Chasing PISA: Rethinking the concept of external examination in the Republic of Macedonia

Gjorgji Dimeski

November 15th, 2017

Introduction

Despite a period of steady investment in the education sector, the Macedonian media and public were recently shocked by a single piece of evidence that raised suspicion about the effectiveness of the policy carried out by the Ministry of Education. Namely, according to the PISA 2015 testing data, about two thirds of Macedonian 15 year old students performed below basic proficiency in all three subjects (science, mathematics and reading) tested by PISA[1]. Moreover, The World Bank report claims that despite having the highest investment in education in the region[2] , the quality of the education process is in decline which is also evident from the comparison of the PISA 2000 and 2015 results.

To put things in context, high scores on PISA are mainly considered as a good indicator of quality education, primarily because the assessments take into consideration skills and knowledge that are relevant for economic growth (Kirsch et al 2013, Bybee 2008, Bybee & McCrae 2011, Nentwig et al 2009). In other words, the skills tested in PISA are translated to economic competencies i.e. competencies that one needs to have in order to succeed in the modern knowledge-based economy. That is why education policy makers generally use PISA (or other ILSAs) as an indicator but nonetheless are always uncertain about the how to act on it i.e. how to adjust the national education system to the standards 'imposed' by PISA. Unfortunately, the academic debate is still not clear on the 'recipe' for achieving education suitable for PISA established benchmarks. Is it clear however, that simply raising the quantity of education through increasing the primary and secondary enrollment rates is not enough, despite its obvious benefits empirically tested in Barro & Lee (1993), Romer (1990) and Mankiw (1992). Recent literature (Hanushek 2013, Hanushek & Kimko 2000) shows evidence on the stock of human capital being dependent of the quality of the schools, with indicative evidence that higher scores on international testing such as PISA are positively associated with some employment sectors important for the growth of the overall economy (Hanushek & Kimko 2000, Murnane et al. 2001, Cheung & Chan 2008). A couple of analysis found that achievements, measured on tests similar to PISA have a clear positive impact on individual productivity and earnings (Hanushek & Woessmann 2008, Murnane et al 2000, Mulligan 1999). However, analyzing specifically the education policy and growth, it is clear that there are no 'one shoe fits all' solutions i.e. country-specific strategies are needed for addressing the country-specific constraints.

With this in mind, the aim of this paper is straightforward; with PISA scores as a common proxy for the quality of the national human capital, the main question is how to craft a policy which will improve the quality of the education system, thus enabling it to provide skills and knowledge favorable to the globalized economic system. The analysis presented below can offer some insights in this regard, drawing an example(s) from the existing literature and contextualizing the lessons in the case of Macedonia, thus providing a general, policy-advising conclusion on how Macedonian education should be guided towards the path of improvement.

A call for reversing the standard input based education policies

Following the global trend of increasing public investment in education (OECD 2017) a lot a policy makers were left scratching their heads after the students' performance on international assessments was unchanged or even worsened in some cases such us Macedonia. The quest to improve the quality of the education system revealed only one thing- the tendency of government policy to typically move

[1] The Programme for International Student Assessment (PISA) is a triennial international survey which aims to evaluate education systems worldwide by testing the skills and knowledge of 15-year-old students.

[2] SouthEast Europe region, including: Macedonia, Albania, Kosovo, Serbia and Montenegro

to what is *thought to be the next best thing* (Hanushek 2003) is often fruitless. This does not mean that providing additional resources to schools doesn't matter *per se*, only that the returns to this kind of investment are often smaller than expected if not nil. The disappointing results of the past generally reflect pursuing of policies for which there is little empirical support, the current policy in Macedonia being one of the examples. Namely, Macedonian education policy is based on resource inputs such as raising the expenditure per student, resulting in smaller classes and/or equipping the classrooms with ICT technologies[3]. Over the decade, the policy was mainly aimed at increasing primary and secondary school enrollment rates[4], with the ultimate goal of higher levels of school attainment which the empirical evidence finds beneficial (Barro & Lee 2001). Nonetheless, the recent student assessment (PISA 2015) has shown than there is a decline in the education outcomes, proving that it is not appropriate to presume that *any spending* on schools is a productive investment. Combing through a vast quantity of empirical research, one might notice that apart from the quality of instructional material and the teaching personnel (Hanushek & Woessmann 2017), there is little sound evidence that any of the typically used, quantitative school inputs such as reducing the class size (pupil-teacher ratio), the characteristics of administration, or the facilities of the school increases the performance of the students (Hanushek & Woesmann 2007).

It turns out that it is *not the money that matter the most, but incentives*. Usually there is a rhetoric following certain policy goal, and the lack of measurements leaves a situation where the investment is continued mostly because of wishful thinking. For example, the arguments for cutting down class sizes usually are permitting more individualized instruction, allowing improved class room interactions, cutting down on pupil disruption etc. Yet, a vast amount of consistent work from Hanushek (2003, 2007, 2008, 2011, 2017) shows that this is simply not true. Instead, it is necessary to focus more on the factors that can be measured and thus *proven* effective. The two factors that more frequently appear to positively influence student outcomes are teacher experience and measures of teacher achievement tests (Hanushek & Woessmann 2007). Both of these factors emphasize incentives not quantitate inputs; From a policy perspective, this means that monitoring mechanism should be favorable for the performance, exercising pressure to all stakeholders to work responsibly.

Moreover, the current system is not capable of concluding precisely what Macedonian students lack in terms of knowledge and/or skills. No mechanism exists to support educators and students with real-time information about the current health of the system or provide feedback about the performance of its student. Education investment policy may seem unmethodical if the policy makers do not know the flaws that are apparent in their education system. A proper analysis of how students acquire their skills and knowledge i.e. decomposing the chain of learning in the existent system is what needs to be done if one desires to construct an educational policy suitable for the standards imposed by the knowledge economy.

Students assessments as a monitoring mechanism

From what was previously stated, it is not surprising that adding resources does not always translate into results, because the link between aspiration (for achievements) and motivation (incentives) is nonexistent, thus rendering policy virtually inept to dictate quality. Economists knew a long ago that accountability increases incentives, hence every structure, be it political or socio-economical has

[3] The biggest government investment was the project ,,Personal computer for every pupil" which equipped the classroom with computers and internet connection, but no evaluation of the project has been made thus making this investment controversial both financially and politically.

[4] In April, 2007 the Law on Secondary Education directed that secondary education is free and mandatory (Article 3) for every citizen of the Republic of Macedonia (Official Gazette No.47/2007)

established its own monitoring structure. One way to make the education system accountable is to impose a monitoring standard that can provide information about the level of student learning and compare the results with previously stated standards. This was the idea behind the external examination, a nation-wide compulsory examination of students in all subject areas, administered by the National Examination Centre in Macedonia. External exit exam systems are a device to increase accountability in the school system that has been repeatedly shown to be related to better student achievement[5] (Bishop et al 2006, Hanushek & Woessmann 2012, Woessmann 2001). After the termination of external examination in June 2017[6], the only nation-wide testing is the State Matura (in many countries known as Matriculation Examination) that assesses students only in 2 compulsory (mother tongue and foreign language or mathematics) and 2 optional[7] subjects. However, the results were not used to evaluate the impact on learning or monitor the students' performance as was initially expected from the external examination. The State Matura is required for enrolling in a university and the grade acquired can influence the chance of getting enrolled, meaning that the incentives for students are there. However, the State Matura as a mechanism for quality monitoring lacks two fundamental pillars of student assessments: 1) incentives for teachers and school administration and 2) the design is framed not to test *learning itself* but knowledge, translated into memorizing hard facts and recalling them during the exam.

The idea of the external examination is establishing a *large scale, system-level monitoring structure* that can provide *reliable comparative information* about school effectiveness. The concept included testing in all subjects, therefore providing comprehensive information about the quality of every component of the educational system; teachers, students, education materials and school administration. This information is further used for improvement, knowing what students/teachers or school administration lacks in terms of quality. A subsequent study of the results gained from the external testing can contribute to the knowledge base on educational effectiveness, observing patterns of relationships between inputs, processes, and outcomes of education.

The concept of national external examination very closely resembles those of International Large Scale Assessments (ILSAs) such as PISA. This resemblance is where one needs to look in order to understand the need of student assessment. The information acquired from ILSAs or external testing allows for *decomposition on the student performance* on three levels: individual, school and system. With this data, a student assessment offers indicators that monitor the functioning, productivity, and equity of education systems and knowledge on factors that determine their effectiveness (Braun 2013). In other words, these assessments provide unique data about what works and what not, and the impact of these factors on economic or social outcomes. The central difference is that the external testing is usually done every year, which enables policy makers to keep track of the educational system more frequently and include some data which are more country specific. In comparison, PISA is done every three years which means that policy makers could use it as a benchmark, a preparation of the national system to keep pace with the PISA requirements. Hence, the triennial PISA results could serve as a standard to evaluate whether the education system reforms done in the previous three years had been successful. Nonetheless, the external examination must be conceptualized in a way that will produce clear policy intentions which will ease its implementation. A swift adoption of the examination by the general public will enable further improvements regarding its purposes; for

[5] For individual positive effects on student learning see Bishop 2006, Holme et al 2010

[6] Official Gazette No.67

[7] While the compulsory subjects testing is more strict and administered by the accredited examiner, the optional subjects are internally tested which means they are administered by the school staff. Anecdotal evidence is showing that the internal testing is very lenient, with reports of cheating running rampant.

example ensuring the quality of the de-regulated system where for-profit institutions (i.e. private schools) enter the market and lessen the growing skepticism about the overall quality of the education system.

Conceptualizing the idea of external testing

Even though the concept of external examination had its obvious benefits, the criticism[8] aimed at its implementation in the case of FYR Macedonia had not been wrong either. Nonetheless, policy makers need to understand that bad implementation does not mean that the external examination is ill-conceived and as such should be repealed. Big lesson is to be learned by the model of the previously implemented testing and the improvements need to be taken specifically in those areas. The external testing in Macedonia had been more of a sanction than evaluate & correct mechanism[9]. The law that enacted the external testing in Macedonia had imposed fines[10] for the teaching staff if there is a difference between the grade given by the professor and the grade that the student achieved on the external examination. In this model, the responsibility for the student performance was only burdening the teacher, without taking into consideration the other factors that determine leaning outcomes such as learning materials, curriculum or school administration. Moreover, the grade from the external testing was included in the diploma that the student receives, thereby included in the overall score of the student. This may have influence on incentives positively by holding students and teachers accountable (Bishop 1997, Bishop 2006, Woessmann 2001), but what was missing is the evaluation done with the information taken from the exam. The low results or the grade deviations may have been due to bad educational materials or to an inappropriately constructed school curriculum, not necessary because of bad teaching staff or unmotivated students. Instead of *analysis that could get to the root of the problem* and act on it, the examination in the case of FYR Macedonia had been rather *used for pointing a finger at someone*. Aside the obvious shortfall of the policy, this also created some political problems with the implementation since everyone would pay the price for the problems in the education system, except the policy makers. Excerpt from the SABER Country Report made in 2012, the year when the external examination was formally launched summarizes:

While policymakers strongly support the External Assessment, educators, students, and parents, as well as some donors, oppose it. Stakeholders have challenged the validity of the External Assessment given its intended use, and expressed concern that its high stakes nature, as well as the format of the assessment instrument (only multiple-choice items), will result in teaching to the test. (Saber Country Report 2012)

Furthermore, the external testing had the same problem as the State Matura; testing knowledge instead of learning. In this case, even the National Examination Centre published the questions on their own website, which caused students to learn to recall memorized facts instead of learning how to apply a specific knowledge in a certain situation. Lastly, there was no mechanism[11] in place to ensure that the

[8] The report from Youth Educational Forum concerning the External Examination states multiple shortcomings in terms of technical issues, legal omissions and problems with the concept of external examination as defined in the Law on Secondary Education

[9] According to the 2012 SABER Country Report the main purpose of the External Assessment is to assess the objectivity of teachers' grading.

[10] Official Gazette 98/2015

[11] The external examination hadn't fully adhered the student performance standards; curriculum quality assessments and student learning performance had not been consistent in terms of all factors included (Education strategy 2018-2025- Draft Version)

external assessment is used in a way that is consistent with its intended purposes and technical characteristics, or to monitor its consequences.

However, the system of external examination as a monitoring procedure has a long tradition of usage (Silver et al 2005) and enjoys a considerable trust in those institutions it has been found (Hannan & Silver 2004, Brandt & Stensaker 2005). The potential of the mechanism lies specifically in the methodology used for data collection and analysis. For example, external examination is a more direct approach providing the examiners with possibilities for direct observations and close involvement in processes directly related to teaching and learning (Stensaker 2008). Given the current policy interest in school performance measured through learning outcomes and the qualifications of graduates, external examination serves as a *tool for analyzing the impact education has on the students*. Moreover, since external examination increases transparency, school staff would be induced to give enhanced learning higher priority. Teachers will assign more homework and parents will demand better science labs and more rigorous teaching. School administrators will be *pressured* to increase the time devoted to examination subjects and hire more qualified teachers. (Bishop 1998). Overall, the policy tools used for evaluation can produce results which by itself will pressurize the policy towards improvement.

Correcting the aforementioned issues should put the external examination back on the right path. It is important for policy makers to remember that the pitfalls of the implementation does not eliminate or diminish the benefits of the policy itself. It is clear how the above explained problems created political difficulty that ultimately lead to termination of the external examination policy. The lesson learned is thus straightforward; the introduction of the newly reformed external examination, with clearly policy stated goals, should be relatively easy from political perspective. The newly established goal should be centered around the idea of performance-tracking and evaluation tool, with the clear aim of achieving better results on PISA. If the policy takes the assumption that PISA scores are an indicator for economic competencies, than national scale, yearly assessment should be conducive to equipping the youth with economic competencies through preparing the education system for better results on the international assessments.

Concluding Remarks

Bad implementation does not mean that a concept is completely wrong. By itself, external examination is an excellent tool for keeping track of performance while also introducing a form of accountability which is favorable for incentives. When asking how education policies in developing countries can create the competencies and learning achievements required for their citizens to prosper in the labor market, the binding constraint seems to be institutional reforms, not resource expansions within the current institutional systems. For educational investments to translate into student learning, all the people involved in the education process have to face the right incentives that make them act in ways that advance student performance. With PISA scores as a benchmark for quality education system, the aim is to introduce reforms that will ensure the progress on the PISA country scoreboard i.e. reforms that will guide the overall system to produce comparable results with other countries. The implicit goal is not PISA by itself, but developing competencies conducive to the newly globalized knowledge economy. The purpose of this paper was to present a conclusion on the correct policy mechanism, hoping to at least spark a discussion in the academic or policy making circles. A deeper analysis however is in order. Drawing from the general idea of this paper, a clear guide of what works and what not is presented, a useful start for adjusting the education to performance-orientation and effectiveness.

References:

1. Barro, R.J. and Lee, J.W., 1993. International comparisons of educational attainment. Journal of monetary economics, 32(3), pp.363-394.

2. Barro, R.J. and Lee, J.W., 2001. International data on educational attainment: updates and implications. oxford Economic papers, 53(3), pp.541-563.

3. Bishop, J.H., 1998. The effect of curriculum-based external exit exam systems on student achievement. The Journal of Economic Education, 29(2), pp.171-182.

4. Bishop, J.H., 2002. What is the appropriate role for student achievement standards?. In Conference Series-Federal Reserve Bank of Boston (Vol. 47, pp. 249-278). Federal Reserve Bank of Boston; 1998.

5. Bishop, J., 2006. Drinking from the fountain of knowledge: Student incentive to study and learn–externalities, information problems and peer pressure. Handbook of the Economics of Education, 2, pp.909-944.

6. Bloem, S., 2015. PISA for low-and middle-income countries. Compare: A Journal of Comparative and International Education, 45(3), pp.481-486

7. Bybee, R.W., 2008. Scientific literacy, environmental issues, and PISA 2006: The 2008 Paul F-Brandwein lecture. Journal of Science Education and Technology, 17(6), pp.566-585.

8. Bybee, R. and McCrae, B., 2011. Scientific literacy and student attitudes: Perspectives from PISA 2006 science. International Journal of Science Education, 33(1), pp.7-26.

9. Cheung, H.Y. and Chan, A.W., 2008. Understanding the relationships among PISA scores, economic growth and employment in different sectors: A cross-country study. Research in Education, 80(1), pp.93-106

10. Ertl, H., 2006. Educational standards and the changing discourse on education: The reception and consequences of the PISA study in Germany. Oxford Review of Education, 32(5), pp.619-634.

11. Greenwald, R., Hedges, L.V. and Laine, R.D., 1996. The effect of school resources on student achievement. Review of educational research, 66(3), pp.361-396.

12. Glewwe, P.W., Hanushek, E.A., Humpage, S.D. and Ravina, R., 2011. School resources and educational outcomes in developing countries: A review of the literature from 1990 to 2010 (No. w17554). National Bureau of Economic Research.

13. Hannan, A. and Silver, H., 2004. ENQUIRY INTO THE NATURE OF EXTERNAL EXAMINING Final report. Higher Education Academy. Retrieved June, 23, p.2006.

14. Hanushek, E.A. and Woessmann, L., 2007. The role of school improvement in economic development (No. w12832). National Bureau of Economic Research.

15. Hanushek, E.A. and Woessmann, L., 2012. Do better schools lead to more growth? Cognitive skills, economic outcomes, and causation. Journal of economic growth, 17(4), pp.267-321.

16. Hanushek, E.A., 2013. Economic growth in developing countries: The role of human capital. Economics of Education Review, 37, pp.204-212.

17. Hanushek, E.A. and Woessmann, L., 2017. School resources and student achievement: a review of cross-country economic research. In Cognitive Abilities and Educational Outcomes (pp. 149-171). Springer International Publishing.

18. Hanushek, E.A. and Kimko, D.D., 2000. Schooling, labor-force quality, and the growth of nations. American economic review, pp.1184-1208

19. Hanushek, E.A. and Woessmann, L., 2008. The role of cognitive skills in economic development. Journal of economic literature, 46(3), pp.607-668.

20. Hanushek, E.A. and Woessmann, L., 2010. The High Cost of Low Educational Performance: The Long-Run Economic Impact of Improving PISA Outcomes. OECD Publishing.

21. Hanushek, E.A., 2003. The failure of input-based schooling policies. The economic journal, 113(485).

22. Holme, J.J., Richards, M.P., Jimerson, J.B. and Cohen, R.W., 2010. Assessing the effects of high school exit examinations. Review of Educational Research, 80(4), pp.476-526.

23. Lang, K., 2010. Measurement matters: Perspectives on education policy from an economist and school board member. The Journal of Economic Perspectives, 24(3), pp.167-181.

24. Lingard, B., Martino, W. and Rezai-Rashti, G., 2013. Testing regimes, accountabilities and education policy: Commensurate global and national developments.

25. Lockheed, M., 2015. Why do countries participate in international large-scale assessments? The Case of PISA.

26. Lockheed, M.E. and Hanushek, E., 1988. Improving Educational Efficiency in Developing Countries: what do we know?[1]. Compare, 18(1), pp.21-38.

27. Mankiw, N.G., Romer, D. and Weil, D.N., 1992. A contribution to the empirics of economic growth. The quarterly journal of economics, 107(2), pp.407-437.

28. Monteils, M., 2004. The analysis of the relation between education and economic growth. Compare, 34(1), pp.103-115.

29. Murnane, R.J., Willett, J.B., Braatz, M.J. and Duhaldeborde, Y., 2001. Do different dimensions of male high school students' skills predict labor market success a decade later? Evidence from the NLSY. Economics of Education Review, 20(4), pp.311-320.

30. Nentwig, P., Roennebeck, S., Schoeps, K., Rumann, S. and Carstensen, C., 2009. Performance and levels of contextualization in a selection of OECD countries in PISA 2006. Journal of Research in Science Teaching, 46(8), pp.897-908.

31. Romer, P.M., 1990, March. Human capital and growth: theory and evidence. In Carnegie-Rochester conference series on public policy (Vol. 32, pp. 251-286). North-Holland.

32. SABER Student Assessment Country Report 2012; FYR of Macedonia

33. Sahlberg, P., 2006. Education reform for raising economic competitiveness. Journal of Educational Change, 7(4), pp.259-287.

34. Sahlberg, P., 2007. Education policies for raising student learning: The Finnish approach. Journal of Education Policy, 22(2), pp.147-171.

35. Sahlgren, G.H., 2014. Incentive to invest? How education affects economic growth.

36. Schleicher, A., 2009. Securing quality and equity in education: Lessons from PISA. Prospects, 39(3), pp.251-263.

37. Stensaker, B., Brandt, E. and Henrik Solum, N., 2008. Changing systems of external examination. Quality assurance in Education, 16(3), pp.211-223.

38. Stensaker, B.R., 2003. Trance, transparency and transformation: the impact of external quality monitoring on higher education. Quality in higher education, 9(2), pp.151-159.

39. Von Davier, M., Gonzalez, E., Kirsch, I. and Yamamoto, K. eds., 2012. The role of international large-scale assessments: Perspectives from technology, economy, and educational research. Springer Science & Business Media.

40. Vegas, E. and Coffin, C., 2012. Education Finance: it's how, not simply how much, that counts.

41. Woessmann, L., 2003. Schooling resources, educational institutions and student performance: the international evidence. Oxford bulletin of economics and statistics, 65(2), pp.117-170.

42. World Bank, Public Expenditure Review: Fiscal Policy for Growth in FYR MACEDONIA