



Centralised Lithuanian Language and Literature Assessments of Secondary School Students: Population Analysis

Rimantas Želvys¹, Saulė Raižienė², Jogaila Vaitekaitis³, Rita Dukynaitė⁴, Audronė Jakaitienė⁵

¹ Vilnius University, Institute of Data Science and Digital Technologies, Akademijos g. 4, LT-08412 Vilnius, Lithuania, rimantas.zelvys@fsf.vu.lt

Vilnius University, Institute of Educational Sciences, Universiteto g. 9, LT-01513 Vilnius, Lithuania, rimantas.zelvys@fsf.vu.lt

² Vilnius University, Institute of Data Science and Digital Technologies, Akademijos g. 4, LT-08412 Vilnius, Lithuania, saule.raiziene@fsf.vu.lt

Vilnius University, Institute of Psychology, Universiteto g. 9, LT-01513 Vilnius, Lithuania, saule.raiziene@fsf.vu.lt

³ Vilnius University, Institute of Data Science and Digital Technologies, Akademijos g. 4, LT-08412 Vilnius, Lithuania, jogaila.vaitekaitis@fsf.vu.lt

Vilnius University, Institute of Educational Sciences, Universiteto g. 9, LT-01513 Vilnius, Lithuania, jogaila.vaitekaitis@fsf.vu.lt

⁴ Vilnius University, Institute of Data Science and Digital Technologies, Akademijos g. 4, LT-08412 Vilnius, Lithuania, Rita.Dukynaite@mif.vu.lt

⁵ Vilnius University, Institute of Data Science and Digital Technologies, Akademijos g. 4, LT-08412 Vilnius, Lithuania, audrone.jakaitiene@mf.vu.lt

Annotation. The aim is to analyse assessments of Lithuanian language and literature for the entire secondary school student population focusing on gender comparison, and development of models for the prediction of achievements. We use the results of the 10th grade test for the period 2011–2015 and the same cohorts in the Matura exam. Gender differences are observed in analysed national centralized assessments. The distribution of the 10th grade test achievements resembles the bell-shaped curve, whereas the Matura – do not.

Keywords: *Lithuanian language and literature, population, 10th grade test, Matura examination, achievement.*

Introduction

External forms of assessment – standardised tests and centrally set examinations (also known as school exit exams, leaving exams, State exams) – are becoming increasingly important across Europe. The Eurydice (2009) report revealed that during the 2008/2009 academic year only the German-speaking communities of Belgium, the Czech Republic, Greece, Wales, and Liechtenstein did not employ national examinations and/or testing in secondary education. Instead, continuous student assessment was carried out internally. The reasons for the introduction of external assessment are: education policy changes leading to further decentralisation of education systems, increasing school autonomy, providing opportunities for school choice, and striving for better teaching quality. The external assessment provides an opportunity to compare student achievement among schools and regions. Most of the countries have established national agencies for educational assessment and evaluation, which are responsible for managing the examination process. In many European countries, national assessment is compulsory. In countries where it is voluntary, the majority of students still tend to participate to receive feedback about their level of achievement. The most frequent examination subjects are native language and mathematics, followed by foreign language and science. Usually, schools can use assessment results at their own discretion, but sometimes they must present assessment results in their internal or external evaluation reports.

Extensive research on external assessment in secondary education started more than two decades ago. In one research study Bishop (1999) found that countries which carry out centralised assessment of student progress, achieve better results in ILSAs (international large-scale student assessment studies). Woessman (2002), conducted a series of research studies on the effects of national assessment and came to the same conclusion. The author analysed the data of TIMSS 2015 and TIMSS-Repeat 1999. The results showed that in countries where centralised examination systems exist students show better results in ILSA when compared to countries without central examinations. The author concluded that the existence of centralised examinations reduces differences among students from different social backgrounds. In a later study Woessman (2005) also analysed IEA Reading Literacy Study 1991 and PISA 2000 data. The study confirmed that the existence of centralised examinations leads to better results of student achievement in ILSAs. However, the effect varies in relation to different ability groups. Centralised assessment has a greater positive influence on high achieving students compared to low achievers. High achievers tend to perform better in order to qualify for further studies in tertiary education, while low achievers have different motivation. Woessman (2005) also noted that regular testing in the lower grades can have a positive impact on the Matura examinations as it provides additional information about student achievement and enables further corrective actions. Woessman (2016) observed that besides the presence of a system of centralised examinations, the existence of private schools, larger classes,

good teaching facilities, a longer study year, and a higher education level of teachers also have a positive impact on student achievement. Providing greater autonomy to schools does not always have positive results in terms of quality when there are no centralised examinations in the country. Fischbach et al. (2013) tried to clarify to what extent student achievement in PISA could predict success in Matura examinations. 1442 secondary school students in Luxembourg participated in the study. The authors used the PISA 2006 data. Research indicated that PISA results allow the prediction of examination outcomes, but the relationship is not very strong and varies from weak to medium. In his latest study Woessman (2018) noted that additional funding or reducing class size does not always lead to expected results. The author also observed that centralised examinations encourage private tutoring, therefore student achievement is not always determined solely by the quality of teaching at school. Together with increased private tutoring and subsequent academic and social segregation (European Commission/EACEA/Eurydice, 2020) there are other unintended consequences of centralised examinations that will be outlined below.

Considered high-stakes¹ testing, national school leaving examinations along with outcomes mentioned above, carry negative unintended outcomes. Findings by Holme et al. (2010) point out that the impact of high school exit exams produces few of the expected benefits and are associated with losses for the most disadvantaged students by increasing drop-out rates and delays in graduation for non-Whites and economically disadvantaged students. Jones (2007) highlights that using tests as a means of holding educators accountable has a negative effect on instruction by narrowing down the curriculum. Also, school exit exams affect student and teacher motivation (remove intrinsic motivation and leave mainly external motivation regarding teaching and learning). Jones (2007) adds that because of these unintended outcomes, at-risk students (in terms of poverty, race, ethnicity, disability, and limited language proficiency) are at an even greater disadvantage. These findings are consistent with the research of Elliot et al. (2018) stating that high scoring PISA and TIMSS Asian countries usually have high-stakes national examination systems in which children are exposed to extreme competition and feel constant high level pressure: “High stakes exit exams leave a large number of students classified as failures with lack of confidence, elevated anxiety and other negative self-concept consequences” (Elliott et al., 2018, 142).

Tampayeva (2015) focused her study on national testing in Kazakhstan. The author noted that assessment reforms in post-soviet countries have their own specificity. In particular, national testing at the end of the secondary school in Kazakhstan was introduced in order to prevent corruption in admission to higher education institutions. In other words, it was not so much an educational, but more a moral problem which the government tried to solve. The author concluded that the assessment reform

¹ A high-stakes test is one that carries important implications for students, teachers, schools, or regions.

contributed to solving the problem of corruption, but did not have positive effects on student achievement. One of the possible reasons is that the community of teachers did not support the introduction of the system of centralised assessment of students in the country. Piattoeva (2015) examined political implications and practical application of the Matura examination results in Russia. In particular, the Matura results are widely used for the evaluation of teachers. The Matura results are treated as indications of teachers' pedagogical professionalism and consequently influence their professional qualification category and the level of payment. However, critics hold teachers' evaluation on the basis of the Matura results responsible for distortions (e.g., corruption, falsification of results, leakage of correct answers, tacit acceptance of cheating, etc.) found in the implementation of the exam.

While acknowledging the disadvantages of external examinations, we, as a team of interdisciplinary scientists, acknowledge, that data gathered from these tests enable us to identify not only educational issues, but social ones as well. A strong advantage of standardised national testing is the possibility to analyse, compare and distinguish discrepancies among different student groups. An example of this kind of research is the work of Schildkamp et al., (2012). These authors looked at the Netherlands secondary education exit examinations which consisted of internal school-based assessment and an external national assessment. Scientists investigated the discrepancy between school and central examination grades for different groups characterised by ethnicity, gender, or socioeconomic status (SES). Subjects covered in this study included Dutch language, modern foreign languages, factual subjects like geography and history, also economic and science subjects. Researchers concluded that the discrepancies for some student groups are too high. For example, girls show a larger overall discrepancy than boys for all subjects, with the greatest gender based discrepancies observed in modern foreign languages and economic subjects. Looking at the discrepancies between school and central examinations, Schildkamp et al. (2012) show, that in general there were higher grades for school examinations than for central examinations. It is interesting to note, that if significant differences between the results of the school examination and central examinations are observed, the Dutch inspectorate considers this a non-acceptable threat to equity (Schildkamp et al., 2012, 230).

Lithuanian researchers look into various aspects of the State Matura examinations in mathematics and informatics (Blonskis et al., 2008; Kaminskienė et al., 2012; Dagiienė et al., 2017), technologies (Numgaudienė & Ramanaukaitė, 2012), history (Arlauskaitė-Bulovienė & Šiaučiukėnienė, 2006), and others. The Matura examinations in Lithuanian language seldom attract the attention of researchers. The studies that we found are mostly focused on the analysis of common mistakes, content or the quality of examination tasks: Jackūnas, 1994; Daujotytė, 1997; Saliene, 2002, 2005, 2013; Nauckūnaitė, et al., 2008; Nauckūnaitė, 2011, 2014; Bredelis et al. 2013; Smetonienė & Petrėnienė, 2016; Tamulionienė, 2018. More attention to the Matura and the 10th grade test result analysis

is given by institutions subordinate to the Ministry of Education, Science and Sports. For example, annual reports “Educational Status Review”² analyse the Matura and the 10th grade test results. The 2016 special issue³ of the Educational Status Review was dedicated solely to student achievements in national examinations. Also, the Ministry issues Educational briefs some of which review the Matura and the 10th grade test results (Vaicekauskienė, 2011; Bakonis et al., 2018; Jevsejevienė, 2019; Bakonis, 2020). Notrimaitė et al. (2012) working at the National examination centre (Lith. NEC) conducted a statistical analysis of the 10th grade Lithuanian language and literature test for the year 2011.

We encounter a problem that there are no studies that compare and analyse the population examination results of both the 10th grade and the Matura exams at student level longitudinally. Thus, the aim of our study is to analyse the Lithuanian language and literature Matura exam and the 10th grade test for the period of five years (the 10th grade test 2012–2016; the Matura 2014–2018), focusing on gender comparison, and the development of models for prediction of achievements.

Current study

This article examines, arguably, the most important centralised student assessments in Lithuanian general education – the 10th grade tests and the Matura exams. Our research focuses on the subject of Lithuanian language and literature (similar analysis of mathematics exams has already been done by Jakaitienė, et al. (2021)).

The 10th grade test (full title – Test of Basic Education Achievement) is designed to provide pupils and schools with information about learning outcomes and help in deciding further pupil learning prospects. The 10th grade Lithuanian language and literature test is intended to assess students’ knowledge, understanding and skills in Lithuanian language, literature, and culture achieved during the implementation of the general program of lower secondary education. One must note that the 10th grade test tasks are prepared centrally, but are assessed by local teachers.

The Matura exam is designed to assess pupils’ competencies and help higher education institutions transparently select prospective students. The Lithuanian language and literature Matura exam must be passed by all students in order to complete the secondary education program (students choose the State Matura examination or the School Matura examination) and receive the Matura certificate. The State Matura Examination is conducted and assessed centrally. Over the analysed period the assessment of both exams was criterion-based.

² NMVA, Švietimo būklės apžvalgos [National School Assessment Agency, Educational Status Review]. Retrieved from <http://www.nmva.smm.lt/politikos-analize/svietimo-bukles-apzvalgos/> [2021-03-03].

³ Lietuva. Švietimas šalyje ir regionuose. 2016. Mokinių pasiekimai. [Lithuania. Education in the country and regions. 2016. Student Achievements]. Retrieved from https://www.smm.lt/uploads/documents/tyrimai_ir_analizes/2019/Lietuva_%C5%A0vietimas%20%C5%A1alyje%20ir%20regionuose%202016_Mokini%C5%B3%20pasiekimai.pdf [2021-03-03].

Research Methodology

Data

We use individual level data for the entire Lithuanian secondary school student population (except vocational schools), who have taken the Matura examinations of Lithuanian language and literature in the period 2014–2018⁴. Along with the results of the Matura exam, we analyse the results of the 10th grade test taken by the same students two years earlier. The data were provided by the Education Management Information System (EMIS, Lith. ŠVIS) of the Ministry of Education, Science and Sports. We analyse the distribution of students' achievements for the 10th grade tests and the Matura examinations, respectively, according to the year of taking the exam, as well as gender, school location, and ownership (state, municipal and private). Differences in achievements of urban and rural schools were compared according to school location in five groups: Vilnius (capital of Lithuania), large cities (Kaunas, Klaipėda, Šiauliai, Panevėžys), cities (15–100 thousand inhabitants), small cities (3–15 thousand inhabitants), rural areas (< 3 thousand inhabitants). In the study, we examine student achievement without analysing the quality and content of exam tasks.

Methods

In this article, we analyse population data where the number of records varies from 17,560 to 37,547 depending on the type of exam and academic year (Table 1). It is important to note that different assessment scales are used for the 10th grade tests and the Matura exams. The 10th grade tests are evaluated on a 10-point scale, while the Matura exam is assessed on a 100-point scale. We present the following descriptive statistics according to selected factors: mean, standard deviation (SD), interquartile range (IQR), minimum, first quartile (Q1), median (Q2), third quartile (Q3), maximum and skewness coefficient (Skew). We report results for population data, therefore all calculated parameters are population parameters for which standardized errors (SE) are presented. In the study, we assess normality using visual representation and the value of skewness coefficient. We start analysis from simple linear dummy (factor variables coded 0/1) regression models for each factor: gender, school location, and ownership. From the latter analysis, we investigate the coefficients of determination, which will allow us to quantify the importance of each factor for Lithuanian language and literature achievements. Next, we will combine these factors with additional student-level context factors (age, social support indicator, special needs indicator, and foreigner status) into multiple linear regression models. Student-level context factors are described together with prediction models. From these models, we will judge about the suitability of the selected variables for predicting

⁴ The data for the implementation purposes of project EFECTAS (<https://www.efectas.projektas.vu.lt>) were obtained in 2019 and does not examine the most recent data.

Lithuanian language and literature achievement for each school year. Beyond this, the ICC intra-school variation coefficient is calculated to estimate the proportion of variance explained by school differences in achievement. All statistical analysis was performed using R version 3.6.3 and RStudio version 1.2.5033.

Research results

In this paper, we analyse the achievements in Lithuanian language and literature in five cycles (Table 1), i.e., students who took the 10th grade test from 2011–2012 to 2015–2016 academic years, and the results of the same cohort participating in the Matura examination two years later. Due to the declining population, we observe a decreasing trend of participation: 1.8–3.1% do not attend or are exempted from the 10th grade test and, respectively, below 0.6% from the Matura examination.

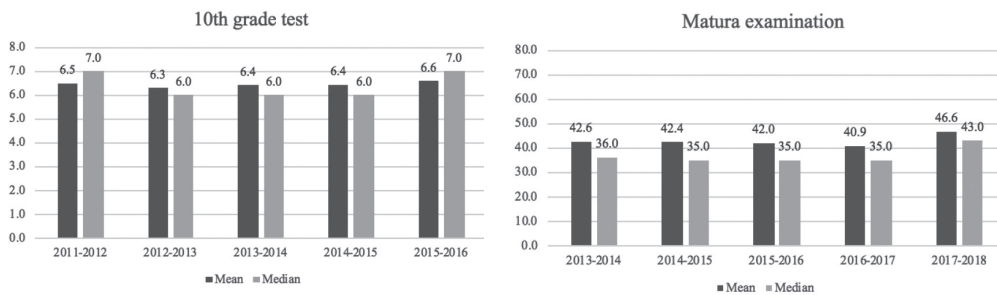
Table 1
Student Population of Analysed Secondary Schools in Lithuania That Took the Lithuanian Language and Literature Matura Examination in the Years 2014–2018

	1 cycle	2 cycle	3 cycle	4 cycle	5 cycle
	2011–2012	2012–2013	2013–2014	2014–2015	2015–2016
The 10th grade test	37547	36577	33039	32015	31148
Participated, N	36500	35449	32050	31246	30581
Participated, %	97.2	96.9	97	97.6	98.2
Annual growth, %		-2.6	-9.7	-3.1	-2.7
Absence, N	1047	1128	989	769	567
Absence, %	2.8	3.1	3	2.4	1.8
	2013–2014	2014–2015	2015–2016	2016–2017	2017–2018
The Matura examination	21245	20190	18510	17918	17560
Participated, N	18869	18226	16717	16079	16116
Participated, %	88.8	90.3	90.3	89.7	91.8
Annual growth, %		-5	-8.3	-3.2	-2
Failed, N	2266	1897	1691	1753	1363
Failed, %	10.7	9.4	9.1	9.8	7.8
Absence, N	110	67	102	86	81
Absence, %	0.5	0.3	0.6	0.5	0.5

The average score of the 10th grade test is always higher than the average score of the Matura examination (divided by 10) (Figure 1). The average score of the Matura examination has some upward bias compared to the median, indicating that high grades elevate results. We observe the mode equal to 6 or 7 for the 10th grade test (Figure 2 and Table 2). The histograms of the Matura examination have a peak between 25–30 points (Figure 2, B). Roughly half of the grades of the Matura examination are below 28–42 points on a 100-point scale (Table 3).

Figure 1

Average and Median Dynamics of the Lithuanian Language Achievements for the 10th Grade Test and the Matura Exams

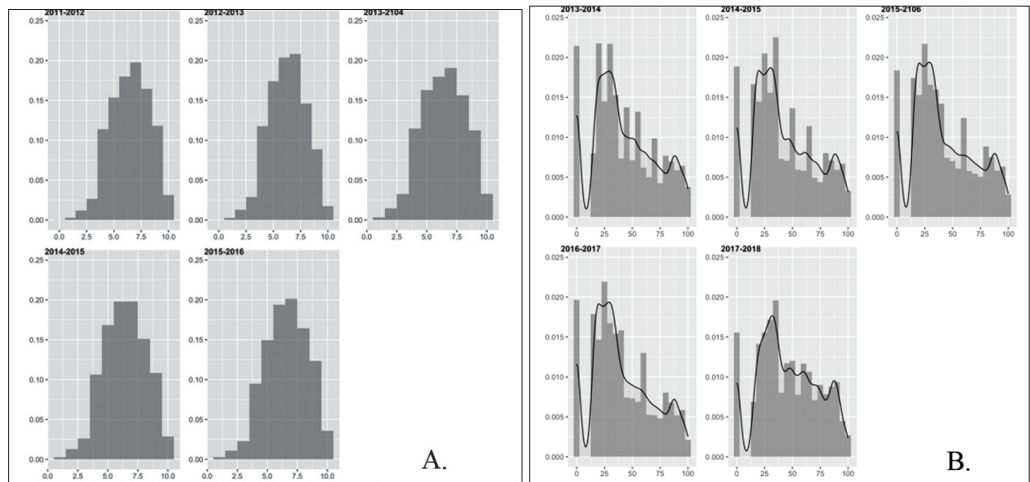


As previously notes, different assessment scales are used for the 10th grade tests and the Matura exams. The 10th grade test is evaluated on a 10-point scale, while the Matura exam is assessed on a 100-point scale. We note that on a 100-point scale there are no observations between 1 point and 15 points of the Matura exam (if a student gets 1 to 15 points – he or she fails the exam). Thus, we are left with a data gap (Figure 2).

We find that the results of the 10th grade test are similar to a normal distribution (Figure 2, A), however the Matura examination does not follow the normal distribution (Figure 2, B). Instead of negative skewness of achievement distribution, we observe a positive skew for the Matura examination, which is contrary to what is expected. Since the distributions of the Matura examination achievements do not correspond to the normal distribution, one should analyse the median as a characteristic of the centre instead of the mean. As a result, we will present both characteristics, but the analysis will be based on the median comparison in further analysis.

Figure 2

The Distributions of Grades for: A. - the 10th Grade Test; B. - the Matura Examination



Although we performed the analysis of students' achievements with respect to gender, school location, and ownership, in the paper we present a detailed analysis of gender differences only, as we detected non-overlapping distributions between boys and girls. The distributions of student examination achievement overlap by school location and ownership.

Table 2
Descriptive Summary of the 10th Grade Test Achievements According to Gender

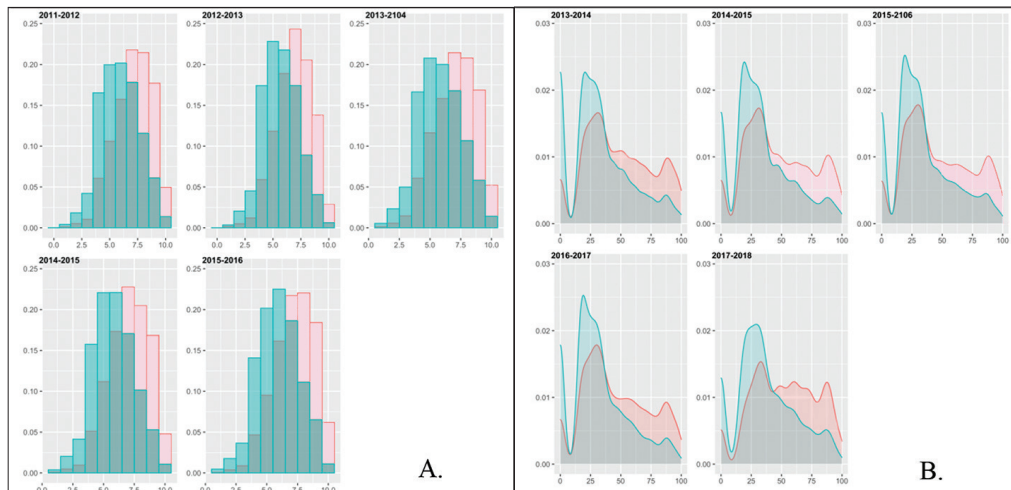
	Mean	SD	IQR	Min	Q1	Q2	Q3	Max	Skew	N
2011–2012										
Total	6.48	1.82	3	0	5	7	8	10	-0.18	36500
Female	7.09	1.68	2	0	6	7	8	10	-0.43	17979
Male	5.89	1.74	2	0	5	6	7	10	0.03	18521
2012–2013										
Total	6.27	1.72	3	0	5	6	8	10	-0.13	35449
Female	6.88	1.6	2	1	6	7	8	10	-0.36	17448
Male	5.69	1.63	2	0	5	6	7	10	0.06	18001
2013–2014										
Total	6.41	1.85	3	1	5	6	8	10	-0.16	32050
Female	7.03	1.72	2	1	6	7	8	10	-0.39	15703
Male	5.8	1.77	2	1	5	6	7	10	0.06	16347

	Mean	SD	IQR	Min	Q1	Q2	Q3	Max	Skew	N
2014–2015										
Total	6.42	1.78	3	1	5	6	8	10	-0.13	31246
Female	7.07	1.64	2	1	6	7	8	10	-0.36	15009
Male	5.81	1.68	2	1	5	6	7	10	0.06	16237
2015–2016										
Total	6.58	1.78	3	1	5	7	8	10	-0.2	30581
Female	7.22	1.64	2	1	6	7	8	10	-0.42	14955
Male	5.96	1.7	2	1	5	6	7	10	-0.03	15626

With respect to the 10th grade test, the median of girls is one unit larger than or equal to the median of boys (see Table 2). About half of the students obtained less than 6 or 7, indicating sufficient knowledge of the Lithuanian language and literature in grade 10. We observe that the distribution of girls' achievements has slightly negative skewness (Figure 3 and Table 2) which indicates that they received higher scores more frequently than boys. The distribution of boys' achievements is along with the normal distribution. The proportion of boys and girls participating in the 10th grade test is almost equal and stable in every academic year analysed. The gender variable explains 10.8%–12.5% of the variation in achievements in the 10th grade test.

Figure 3

The Distributions of Grades With Respect to Gender for: A. - the 10th Grade Test; B. - the Matura Examination. The Distribution for Boys is in Green, as for Girls in Red



Regarding the Matura examination, the gender gap in achievement is even more pronounced. The median of girls is 44.9%–64.3% higher than the median of boys (see

Table 3). The gap is more visible comparing Q3 between genders. Q3 of boys is remarkably like Q2 of girls. The variable gender explains 6.0%–8.2% of the variation in achievement in the Matura examination. The proportion of boys (40%) and girls (60%) participating in the Matura examination is very nearly stable in every academic year analysed.

Table 3
Descriptive Summary of the Matura Examination Achievements According to Gender

	Mean	SD	IQR	Min	Q1	Q2	Q3	Max	Skew	N
2013–2014										
Total	42.59	27.43	39	0	22	36	61	100	0.36	21135
Female	48.90	27.25	43	0	28	46	71	100	0.17	12620
Male	33.23	24.90	28	0	18	28	46	100	0.67	8515
2014–2015										
Total	42.35	27.19	40	0	23	35	63	100	0.41	20123
Female	48.21	27.38	44	0	27	45	71	100	0.2	12074
Male	33.57	24.40	31	0	17	29	48	100	0.74	8049
2015–2016										
Total	41.99	27.13	41	0	21	35	62	100	0.44	18408
Female	47.47	27.50	44	0	26	42	70	100	0.24	10934
Male	33.98	24.47	29	0	17	29	46	100	0.73	7474
2016–2017										
Total	40.92	26.59	37	0	21	35	58	100	0.46	17830
Female	46.81	26.94	43	0	27	42	70	100	0.25	10392
Male	32.68	23.75	29	0	17	29	46	100	0.76	7438
2017–2018										
Total	46.61	26.45	40	0	27	43	67	100	0.14	17477
Female	53.01	25.81	41	0	33	52	74	100	-0.08	10199
Male	37.64	24.66	33	0	21	33	54	100	0.48	7278

To understand the driving factors behind the results of the Lithuanian language and literature, one might want to develop models that explain the variation in achievements. Therefore, we construct multivariable linear regression models for each exam and academic year separately (Table 4 and Table 5). Age, gender, school location, school ownership, social support indicator, special needs indicator, and foreigner status are explanatory variables for the achievements of the Lithuanian language and literature. The latter models with high goodness-of-fit could serve for the prediction of achievements and could be valuable in personalised education.

We observe the negative association between the age of students and achievements in the Lithuanian language and literature. The median age of students is equal to 16 years (min 13, max 70) for the 10th grade test and, respectively, the median is equal to 18 years (min 15, max 46) for the Matura examination. Postponing the examination time for either the 10th grade test or the Matura examination generally diminished the achievement levels.

As discussed above, the gender gap is pronounced in favour of females in the results of both examinations while controlling for other independent variables. Gender has the largest effect from all explanatory variables for the Matura examination achievements. The 10th grade and the Matura achievement levels are slightly higher in urban areas than in rural areas having other independent variables fixed. We estimate the largest, yet small, positive effects for students from Vilnius schools compared to other locations. Attendance at a private school leads to higher achievements in both assessments. Also, in both tests, the achievements of students with Lithuanian citizenship are higher compared to students with foreign citizenship.

The achievements are marginally lower for students that need social support. The indicator according to the Law on Social Assistance to Pupils distinguishes two forms of social support for learners: the provision of free school meals (breakfast, lunch, dinner, and meals in summer camps organised by schools); and the provision of basic school supplies. Pupils have the right to free school meals and support for the purchase of basic school supplies if the average income for family members is less than 1.5 of the state-supported income. Other cases (related to sickness, accident, loss of the breadwinner, provision of assistance to a pupil of disabled parents or from a family with three or more children, etc.) are subject to the decision of a municipality's council (Eurydice, 2020). We had no possibility to differentiate between these two forms of social support.

The special needs indicator is an independent variable which represents students with special needs that are provided with complete or partial inclusion (in regular classes or special classes within mainstream municipal schools). We estimate the strongest negative association between the achievements and special needs variable for the 10th grade test. For the results of the Matura examination, it is of smaller importance compared to gender and foreign citizenship variables.

Overall, all selected explanatory variables explain 19.8%–24.7% of the variation for the 10th grade test achievements, and, respectively, 8.8%–10.9% of the variation for the Matura achievements. We also calculated the intra-school variation coefficient (ICC) to estimate what proportion of achievement variance is explained by school differences. The differences between schools explain 26–28% of the variation in the 10th grade test achievements and up to 18 percent of the variation in the Matura examination achievements.

Table 4

Result From Multivariable Linear Regression for the 10th Grade Test Achievements (1–10 Grade Scale)

Independent variable	Category	2011-2012		2012-2013		2013-2014		2014-2015		2015-2016	
		β	SE	β	SE	β	SE	β	SE	β	SE
Age		-0.142	0.005	-0.094	0.004	-0.12	0.004	-0.119	0.004	-0.130	0.004
Gender	Female	Reference category									
	Male	-1.108	0.017	-1.121	0.016	-1.14	0.019	-1.160	0.018	-1.150	0.018
School location	Rural area	Reference category									
	Vilnius	0.548	0.028	0.395	0.027	0.466	0.030	0.459	0.029	0.384	0.029
	Large cities	0.456	0.024	0.309	0.024	0.376	0.027	0.272	0.026	0.321	0.026
	Cities	0.134	0.027	0.062	0.026	0.046	0.029	0.038	0.029	0.130	0.028
	Small cities	0.297	0.026	0.133	0.025	0.221	0.028	0.229	0.028	0.240	0.028
School type	Municipal	Reference category									
	State	0.271	0.065	0.095	0.060	0.013	0.067	0.099	0.062	0.230	0.060
	Private	0.705	0.069	0.807	0.064	0.904	0.065	0.997	0.057	0.947	0.058
Social support indicator		-0.549	0.024	-0.521	0.023	-0.600	0.028	-0.568	0.028	-0.590	0.029
Foreigner status		-0.386	0.172	-0.177	0.138	-0.640	0.151	-0.548	0.132	-0.650	0.115
Special needs indicator		-1.844	0.041	-1.663	0.042	-1.710	0.042	-1.390	0.035	-1.702	0.035
Adjusted R ² , %		20.8		19.8		20.7		22.5		24.7	
ICC		26.0		26.3		27.8		27.9		27.7	

SE – standard error of β .

School location: Vilnius (capital of Lithuania), large cities (Kaunas, Klaipėda, Šiauliai, Panevėžys), cities (15–100 thous. inhabitants), small cities (3–15 thous. inhabitants), rural area (<3 thous. inhabitants).

Table 5
Result From Multivariable Linear Regression for the Matura Examination Achievements (0–100 Grade Scale)

Independent variable	Category	2013-2014		2014-2015		2015-2016		2016-2017		2017-2018	
		β	SE	β	SE	β	SE	β	SE	β	SE
Age		-1.98	0.26	-2.09	0.26	-2.44	0.38	-1.89	0.31	-1.8	0.33
Gender	Female	Reference category									
	Male	-15.96	0.37	-15.00	0.37	-13.80	0.39	-14.43	0.38	-16.00	0.38
School location	Rural area	Reference category									
	Vilnius	9.81	0.61	12.60	0.62	10.30	0.66	9.46	0.66	7.50	0.66
	Large cities	9.01	0.58	7.53	0.58	8.02	0.62	7.99	0.63	8.25	0.62
	Cities	5.15	0.61	5.79	0.62	6.24	0.66	5.44	0.67	5.82	0.66
	Small cities	5.08	0.60	3.52	0.61	4.14	0.65	4.57	0.66	4.49	0.65
School type	Municipal	Reference category									
	State	-1.44	1.35	4.40	1.33	-0.19	1.39	2.16	1.33	1.48	1.24
	Private	11.46	1.17	13.00	1.09	12.00	1.17	13.57	1.07	15.20	1.09
Social support indicator		-4.80	0.65	-5.31	0.7	-5.36	0.8	-5.92	0.88	-5.20	0.94
Foreigner status		-6.34	4.34	-4.78	3.35	-18.7	4.21	-6.98	3.12	-7.30	2.89
Special needs indicator		-11.48	2.75	-6.01	1.51	-8.05	2.86	-5.69	3.14	-10.00	2.87
Adjusted R ² , %		10.5		10.8		8.8		9.8		10.9	
ICC		16.1		16.1		13.2		14.1		18.0	

SE – standard error of β .

School location: Vilnius (capital of Lithuania), large cities (Kaunas, Klaipėda, Šiauliai, Panevėžys), cities (15–100 thous. inhabitants), small cities (3–15 thous. inhabitants), rural area (<3 thous. inhabitants).

Discussion

In this paper, we analysed achievements of the Lithuanian language and literature of the 10th grade test and the Matura exam in five cycles. We focused on gender comparison, and developed linear regression models for the prediction of achievements. From the analysis, we observe that the 10th grade test is similar to a bell-shaped curve while the Matura exam does not follow the normal distribution. Royal and Guskey (2015) remind us, that the bell-shaped curve is largely based on the belief that intelligence test scores look like a normal distribution and thus, grade distributions on tests must also resemble a bell-shaped curve. As an illustration of this statistical theory Royal and Guskey (2015) suggest the metaphor of a crop yield: if nothing in nature intervenes, one could imagine that crop yields sometimes are high and sometimes – low, but usually – average. If someone intervenes, say, by adding fertiliser, the distribution of results is likely to be different. Fertiliser would generate high crop yield, thus changing the shape of the bell curve to be negatively skewed. Teaching could be compared to a fertiliser intended to help students grow. Thus, criterion-referenced assessments, unlike norm-referenced ones, should not resemble a normal distribution. All examinations have used criterion-referenced assessment systems and achievement distribution might be negatively skewed, i.e., a mean should be smaller than the median.

We note that the Matura exam is assessed on a 100-point scale and there are no observations between 1 point and 15 points (if a student gets 1 to 15 points – he or she fails the exam). Thus, we are left with a data gap. We detect large numbers of students classified as failures, which is not transparent or fair to students, or appropriate from the perspective of modern educational sciences. Also, distinct scales complicate comparison between exams. Additionally, an analysis of discontinued distribution is complicated.

This study has identified that girls demonstrate better results than boys in both the 10th grade and the Matura Lithuanian language and literature assessments. Furthermore, boys were overrepresented as low-achievers, and girls as high-achievers. Our results are in line with the ILSA studies, which also show that in Lithuania, as in many other countries, girls' reading proficiency is higher than boys' for both fourth graders (Mullis et al., 2017) and for the fifteen years olds (OECD, 2019). Although it is acknowledged that males and females do not differ in intelligence (Halpern, 2000), they do differ in reading and writing skills in a variety of ways during different levels of schooling. This study has assessed, that the gender gap in reading and writing is stable over the analysed period of five years (2012–2018). The results of the PISA survey show that in Lithuania from 2009 to 2018 the gender gap in reading narrowed, due to improved boys reading abilities while girls' scores did not change (OECD, 2019). The results of our study do not confirm this. A similar conclusion was made by Reilly et al. (2019) in a meta-analysis investigating the effect of gender on U.S. students' reading and writing achievement from the National Assessment of Education Progress (NAEP) for the period 1988–2015.

Thus, while recognising the universality of the gender gap in reading and writing, it is necessary to further examine the biological and sociocultural factors that contribute to gender differences and to seek educational interventions that would effectively provide opportunities to improve reading and writing proficiency for both genders leading to equity in education.

The current study revealed that achievement distributions are overlapping for different school locations. However, ILSA studies show a significant urban/rural gap. Zabulionis (2020) noted that the difference in results between urban and rural schools in PISA 2018 was as large as 60 points. During the entire period of Lithuania's participation in the PISA study (2006–2018) achievements of students in rural schools were the lowest and did not change over time. Zabulionis (2020) concludes that, according to the PISA study, the gap between urban and rural schools is increasing. The policy paper of the Ministry of Education and Science (2011) also noted differences in student achievement in urban and rural schools during the 2003, 2005 and 2007 national testing of 8th grade students. However, since 2003 the results of students in cities and regional centres remained more or less at the same level while the results of students in small towns and villages improved (MoES, 2011). We demonstrate that ILSA studies and national examination and testing may reflect different tendencies as they follow different methodological approaches. We tend to agree with Reeves and Bylund (2005) that educational research does not provide clear evidence that rural schools are inferior to urban schools, so there is a need for more research on urban-rural differences in other school quality factors (Othman & Muijs, 2013).

For achievements in the Lithuanian language and literature, gender had the highest prognostic value in the compiled prognostic models considering age, school location and ownership, social support indicator, special needs indicator, and foreigner status. However, it should be noted that all selected explanatory variables explain up to $\frac{1}{4}$ of the variation for the 10th grade test achievements, and half less of the variation for the Matura achievements. This indicates that the developed multiple linear regression models, despite embedding important variables, lack precision for the prediction of Lithuanian language and literature achievements and especially for the Matura examination. Thus, we need more student-level variables (such as cognitive abilities, motivational aspects, variables reflecting social, economic, and cultural status) to explain the variation in achievement.

In addition, one might consider the hierarchical structure of educational data while the students are nested in classes and the classes in schools. The results of the 10th grade and the Matura Lithuanian language and literature exams confirm that students' achievements are partially related to schools. This is in line with results from PISA studies (Brunner et al., 2018) and could reflect school differences in student composition, school policies on instruction or resources (OECD, 2006). It is interesting that the differences between schools explain more variation of the 10th grade test than of the Matura examination achievements (respectively, 26.0%–27.9% variation of the 10th grade test and

13.2%–18.0% of the Matura examination). This difference can be accounted for in part by the differences in organisation of these exams: the 10th grade test tasks are prepared centrally, but are assessed by local teachers. It might be that the greater centralisation of exams could diminish the influence of school factors on the results of exams in Lithuania. However, more evidence is needed about the impact of different forms of exam organisation on exam achievement.

Conclusions

In this article, we examine the results of the 10th grade test and the Matura exams without analysing the content and quality of the exam tasks. From the study, we see somewhat different tendencies in national centralised examination results and ILSA studies. A gender gap in achievements is observed in both ILSA and national assessments; however, the urban/rural gap is present in ILSA only. We analyse the national assessment data for the entire Lithuanian population of secondary school students that have no sampling errors. Both the analysed centralised assessment examinations serve different purposes and rely on different methodological principles, however each examination provides valuable information about students' literacy and should be used for improvements in educational effectiveness and policy development.

Acknowledgments

This research is funded by the European Social Fund according to the activity Improvement of researchers' qualification by implementing world-class R&D projects of Measure No.09.3.3-LMT-K-712. The project No. DOTSUT-39 (09.3.3-LMT-K-712-01-0018) / LSS-250000-57.

References

- Arlauskaitė-Bulovienė, V., & Šiaučiukėnienė, L. (2006). Interpretacinių gebėjimų ugdymo rezultatų įvertinimas mokantis istorijos. *Pedagogika*, 84, 101–107.
- Bakonis, E. (2020). Lietuvos miestų ir kaimų mokinių pasiekimai: skirtumai, jų priežastys, galimi įveikos būdai (2020 m. rugsėjis) [Achievements of Lithuanian urban and rural students: differences, their causes, possible ways of overcoming them (September 2020)]. Retrieved March 3, 2021, from <https://www.nsa.smm.lt/wp-content/uploads/2020/12/Miesto-ir-kaimo-mokiniu-pasiekimai.pdf>.

- Bakonis, E., Balevičienė, S., Gražytė-Skominienė, A., Jevsejevienė, J., Kalvaitis, A., Mikėnė, S., Paurienė, L., & Zablackė, R. (2018). *Lietuvos tautinių mažumų švietimo būklės analizė* (2018 m.) [Analysis of the state of education of Lithuanian national minorities]. Retrieved March 3, 2021 from <https://www.nsa.smm.lt/wp-content/uploads/2020/07/Lietuvos-tautiniu-mazumu-svietimo-bukles-analize-2018-m..pdf>
- Bishop, J. H. (1999). Are national exit examinations important for educational efficiency? *Swedish Economic Policy Review*, 6 (2), 349–401.
- Blonskis, J., Dagienė, V., & Burbaitė, R. (2008). Lyginamoji informacinių technologijų valstybinių brandos egzaminų analizė. *Lietuvos matematikos rinkinys*, 48, 92–98.
- Bredelis, K., Lesauskienė, A., Nauckūnaitė, Z., Pociūtė-Abukevičienė, D., & Šventickienė, A. (2013). *Lietuvių kalbos ir literatūros brandos egzamino analizė* [Analysis of Lithuanian language and literature maturity exam], Retrieved February 21, 2021 from: https://www.nec.lt/failai/3738_2013_m_LT_VBE_analize.pdf
- Brunner, M., Keller, U., Wenger, M., Fischbach, A., & Lüdtke, O. (2018). Between-school variation in students' achievement, motivation, affect, and learning strategies: Results from 81 countries for planning group-randomized trials in education. *Journal of Research on Educational Effectiveness*, 11(3), 452–478.
- Dagienė, V., Vinikienė, L., Sutkutė, E. (2017). Pagrindinio ugdymo Lietuvos mokinių matematinių gebėjimų tyrimas. *Informacijos mokslai*, 79, 17–29.
- Daujotytė, V. (1997). Tema ar / ir tekstas? *Mokykla*, 10, 6–11.
- Elliott, J., Stankov, L., Lee, J., & Beckmann, J. F. (2019). What did PISA and TIMSS ever do for us?: The potential of large scale datasets for understanding and improving educational practice. *Comparative Education*, 55(1), 133–155.
- European Commission/EACEA/Eurydice. (2020). Equity in school education in Europe: Structures, policies and student performance. *Eurydice Report*. Luxembourg: Publications Office of the European Union.
- Eurydice (2009). *National testing of pupils in Europe: Objectives, organisation and use of results*. Brussels: Education, Audiovisual and Cultural Agency.
- Eurydice (2018). Assessment in general upper secondary education, Lithuania. Retrieved February 27, 2021 from https://eacea.ec.europa.eu/national-policies/eurydice/content/assessment-general-upper-secondary-education-31_en?2nd-language=lt
- Fischbach, A., Keller, U., Preckel, F., & Brunner, M. (2013). PISA proficiency scores predict educational outcomes. *Learning and Individual Differences*, 24, 63–72.
- Halpern, D. F. (2000). *Sex differences in cognitive abilities* (3rd ed.). Mahwah, NJ: Erlbaum.
- Holme, J. J., Richards, M. P., Jimerson, J. B., & Cohen, R. W. (2010). Assessing the effects of high school exit examinations. *Review of Educational Research*, 80(4), 476–526. doi:10.3102/0034654310383147
- Jackūnas, Ž. (1994). Lietuvių kalba reformuotoje Lietuvos mokykloje. *Gimtasis žodis*, 6, 1–3.

- Jakaitienė, A., Želvys, R., Vaitekaitis, J., Raižienė, S., Dukynaitė, R. (2021). Centralised mathematics assessments of Lithuanian secondary school students: population analysis. *Informatics in Education*. In press.
- Jevsejevienė, J. (2019). Kaip vertinti abiturientų mokymosi rezultatus (2019 m. lapkritis), [How to assess graduate learning outcomes (November 2019)]. Retrieved March 3, 2021, from <https://www.nsa.smm.lt/wp-content/uploads/2020/07/Kaip-vertinti-abiturientu-mokymosi-rezultatus.pdf>
- Jones, B. D. (2007). The unintended outcomes of high-stakes testing. *Journal of Applied School Psychology*, 23(2), 65–86.
- Kaminskienė, J., Rimkuvienė, D., & Raškinienė, D. (2012) LŽŪU pirmo kurso studentų matematinio pasirengimo analizė. *Pedagogika*, 105, 53–60.
- Mullis, I. V., Martin, M. O., Foy, P., & Hooper, M. (2017). *ePIRLS 2016: International results in online informational reading*. Boston College: International Association for the Evaluation of Educational Achievement.
- Nauckūnaitė, Z. (2011). Lietuvių kalbos ir literatūros dalyko samprata: vidurinio ugdymo kaitos diskursas. *Acta Paedagogica Vilnensia*, 27, 102–114.
- Nauckūnaitė, Z. (2014) Lietuvių kalbos ir literatūros ugdymo siekiniai ir jų įgyvendinimo retrospektyva. *Acta Paedagogica Vilnensia*, 33, 17–32.
- Nauckūnaitė, Z., Šventickienė, A., & Kubilinskiene, D. (2008). *Dalykinė analizė, 2008: Lietuvių gimtosios kalbos 2008 m. valstybinio brandos egzamino teksto rašymo darbų vertinimo dalykinė analizė*. Vilnius: NEC.
- Notrimaitė, G., Bigelienė, D., Dargytė, J., Melnikė, E., & Gudynas, P. (2012). 2011 metų pagrindinių ugdymo pasiekimų patikrinimo rezultatų analizė. Lietuvių kalba (gimtoji). Vilnius: NEC. https://www.egzaminai.lt/failai/2483_LT-G_PUPP_ANALIZE_2011.pdf
- Numgaudienė, A., & Ramanauskaitė, A. (2012). Technologijų mokomojo dalyko ir technologijų brandos egzamino svarba ateičiai: vidurinės ir profesinės mokyklos mokinių požiūris. *Profesinis rengimas: tyrimai ir realijos*, 22, 92–103.
- OECD (2006). *Between- and within-school variation in the mathematics performance of 15-year-olds. Education at a glance 2006: OECD indicators*. Paris: OECD Publishing. <https://doi.org/10.1787/eag-2006-6-en>.
- OECD (2019). *PISA 2018 results (volume II): Where all students can succeed*, PISA. Paris: OECD Publishing <https://doi.org/10.1787/b5fd1b8f-en>.
- Othman, M., & Muijs, D. (2013). Educational quality differences in a middle-income country: the urban-rural gap in Malaysian primary schools. *School Effectiveness and School Improvement*, 24(1), 1–18.
- Piattoeva, N. (2015). Elastic numbers: national examinations data as a technology of government, *Journal of Education Policy*, 30 (3), 316–334.
- Reeves, E. B., & Bylund, R. A. (2005). Are rural schools inferior to urban schools? A multilevel analysis of school accountability trends in Kentucky. *Rural Sociology*, 70(3), 360–386.

- Reilly, D., Neumann, D. L., & Andrews, G. (2019). Gender differences in reading and writing achievement: Evidence from the National Assessment of Educational Progress (NAEP). *American Psychologist*, 74(4), 445–458. <https://doi.org/10.1037/amp0000356>
- Salienė, V. (2002). Lietuvių kalbos ugdymo didaktinių nuostatų ir programos 11–12 klasėje turinio kaita ir dermės galimybės. *Žmogus ir žodis*, 1, 76–83.
- Salienė, V. (2013). Kalba kaip socialinis kultūrinis reiškinys pagrindinės ir vidurinės mokyklos gimtosios kalbos ugdymo sistemoje. *Pedagogika*, 109, 105–109.
- Salienė, V. (2005). Dvyliktos klasės mokinių lietuvių kalbos žinių ir gebėjimų analizė. *Pedagogika*, 76, 100–106.
- Schildkamp, K., & Rekers-Mombarg, L. T., & Harms, T. J. (2012). Student group differences in examination results and utilization for policy and school development. *School effectiveness and school improvement*, 23(2), 229–255.
- Smetonienė, I., & Petrėnienė, O. (2016). Lietuvių kalbos ir literatūros valstybinio brandos egzamino rašinių stiliaus problemos. *Žmogus ir žodis*, 18(1), 30–45.
- MoES (2011). *Kokie ugdymo skirtumai kaimo ir miesto mokyklose? Švietimo problemos analizė*. Vilnius: Švietimo aprūpinimo centras.
- Tampayeva, G. Y. (2015). Importing education: Europeanisation and the Bologna process in Europe's backyard – the case of Kazakhstan. *European Educational Research Journal*, 14(1), 74–85.
- Tamulionienė, A. (2018). Skyrybos ypatumai lietuvių kalbos ir literatūros valstybinio brandos egzamino darbuose. *Bendrinė kalba*, 91, 1–24.
- Vaicekauskienė, L. (2011). Tautinių mažumų švietimo politika: kur einama? (2011 m. lapkritis) [Minority education policy: Where to go?] Retrieved March 3, 2021, from <https://www.nsa.smm.lt/wp-content/uploads/2020/07/Tautiniu-mazumu-svietimo-politika-Kur-einama-2011-lapkritis.pdf>
- Woessmann, L. (2002). How central exams affect educational achievement: International evidence from TIMSS and TIMSS-Repeat. *Harvard University Program on Education Policy and Governance Working Paper No. PEPG/02-10*.
- Woessmann, L. (2005). The effect heterogeneity of central examinations: evidence from TIMSS, TIMSS-Repeat and PISA. *Education Economics*, 13(2), 143–169.
- Woessmann, L. (2016). The importance of school systems: Evidence from international differences in student achievement. *Journal of Economic Perspectives*, 30(3), 3–32.
- Woessmann, L. (2018). Central exit exams improve student outcomes. *IZA World of Labor*, 419, 1–10.
- Zabulionis, A. (2020). *Tarptautinio švietimo tyrimo OECD PISA Lietuvos ir kaimyninių šalių tikslinė antrinė analizė*. Vilnius: Nacionalinė švietimo agentūra.

Centralizuoti vidurinės mokyklos mokinių lietuvių kalbos ir literatūros vertinimai: populiacijos analizė

Rimantas Želvys¹, Saulė Raižienė², Jogaila Vaitekaitis³, Rita Dukynaitė⁴, Audronė Jakaitienė⁵

¹ Vilniaus universitetas, Duomenų mokslo ir skaitmeninių technologijų institutas, Akademijos g. 4, 08412 Vilnius, rimantas.zelvys@fsf.vu.lt

Vilniaus universitetas, Ugdymo mokslų institutas, Universiteto g. 9, 01513 Vilnius, rimantas.zelvys@fsf.vu.lt

² Vilniaus universitetas, Duomenų mokslo ir skaitmeninių technologijų institutas, Akademijos g. 4, 08412 Vilnius, saule.raiziene@fsf.vu.lt

Vilniaus universitetas, Psichologijos institutas, Universiteto g. 9, 01513 Vilnius, saule.raiziene@fsf.vu.lt

³ Vilniaus universitetas, Duomenų mokslo ir skaitmeninių technologijų institutas, Akademijos g. 4, 08412 Vilnius, jogaila.vaitekaitis@fsf.vu.lt

Vilniaus universitetas, Ugdymo mokslų institutas, Universiteto g. 9, 01513 Vilnius, jogaila.vaitekaitis@fsf.vu.lt

⁴ Vilniaus universitetas, Duomenų mokslo ir skaitmeninių technologijų institutas, Akademijos g. 4, 08412 Vilnius, Rita.Dukynaitė@mif.vu.lt

⁵ Vilniaus universitetas, Duomenų mokslo ir skaitmeninių technologijų institutas, Akademijos g. 4, 08412 Vilnius, audrone.jakaitiene@mf.vu.lt

Santrauka

Tyrimo tikslas yra išanalizuoti vidurinių mokyklų mokinių populiacijos (nėra imties sudarymo paklaidų) lietuvių kalbos ir literatūros rezultatus, dėmesį skiriant lyčių skirtumams, ir sudaryti modelius pasiekimams prognozuoti. Straipsnyje nagrinėjami Pagrindinio ugdymo pasiekimų patikrinimo (toliau – PUPP) ir Valstybinių brandos egzaminų (toliau – VBE) rezultatai, neanalizuojant egzamino užduočių turinio ir kokybės. Nagrinėti visų Lietuvos vidurinių mokyklų mokinių, kurie 2014–2018 m. laikė egzaminus, individo lygmens duomenys. Kartu su brandos egzamino rezultatais analizuojami tų pačių mokinių PUPP pasiekimų rezultatai.

Nustatytos skirtingos nacionalinių centralizuotų ir tarptautinių mokinių pasiekimų tyrimų rezultatų tendencijos. Lyčių skirtumas stebimas ir tarptautiniuose, ir nacionaliniuose tyrimuose, o miesto ir kaimo vietovių skirtumai – tik tarptautiniuose tyrimuose. PUPP pasiekimų skirstinys yra artimas varpo formos kreivei, o VBE neatitinka normaliojo skirstinio, pasižymi teigiama asimetrija. VBE egzaminų skalėje yra trūkis (laikoma, kad jei mokyns gauna nuo 1 iki 15 balų, jis neišlaiko egzamino ir vertinimas lygus 0), tai komplikuoja analizę, apsunkina palyginimą ir nėra teisinga edukologinės perspektyvos požiūriu.

Centralizuoti mokinių vertinimai turi skirtingus tikslus, remiasi skirtingomis metodologijomis, tačiau kiekvienas iš jų suteikia vertingos informacijos apie mokinių gebėjimus ir turėtų būti naudojamas ugdymo efektyvumui gerinti ir švietimo politikai tobulinti.

Esminiai žodžiai: *lietuvių kalba ir literatūra, populiacija, PUPP, VBE, pasiekimų patikrinimas.*

Gauta 2021 03 04 / Received 04 03 2021
Priimta 2021 03 29 / Accepted 29 03 2021