THE IMPACT OF SCHOOL CONTEXTS ON STUDENT OUTCOMES AND ON CURRICULUM EVALUATION IN HIGH SCHOOLS OF KORÇA REGION

(Comparative study of the State Matura results and the internal assessment of high school students in the region of Korça)

Abstract

Assessment of students' achievements interacts with the evaluation of the school curriculum. It is one of the curricula's components that has a great influence on student outcomes. The survey tries to debate the impact of assessment in gymnasiums with different characteristics in the results of Matura exams. ANOVA analysis of the effects of internal assessment and external assessment, presents one aspect of the impact of school context on student outcomes.

The analysis concludes that public urban Gymnasiums have, the more sustainable internal assessment, compared with public non-urban Gymnasium and on high schools with fewer than 100 pupils.

The results of the survey can be used to restructure the high school system based on the needs and possibilities of the school's community.

Keywords

Student assessment, State Matura exams, curriculum evaluation, high schools, Korça región.

Theoretical Framework

The assessment is one of the most important school curriculum components that has a great influence on student achievement. Assessment has three roles or functions (Nagy, 2000). The role with the longest history is that of gatekeeping (National Commission on Testing and Public Policy, 1990), in which assessment determines who is given a privilege such as admission or graduation. This function has recently been added the role of ensuring accountability (Darling-Hammond & Ascher, 1991), in which assessment is used to determine if schools are functioning well.

The third role is that of instructional diagnosis (Levesque, Bradby, & Rossi, 1996), in which assessment is used to determine what students do and do not know, and what to do about it. These three functions have not simply replaced each other; rather, additional requirements have gradually been added to the expectations held for external assessment.

Assessment as instructional diagnosis

Mehrens (1998) notes that most writing on effects of assessments is data-free rhetoric. He further points out the difficulties of doing research that would examine the effects of assessment on instruction.

Based on limited evidence, he concludes that if the stakes are high enough and teachers see the material assessed to be appropriate, they will shift instruction to cover test content. If not, the impact of the assessment will not be obvious.

There are reasons why it is difficult to apply large-scale assessment results to instructional diagnosis. Serafini (2001) argues from a constructivist viewpoint that top-down accountability is so fundamentally different from providing data for instructional diagnosis that educators should not expect

any classroom effects. He continues to advocate replacing large scale assessment with "assessment as inquiry", much in line with Darling-Hammond and Ascher's views (Darling-Hammond & Ascher, 1991). Such assessment would be done in the classroom and would focus on determining specific reasons for student misunderstanding or lack of skill. He argues as if teachers had some say in whether they engage in top down accountability practices; in a climate of top-down accountability, this is not the case.

Relationship of external assessment of the curriculum evaluation

Curriculum evaluation is an attempt to toss lightly on two questions: Do planned courses, programs, activities, and learning opportunities as development and organized actually produce desired results? How can the curriculum offerings best be improved?

Evaluation is comparing a student's achievement with other students or with a set of criteria. Effective assessment and evaluation are incorporated into all facets of the curriculum, providing both teachers and students with relevant and useful data to gauge progress and determine the effectiveness of materials and processes. We can define the curriculum evaluation as the judgment of the merit and worth of a program of studies, a field of study, or a course of study. (Guba and Lincoln, 1981)

The evaluation of the written curriculum and its effectiveness has been ever more dependent on the taught curriculum. To be effective, the written curriculum needs to be the "taught curriculum". Evaluation is comparing a student's achievement with other students or with a set of criteria.

Pre-university education system in Albania, performs external assessment of student achievements through examinations agency NAE (National Agency of Examination)(Regulation, 2013).

NAE provides measurement and certification of high school student achievements through mature exams called the State Matura (SM). State Matura is a system regarding final exams that young adults (aged 18 or 19) take at the end of their secondary education in Albania. State Matura consists of two required exams (written Mathematics and written Literature) and two other exams. These last two exams are chosen by the student from a pool of subjects (Chemistry, Biology, Physics, English, History-Geography and Sociology-Economy-Philosophy). (Law, Nr. 69/12, 2012)

According to the new pre-university education Law, the high school system in Albania consists of various high school profiles.

The following table shows the statistics for the distribution of secondary education institutions in the district of Korca.

Table 1.0

Type of high school	Number of students	Nr. of institutions	Number of students	Percentage %	Number of high school graduates	Percentage %
Public Gymnasium/urban areas	Until 100	0	1802	34	667	44
	100 – 300	2				
	300 – 600	1				
	Over 600	2				
Public Gymnasium/rural areas	Until 100	6	1111	21	349	23
	100 – 300	4				
	300 – 600	0				
	Over 600	0				
Non public Gymnasium	Until 100	1	512	10	151	10
	100 – 300	2				
Part-time Gymnasium	100 – 300	2	540	10	106	7
Vocational high school	Until 100	1	1260	25	222	16

100 – 300	3				
300 – 600	2				
Total	25	5225	100	1495	100

SOURCE: Statistical Office of the Regional Education Directorate Korce, 2013

Objectives:

In this context, the study aims to describe how the specific conditions of schools affect school curriculum evaluation, and specifically the assessment of student achievement.

The goal is to answer the following questions:
- Does it impact in students' achievements, the school's context?
- How affects the school size in assessment of students?
- Does exist any difference between the assessment of students in high schools on rural schools and in urban areas?

Method

Participants

The study will look at the relationship between the internal assessment as a component of curriculum evaluation in high school (gymnasium) and the results of Matura exams. Participants in the study are high school graduate students in academic year 2012 - 2013 that are assessed on the exams of State Matura. To realize this study will be used the results of Matura exams of high school students in Korca district 2013 and the results scores of students in high school in subject-matter: Albanian language and the Literature and Mathematics, as well as the average score of students in high school. Albanian Language and Literature and Mathematics are assessed as compulsory exams to Matura. (Instruction Nr. 35, 2013)

The tables below show the statistics of the results of internal evaluation of students in the school and external assessment of their state exams. (Book of transparency, 2013)

```
INTERNAL ASSESSMENT
Annual Score (Albanian Language and the Literature) (D1)
Annual Score (Mathematics) (D2)
Annual Average Score (AAS)
```

Table 2.1

	N	SCORES		Mean		Std. Deviation	Variance
	Statistic	Minimum	Maximum	Statistic	Std. Error	Statistic	Statistic
D1	1495	4,50	10,00	6,6752	,04422	1,70977	2,923
D2	1495	4,50	10,00	6,2931	,04434	1,71424	2,939
AAS	1495	4,52	10,90	7,1074	,04176	1,61458	2,607
Valid N (listwise)	1495						

```
EXTERNAL ASSESSMENT (SM)
SM exam score - Albanian Languages and the Literature (D1)
```

3

```
Maths exam score (D2)
The average score of the SM exam (including Z1 + Z2)
```

Table 2.2

	N	Minimum	Maximum	Mean	Std. Deviation	Variance
Exam scores D1	1495	4,00	10,00	6,9905	1,46455	2,145
Exam scores D2	1495	4,00	10,00	6,0242	1,54079	2,374
Average Score Exam (ASE)	1495	4,00	9,93	6,6511	1,25556	1,576
Valid N (listwise)	1495					

Procedure

The study will compare the differences of average student scores in subjects D1 and D2 and the average scores of Matura exams, as well as annual average scores of students with average scores of the exams (D1 + D2 + Z1 + Z2[1]).

To compare these results, is chosen to use ANOVA statistical analysis.

-Internal assessment of students (that is made by teachers in school - Annual Score (Albanian Language and the Literature) (D1); Annual Score (Mathematics) (D2); Annual average score (AAS);

-        ***External assessment*** (that is realized by NAE, (Regulation, 2013) which organizes the State Matura exams – State Matura exam score - Albanian Languages and the Literature (D1); State Matura Maths exam score (D2); and the average score of the State Matura exam (including Z1 + Z2).

Analysis of these indicators is based on the instruction and the regulation of Matura. This regulation defines the compulsory exams and the optional exams.

Results

Through ANOVA analysis test Oneway, are observed differences in means of compulsory Matura exams (D1 and D2) and the annual assessment of respective subject-matters, depending on the type of high school and number of students of the school.

Oneway analysis of variance test the equality of population means and allow us to examine differences among means using multiple comparisons. Multiple comparisons of means allow you to examine which means are different and to estimate by how much they are different.

Although there is no a categorization of schools according to their numbers, in many studies, are defined various categories. Ornstein (1990), categorize small schools from 495 to 1280, whereas in Lee's 2000 study, the small school category included only schools under 300 students and Johnson's 2006 study classified such as being below 200 students. In our study, schools are divided by the number of students in four categories, based on instruction 21, for labor rates and the number of students in class. (Instruction Nr. 21, 2010)

a) Less than 100 students
b) From 100 to 300 students
c) More than 300 students

---

[1] Z1 and Z2 are the optional exams of State Matura

Through Oneway analysis were observed differences between 3 groups:

    a) Public Gymnasium / urban area and Public Gymnasium / rural area

    b) Public Gymnasium / urban area and Nonpublic Gymnasium

    c) High Schools with fewer than 100 students and High schools with 300-600 students.

For the results see the Means Plot in Appendix.

Why ANOVA and not T-test?

1. Comparing 3 groups using t-tests would require that 3 t-tests be conducted. Group a vs. Group b, Group a vs. Group c, and Group b vs. Group c. This increases the chances of making a type I error. Only a single ANOVA is required to determine if there are differences between multiple groups.
2. The t-test does not make use of all of the available information from which the samples were drawn. For example, in a comparison of Group a vs. Group b, the information from Group c is neglected. An ANOVA makes use of the entire data set.
3. The ANOVA is a statistical technique which compares different sources of variance within a data set. The purpose of the comparison is to determine if significant differences exist between two or more groups.

Oneway ANOVA can be used when the data can be partitioned into more than two groups, additional comparisons can be made. This might involve one aspect or dimension, for example four groups each representing a region of the country. Or the groups might vary along several dimensions, for example eight groups each composed of a gender (two categories) by region (four categories) combination. In this latter case, we can ask additional questions:

(1) is there a gender difference?

(2) is there a region difference?

(3) do gender and region interact?

Analysis

ANOVA analysis results through one way test indicates that the differences between the averages of internal evaluation (exam scores on subjects Albanian Language and the Literature (D1) and Mathematics (D2)), and the external evaluation of students (compulsory exam scores to Matura), are significant.

1- Public gymnasium in urban areas has a smaller difference in relative value between the internal assessment and the external assessment. Public gymnasium in urban areas has a positive difference compared with the external assessment of internal evaluation.

    This shows that the scores of students in the exam have higher relative value compared with the scores of students assessed by teachers in school. The highest difference is in the case of D1 (Albanian Language and the Literature), respectively ($\alpha = + 0.8376$)

2- Public gymnasium in rural areas has a greater difference in the relative value between the internal assessment and the external assessment. This indicates that the score of students in examinations is lower in relative value compared with the outcomes of students judged by the school. The difference appears higher in mathematics (D2 - Exam scores - Annual scores), ($\alpha = - 0.5963$).

3- Nonpublic gymnasium has a negative difference between the external assessment and the internal assessment.

D1 (Exam score - Annual Score) α = - 0.3553

D2 (Exam score - Annual Score) α = - 0.2397

Nonpublic gymnasium has the highest positive difference between the internal assessment and the external assessment:

(Average score exams & Annual Average Score α = + 0.9962)

Table 3.1 - Differences between external assessment and internal assessment of students

Type of High School	Exam score – Annual Score		Annual Average score – Exam Score		Annual Average Score – Average Exam Score
	D1	D2	D1	D2	
Urban public High school	0,8376 ⇑	0,0669 ⇑	-0,1372 ⇑	0,8529 ⇓	0,3452 ⇓
Rural Public High school	-0,0159 ⇓	-0,5963 ⇓	0,3123 ⇓	1,4376 ⇓	0,7485 ⇓
Nonpublic High School	-0,3553 ⇓	-0,2397 ⇓	0,7551 ⇓	1,3074 ⇓	0,9962 ⇓

The findings of the study, according to the number of students are as follows:

1- Schools with the number of students up to 100, have negative difference between external evaluation and internal evaluation. The Alpha indicator for differences between the annual average scores and the average exam scores is (α = -0,9461)

2- Schools with the number of students from 300 to 600, have relatively sustainable difference. Alpha index is smaller compared to the number of students in schools with more or less than interval 300-600. The smaller alpha indicator is the D2 difference (Exam scores - Annual Score).

3- Schools with the number of students over 600 have positive difference between the external evaluation and the internal evaluation of field of Albanian language and the literature and Mathematics. Comparison of average annual values scores with average exam scores shows that internal assessment is higher in relative value compared with the external evaluation is (α = +0,4228).

Table 3.2 - Differences between external assessment and internal assessment of students

Size of High School	Exam score – Annual Score	Annual Average score – Exam Score	Annual Average Score – Average Exam Score

	D1	D2	D1	D2	
Less than 100 students	-0,3335 ⬇	-0,8047 ⬇	0,6144 ⬇	1,6611 ⬇	0,9461 ⬇
From 100 – 300	0,0799 ⬆	-0,3713 ⬇	0,1732 ⬇	1,2101 ⬇	0,6011 ⬇
More than 300 students	0,4897 ⬆	0,1089 ⬆	-0,311 ⬆	0,5090 ⬇	0,4228 ⬇

Discussion:

Assessment date is helpful to the extent that they provide relevant, valid, timely, and useful information about how much individual students are learning and how well schools are attending them. Oakes (1989) notes that school context indicators can provide clues about why schools achieve the outcomes they get along. According of another study, one of the great dangers of indicators of school context is the ease with which can be misunderstood. One way to create fair comparison is to develop longitudinal measures of student progress or to compare schools with "like" schools. (Darling-Hammond & Ascher, 1991)

However, analysis of the difference between external evaluation and internal evaluation, of high school student achievements, shows a general trend of how has developed the curriculum on school and especially how can affect on the assessment of students the school-specific conditions, such as location, the number of students and the type of school.

Analysis of test results through one-way, showed that there is a relation between the type of school and student achievement on standardized assessment tests.

Students of public schools in urban areas, have external assessment indicators approximately or with very little change in the relative value compared with internal assessment.

Students of public high school in rural areas and nonpublic high schools, have external assessment index lower compared to internal assessment.

The analysis demonstrates that school size (number of students at school) effects on student achievements. Schools with less than 100 students have lower scores of students in exams compared with the results of internal evaluation. Schools with the number of students with more than 300 students (300-600 students), have indicators of external evaluation in generally more sustainable and higher than indicators of internal assessment.

Considering the data in Table 1.0, we see that 60% of high schools in urban areas have over 300 students, while 60% of high school less than 100 students, are situated in rural areas. Similarly, all non-public high schools have fewer than 300 pupils.

Considerable research has been conducted on the relationship between the size of high schools and students' performance on standardized tests.

Lindalh & Cain (2012), in their study showed a relationship between student achievement and school size. The correlations in that study between high school size and student performance on standardized tests were higher for students from higher socioeconomic backgrounds. (Cannel, 1988) database has concluded that students in high schools with between 600 and 900 students scored better in both reading and math. In smaller schools, students performed less well;

Gardner, Ritblatt, and Beatty in their 2000 study of large and small high schools in California, found that larger, high schools have higher academic achievement on the Scholastic Aptitude Test than do smaller high schools. Lee's (2000) national survey concluded that school size makes a bigger difference in

schools enrolling students of low socioeconomic status than I those serving students of higher socioeconomic status.

The study showed that high schools in urban areas with more than 300 students have higher scores on exams and a positive difference between external evaluation and internal assessment; high schools in rural areas with less than 100 students have a negative difference, hence the non-public high schools.

The survey indicates that internal assessment on Albanian Languages   and the Literature and Mathematics is higher in high schools in rural and non-public high schools.

In high schools in urban areas, internal assessment is lower and scores of students in exams are higher. These results show that in rural schools, teachers are more "generous" in their assessment and as it says Nagy (1984) is used to maximize the student 'chances of producing something of their lives.

The results of the national assessment of students show two tendencies:

a) In high schools in urban areas, teachers are "stingy" in evaluation of students. This phenomenon appears in the lower results of this assessment compared with the values the State Matura results.

b) In high schools in rural areas, teachers are more "permissive". Values of internal assessment results are higher than the values of the Matura exam results.

One factor that affects this phenomenon has to do with the "pressure" that exert national exams in school programs (Goodlad, 1984), especially in high schools in urban areas, where the number of students is greater and the competition is great for high achievement.

Another factor has to do with other aspects of the school context, such as access to knowledge and professional level of teachers. (Oakes, 1989).

Another interpretation of this event is that the increased breadth of curriculum offerings in larger schools allows a higher part of teachers to remain teaching exclusively in their fields of certification, whereas in smaller high schools, there may not be enough sections needed in a given discipline are to occupy the teacher full-time, forcing that teacher to teach in another area in which he or she is not "highly qualified". (Lindahl & Cain, 2012)

Conclusions:

As a conclusion, we can say that the context of the school and its number of students, have an impact on student achievement.

In high schools in urban areas with a number of students over 300, students are assessed in the most correct way, compared with High school students in rural areas and nonpublic schools.
In high schools in rural areas, with a number less than 100 pupils, students are assessed ineffective and unfair. Their results graduation exams have significant differences compared with the results of internal assessment.

The results of the study suggest that the organization of high schools in the district of Korca, it is necessary to consider the context of the school, their location and the number of students.

High schools in rural areas with less than 100 students, are not effective in achieving the school curriculum. Scores of students and their achievements are affected by the conditions of the school and a number of students. Organizing schools to institutions with a number of 300-600 students appear to be more effective.

REFERENCES

Cannell, J. J. (1988), Nationally Normed Elementary Achievement Testing in America's Public Schools: How All 50 States Are Above the National Average. Educational Measurement: Issues and Practice, 7: 5–9. doi: 10.1111/j.1745-3992.1988.tb00424.x

Darling-Hammond, L., & Ascher, C. (1991). *Creating accountability in big city schools* (Urban Diversity Series No. 102). New York: National Center for Restructuring Education, Schools, and Teaching. (ERIC Document Reproduction Service No. ED334339) Retrieved October 10, 2013 from http://files.eric.ed.gov/fulltext/ED334339.pdf

Gardner, P. W., Ritblatt, S. N., & Beatty, J. R. (2000). *Academic achievement and parental school involvement as a function of high school size.* High School Journal, 83(2), 21-27.

Goodlad, John I.(1984). *A place called school.* New York: McGraw-Hill

Guba, E. & Lincoln, Y. (1981). *Effective evaluation.* San Francisko: Jossey-Bass.

Johnson, J. (2006). *More doesn't mean better.* Randolph, VT: The Rural School and Community Trust Policy Program.

Lee, V.E., Smerdon, B. A., Alfeld-Liro, C., & Brown, S.L.(2000). *Inside large and small high schools: Curriculum and social relations.* Educational Evaluation and Policy Analysis, 22(2).

Levesque, K., Bradby, D., & Rossi, K. (1996). Using data for program improvement: How do we encourage schools to do it? *CenterFocus, 12.* In: Nagy, Ph. *The three roles of assessment : Gatekeeping, Accountability and Instructional Diagnosis* (Canadian Journal of Education 25, 4 (2000)
　　　Retrieved October 2, 2013 from http://ncrve.berkeley.edu/CenterFocus/CF12.html

Mehrens, W. A. (1998). Consequences of assessment: What is the evidence? Educational Policy Analysis Archives, 6(13).
　　　Retrieved October 2, 2013 from http://onlinelibrary.wiley.com/doi/10.1111/j.1745-3992.1989.tb00304.x/pdf

MoES (2013) Book of transparency 2013
　　　Retrieved November 30, 2013 from
　　　　http://arkiva.mash.gov.al/File/matura/2013/libri-transparences/publikimi-l-t-2013.pdf
　　　　http://arkiva.mash.gov.al/File/matura/2013/libri-transparences/KORCE.pdf

MoES (2010). *Instruction Nr. 10, date 23.07.2010, For labor rates and the number of students in class*

MoES (2012). *Instruction Nr. 35, date 27.11.2012, For Development of the State Matura exams 2013*

MoES (2012). *Law nr. 69 21.06.2012 for Pre_University Education System*

MoES (2013). *The regulation for the process of State Matura 2013 in the Republic of Albania.* Tirana, 07.01.2013

Nagy, P. (1984). *An examination of differences in high school graduation standards.* (Canadian Journal of Education, 9, 276-297.)

Nagy, Ph. (2000) *The three roles of assessment : Gatekeeping, Accountability and Instructional Diagnosis* (Canadian Journal of Education 25, 4 (2000)
　　　Retrieved October 2, 2013 from http://ncrve.berkeley.edu/CenterFocus/CF12.html

National Commission on Testing and Public Policy. (1990). *From gatekeeper to gateway: Transforming testing in America.* Boston: Boston College, National Commission on Testing and Public Policy.

Oakes, J. (1989). *What educational indicators? The case for assessing the school context. (*Educational Evaluation and Policy Analysis, 11(2), 189-199).
　　　Retrieved November 23, 2013 from
　　　http://gse.buffalo.edu/fas/yerrick/ubscience/UB_Science_Education_Goes_to_School/21C_Literature_files/Oakes,%201989.pdf

Ornstein, A. C. (1990). *How BIG should schools and districts be?* Education Digest, 56(2), 44-47.

Serafini. F. (2001). *Three paragigm of assessment: Measurement, procedure, and inquiry.* (The Reading Teacher, 54, 384-393.)
　　　Retrieved December, 14, 2013 from http://www.frankserafini.com/publications/serafini-3-parad-assess.pdf

WEBLIOGRAPHY

"Assessment & Evaluation"
http://www.teachervision.fen.com/assessment/new-teacher/48353.html

Curriculum audit
http://www.pdkintl.org/curriculum/curriculumAudits.html

Curriculum evaluation
www.mcrel.org
www.ascd.org

National Center for Evaluation Statistics
http://nces.ef.gov

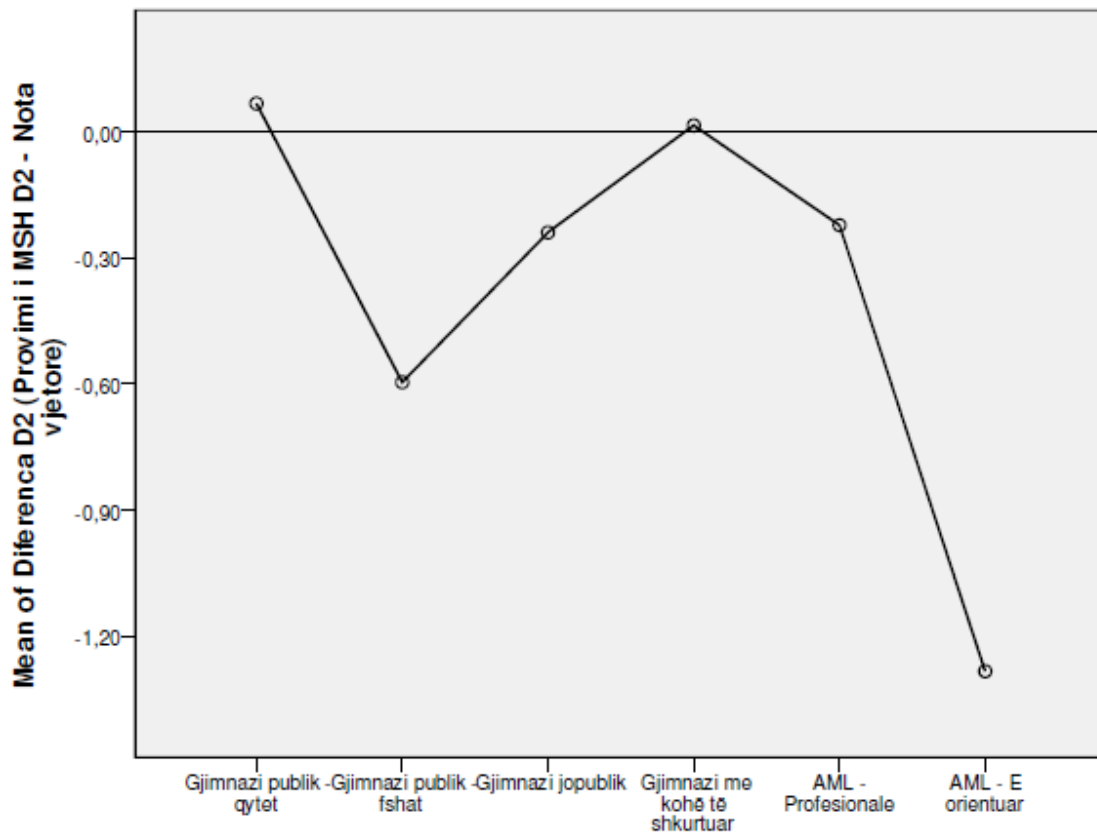American Educational Research Association
http://www.jstore.org

Type of high school

MEANS PLOT
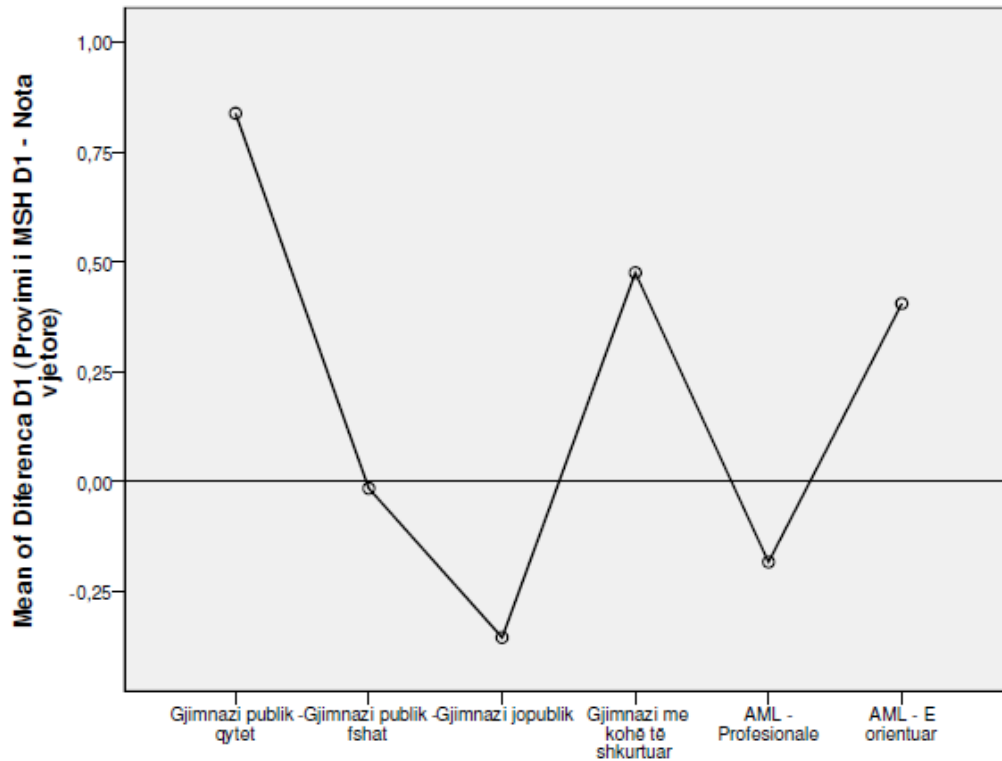
Graph 1.1 – Difference D1 (Exam score – Annual score)

Graph 1.2 – Difference D2 (Exam score – Annual score)
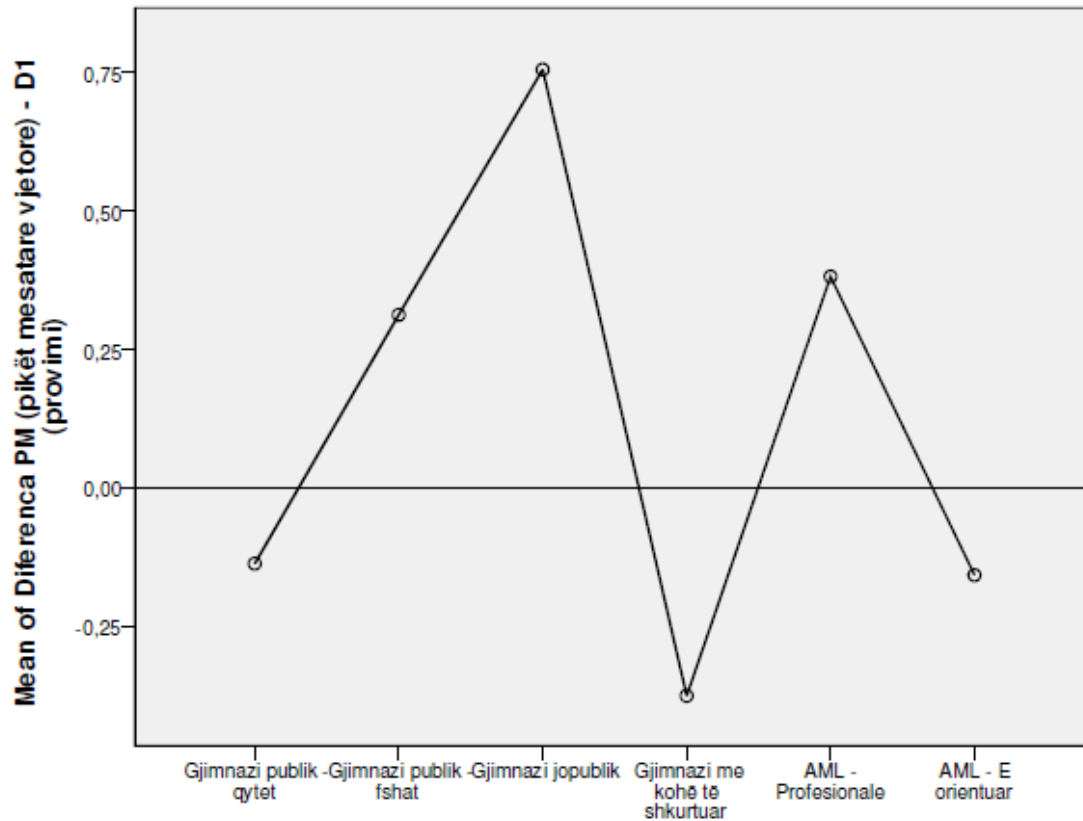
Graph 1.3 – Difference D1 (Annual score – D1 Exam score)
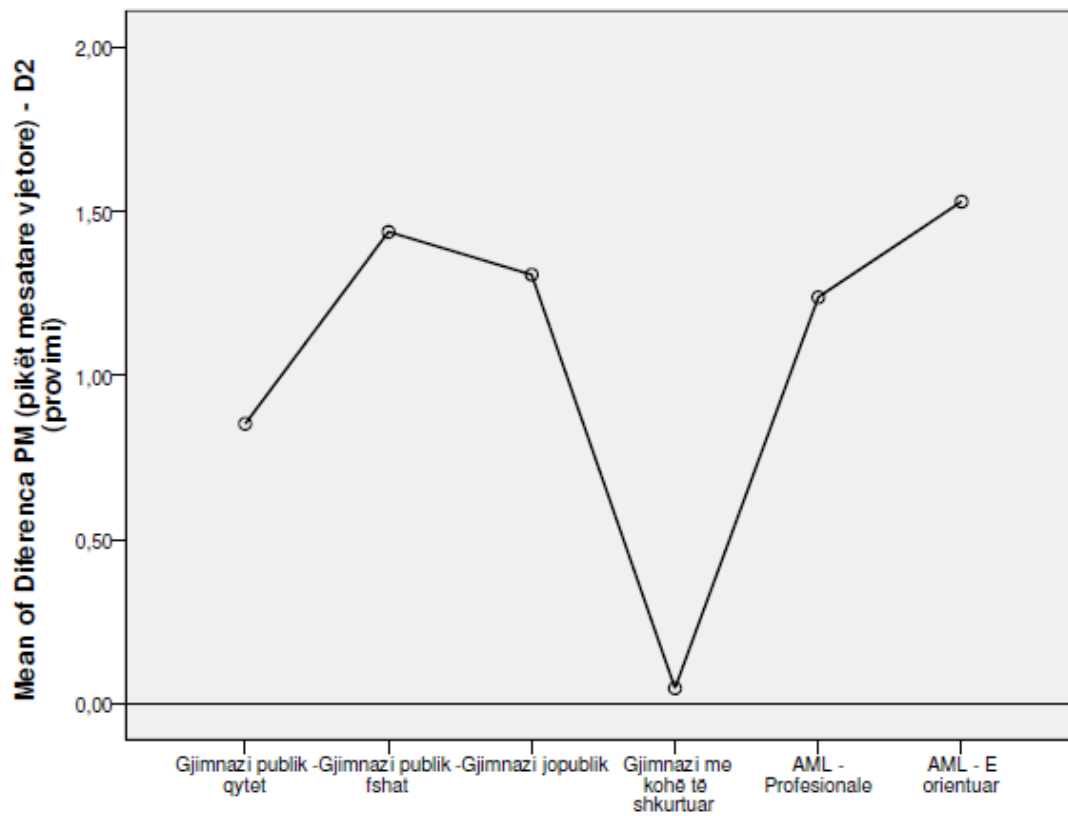
Graph 1.4 – Difference D2 (Annual score – D2 Exam score)
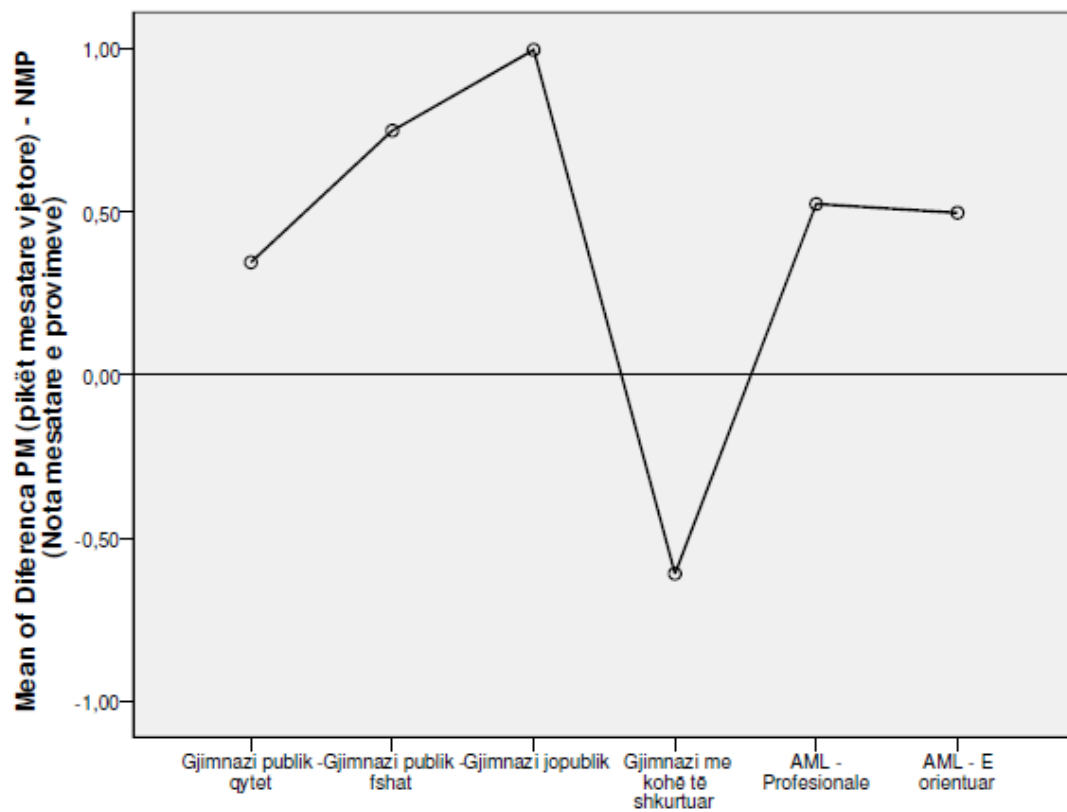
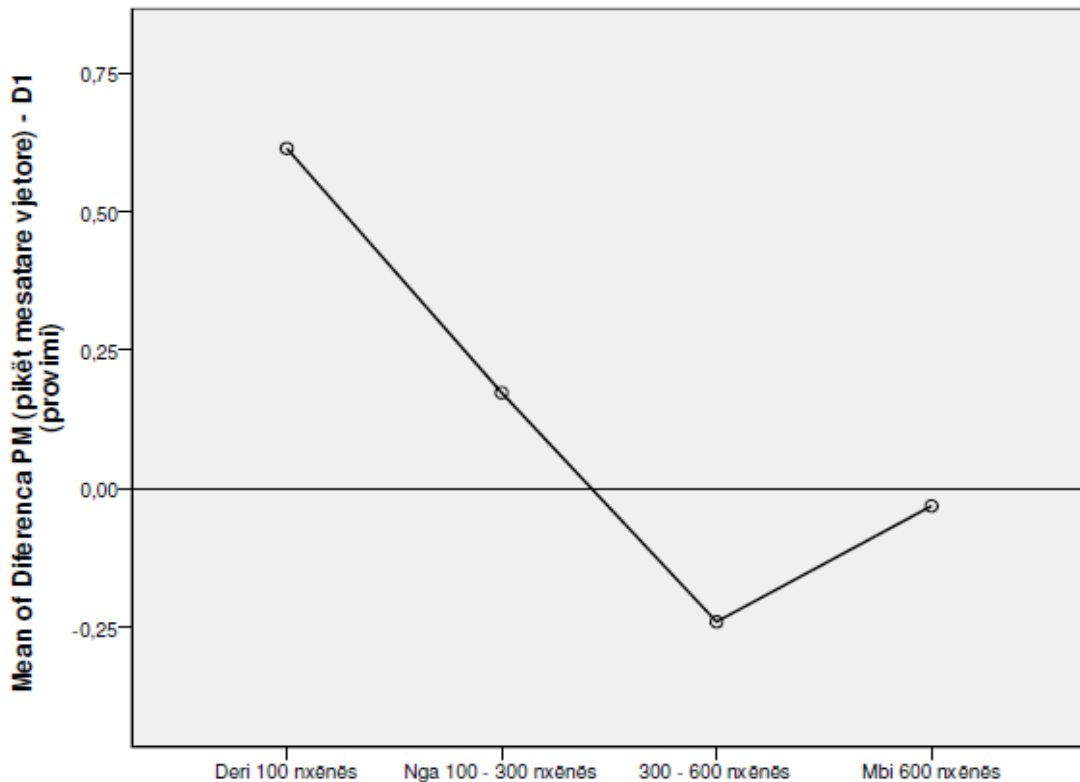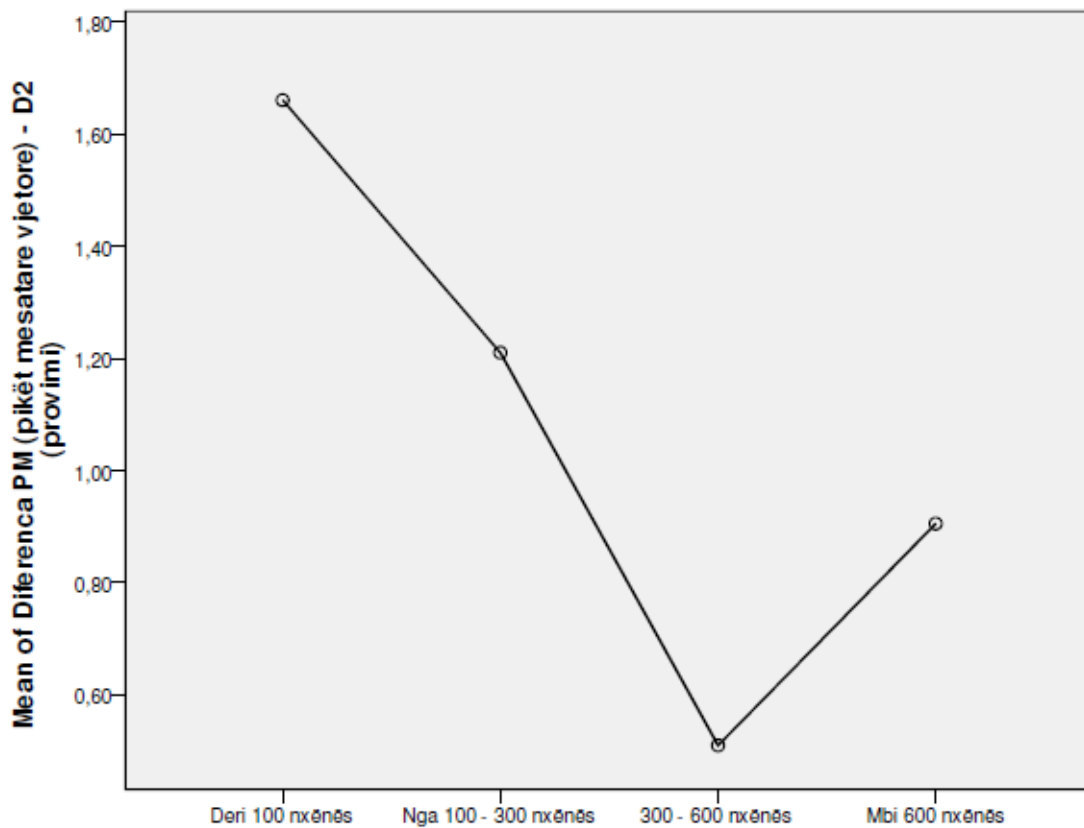Graph 1.5 – Difference (Annual average score – Exam average score
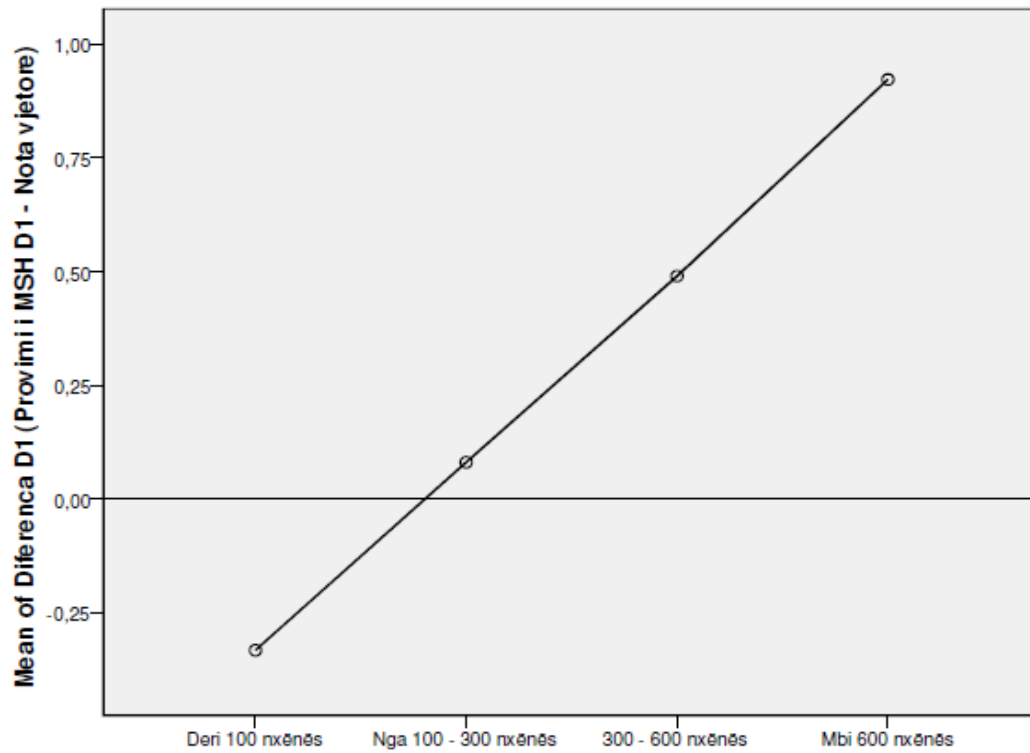
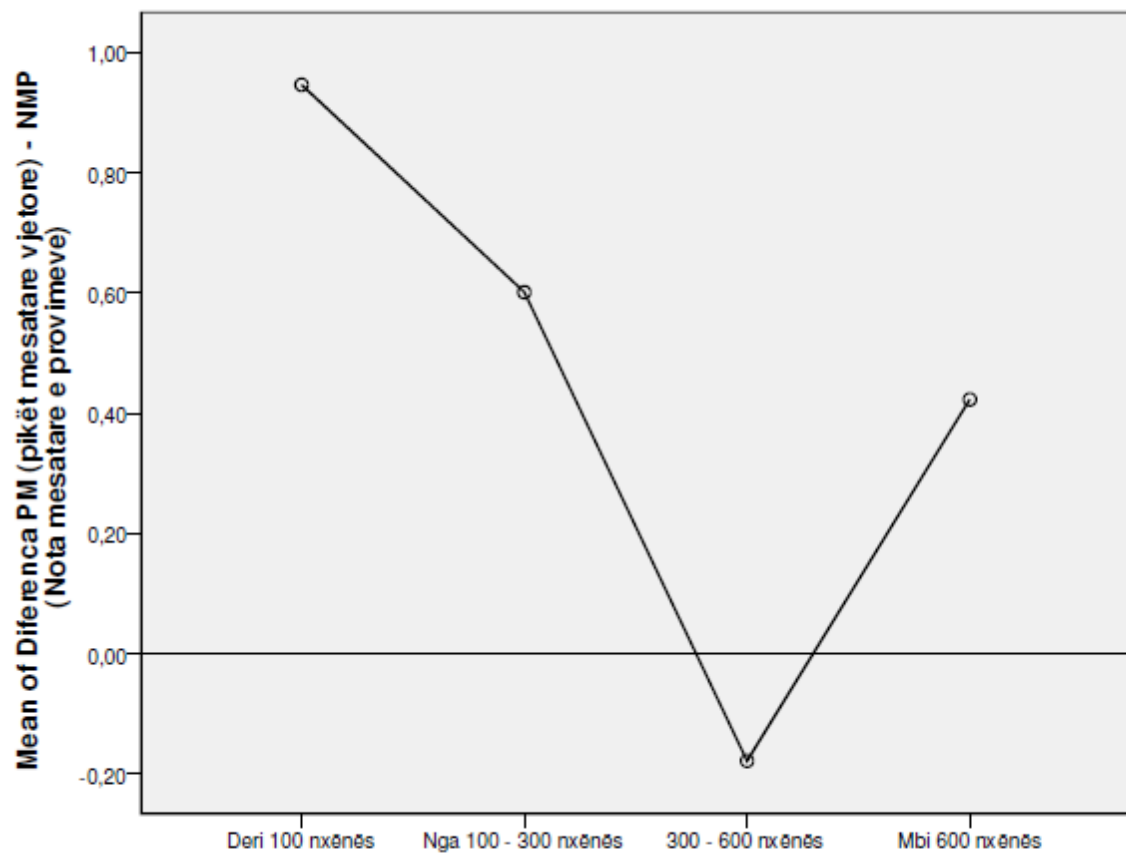Graph 2.1 – Difference D1 (Exam score – Annual score)
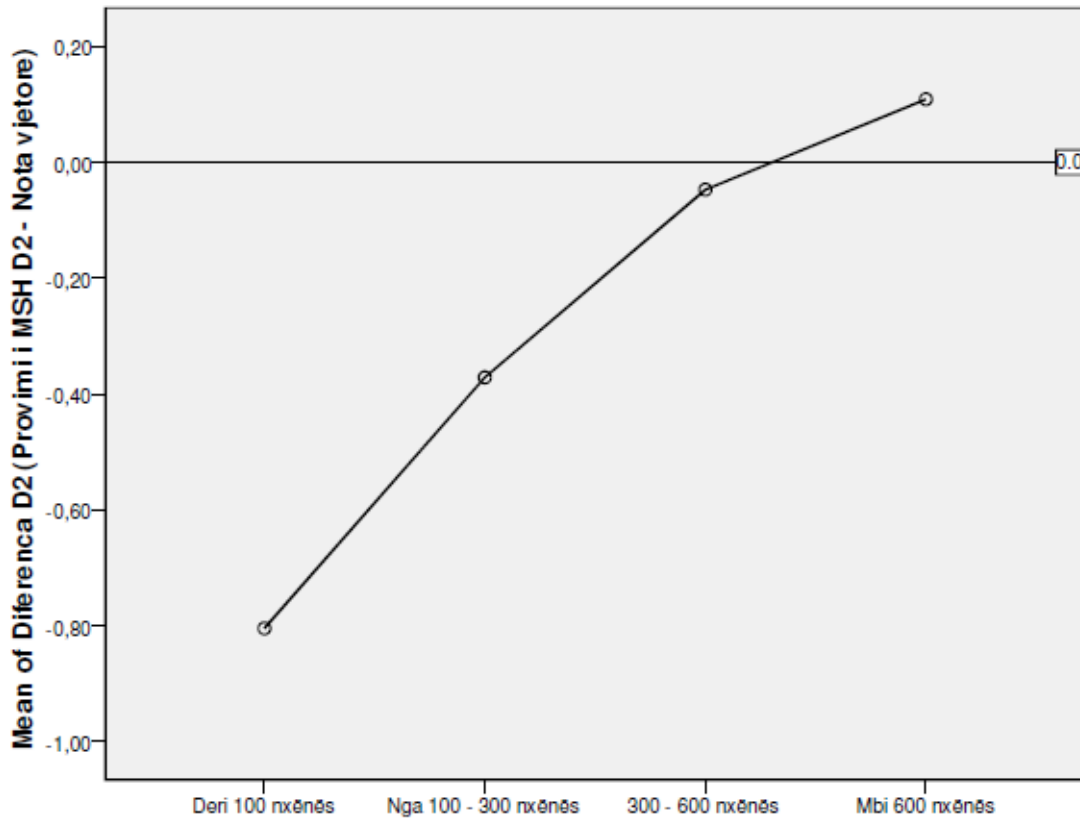
Graph 2.2 – Difference D2 (Exam score – Annual score)

Graph 2.3 – Difference D1 (Annual score – D1 Exam score)

Graph 2.4 – Difference D2 (Annual score – D2 Exam score)

Graph 2.5 – Difference (Annual average score – Exam average score