Value-added indicators for Polish schools: results of their validity analysis

Aleksandra Jasińska
Educational Research Institute
Poland

Abstract

For nearly a decade, the methodology of the evaluation of schools based on their educational value-added has been developed and popularised in Poland. So far, we have developed the so-called Educational Value-Added (EVA) indicators for lower-secondary and upper-secondary schools. In order to improve our models and assess validity of the value-added indicators for Polish schools, we had also done longitudinal researches on the random samples of schools. In the article I will present main results of one of these studies, which were conducted in Polish lower-secondary schools. I will focus on the issue of dependence between educational value-added and students characteristics that are not influenced by the school (like socio-economic status and level of intelligence). Selected factors that are connected with variance of the value-added indicators will be also shown.

Keywords: value-added, validity, contextual variables, intelligence increments, school effectiveness

1. Introduction

Since 2002, the external examinations system has been present in the Polish education system. Its emergence made it possible to develop teaching evaluation methodology. In the period of 2005-2007, research activities were launched in order to develop the Educational Value-Added Models for Polish schools. As a result of works of the methodological team, EVA indicators for lower-secondary schools were created, and potential users of this method were trained. The year 2007 saw the launch of a research project whose goal was to develop methodology for applying external examination scores to the evaluation of teaching effectiveness[1]. The team's works focus on the creation and testing of Value-Added Models, monitoring how external examinations are used to teaching evaluation, development of training materials and strategies for the users of the EVA method. To date, EVA models for lower secondary schools and upper secondary schools have been developed[2]. Studies whose purpose is to improve the methodology for calculating the EVA indicators and their validity assessment constitute an important part of the project. As part of the project, three longitudinal studies are being conducted, each of which corresponds to a subsequent education level. These are: the study in primary schools, study in lower secondary schools and study in upper secondary schools. The study in primary schools was planned in order to evaluate the capacity to assess the value-added indicators at this stage of education. This article presents selected results of the longitudinal study carried out in lower secondary schools.

The volume of this text requires generality of description and selectiveness in presentation of analyses results. An in-depth analysis will be presented in the study report which will be published by the end of 2013 on the project website: www.ewd.edu.pl.

[1] The project is co-financed with the European Union funds under the European Social Fund.
[2] More information you can find at the website: http://ewd.edu.pl/educational-value-added-in-poland/

2. Method

2.1. Concept of VAM validity study

It is possible to distinguish three most important problems around which the study of the VAM validity is concentrated. The first one is quality of the data used for indicators calculation. Of course, quality of examination tests is of greatest importance. Can they be regarded as an accurate and reliable tool for measuring school achievements? If not, neither statistical method will produce reliable results. Owing to an additional study conducted, additional data that allow to verify the accuracy of examination scores as a measure of learning outcomes have been collected. Another problem is related to the choice of an adequate Educational Value-Added statistical model. Such that would take into account all necessary variables and model correctly the relationships between them. The choice of a method for assessing school performance is also of significance. The aim of the study being discussed was also to analyse how adding new contextual variables to the model may improve the value of the indicators being calculated. The last problem in the examination of VAM validity is related to a question whether we can interpret the calculated the EVA indicators as a measurement of school effectiveness. This question may be divided into two, more detailed questions: (1) How much does EVA depend on factors beyond the schools' control that influence school achievements? (2) Is EVA connected with other commonly acknowledged criteria of schools' performance and effectiveness? The first of these questions is related to the postulate of EVA independence of such personal traits of a student as e.g. intelligence, socio-economic status, prior and extracurricular educational experience and social and economic characteristics of the local environment in which school is located. Since school does not have an influence on the characteristics of students and local environment, then in the case of strong links with these variables, EVA could not be regarded as an indicator of teaching effectiveness. It would rather be an indicator of a school's social composition. The second question is related to the concept of criterion-related validity. The study has taken into account many external criteria of EVA validity, including professional development of teachers, teaching methods, leadership and management, and also increment of students' intelligence. The article presents selected analyses results referring to the two latter questions.

2.2. Data

The data used in the analyses have been collected during the longitudinal study in Polish lower secondary schools carried out as part of the project "Research project for the development of educational value-added (EVA) models". The study was conducted on a random, nationwide sample of 150 schools (in these schools 291 classes and 5 249 students were examined). It began in the school year 2010/11 and covered first-grade students in lower secondary schools. It consisted of three stages during which the measurements covered the same students, their parents, teachers and school headmasters. The information collected was supplemented with data from the external examinations system: with the results of the examination after primary school and the results of the examination at the end of lower secondary school. Based on data from the external examinations system, the EVA indicators were calculated for the examined schools. In the presented analyses, data collected during the first and third stage of the study were used.

The level of students' intelligence was tested twice: in the first and in the third grade using the Raven Matrices Test. Data on the socio-economic status were collected by questionnaire survey among parents of first-grade students in lower secondary school. Education of the parents has been expressed in average years of education necessary to obtain a given level of education.

2.3. Data Analysis Scheme

Lower secondary school examination scores have been scaled with the two-parameter IRT model and standardized so that the mean student population is 100, and the standard deviation is 15.

Intelligence test results of both measurements have been scaled together with the two-parameter IRT model, so that they are expressed on a common scale with a mean 0 and a standard deviation of 1 in the population of first-grade students in lower secondary school.

The EVA indicators have been calculated as the Bayesian predictions of random effects of schools in the two-level mixed-effects model in which the dependent variable was the result of the examination at the end of lower secondary school, and the independent variables were the results of the examination after primary school, gender, and certificates of dyslexia.

Average intelligence increments of students in schools have been calculated using an analogical Value-Added model. Effects for schools were calculated as the Bayesian predictions of random effects of schools in the two-level regression model where the dependent variable was the test result in the third grade, and the independent variable was the test result in the first grade.

3. Results

3.1. How much EVA depend on factors beyond the schools' control?

Student's school achievements are influenced by many factors. One of the strongest and most stable predictors of school success is intelligence[3]. Regardless of the importance of genetic factors, also the relevance of a student's family and origin for his or her school achievements is visible[4]. Examination scores of a school will therefore provide an evidence of what kind of students (of what characteristics) attend the institution. The measurement of the education effectiveness (such as EVA) should be an evidence of school's contribution to students' achievements. It should be independent (or dependent to a much lesser extent than the rough results) of matters on which neither school has any influence, e.g. of students' potential. This postulate means that each school should have the same chance for high EVA regardless the characteristics of students attending a given school.

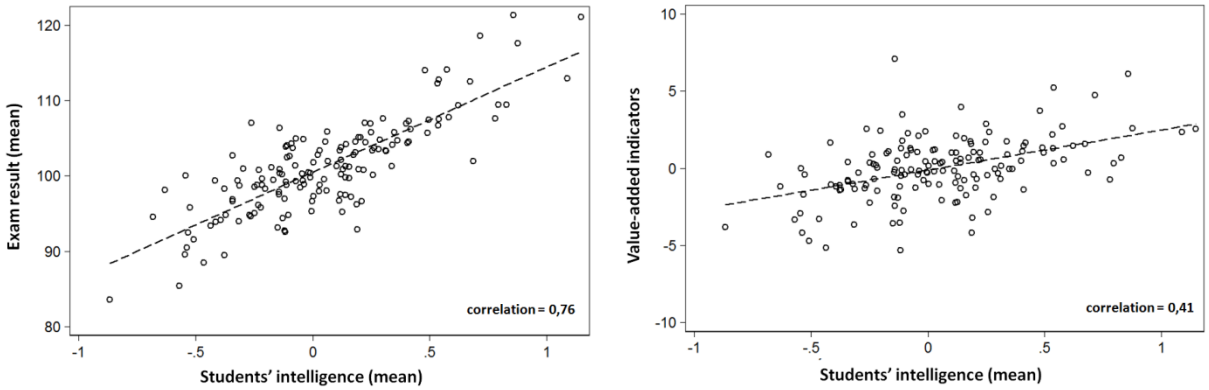
The following pictures compare the strength of the relationship between school examination scores and EVA with selected student characteristics relevant to his or her achievements: intelligence and one of the students' family status indicators: education of parents. The results have been presented for one of the lower secondary school examinations (math & science), because for the second one (humanities) the dependencies are similar.

---

[3] e.g.: Deary, I.J., Strand, S., Smith, P., Fernandes, C. 2007. Intelligence and educational achievement. In *Intelligence*, 2007, No. 35(1), pp. 13–21.; Duckworth, A. L., Quinn, P. D., Tsukayama, E. 2012. What No Child Left Behind leaves behind: The roles of IQ and self-control in predicting standardized achievement test scores and report card grades. In *Journal of Educational Psychology,* 2012, No. 104(2), pp. 439–451; Heaven, P. C. L., Ciarrochi, J. 2012. When IQ is not everything: Intelligence, personality and academic performance at school. In *Personality and Individual Differences*, No. 53(4), pp. 518–522.

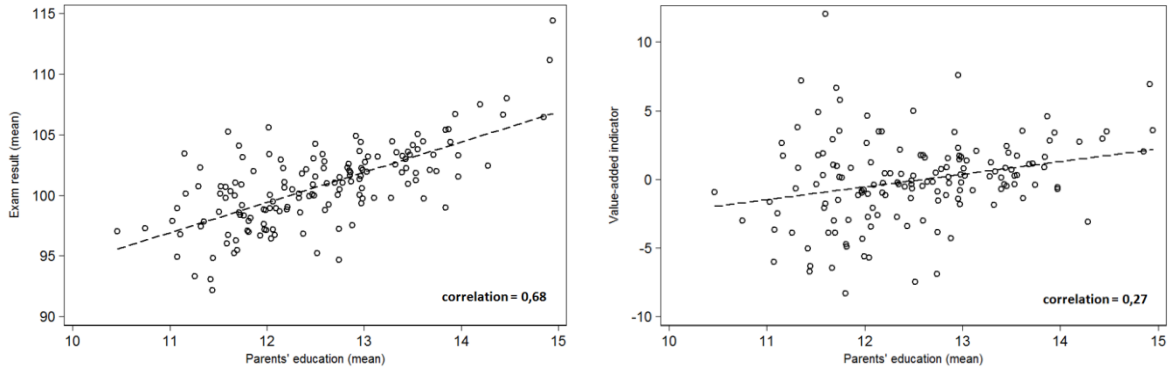
[4] Review of studies on the status determination of school achievements and discussion of the mechanisms that explain the observed dependencies with particular emphasis on Polish education may be found for example in: Dolata, R. 2008. *Szkoła, segregacje, nierówności*. Warszawa: Wydawnictwa Uniwersytetu Warszawskiego. ISBN 978-83-235-0434-4.; Dolata, R., Jakubowski, M., Pokropek, A. 2013. *Polska oświata w międzynarodowych badaniach umiejętności uczniów PISA OECD. Wyniki, trendy, kontekst i porównywalność*. Warszawa: Wydawnictwa Uniwersytetu Warszawskiego. ISBN 978-83-235-1011-6.

The graphs in Picture 1 show the dependencies between the level of students' intelligence (horizontal axis) and the results of school examinations (vertical axis, picture on the left) and the dependencies between the level of intelligence of students and the Educational Value-Added (picture on the right). Each dot represents one school. We can see a very strong correlation between school examination scores and students' average intelligence. The correlation coefficient for these variables is 0.76. This means that school attended by students with low level of intelligence have virtually no chance for very good examination scores.



Picture 1. Relationship between students' intelligence and schools' exam scores (left panel) or value-added indicators (right panel)

However, looking at the EVA dependency of the average students' intelligence level (picture on the right), we can observe that it is much weaker. The correlation coefficient is 0.41. This dependency is positive and statistically significant, but not as strong. Therefore, even those schools that work with students with lower intelligence level may achieve positive EVA.



Picture 2. Relationship between parents' education and schools' exam scores (left panel) or value-added indicators (right panel)

The analysis of the dependencies between parent's education and school examination scores (picture on the left above) or EVA (picture on the right above) leads to analogical conclusions. Examination scores are much more strongly related to parent's education (correlation coefficient is 0.68) compared with EVA (correlation is 0.27).

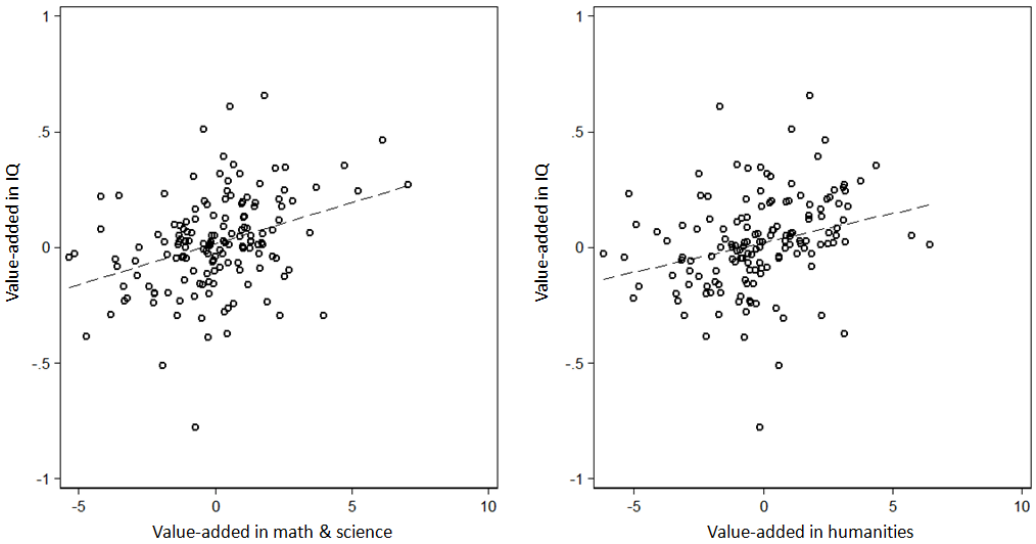
3.2. Is EVA connected with other commonly acknowledged criteria of schools' performance and effectiveness?

EVA method will prove to be valid (according to the criterion-related validity approach), if schools with high EVA are those which, according to our normative knowledge, "do well". One of the criteria of teaching effectiveness may be students' intelligence increment. The study hypothesized that effective schools contribute to faster increment in students'

intelligence. Not only do they prepare students for examinations, but also help them practice the use of cognitive resources[5].

In order to claim that school may influence students' intelligence increment, we must first observe the differences in the size of these increments among schools. In order to verify this hypothesis, intelligence increment for each student was calculated as a difference between the result of the measurement in the third and first grade. Then individual variance of increments was decomposed on the portion of variance that may be attributed to the school division and on the portion of variance within schools. The two-level regression model with the random effects was used for this purpose. The results of this analysis have shown that there are significant differences in the size of intelligence increments between schools. School division explains 13.2% of intelligence increments variance. This means that, to such extent, the chance for a high intelligence increment in a given student depends on a school that he or she attended.

Most important question, however, is whether the school with higher intelligence increments are those with higher values of the EVA indicator. In other words, whether we observe a positive relationship between average rate of students' intelligence increments and EVA? The scaterplots pictures below summarise the relationship between intelligence increments (vertical axis) and EVA for both parts of the examination (horizontal axis). One dot represents one school. We may observe not very strong, but a positive relationship. The correlation coefficient for the math and science part is 0.32 and for the humanities it is 0.24. These correlations are statistically significant at $p<0.05$.



Picture 3. Relationship between intelligence increments (vertical axis) and EVA (horizontal axis)

4. Conclusions

The presented analyses results have shown that EVA for Polish schools is much less dependent on the characteristics of students, such as intelligence and socio-economic status than examination scores. It has also shown that Polish lower secondary schools differ in terms of the size of students' intelligence increments. It has also shown the existence of a positive dependency between the average rate of students' intelligence increments and EVA. These results confirm the validity of the EVA indicators for Polish lower secondary schools.

[5] Compare with the processual approach to intelligence: Nęcka, E. 2003. *Inteligencja: geneza, struktura, funkcje*. Gdańsk: Gdańskie Wydawnictwo Psychologiczne. ISBN 83-89120-08-9.; Nęcka, E. 2009. *Inteligencja jest procesem*. In Kozielecki, J. (Ed.), *Nowe idee w psychologii: psychologia XXI wieku* (pp. 21–39). Sopot: Gdańskie Wydawnictwo Psychologiczne. ISBN 978-83-7489-237-7.