Educational Resources: Saturation or Satiety?

Agranovich M.

Mark Agranovich
Candidate of Sciences in Economics, Director of the Research Center for Monitoring and Statistics of Education, Federal Institute for the Development of Education, Russian Presidential Academy of National Economy and Public Administration. Address: Bld. 9-1, Chernyakhovskogo St, 125319 Moscow, Russian Federation. Email: agranovich-ml@ranepa.ru

Abstract. The article explores the methodological issues of education system evaluation. Such evaluation, based on the comparative analysis of national and regional education systems, is an important source of information for education policy design and implementation through educational development programs. The existing approaches to and methods of education system evaluation are discussed. It is shown that they are based on internal indicators, not those perceived by end users or the effects external to the system. Even though there have been some objective reasons for using such approach, it remains unclear to what extent its results reflect the educational outcomes for the end users—that is, individuals, society, and economy? Statistical analysis methods are applied to OECD education statistics to investigate the relations between the education indicators characterizing the level of educational attainment, education accessibility, and the amount and quality of the resources involved, on the one hand, and the outcome effects for individuals, society, and economy. Where such relations are observed, they tend to be non-smooth and only manifest themselves up to a certain point in the vast majority of cases. Such cessation of growth in the outcome indicators that happens after achieving certain levels of resources involved, educational attainment, and other education indicators can be described as oversaturation or satiety effect. Inferences about the limitations and conditions of applying education indicators in education system evaluation are drawn from the findings. Keywords: education system evaluation, education indicators, external effects, outcome indicators for end users.

DOI: 10.17323/1814-9545-2019-4-254-275

Researchers have been interested in evaluating education systems and their learning outcomes for decades. Such evaluation is vital to understand how adequately education policy and education system respond to social, economic, and individual needs. But what is actually “good education” or a good education system? What criteria and, consequently, indicators can be used to evaluate an education system, its performance and evolution dynamics?

It would seem natural to link education system learning outcomes to the goals set. However, the statutory education objectives
officially declared in the National Doctrine of Education in the Russian Federation\(^1\), the Key Areas of Long-Term Socioeconomic Policy of the Government of the Russian Federation (The Gref Program, 2000), and the Education Development State Program 2013–2020\(^2\) are too general to build a system of evaluation criteria and indicators on their basis.

Another possible way of evaluating an education system is to see whether its learning outcomes match the demands of society, economy, and individuals/families. Yet, this option is fraught with two hindrances. First, such demand is only formulated (if at all) in a very general, non-operational form, being hard to convert into criteria and indicators. Second, learning outcomes, except in professional development courses, are achieved with a significant time lag, so the demand identified today may become irrelevant by the time current students graduate.

Impossibility or inability to provide objective- or demand-based evaluation of education systems along with a number of other challenges induced by the intrinsic characteristics of education that will be discussed below necessitate searching for proxy ways of building education system evaluation. The popular quality–accessibility–effectiveness triad has not solved the problem completely, faced with the predictable lack of uniform understanding of what quality is and how effectiveness should be measured.

Today’s most common approach to education system evaluation makes allowance for system size (educational attainment), academic achievements, and resources involved. Using such criteria and relevant indicators is a desperate measure, since they characterize intermediary, intersystem results instead of end-user ones. The approach based on those characteristics assumes the following:

- The higher the level of educational attainment, the higher the output in terms of economic growth and employment;
- We know which knowledge and competencies will be in demand in after-school life;
- The more resources involved in education and the higher their quality, the better the final outcomes of system functioning.

This article uses statistical analysis of international and regional statistics in an effort to measure the extent to which the assumptions above are plausible, i.e. whether it is true that the amount and quality of educational resources, educational attainment, and student achievement are positively related to end-user satisfaction with education system outcomes.


\(^2\) Approved by Governmental Resolution No.295 of April 15, 2014.
The possibility of using other indicators external to education systems to measure the relationship between education characteristics and its potential outcomes for the economy and society is also discussed in the article.

### 1. Education System Evaluation Indicators and System Outcomes

Three categories of indicators are used in various combinations today to evaluate education systems, based on whether they characterize (a) student achievement, (b) education system size and education accessibility, and (c) resources involved in education.

To what extent are those indicators related to the external effects of education for individuals, society, and economy? Is outstanding academic achievement related to successful graduate socialization, high levels of educational attainment to economic growth and labor productivity, and large amounts of resources involved in education to better system outcomes?

To answer this question, we are comparing the education system indicators to the end-user outcomes of education for individuals, society, and economy. We are using economic growth and labor productivity as outcome indicators for economy and the Gini coefficient as an outcome indicator for society. Of course, these two indicators are largely affected by other factors as well, but positive economic impact and reduction of social inequality are key objectives of education systems, so relations to education are expected to be observable.

As for end-user outcomes for individuals, adaptation to after-school life, i.e. successful transition to the labor market or to the next education level, is a critical one. The success of this transition is measured based on the “share of youth not in employment, education or training” (NEET) [OECD 2018]. This measure is external to the education system and describes the integration of graduates into socioeconomic life. The relationship between NEET and education has been investigated in a number of studies, including the longitudinal Effective Pre-School, Primary and Secondary Education survey (EPPSE 3–16+) administered by the University of London’s Institute of Education [Siraj at al. 2014]. Another study, produced by Sarah Gracey and Scott Kelly, demonstrates that the content of lower secondary curriculum as well as early school leaving have a critical impact on NEET [Gracey, Kelly 2010].

National and regional NEET rates are certainly affected by other factors as well, unemployment in particular, but they measure graduate socialization, which is bound to be a educational outcome, at least to some extent. We are using NEET as an external measure of school education quality among 15- to 19-year-olds. Education system outcomes are more than likely to have an impact at later stages of human life as well, but the longer the time lag, the more difficult it is to separate the influence of education from that of other factors affecting ca-
reer and social life. For that reason, analysis in this article is restricted to the age group specified above.

1.1. Integrated Assessment

Pearson’s Global Index of Cognitive Skills and Educational Attainment\(^3\) and UNESCO’s Education for All (EFA) Development Index\(^4\) are the most famous integrated systems of education assessment today.

The Global Index of Cognitive Skills and Educational Attainment ranks countries based on PISA, PIRLS, and TIMSS scores, literacy rate, and secondary and tertiary enrollment rates.

The EFA Development Index is compiled of primary education enrollment, adult literacy, and mean male the gender parity indexes (GPI) for primary education, secondary education and adult literacy, and primary school graduation rate. However, EFA estimators apparently use poor-quality data and/or imperfect methodology, otherwise it is hard to explain why Russia is ranked below Kirgizia and Tajikistan, being essentially behind the latter in primary education enrollment and primary school completion rates\(^5\). Yet, some comfort can be taken in the fact that education systems of several advanced economies, including South Korea, the United States, and Israel, are ranked even lower.

Integrated education assessments have been performed in some countries, including Czech Republic [OECD 2012b] and New Zealand [OECD 2010], within the framework of the OECD Education Statistics program. Those assessments used a wide array of indicators largely based on two types of data, (a) statistics submitted by educational institutions and (b) national performance in PISA, PIRLS, and TIMSS.

The key indicators of Russia’s Education National Priority Project\(^6\) also belong to the domains of academic achievement and enrolment in education, describing the performance of Russian school students in international student assessments and their participation in extracurricular activities.

A noteworthy attempt to develop a theoretical framework for integrated assessment of education systems in Russia was made by Al-


\(^6\) Passport of National Priority Project “Education” was approved by the President of the Presidential Council for Strategic Development and National Priority Projects (Protocol No.16 of December 24, 2018).
exander Novikov and Dmitry Novikov [2006]. The authors identified groups of users of assessment results, from students to society to enterprises, and assigned a set of criteria to each of the groups. Unfortunately, those sets of criteria have not been justified scientifically, the authors admitting that nearly each of them “has not been formalized so far”.

1.2. Student Achievement and Its Trends

Student performance in national and international educational assessments is the most popular measure of education system evaluation. However, validity of such approach is open to question. First, that same PISA showed that socioeconomic background explained the largest amount (19 percent) of the within-school variance in student performance [OECD 2004], which means that it is not educational effort but parental education and family income that this variance depends on. High scores in the Unified State Exam (USE) are largely the product of tutoring and extracurricular courses rather than school effectiveness. What proportion of learning outcomes is actually owed to school, teachers, program, and curriculum? If it is less than a half, then whom and what do we evaluate? Second, and this seems to be an even stronger reason for doubting the feasibility of student achievement as a criterion for education system evaluation. Education is a unique sector where those who decide what to teach and how to assess learning outcomes are the same people who actually perform evaluation and use its results to make inferences about system effectiveness and development. Yet, there has been no evidence that the educational objectives and learning outcomes reflected in curricula or, indirectly, national and international assessment tests, will pave the way to success in adult life. Neither is there evidence that high USE or PISA scores correlate positively with self-fulfillment and success in after-school life or satisfy the demand of end users, i.e. individuals, society, and economy. Particularly, it concerns the prospective demand, since it is not until tomorrow that the demand for and use of today’s education outcomes can be assessed. Nonetheless, the results of international student assessments are still basically the only measure of education system performance available today.

The relationship between student achievements and graduate socialization is ambiguous. Figure 1 demonstrates PISA scores and NEET values across countries. As youth unemployment varies significantly across the sample—from 2.3% in the Czech Republic to 19.0% in Greece—the NEET values have been adjusted to eliminate the impact of this factor.

As seen from the graph, no correlation between PISA scores and NEET is observed beyond the threshold of 470 scores.
Educational attainment is an important indicator in most methods of education system evaluation. By tacit agreement, higher percentage of educated population is expected to correlate with higher levels of national and personal income. With regard to correlation between educational attainment and economic growth, this assumption is confirmed with a correlation coefficient of 0.47 (Figure 2).

However, it cannot be claimed that high economic growth follows from high educational attainment rate and not vice versa. Besides, a
strong positive correlation is only observed up to approximately the 45th percentile of the population with tertiary diploma. Beyond that, other factors apparently come into play, as per capita GDP is higher in the countries with tertiary attainment rates of 40–45% than in those with higher rates.

It could be suggested that educational attainment is related to labor productivity at the national level. This hypothesis was investigated, in particular, using the 1990–2007 data on South Korea, where a significant relationship between the change in the rate of people with higher education among the economically active population and sales per person was revealed [Jung Dae Bum 2015].

Here, however, just as in the previous case, the positive correlation assumption is only confirmed up to a certain point. As seen in Figure 3, labor force productivity is lower in the countries with the highest proportion of people with tertiary education (above 46%) than in those with less educated population (40–45%).

This is where it would be appropriate to recall Eric A. Hanushek’s work demonstrating that economic growth is affected more by education quality than by the enrollment and attainment rates [Hanushek, Woessmann 2007].

Results of the Programme for the International Assessment of Adult Competencies (PIAAC) [OECD 2016b] contributed to the evidence for the prevalent role of education quality and curriculum as compared to attainment indicators. A comparison of average reading and mathematical literacy rates to those of educational attainment across the OECD countries revealed, again, that an increase in educational attainment correlates significantly positively with function-
al literacy but only at small scales, losing its effects on both reading and mathematical literacy after reaching a certain threshold (Figure 4).

Institutional selectivity is another reason why the influence of educational attainment on labor productivity and individual indicators of socioeconomic status has been put into question [Torche 2011].

Reduction of social inequality is another socioeconomic effect allegedly related to education. It is widely believed that better education accessibility may gradually reduce the social inequality gap. A number of studies demonstrate the role of education as a driver of social mobility; in particular, Fabian T. Pfeffer and Florian Hertel have found that the impact of socioeconomic status on social mobility goes down with increasing educational attainment [Pfeffer, Hertel 2015]. The income Gini coefficient is a recognized indicator of social inequality. Education accessibility is measured using the participation, graduation, and attainment rates. Figure 5 presents correlation between Master’s degree graduation rates and Gini coefficient value by countries. For the whole sample the correlation coefficient between the two variables is statistically significant (−0.59) across the 39 countries for which data is available, confirming the assumption that income disparities diminishes as education grows more accessible. Yet, the correlation disappears as soon as the graduation rate exceeds 22%.

A similar trend is observed when the Gini coefficient is linked to educational attainment (Figure 6), correlation between the indicators disappearing or changing its sign as soon as the percentage of people with tertiary education exceeds 31%.
Along with attainment rate and student achievement, indicators reflecting changes in the amount or quality of resources involved in education are also used when measuring progress in the development of education and education systems at international and national levels. Common examples include such indicator of Sustainable Development Goals in education (SDG 4) as "percentage of teachers who
received in-service training in the last 12 months” or the Education National Priority Project indicator “participation rate of children aged from 5 to 18 in extracurricular activities”.

Clearly, higher participation rate in extracurricular activities, or higher teacher salaries, or even higher percentage of teachers who received in-service training cannot serve as targets of an education development program. Those are outcomes within the system. Program targets and relevant target indicators should be described in terms of effects for the external user. For example, it can be academic and social success in primary school for preschool education, increase in the percentage of economically active youth and reduction of youth unemployment for vocational education, etc. However, using such indicators involves considerable difficulties.

The difficulties associated with using outcome indicators are well-known. They include, first of all, delayed effects of education on socioeconomic status, social life, career and social success of an individual. The end results of innovations, managerial and economic initiatives in education often become visible beyond the program’s planning horizon—and even then, they are usually distorted by external factors.

Furthermore, the program reporting systems are very specific, which cannot be neglected. In Russia, programs have no right to remain unfulfilled. That is why developers and implementers refrain from using indicators they cannot influence directly, which is exactly what outcome indicators are. Unlike final outcome indicators and end-user effects, indicators of resources involved entirely depend on program implementers and funds availability, making it much easier to achieve the goals set.

For the reasons stated above, indicators of resource quality and quantity are used as a proxy measure. Investments in education have been traditionally viewed as a way of improving its quality and accessibility, better teacher competence being expected to enhance student performance. The interrelation between results and spending has been represented as a saturation curve and searching for a reasonable limit to resource intensification has been widely discussed in the advanced and most of the middle-income economies. In the developed countries, a trend has even emerged towards reducing somewhat the relative education costs in one form or another, increasing the average class size in particular. However, analysis of international statistics shows that the relations between investments and outcomes in education are non-smooth and can rarely be described using the saturation curve.

The choice of indicators reflecting quantitative and qualitative changes in resources (similar to those mentioned above) as program

target indicators is fairly explainable. It is based on the assumption that an increase in the amount or quality of resources involved in education leads to improvements in education quality and/or accessibility. However, resource-based indicators can only be used as target indicators of national and international programs in case there is a proved causal relation between resources and final outcomes, which is certainly not always the case, no matter how obvious it might seem. With respect to all the SDG 4 indicators designed to increase the quantity or quality of resources involved in education, no statistically significant correlation has been found with the results of international education assessments [Agranovich 2017]. In some instances, the relation between expenses in the broadest sense and learning outcomes is observed up to a certain point, beyond which the correlation between the size of a country’s investments in education and its performance in international student assessments is close to zero.

Let us consider some examples

When developing the SDG 4 indicator “average teacher salary relative to other professions requiring a comparable level of qualification” [Statistical Commission of the United Nations Economic and Social Council 2016], the authors (UNESCO Institute for Statistics) premised that higher salaries would attract or allow selection of top professionals, thus enhancing education quality. However, PISA results raise doubts as to the direct relationship between teacher pay and learning outcomes. Figure 7 shows the correlation between relative lower secondary school teacher pay and average performance in PISA 2015 across the countries.

The graph clearly demonstrates the absence of relationship between relative teacher pay in Russia and performance of 15-year-old Russian students in the PISA assessment. Unfortunately, data on teacher salaries is only available for 28 countries participating in PISA, which is not enough for an adequate correlation analysis, yet the correlation coefficients estimated using this data (–0.18 for reading and –0.13 for mathematics) definitely indicate that there is no positive correlation between the size of teacher pay and the learning outcomes.

Similarly, no relationship was found between relative teacher pay and NEET (adjusted). In addition to being statistically insignificant (0.11), the correlation coefficient between the two indicators has a positive sign, meaning that higher teacher salaries leads to the higher NEET.

Correlation analysis between the indicator “percentage of teachers who received in-service training in the last 12 months by type of training” and performance in PISA 2015 across the countries (Figure 8) also finds no positive relationship between the two variables.

The correlation coefficients between the percentage of teachers who received in-service training and PISA performance on the reading and mathematical literacy scales estimated for the 28 countries for which data is available are 0.07 and 0.01, respectively. Other-
wise speaking, there is no statistically significant relationship between these two indicators. It may be suggested that the content, method, and duration of teacher training are at least as important indicators as the very fact of training experience.

No relationship between teacher participation in in-service training and NEET as a characteristic of graduate socialization has been found, either (Figure 9).
In addition to the above examples of no expected correlation between invested resources and educational outcomes, two more arguments will be given below to support the idea that caution should be exercised when using resources as an indicator in education quality evaluation.

Figure 10 presents data on time devoted to mathematics in primary school curricula and four-graders’ TIMSS scores by country [Mullis et al. 2016]. It is clearly seen from the graph that time spent on ear-
ly numeracy learning activities is not related to TIMSS performance in any way.

The other argument is borrowed from PISA in Focus Notes [OECD 2012c]. As can be seen in Figure 11, larger amount of resources (in this case, average spending per student aged 6 to 15) is related to better student performance up to a certain point. Beyond that point, however, the correlation between investments and learning outcomes is not observed.

2. How Education is Perceived by the Population and the Parties Involved

Satisfaction with education occupies a special place in education system evaluation. In particular, this indicator is used to assess performance of executive authorities.

The Russian Education National Priority Project includes indicators of employers' satisfaction with graduate skills with a breakdown by college and higher education programs.

---

It is not until monitoring instruments are developed (integration expected in 2022) that the use of this indicator for education system evaluation can be considered. Yet, a lot of studies in Russia [Selivertova 2018] and other countries (e.g. [Succi, Canovi 2019]) indicate that graduates’ professional skills are not the top-priority recruitment criteria for employers.

Using the results of population surveys for education system evaluation also appears to be irrelevant. Our findings [Agranovich 2010] show that satisfaction with the education system does not correlate with such indicators important for families as participation in preschool education or regional average USE score. Rather, what is observed is that population’s perceptions of the education system are related to regional economic growth indicators, unemployment, and household income. Therefore, it is rather general public mood than attitude toward education that is described by such surveys.

In fact, opinions of the parties involved in the educational process (students, parents, school teachers and administrators) [Avraamova et al. 2015] that are revealed as a result of population surveys can tell more about the respondents themselves than the educational situation.

3. Conclusion

Analysis of the education system evaluation and outcome indicators presented in this article indicates that an increase in educational attainment or amount and quality of educational resources only promotes improvements in the outcomes for end users (individuals, society, and economy) up to a certain point. In other words, there is no evidence to support the hypothesis that the saturation curve can describe the relationship between resources, student achievement, and educational attainment, on the one part, and education system performance, on the other part. There are points of satiety, beyond which resource intensification does not improve the outcomes anymore.

Consequently, the existing approaches to developing the indicators for evaluating education systems and their progress have to be revisited, both in general and with regard to education development program monitoring.

First, indicators of final/end-user effects should be preferred over intermediary/within-system indicators. Educational outcomes for individuals manifest themselves in the successful socialization in life after school. Of all the existing indicators of socialization, “youth not in employment, education or training” (NEET) is most workable one. Of course, this indicator is affected by other factors as well, unemploy-

---

ment in particular, but it measures graduate socialization, which is bound to be a learning outcome, at least to some extent.

Second, where using external indicators to evaluate education systems or monitor achievement of education development program goals has some objective limitations, education indicators can only be used provided there is a proven positive correlation between resource intensification and end-user outcome indicators, a correlation between resource characteristics and external effects provided by the education system. Proxy measures should also be designed to evaluate the most advanced education systems, as capacity-based indicators are ineffective as a measure in this case, which has been proved by the analysis above.

With regard to the design of national or regional education development programs, using indicators that reflect quantitative or qualitative improvements in resources involved in education as well as quantitative education system characteristics (participation rate by level/program, educational attainment, etc.) as target indicators is fraught with actually failing to achieve the program goals, because an increase in the quantity or quality of resources may have no impact on the final outcome beyond a certain threshold. Moreover, there is a real risk of using ineffectively the funds allocated for resource intensification.

References


Agranovich M
Educational Resources: Saturation or Satiety


OECD (2012c) PISA in Focus 2012/2. Paris: OECD. http://dx.doi.org/10.1787/5k-9fhmfzc4xx-en


OECD (2016b) PISA 2015 Results in Focus. Paris: OECD.


