

Comparing Elephants and Bananas in Educational Achievements: What Do Data Reveal?

Tihomira Trifonova
Centre Immigration and Integration

The problem of functional illiteracy emerged in the Bulgarian society when the results of the OECD's Programme for International Student Assessment (PISA test) were announced. Until then it was not aware of such a deficit. The national external assessment of students' educational achievements did not give any signs of a pervasive systemic deficit in the Bulgarian education. A comparison of the two tests' scores however reveals a considerable discrepancy. Looking at the tests' metadata, it becomes obvious the comparison is between elephants and bananas and that explains the inconsistency. This paper compares the two datasets of students' scores and the tests' measuring methodologies. It further concludes that the national system fails to adapt to the needs of a changing society. However, it has an important ally in the face of the civil society, which provides its own resources to satisfy learning needs.

Keywords: functional literacy, educational achievements, assessment

INTRODUCTION

Frameworks that conceptualize quality of education are necessarily value-based. Looking at the most dominant among them, we can differentiate between the humanist/progressive approach, typical with its focus on the development of the individual and on collective human development for social change, and the economist approach, focused on efficiency and effectiveness and the achievements of learning outcomes at reasonable cost. In 2005 the United Nations drew attention to the importance of quality in education, issuing its Global Monitoring Report: The Quality Imperative, which reviews research evidence on the multiple factors determining quality and maps out key policies for improving the teaching and learning process, especially in low-income countries.

Comparative educational scholars emphasize the importance of asking critical questions about quality education: how to access it and for what purposes; what contextual and cultural factors are important; what implementation issues and local-level realities must be considered for true understanding of standardized assessment results; what content, skills and values are necessary and desirable ingredients (Bray, Adamson and Mason 2014). Critiques of narrow interpretations of standardised assessment data contrast with research-based evidence that participation in large scale assessments such as PISA, TIMSS and PIRLS can indeed be beneficial to identify needed reform refinements and implementation shortcomings (Popham 2000).

Education is universally essential for the economic, social, and human development. It is subjected to the results of progress that it has engendered itself, both regarding content and methods and established aims. With the aim of contributing to the debate, this paper explores two assessments – one national and

one international – of educational achievements in reading/literature at the same level of education to analytically compare them. The results of international and national assessments in education are expected to be comparable and completing, since international assessments are meant to assist reforms in education. In this national case they appear not to be, and the achievements deviate very significantly both in terms of substantive results and as trends.

THE FUNCTIONAL ILLITERACY PROBLEM

Low literacy has many consequences: it affects people in their daily lives negatively, it can jeopardize their future and it has a significant effect on society, both socially and economically. On individual level low literacy means, perhaps not exhaustively, limited ability to obtain and understand information; much higher probability for unemployment; absence of advanced opportunities and lower income; low-quality jobs and little access to professional development; precarious financial situation; low self-esteem, often accompanied by isolation; and it also has an impact on health. On societal level it means overall less competitive economy and increased unemployment and welfare payments, slower GDP growth rate, workforce less competitive in the global knowledge economy and weak community involvement and civic participation.

The ability to function effectively in one's community is referred to by UNESCO as functional literacy. According to the definition, it involves the skills needed to cope with the circumstances of life. Functional literacy goes beyond the mere capacity to read and write, since a person may be able to read the words in a text and not grasp its content. By analogy, students may be able to write sentences, but not express complete, cogent thoughts. Functional literacy is the basis for a functional society. Reversely, widespread functional illiteracy will mean living in a dysfunctional society. To be able to establish whether or not literacy is truly functional, it has to be measured at different levels of proficiency. The UNESCO, the OECD and the IEA measure literacy and other key knowledge skills of children, young adults, and adults through large-scale, international assessments of strengths and weaknesses in different countries. These international test types generally measure literacy and numeracy skills in a variety of ways, including mapping the whole literacy spectrum and grouping the performance and the abilities into discrete levels.

RESEARCH METHODOLOGY

In this analysis we look at two datasets – one national and one international - containing the results of educational achievements' assessment in reading/language and literature. They are for the years 2015 and 2018, the last two PISA waves, since those are the only publicly available national PISA test results to compare with the results of the national external assessment after grade 7, i.e. at the same level of schooling. We also look at the data sets' metadata, which reveal the baseline logic of what and how should be tested.

PISA conceives of reading literacy as a broader set of competencies that allows readers to engage with written information, presented in one or more texts, for a specific purpose (RAND Reading Study Group and Snow, 2002; Perfetti, Landi and Oakhill, 2005). To engage with what they read, readers must understand the text and integrate it with their pre-existing knowledge. Evident from the methodological description is that a very important characteristic of the test is its future orientation. The highest-level process identified by the PISA 2018 reading literacy framework is "evaluating and reflecting". Comprehension is the core of the process since it involves processing operations like combining information and making sense of the text through construction of mental patterns, showing how the text relates to reality. Assessment is a new component, which adds the idea of reading as a targeted activity, involving judgments. This skill is traditionally related with critical thinking and is particularly salient nowadays.

The national external assessment at grade 7 in Bulgarian language and literature (corresponding to reading comprehension) is a form of test to evaluate educational results, which are then used by the secondary level schools as an entry score. The maximum result is 100 points. The assessment follows standardized criteria and every task can generate a set number of points, corresponding to its specifics, level

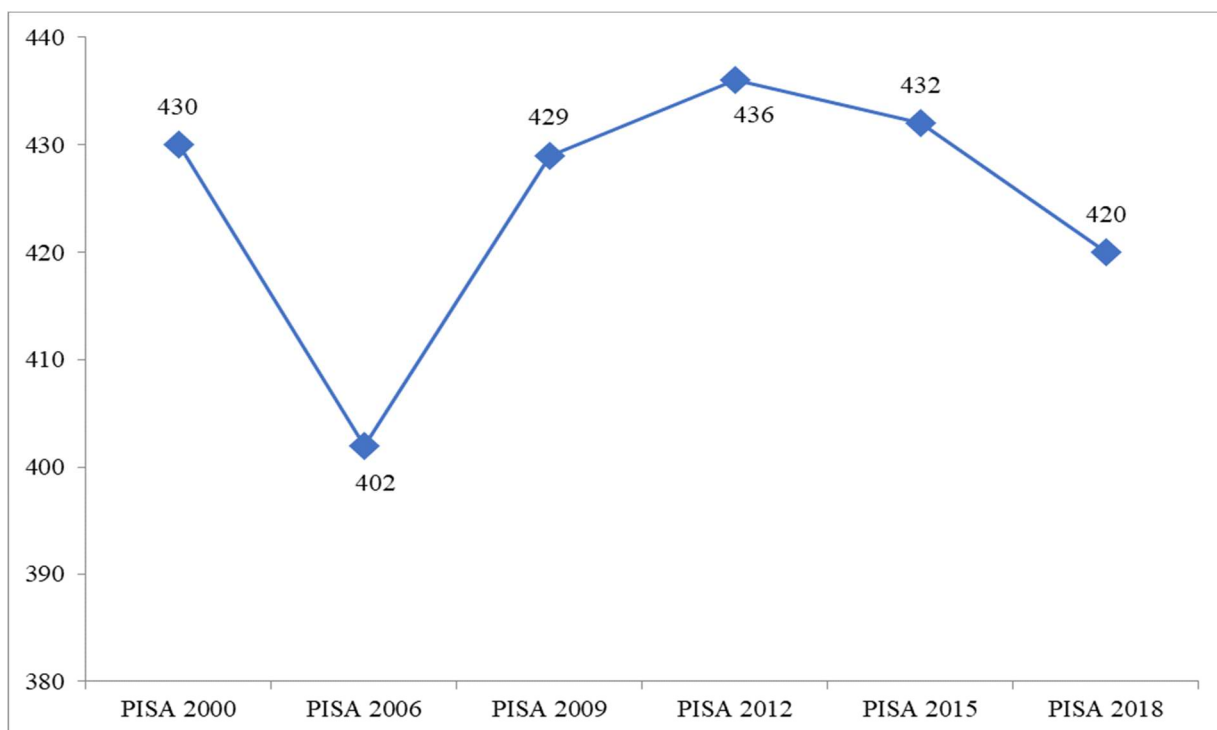
of difficulty and the solution logic. The result is expressed in total number of points. The Bulgarian grading system uses the following scale: up to 17.5 points means a poor score; 18–31,5 points means a pass; 32–58,5 points are classified as a good score; 59–86 points mean a very good score and 86,5–100 points are the excellent score.

RESEARCH RESULTS

A distinctive feature of PISA scores is that they do not have a substantive meaning. Instead, they are set in relation to the variation in results, observed across all the participants in the test. The results are scaled to approximately fit normal distributions, with means around 500 score points and standard deviations around 100 score points. To assist their interpretation in substantive terms, PISA scales are divided into proficiency levels. Each proficiency level corresponds to a range of about 80 score points; hence, the differences can be interpreted as the difference between successive proficiency levels. For the first time in 2018, the reading test was computer-based. The first indicator is the average score, calculated on the basis of the individual results of all the students from the nation. The average country score is analysed versus the average OECD countries' score, which in 2018 was 487 points.

The average score of the Bulgarian students in 2018 was 420 points versus the average OECD countries' score of 487 points. This result places Bulgaria last among the EU and EEA member states. The figure below, published in the OECD report, shows the average scores of the Bulgarian students in the waves held between the years 2000 and 2018. Evidently, after the year 2012 the average scores of the Bulgarian students are steadily deteriorating.

FIGURE 1
AVERAGE SCORES OF BULGARIAN STUDENTS



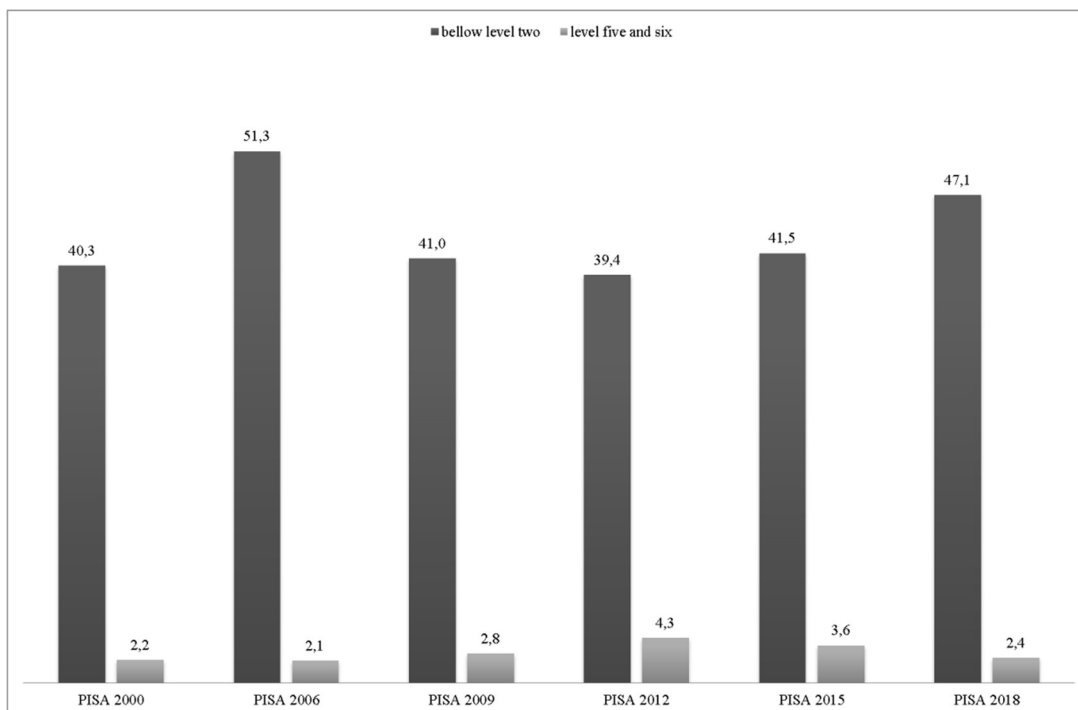
Source: OECD/PISA

To interpret the substantive meaning of these results we should use the average country score indicator. When compared, the 2015 proportion of 432/493 is very similar to the 2018 420/487, so we can assume

negligible difference in the performance. Thus, score itself is not the most disturbing fact. Traditionally, level 2 of the 6-level descriptive scale is defined in PISA as a critical threshold in the students' learning achievements, meaning that students whose results are below level 2 have significant omissions in their learning, qualified as functional illiteracy.

The next figure shows that the proportion of Bulgarian students below the critical level 2 threshold is quite large in all stages of the test. The students in the OECD countries with scores below level 2 for instance form just 22.7% versus 47.1% for the Bulgarian group in the latest assessment. At the same time, the proportion of students whose results are at the highest levels 5 and 6 is small and steadily going down after 2012.

FIGURE 2
PROPORTION OF BULGARIAN STUDENTS WITH RESULTS BELOW LEVEL 2 AND AT LEVELS 5 AND 6



Source: OECD/PISA

The conclusion is a steadily high proportion (close to 50%) of functionally illiterate students and an insignificant proportion of students with maximum achievements (less than 5%). The 2018 scores are similar to the 2009 scores, which tells us that no meaningful change has occurred in that period.

It further causes anxiety that Bulgaria is the only country where the difference in the learning results of the students in the highest and lowest quartile of the economic, social and cultural status is shrinking as a result of deteriorated scores of the students with the high socio-economic status. At the same time in Bulgaria the share of students with less favourable socio-economic status who achieve high results in reading (levels 5-6) is among the lowest of all the countries included in the 2018 test wave – 6% versus an OECD average of 11%. Also, the results show great heterogeneity in the quality of the education offered by the different schools. Bulgaria is one of the countries with the greatest variation in the “reading” results among the schools – 54.7%. Particularly worrying is also the academic segregation: the concentration of low achievers in certain schools and of high achievers in other schools is especially high (in fact among the highest).

This discovery instigated widespread public debate, in which a variety of explanatory opinions were voiced. Some blamed the conceptually outdated educational system (meaning its content); others the inadequate preparation of teachers (meaning the method); and still others the uninterested students and teachers (meaning the motivation). Many analysts on the education field define the Bulgarian learning paradigm as still focused on an educational model that is typical of the beginning of last century, which features a high level of standardization, formalism, unification and accuracy, teaching factual academic knowledge and formation of a narrow set of very specific skills. The described model is fit for a static system where the knowledge is fixed and once for a lifetime. In clear contradiction to it, there is a nearly universal consensus that in today's dynamic reality educational institutions should create skills for continuous learning. Both the empirical and the literature observations give us grounds to assume that the national system of education is lagging behind the societal development (confirmed by a series of national and international reports) and fails to adapt to the demands of the modern society.

The announced (Ministry of Education, Directorate Educational Programmes and Educational Content) average score in BLL (Bulgarian language and literature, corresponding to the reading ability in PISA) for the year 2015 was 58.68 (out of 100), which in the Bulgarian grading system is equal to a good-very/good score. Our data map shows a different picture - in 2015 the average score in NEA/BLL in the vast majority of the 265 municipalities was below 50 points, with roughly 2% of them having scores around the 50%. In 2018 again the average score is predominantly below 50 points with only one municipality as an exception.

FIGURE 3
AVERAGE NATIONAL EXTERNAL EVALUATION SCORES BY MUNICIPALITY - 7-TH
GRADE: 2015 / 2018

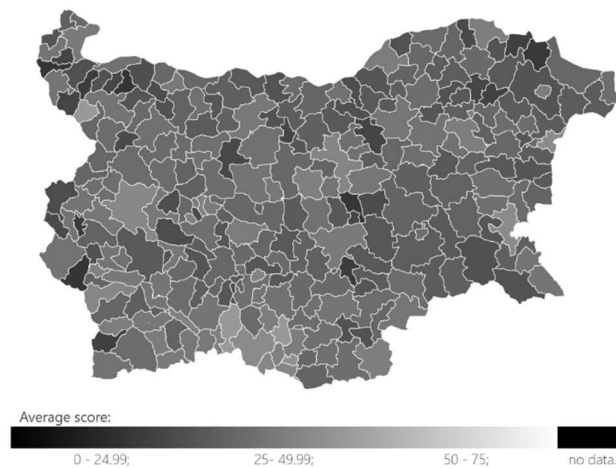


FIGURE 4
AVERAGE NATIONAL EXTERNAL EVALUATION SCORES BY MUNICIPALITY - 7-TH
GRADE: 2015 / 2018

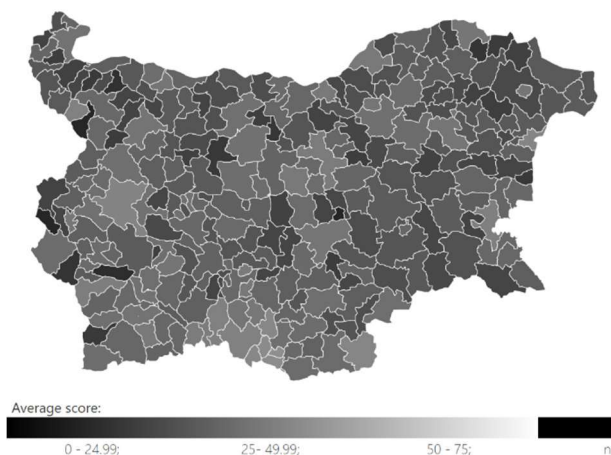


TABLE 1
COMPERISON OF AVERAGE SCORES

<u>Year</u>	<u>PISA</u>	<u>NEA</u>
2018	420/487	51.98/100
2015	432/493	34.10/100

In other words, looking at the year 2018 an average score of 51.98 points will be a good result in the Bulgarian interpretation, while the 420-average score in PISA suggest functional illiteracy. It is important to point out that in most grading systems worldwide a score of 50 and below is generally considered a failure, that is, below acceptance standards. There are few exceptions (Ireland, Latvia, Lithuania in the EU, certain countries in Africa and Asia) of grading systems where scores between 40 and 50 points are considered a pass, defined as “minimally accepted”. Compared to previous years, the 2015 results of the national external assessment (NEA) demonstrate a downward trend. Achievements of the students of the cohort in this year were worse compared to their NEA results after grade 4.

CONCLUSIONS

The objectives of NEA, as formulated by the Ministry of Education and Science, are to *diagnose the individual progress and the educational needs of the students; monitor the educational process for the purposes of quality improving policies and measures; establishing the level of achievement of the expected results as set out in the training curriculum and in the state educational standard of comprehensive education*. The tasks in the test are based on the mandatory learning content. As stated in the objectives, the national external assessment data are meant to be used at system and school level to reform the school curricula, plans and content, and for readjusting teaching methods with the aim of increasing the quality of education. In reality however, the published data on the results of the national external assessment in Bulgaria are a real challenge to use. They are of low quality and not fit for processing due to numerous omissions and inconsistencies. The absence of codes attributed to the schools until 2017 requires a disproportionate effort to make them usable and even after that moment they are very unreliable because the names of many schools change and are either not written correctly, or on occasions not at all. The figures included here are derived from data subject to considerable cleaning and processing and exclusion of all entries (schools) that cannot be identified. Secondly, the educational administration is not interpreting and

analysing the results. This is partly done by the Institute for Research in Education, an independent analytical centre working to help overcome the deficits. The Institute’s mission states that governance in education needs creation and intensive exchange of knowledge.

So how do the two sets of results compare and what do we learn from them? Firstly, the two sets of scores are incomparable due to the large discrepancy in the testing method and content of the learning achievements. The international test is visionary: it tells us where the country stands globally in terms of adaptability of a projected future society to the development dynamics based on degrees of proficiency and skills acquisition. The national assessment is myopic: it tells us what the students have learnt or not of the standard educational content. Secondly, the national results grading system does not correspond to the prevailing standards internationally and cannot provide a reasonable level of assurance that the measurement is adequate. Although they are basically testing one and the same thing – literacy and reading ability – these scores are in no way comparable.

Overall, the comparison discloses that the educational system fails to perform. It produces certain outstanding results on individual level, evidence of which is the continuous success of Bulgarian students at international Olympiads and the very good representation of Bulgarian students in prestigious universities. But clearly it does not perform enough to reproduce a well-functioning society. Indeed, this analysis is only based on a single competence, that of reading comprehension, and is thus limited. Therefore, it is important to point out here that those are actually the highest achievements of the Bulgarian students. The results in math and science are even worse.

The data, national and international, are consistent in one indicator – both show that the size of the place of living is a factor of quality of education and academic success. We see the highest scores in the largest settlements (cities with over 1 million and between 100 thousand and 1 million residents) and significantly lower in the smaller ones, descending along with the size of the settlement. The PISA test shows the same distribution of the results across the places of residence of different size.

FIGURE 5
RESULTS BY LOCATION SIZE AND TYPE OF SCHOOL

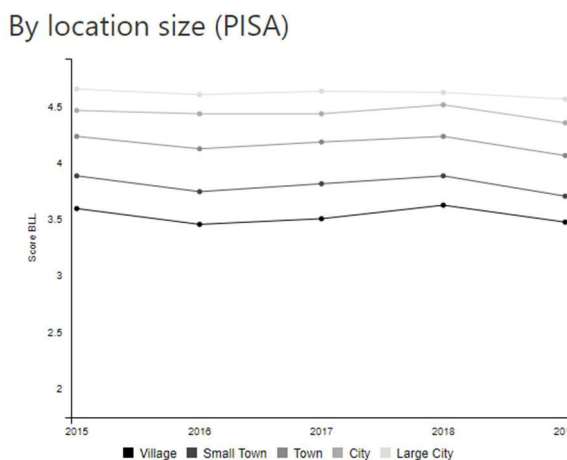
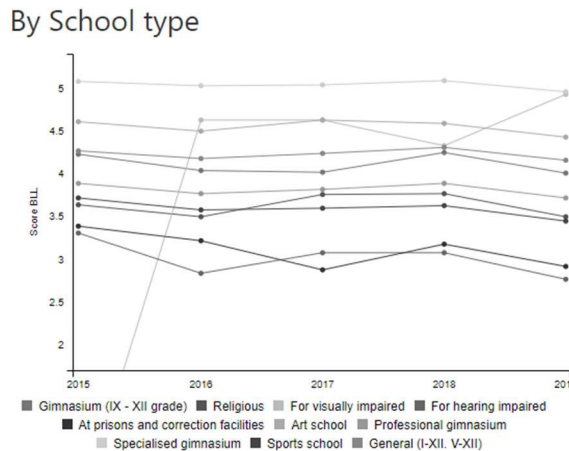


FIGURE 6
RESULTS BY LOCATION SIZE AND TYPE OF SCHOOL

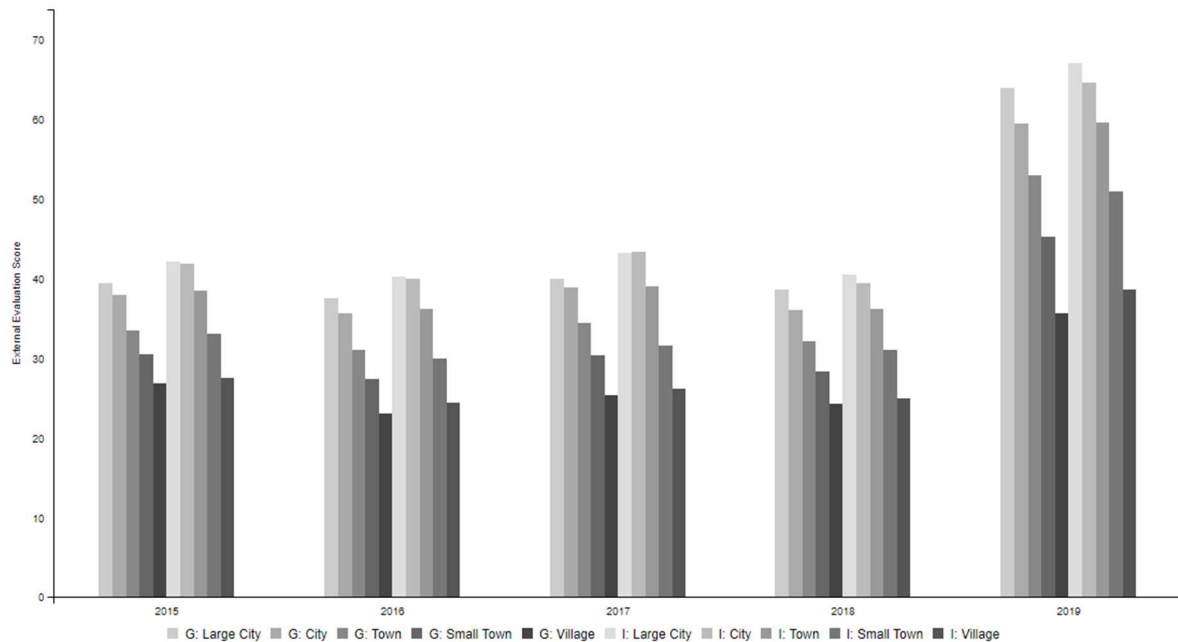


As seen on the above figure, the type of school also matters and here the data reflect the general trend of specialised (in math, science or various languages) comprehensive secondary schools demonstrating the best performance. An interesting fact worth researching is the exceptional performance of schools for visually impaired students.

It is largely believed that innovation has a role to play in providing good quality education and creating opportunities to develop a favourable learning environment and flexible approaches to teaching. This is also suggested by the European Commission and OECD, which states that “*In addition to quality learning, contemporary education needs to equip learners with transversal skills such as creativity, critical thinking and problem solving, as well as resilience and the ability to adapt to change*”. This assertion is checked here through comparison of the results of the so-called innovative schools with those of the traditional ones.

In the words of the Ministry of Education, innovative schools are a model of building the new educational paradigm for improving academic results and boosting critical thinking and creativity through innovative educational processes, teaching methods, school leadership and educational strategies. It is a bottom-up initiative where schools develop the innovative methods and are expected to disseminate and make them available to other schools. To qualify as innovative, a school should have the following characteristics: innovative learning content, curricula, and training plans; new learning standards are introduced, and students are trained to develop critical thinking; innovative teaching methods and supportive learning environment; innovative management and academic process organization; innovative teacher qualification and cooperation. The aim of the innovation should be increased quality of the education.

FIGURE 7
PERFORMANCE OF INNOVATIVE SCHOOLS BY LOCATION SIZE (PISA): BLL 2015-2019



Innovative schools have higher scores in places of residence of all sizes and the trend is sustained through the years for which data is available. The differences in the achievements between regular and innovative schools are the smallest in the smallest settlements but we should keep in mind that innovative schools there are very few. In conclusion, innovation is definitely a factor of quality in education.

There is also the question of an elaborate network of private lesson providers, which make it hard to distinguish between effective schooling methods and external sources. Various extra-school learning channels with high efficiency and popularity are also provided by the civil sector.

THE MEANING OF A LEARNING SOCIETY

It is said that the learning society is a means to achieve the knowledge economy and is potentially the answer to most problems of the modern world. The knowledge economy is increasingly dependent on knowledge workers rather than traditional manufacturing industries, which needs the workforce to become adaptable, particularly in the light of new technological developments. The White Paper “Teaching and Learning: towards the learning society” (1995) of the European Commission treats capital investment and investment in training on an equal basis. A learning society regards the actual process of learning as an activity and not a place, meaning that it is taken outside regular educational institutions, decentralized and deregulated. The accumulation of interactions between individuals becomes social capital for the learning society to develop. It needs motivated, engaged learners, prepared to conquer the unforeseen challenges of today and tomorrow; recognizes that people learn differently and strives to meet those needs; cultivates and embraces new learning providers from all sectors.

The learning society is a goal, aligned with the two dominant conceptualisations of education that we described at the beginning of this paper – human development for social change, and efficiency and effectiveness. The progress toward this goal is heavily dependant on the quality of education. The Annual Reports of the Ministry of Education on the implementation of the National Life-Long Learning Strategy do not contain any analysis of the indicator for reducing the proportion of students with low educational achievements at the age of 15. This is another evidence of the futility of any on-going reforms, assuming any are being implemented.

It becomes clear though that initiatives of the “innovative school” type which prove effective need serious attention.

REFERENCES

- Bråten, I., Strømsø, H., & Britt, M. (2009). Trust Matters: Examining the Role of Source Evaluation in Students’ Construction of Meaning Within and Across Multiple Texts. *Reading Research Quarterly*, 44(1). <http://dx.doi.org/10.1598/rrq.44.1.1>
- Britt, M., Rouet, J., & Durik, A. (2017). *Literacy Beyond Text Comprehension*. Routledge. <http://dx.doi.org/10.4324/9781315685682860>
- Bray, M., Adamson, B., & Mason, M. (Eds.). (2014). *Comparative education research: Approaches and methods*. Hong Kong: Comparative Education Research Centre and Springer. ISBN: 978-988-17852-8-2
- OECD. (2019). *PISA 2018 Results (Volume I). What Students Know and Can Do*. PISA, OECD Publishing, Paris. <http://dx.doi.org/10.1787/5f07c754-en>
- Perfetti, C., Landi, N., & Oakhill, J. (2005). The Acquisition of Reading Comprehension Skill. In M. Snowling & C. Hulme (Eds.), *The Science of Reading: A Handbook*. Blackwell Publishing Ltd. Oxford. <http://dx.doi.org/10.1002/9780470757642.ch13>
- Popham, J. (2000). *Modern Educational Measurement: Practical Guidelines for Educational Leaders*. ISBN-13: 978-0205287703
- RAND Reading Study Group, & Snow, C. (2002). *Reading for Understanding: Toward an R&D Program*. Reading Comprehension, Santa Monica, CA. Arlington, VA, Pittsburgh, PA. Retrieved from <https://www.jstor.org/stable/10.7249/mr1465oeri>
- UN. (2005). *Global Monitoring Report: The Quality Imperative*.
- van den Broek P., Bohn-Gettler, C.M., Kendeou, P., Carlson, S., & White, M.J. (2011). When a Reader Meets a Text: The Role of Standards of Coherence in Reading Comprehension. In M.T. McCrudden, J.P. Magliano, & G.J. Schraw (Eds.), *Text relevance and learning from text*. Information Age Publishing.