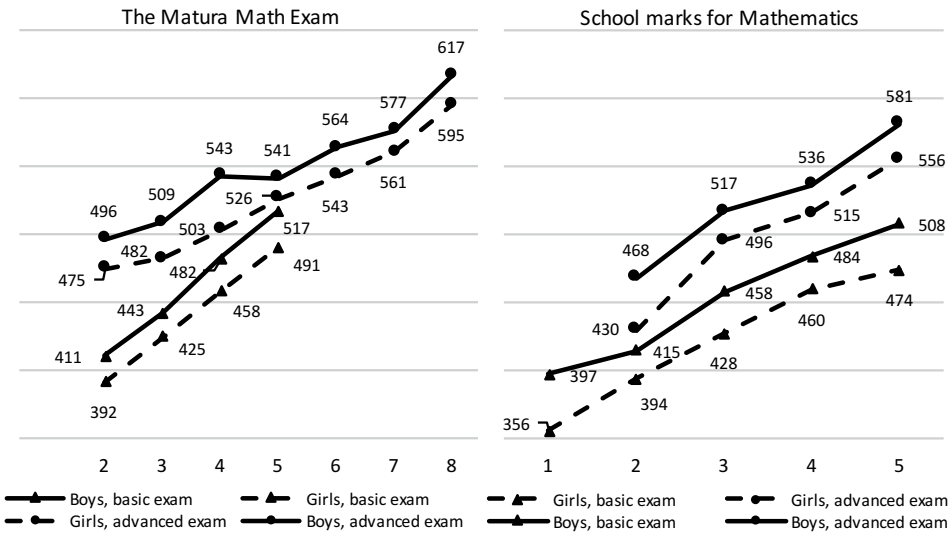


Figure 4 TIMSS Advanced achievement by the Matura exam and school mathematics marks, Grade 13. **Note:** Differences are not significant only between girls with school marks 4 and 5 taking the basic level, and between girls and boys taking the advanced level of the math exam and graded by 2 and 3 or 3 and 4.

TIMSS Advanced scores by national marks given at the Matura exam and by teachers in classes of upper secondary schools (Figure 4) describe the inconsistent links between both scores. Although having the same national score, boys reached higher TIMSS Advanced score than girls. Or, girls with similar TIMSS Advanced score as boys were given higher national mark than boys. Additionally, students with the same TIMSS Advanced score who choose the BL of the Matura math exam reached higher final national mark than students who choose the AL of the Matura math exam. For example, girls with TIMSS Advanced achievement of 475 points reached 2 points if taking the AL but 4 or 5 points if taking the BL. Boys with 517 points from TIMSS Advanced reached 5 points at BL of Matura exam but achieved 3 or 4 points if taking the AL. Results show that marking at the basic and advanced level of the Matura math exam is not completely consistent. The same pattern is seen from the comparison by school marks. Girls with TIMSS Advanced achievement of about 430 points are given mark 3-good if they intend to take basic level of the national exam and mark 2-satisfactory if they choose advanced level of the exam. Results are problematic. Regular mathematics course is following the same standards and curriculum and students are taught in mixed classes of students that choose any level of exam. Therefore, students' marks for the same math knowledge should be the same for both genders and student of both Matura levels.

5 Searching for explanations of gender gaps

From the comparison of grades and TIMSS Advanced achievement it seems that national assessments do not recognize some knowledge of boys. The Table 8 of TIMSS Advanced scores from each cognitive domain clearly shows the largest gender differences between equally marked boys and girls in reasoning. While there are no differences by gender between students of the advanced level of the Matura Exam for knowing and applying domains (scores 6, 7 and 8), in reasoning, boys with any score from the Matura math exam outperformed girls with the same score at TIMSS advanced. Also, boys with lower scores (5 or less) at Matura outperformed girls in all three cognitive domains on TIMSS Advanced.

Table 8 Gender differences of achievement in cognitive domains by the Matura exam scores, TIMSS Advanced, Grade 13.

Nat. exam score	Mean achievement – reasoning					Mean achievement – applying					Mean achievement – knowing				
	Boys	(SE)	Girls	(SE)	Diff.	Boys	(SE)	Girls	(SE)	Diff.	Boys	(SE)	Girls	(SE)	Diff.
1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
2	402	(8.3)	373	(8.6)	29*	420	(6.9)	402	(6.6)	18*	420	(7.8)	402	(5.1)	18*
3	433	(6.2)	407	(6.1)	26*	452	(5.2)	435	(4.3)	16*	450	(7.4)	431	(4.6)	19*
4	475	(5.6)	439	(5.0)	37*	493	(5.8)	469	(3.9)	24*	491	(8.2)	468	(3.2)	23*
5	514	(8.0)	481	(5.9)	33*	533	(7.0)	509	(4.9)	23	531	(6.9)	510	(4.8)	21*
6	555	(13.1)	521	(6.3)	34*	569	(11.9)	551	(8.9)	19	568	(9.7)	551	(5.5)	17
7	570	(7.6)	544	(7.4)	26*	581	(8.8)	567	(5.5)	14	579	(9.9)	573	(6.5)	6
8	617	(9.2)	585	(8.6)	33*	623	(7.9)	602	(11.0)	21	621	(8.6)	604	(7.2)	17

* Difference is statistically significant at the 0.05 level.

These results suggest that the Matura from mathematics measures similar knowledge between boys and girls in knowing and applying domain at advanced level of exam. But it does not recognize the knowledge of boys in reasoning that was required by items in TIMSS Advanced assessment. Even worse, at BL it seems that Matura does not see and grade the mathematics knowledge of boys that was measured by TIMSS Advanced in all three cognitive domains. The findings warn that there are conceptual differences in the content domains of both tests.

Similar, in Grade 8 students with the highest excellent teacher's mark show no gender difference in TIMSS achievement across cognitive domains. However, among student with the middle mark "good", boys outscored girls in all three domains. Therefore, we conclude that assigning marks in Slovene schools is not well focused to recognise and award the intermediate knowledge of boys. Giving lower marks leads into less opportunities for placement into more demanding upper secondary schools for boys as it influences admission where they require high marks from el-

94 Table 9 Gender differences of achievement in cognitive domains by the teacher marks, TIMSS, Grade 8.

School mark	Mean achievement – reasoning				Mean achievement – applying				Mean achievement – knowing						
	Boys	(SE)	Girls	(SE)	Diff.	Boys	(SE)	Girls	(SE)	Diff.	Boys	(SE)	Girls	(SE)	Diff.
1	438	(10.6)	431	(14.3)	7	442	(10.0)	427	(11.8)	15	437	(9.1)	430	(12.8)	7
2	449	(5.8)	441	(6.1)	8	455	(3.8)	441	(4.8)	14*	454	(4.0)	446	(4.3)	8
3	495	(4.3)	482	(4.3)	13*	497	(3.3)	482	(3.3)	14*	498	(3.4)	486	(3.8)	12*
4	537	(4.1)	531	(3.7)	6	535	(3.5)	527	(3.1)	8*	539	(4.0)	535	(3.6)	4
5	592	(4.9)	586	(4.4)	6	585	(3.6)	578	(3.7)	7	587	(4.3)	586	(3.9)	1

* Difference is statistically significant at the 0.05 level.

Note about marks: 1 is unsatisfactory; 2 is satisfactory, 3 is good, 4 is very good, 5 is excellent.

elementary schools. From PISA 2015 data for Slovenia which covers all programs of upper secondary school students, we observe that boys systematically enter less demanding programs in notable higher percentages than girls, that ends with large 60% of female population in gymnasias.

When searching for reasons of differences between school marks and TIMSS achievement we tested the links between both scores and many background factors. Opposite to the case of Sweden, socioeconomic status, measured in TIMSS on scale of home educational resources, including material sources and parental education, in Slovenia is not linked to the inconsistencies in marking. Correlations with this and two main students' attitudes are listed in Table 10.

Table 10 Correlations of school marks and TIMSS achievement by gender, Grade 8.

	Correlation with school marks				Correlation with TIMSS			
	Boys		Girls		Boys		Girls	
	Corr.	SE	Corr.	SE	Corr.	SE	Corr.	SE
Home educational resources of student	0.33	(0.02)	0.32	(0.03)	0.34	(0.02)	0.34	(0.03)
Student's liking mathematics	0.39	(0.03)	0.42	(0.02)	0.30	(0.03)	0.31	(0.03)
Student's perception of engaging teaching	0.26	(0.03)	0.28	(0.02)	0.21	(0.03)	0.19	(0.02)
TIMSS score	0.65	(0.02)	0.68	(0.02)				
TIMSS score for knowing	0.63	(0.02)	0.67	(0.02)	0.86	(0.01)	0.86	(0.01)
TIMSS score for applying	0.62	(0.02)	0.66	(0.02)	0.87	(0.01)	0.87	(0.01)
TIMSS score for reasoning	0.61	(0.02)	0.63	(0.02)	0.85	(0.01)	0.85	(0.01)

Correlations do not differ by gender. Correlations of liking mathematics and engaging teaching of mathematics with school marks are higher than with TIMSS scores for both genders and confirm the importance of school marks for motivation of learn-

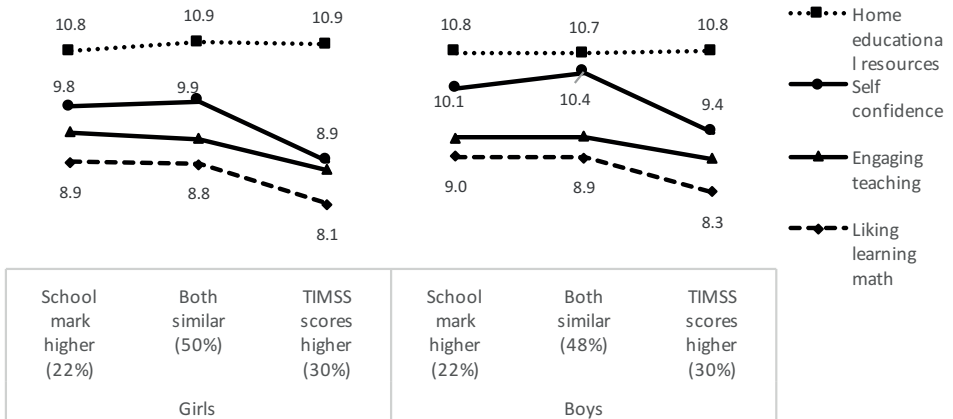


Figure 5 Background factors of learning mathematics with real and estimated marks according to TIMSS by gender, Grade 8.

ing mathematics. The differences in pattern of attitudes among girls and in among students who are marked in school differently as it would be expected from their TIMSS score can be observed in Figure 5.

Across groups of students with higher, similar or lower school mark than expected from TIMSS score, there is almost no difference in their educational home support. We may conclude that socioeconomic status of students does not impact significantly the marking of students in schools. But other three attitudes are dropping from students who have higher school mark than estimated from TIMSS achievement to students with lower school mark than estimated from TIMSS. The decreases are larger among girls, but self-confidence and liking of learning are of significant sizes also for boys. We may not discuss the linkage as causal relations but see that students who were getting lower marks in schools than they demonstrated in TIMSS assessment are less confident and like learning mathematics less than others. Low marks decrease the confidence, but less confident students do not show all their strength in assessments and therefore get lower marks. In Grade 8, there are significant gender differences in valuing mathematics – 8.86 for girls (*SE* 0.04) vs 9.07 (*SE* 0.05) for boys; and in self-confidence – 9.66 (*SE* 0.05) for girls vs 10.04 (*SE* 0.05) for boys, but not in liking learning mathematics (not significant difference of 0.05). Although different attitudes may affect teachers’ subjective perceptions (Voyer & Voyer, 2014) it seems that in Slovenia teachers’ higher marks for girls are not in accordance to girls’ low attitudes. Some information about girls’ reaction to tests provide the assessed effort put into the test among participating students in TIMSS. It was measured with five questions developed to help especially in analysing large scale assessments that have no consequences for individual students (Eklof, 2006). The question asked student how much they agree with the following statements: (a) I gave my best effort on this test, (b) I did not give this test my full attention while completing it, (c) I tried less hard on this test than

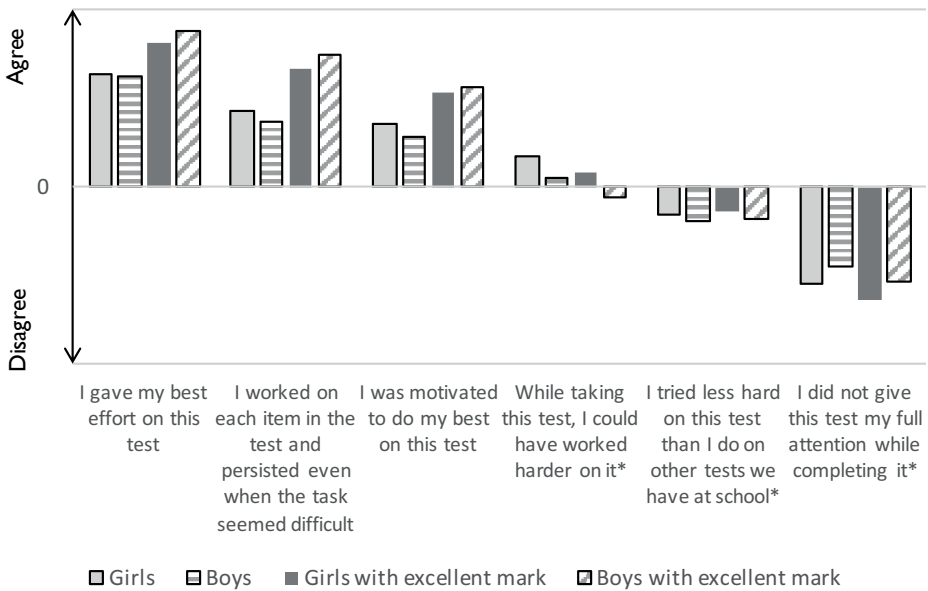


Figure 6 Effort put in TIMSS test by gender, Grade 8.

I do on other tests we have at school, (d) I worked on each item in the test and persisted even when the task seemed difficult, (e) I was motivated to do my best on this test, (f) While taking this test, I could have worked harder on it; with answer agree a lot, agree a little, disagree a little, disagree a lot. When interpreting statistics, the order of answer categories for (b), (c) and (f) should be turned. From the comparisons of answers per gender (Figure 6) it is evident that in both populations girls admitted that they have put less effort into solving the test than boys, especially among the higher achievers. Although differences are very small, girls in Grade 8 and Grade 13 (Figure 7) admitted they try more for school tests than they did for TIMSS. In general, girls with excellent marks were less motivated for TIMSS than boys with excellent marks. These data therefore support the hypothesis that Grade 8 and Grade 13 girls, especially excellent, try harder for tests with consequences for their promotion in school and boys try as hard as they can even for tests without consequences for their schooling.

In the Grade 13, the most effort was reported by boys of advanced level of national exam who also achieved the highest score and the least effort was reported by girls of basic level of national exam. Results for both populations together support the possible explanation for gender differences. Girls achieved higher scores on national exam because they tried and worked harder than for TIMSS test while middle achieving boys (basic level of NE) did not try for TIMSS tests as well as they do not try very hard for school tests either.

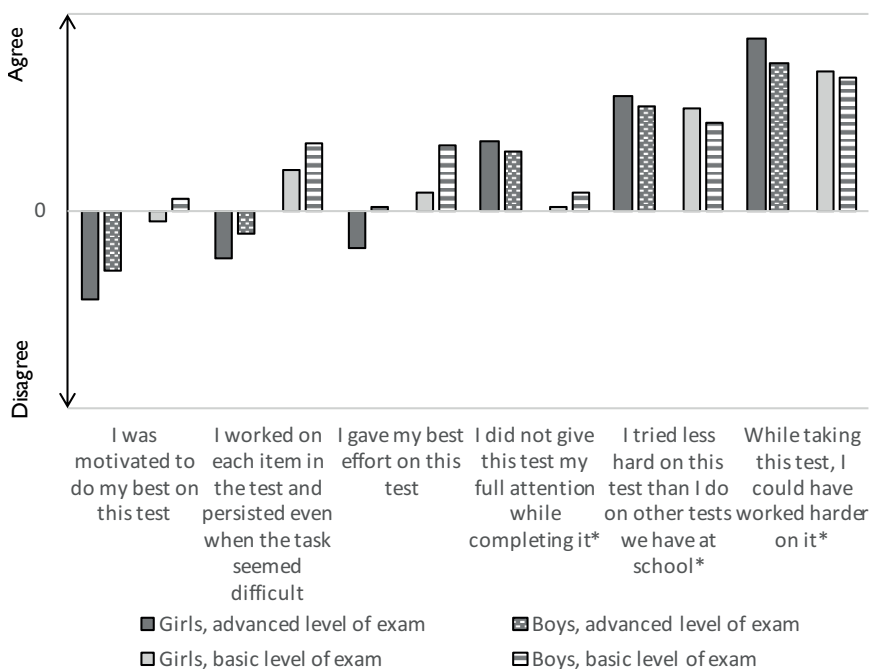


Figure 7 Effort put in TIMSS Advanced test by gender, Grade 13.

6 Conclusion

The results of our study provide some new information about the mathematics achievement of pre-university students in Slovenia and answers to our research questions. The answer to the first question is no, the highest achieving students with regard to the national scores, teacher marks and international assessments outcomes in both observed populations are found to be different groups. In general, Matura scores and school marks were not found to be consistent with TIMSS Advanced achievement, gender gap in Matura is in favour of girls and gender gap in TIMSS Advanced is in favour of boys. The results are consistent with findings in other studies. The Matura math exam in Grade 13 was found to measure similar knowledge between boys and girls who take the AL of exam in domain of knowing and applying but it does not recognize some knowledge of boys in mathematical reasoning. Among lower achievers, the national exam does not recognize and grade part of the mathematics knowledge of boys from all cognitive domains. The main difference between both assessments is oral questioning present in Matura but not in TIMSS. As the results of oral part of Matura significantly favour girls we believe that oral questioning is important factor which is also linked to the fact that girls are given relatively higher school grades. The essential part of school marks in Slovenia

98 is teacher's oral assessment of each individual student. The finding helps to explain gender difference between international and national assessment but also provides ideas for improvement of the Matura math exam. It suggests that mathematical knowledge of middle achieving boys, especially mathematical reasoning, would need to be better recognized by the national examination as well as by teachers if we want them to be similar to TIMSS. The changing gender gap could be partly described by the reported effort put into solving the TIMSS tests. Girls more often than boys reported that they didn't work as hard on TIMSS test and that they put slightly more effort into the national exams. The mediating effect of effort could explain higher school marks of girls, especially of high achievers, and similar or higher TIMSS achievement of boys. The measurement of effort supports the hypothesis that the larger problem of marking in school is the undermarking of boys (who most likely do not try to show all their knowledge in any test) than overmarking of girls. Next, the marking was found important as a factor of motivation which is a large problem in Slovenia, demonstrated with the very low or lowest mean national scores on any motivation scales in international comparisons and with still declining trends. In particular, in Grade 8, the motivation for learning mathematics is found to be more strongly linked to marks than to the international measurement of knowledge. School marks from mathematics contribute to the opportunities for student's placement in upper secondary school. Therefore, the above mentioned problematic side of marking could also contribute to the lower percentages of boys than girls choosing higher demanding upper secondary school programs and therefore it could be one of the reasons for observed lower overall achieved upper secondary education of boys compared to girls in Slovenia. The evaluation of marking system, especially with regard to middle achieving boys, is clearly needed.

With regard to known results from research literature, gender differences of achievement in Slovenia do not differ much from other countries. They show almost similar achievement in Grade 8 and higher boy's achievement in Grade 13 while school marks were found to be higher for girls, similar to already found pattern in some other countries. By linking TIMSS achievement and school marks we were able to recognise some problems of marking students on national level, after taking into account that differences in students' individual characteristics contribute to a significant extent to gender differences in any school performance.

The data used for this study are limited to the international databases of TIMSS and TIMSS Advanced with few additional answers to the national questions for students. Therefore, they cannot provide all needed information for an extended study of individual or group student characteristics on achievement or on the gender gap. The study found some basic facts which will be studied further, most likely together with teacher characteristics and ways of their assessments of student knowledge in mathematics and science.

In this research girls were found to be somewhat better adapted to today's school environments, as research literature suggests, most likely because of their better verbal intelligence, higher agreeableness, stronger self-discipline, as well as certain

aspects of their motivation (Spinath, Eckert, & Steinmayr, 2014). In light of these specific differences, it could be expected that changing certain aspects of school environments with regard to stimulating boys' motivation and engagement might help boys to better succeed in school and, thus, reduce our national educational inequality.

References

- Baye, A. & Monseur, C. (2016). Gender differences in variability and extreme scores in an international context. *Large-Scale Assessment in Education* 4(1). <https://doi.org/10.1186/s40536-015-0015-x>
- Benko, D., Erker, J., Hvastija, D., Jan, M., Miler, A., Robnik, A., Škof, M., & Žerovnik, J. (2015). *Predmetni izpitni katalog za splošno maturo – matematika* [Standards for the National Examination – Mathematics]. National Examinations Centre, Ljubljana. Retrieved from <http://www.ric.si/mma/201720M-MAT-2017/2015083113005248>.
- Bergold, S., Wendt, H., Kasper, D., & Steinmayr, R. (2017). Academic competencies: Their interrelatedness and gender differences at their high end. *Journal of Educational Psychology*, 109(3), 439–449. <https://doi.org/10.1037/edu0000140>
- Burusic, J., Babarovic, T., & Seric, M. (2012). Differences in elementary school achievement between girls and boys: Does the teacher gender play a role? *European Journal of Psychology of Education*, 27(4), 523–538.
- Cankar, G. (2010). PISA 2006 main findings in Slovenia and Simpson's paradox. V: XIV. IOSTE Symposium, International Organization for Science and Technology Education, June 13–18, 2010, Bled, Slovenia. In S. Dolinšek (Ed.), *Socio-cultural and human values in science and technology education: proceedings*. Institute for Innovation and Development of University. Ljubljana.
- Cankar, G. (2016). Governing the transition to higher levels of education and differences between achievement and school grades by gender. *The Journal of Educational Research*, 8(2), 60–86.
- Ceci, S. J., Williams, W. M. & Barnett, S. M. (2009). Women's underrepresentation in science: Sociocultural and biological considerations. *Psychological Bulletin*, 135(2), 218–261.
- Eklöf, H. (2006). Development and validation of scores from an instrument measuring student test-taking motivation. *Educational and Psychological Measurement*, 66 (4), 643–656.
- Else-Quest, N. M., Shibley Hyde J., & Linn, M.C. (2010). Cross-national patterns of gender differences in mathematics: A meta-analysis. *Psychological Bulletin*, 136(1), 103–127.
- Holman L., Stuart-Fox D., & Hauser C. E. (2018). The gender gap in science: How long until women are equally represented? *PLOS Biology* 16(4), e2004956. <https://doi.org/10.1371/journal.pbio.2004956>
- Houtte, M. V. (2004). Why boys achieve less at school than girls: The difference between boys' and girls' academic culture. *Educational Studies*, 30(2), 159–173.
- Hyde, J. S. (2005). The gender similarities hypothesis. *American Psychologist*, 60(6), 581–592.
- IEA (2017a). *TIMSS Advanced SPSS Data*. TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College. Retrieved from <https://timssandpirls.bc.edu/timss2015/advanced-international-database>.
- IEA (2017b). *TIMSS SPSS Data*. TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College. Retrieved from <https://timssandpirls.bc.edu/timss2015/international-database>.
- Klapp Lekholm A., & Cliffordson C. (2008). Discrepancies between school grades and test scores at individual and school level: effects of gender and family background, *Educational Research and Evaluation*, 14(2), 181–199. <https://doi.org/10.1080/13803610801956663>

- 100 Lindberg, S. M., Hyde, J. S., Petersen, J. L. & Linn, M. C. (2010). New trends in gender and mathematics performance: A meta-analysis. *Psychological Bulletin*, 136(6), 1123–1135.
- Mullis, I. V. S. & Martin, M. O. (Eds.). (2013). *TIMSS 2015 assessment frameworks*. TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College. Retrieved from <http://timssandpirls.bc.edu/timss2015/frameworks.html>.
- Mullis, I. V. S., Martin, M. O., Foy, P., & Hooper, M. (2016a). *TIMSS Advanced 2015 International Results in Advanced Mathematics and Physics*. TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College. Retrieved from <http://timssandpirls.bc.edu/timss2015/international-results/advanced>.
- Mullis, I. V. S., Martin, M. O., Foy, P. & Hooper, M. (2016b). *TIMSS 2015 International Results in Mathematics*. Boston College, TIMSS & PIRLS International Study Center. Retrieved from <http://timssandpirls.bc.edu/timss2015/international-results>.
- Nosek, A. B., Smyth, F. L., Sriram, N., Lindner, N. M., Devos, T., Ayala, A., ... , Greenwald, A. G. (2009). National differences in gender-science stereotypes predict national sex differences in science and math achievement. *Proceedings of the National Academy of Sciences*, 106, 10593–10597. <https://doi.org/10.1073/pnas.0809921106>
- OECD (2015). *The ABC of gender equality in education: Aptitude, behaviour, confidence*. Pisa, OECD Publishing. <https://doi.org/10.1787/9789264229945-en>
- Rutkowski, D., Rutkowski L., & Plucker, J. A. (2012). Trends in education excellence gaps: a 12-year international perspective via the multilevel model for change. *High Ability Studies*, 23(2), 143–166. <https://doi.org/10.1080/13598139.2012.735414>
- Spinath, B., Eckert, C., & Steinmayr, R. (2014). Gender differences in school success: what are the roles of students' intelligence, personality and motivation? *Educational Research*, 56(2), 230–243. <https://doi.org/10.1080/00131881.2014.898917>
- Steinmayr, R. & Spinath, B. (2008). Sex differences in school achievement: What are the roles of personality and achievement motivation? *European Journal of Personality*, 22(3), 185–209.
- Voyer, D., & Voyer, S. D. (2014). Gender differences in scholastic achievement: A meta-analysis. *Psychological Bulletin*, 140(4), 1174–1204.

Barbara Japelj Pavešić
Educational Research Institute
Gerbiceva 62
1000 Ljubljana
Slovenia
Barbara.Japelj@pei.si

Gašper Cankar
National Examinations Centre
Kajuhova ulica 32u
1000 Ljubljana
Slovenia
Gasper.Cankar@ric.si