polish 3(187)'14
sociological
review
ISSN 1231 – 1413

ALICJA ZAWISTOWSKA
University of Białystok

The Black Box of the Educational Reforms in Poland: What Caused the Improvement in the PISA Scores of Polish Students?

Abstract: On the basis of the existing research and available theories this paper attempts to determine whether there is any causative relation between the educational reform introduced in Poland in 1999 and the improvement of Polish pupils' scores in the PISA study. Analysis suggests that the element of the reforms most likely to have contributed to the improvement of Poland's PISA scores was the introduction of externally evaluated examinations with a multiple-choice format. There are also two strong methodological conclusions that can be drawn from this research: (1) given the complexity of the reform, its aims and outcomes should be 'decomposed' for more detailed analysis, and (2) given the cumulative outcomes of the reform, PISA scores between 2000 and 2012 should be explained by a different set of hypotheses for each round.

Keywords: PISA, reform, education, Poland, central examinations.

Introduction

In the first report of the Programme for International Student Assessment (PISA), published in 2000, fifteen-year-old Polish students were ranked significantly below the OECD average. In the next three rounds of the assessment, the results were gradually improved, and in 2012 Polish students were ranked just below those EU countries whose educational systems have been recognized as the most effective (for instance, Finland or Estonia). During this period, Poland's educational system was being fundamentally transformed. The reform that began in 1999 was the first such deep intervention in the educational system from the moment of its reconstruction at the beginning of the 1990s. The coincidence of these two phenomena—the educational reform with the improvement in PISA scores—is an enticement to conclude that there is a causal relationship between the two. Unlike in the case of experimental-design studies, which allow a researcher to control the impact of external factors more strictly, establishing a casual relationship between the 'objects' of social research is often more challenging. This is especially the case when the nature of these objects is more complex, usually traversing the micro and macro levels, and is irreducible to a single variable. Educational reforms belong to this category of issue. Checking whether the textbook criteria of causality have been fulfilled is additionally hampered by the fact that while social interventions have the potential to change the environment, the

environment where the intervention is applied changes itself. As one researcher into the educational reforms has pointed out—government interventions change schools, but at the same time schools shape the ultimate goals of the reform (Tyack, Cuban 1997: 60). What is meant is that both spontaneous and expected changes could affect the final outcome of a reform (Levin 2001). To borrow a metaphor from behavioural psychology, it could be said that implementation of the reform is like a 'black box'—it contains a full record of the process underway, but this information is inaccessible to the outside observer. The researcher can at most recreate the relationships between cause and effect based on an analysis of changes in the environment being transformed. The present article sheds more light on this issue. I am interested in the connections between certain elements of the reform of the educational system begun in 1999 in Poland and the improvement of students' scores in the 2000–2012 PISA tests. Referring to existing research, and where possible, to theory, I concentrate on the potential effect of four main changes: the introduction of lower secondary schools, reform of the system of teachers' promotions and remuneration, introduction of central examinations and core curriculum, and decentralization. More generally, I am interested in those aspects of the reform that have had an impact on daily school practice, as well as those that constitute the broader context of the school system's functioning. I will pass over, however, the issue of measuring students' ability and detailed analysis of the PISA results, which are the object of separate, in-depth analyses undertaken by other researchers (PISA 2012, Dolata, Jakubowski, Pokropek 2013).

The Reform of 1999

Introduced in September 1999, the reform of the educational system was intended to increase the number of students in secondary school and higher education, equalize education opportunities, and improve the quality of teaching. The idea was to achieve these aims through the introduction of parallel changes in almost every dimension of the education system: beginning with the transformation of education's structures and finance policies, through the introduction of new core curriculum and methods of appraising students' abilities, to reforms in the system of promoting teachers. The main organizational change was the introduction of the three-year lower secondary school (*gymnasium*) preceded by a six-year elementary school and followed by a three- or four-year upper secondary school (after the reform, the educational structure is 6 + 3 + 3). The policymakers intended these two latter stages of schooling to be complementary in terms of curriculum.[1] All these changes created a need to prepare a new core curriculum more suited to the new structures and not overlapping at the various levels. In accordance with the new curriculum, students are required to obtain greater independence in thinking, improved ability to associate information from various fields, and strengthened analytical thinking. The general idea of the policymakers was to depart from teaching of an encyclopaedic nature, focused on the memorization of

[1] The reform reflects the school system existing in Poland before the Second World War.

facts and their repetition, toward greater reflexivity. Contrary to previously, the core curriculum of 2009 set forth educational aims and the competencies students were to possess at the end of a given level in the language of specific requirements. Changes in the core curriculum were strongly related to the reform of exit exams—since the reform students have taken a central, external examination at the end of each stage of education: at the end of elementary school, lower secondary school, and upper secondary school.

The reform of 1999 opened a new chapter in the history of Polish education, although it was itself a result of the political and economic transformations that began in the early 90s. The rapid transition to democracy 'freed' education from the limitations imposed on it by the communist system, where the ambition had been full control and top-down guidance of the education process. The changes undertaken at the beginning of the 90s were consistently intended to improve the quality of education by introducing new curricula (followed by liberalization of the textbook market and greater freedom for teachers in shaping the educational process), the most rapid possible dismantling of the structures of the communist regime, and the gradual transfer of schools from the central to local powers (Herczyński 2012, Śliwierski 2009). However, the time for a deeper transformation of the educational system came ten years after the democratic breakthrough.

At first the reform of 1999 provoked considerable controversy among teaching staff, parents, and the general public. The idea of lower secondary school was the subject of particular criticism; its introduction was supported by 51% of Poles in 1999 (CBOS 1999). In later years, opinion on the *gymnasium* somewhat stabilized, although the idea of returning to the 'old' system with only two levels of schooling reappears almost every time the reform's originators debate the problem of the quality of education with their political opponents. It is significant that the remaining elements of the reform, in spite of their extensiveness, rarely appear as subjects of the public discourse.

Main Trends in PISA's Test Results

The PISA study has been conducted every three years since 2000 with the aim of assessing the competencies of fifteen-year-olds in reading, mathematics, and science. Each round of the study concentrates on the detailed diagnosis of one of these areas, while treating the remainder more superficially. Therefore, reading ability is best compared by the years 2000 and 2009, while mathematical skills by the tests conducted in 2003 and 2012. It is worth remembering that although in every cycle the identical method of scaling results is retained, only a part of the questions have the same content (PISA 2012). The introduction of new questions in each round resulted from changes in the conceptualization of abilities: for example in addition to the reading of traditional texts in 2000, the 2009 test also included analysis of electronic texts and diagnosis of the reader's engagement and self-awareness (PISA 2009).[2]

[2] Details of the methodological issues can be found on the PISA OECD web page.

Graph 1 presents the difference in the average results for mathematics and reading in Poland and the OECD countries (represented by grey bars) in the context of the main changes in Polish education. A growth in scores can be observed from the first cycle of the study, conducted in 2000, although the pattern of the increase is not the same in both domains. Scores in mathematics rose to a small degree in the first three rounds and only in the last cycle did the Polish average (518 points) exceed the OECD average (494 points). Improvement in reading scores was, however, more systematic and occurred at once after the first PISA cycle. In Poland in 2000, the average was 479 points, while ten years later it was 500, meaning that this result was significantly higher than the OECD average. The growth in the average also affected the distribution of results, which inclined toward higher scores, mainly by diminishing the category of students with lower results (Table 1).

Table 1

Changes in the Percentage of Poorest and Best Students in Reading and Mathematical Literacy in the PISA Study

	Level 2 and below*	Level 5 or higher**
Mathematics 2003	22	10
Mathematics 2012	14	17
Difference	−6	+6
Reading 2000	23	6
Reading 2009	11	10
Difference	−12	+4

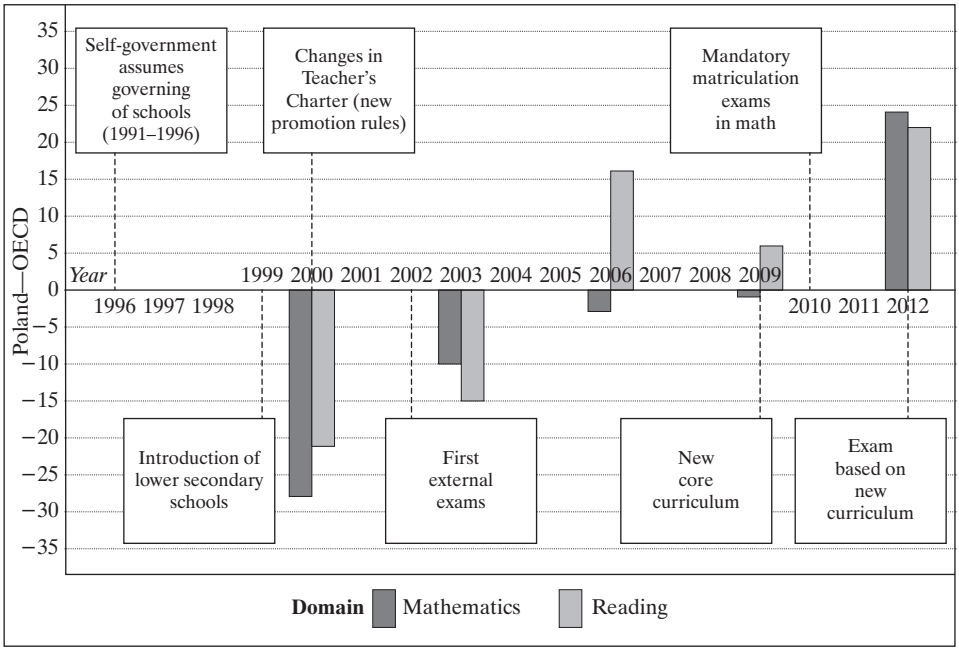
Source: *PISA 2012 Results: What Students Know and Can Do* (Volume I), OECD 2013.

* Level 2 and below in the PISA classification (less than 420 points).
** Level 5 or higher in the PISA classification (606 points or more).

'Correlating' the sequence of changes in the school system with the improvement of PISA results in Graph 1 does not resolve the problem of causality, but allows certain conjectures to be made. Above all, it is doubtful whether extensive reform—like the one discussed here—could have brought lightning results. Change in any system—and particularly a system whose existence is based on the interplay of actors with sometimes contradictory interests—requires time before producing a 'new equilibrium'. Graph 1 makes clear, however, that the full effects of the reform could only be observed after 2009, when the earlier structural and organizational changes were supplemented by the new curriculum. Accordingly, the results of 2012 were the first to reflect the full impact of all the changes. This is also the year in which, for the first time, students took the upper secondary exit exam adapted to the new curriculum (the one close to the PISA tests in terms of aims). It is not without significance, particularly for the improvement of PISA results in mathematics, that in 2010 mathematics was once again (after 25 years) an obligatory subject of the secondary-school exit exam. From the perspective of a *gymnasium* student, the mathematics of the upper-secondary-school exit examination may seem a problem of the fairly distant future, but

Graph 1

Improvement in PISA Scores (the difference between Poland's average and the average for OECD countries) in the Context of the Main Changes to Poland's Educational System



already at this stage it could have motivational importance. Stable results in future PISA tests might be taken as a partial confirmation of this hypothesis.

It is worth noting in the margins that not all categories of students improved their scores in the PISA study to the same degree. After taking into account the socio-economic statuses of students and how they changed over the course of the five PISA rounds, it emerges that the beneficiaries of this success were, in particular, the poorer performing students and girls, while boys and better performing students were encompassed to a lesser degree (see Dolata, Jakubowski, and Pokropek 2013). Furthermore, as a result of the accumulation of various processes occurring in the educational system in recent years—linked directly with the reform or with broader social changes—variation in the results of exit examinations at lower secondary schools has been growing and this process of differentiation is becoming especially significant in large cities (Dolata 2011). It might be recognized as a symptom of the increasing role of horizontal inequalities in education (which means there is a stronger selection process between schools at a given level than between levels).

The Effects of Individual Components of the Reform

In analysing the effects of the reform it must be assumed that not all the reform's components had an equal impact on the final outcome. In this respect, the four main

elements of the reform will be discussed separately below: (1) the change in organizational structure, (2) the introduction of external examinations and new curriculum, (3) the reform of the system of teacher promotions and remuneration, and (4) the decentralization of the educational system. Their potential impact on the scores of students in PISA tests will be analyzed in reference to other studies.

The New Structure of Education

When the first PISA study was conducted in 2000, fifteen-year-old Polish students had already been divided among three types of secondary schools: general, technical, or vocational. These schools, as is typical of educational systems with early selection thresholds, varied considerably in terms of their students' abilities, social background, and the content of their core curriculum. The gap in skills was especially significant between students of vocational and general schools. The former were often chosen by students from low social backgrounds and with the lowest performance in elementary school, while the latter attracted students who had better scores and intended to continue their education at the university level. As the PISA 2000 showed, it was students from the vocational schools who constituted the decided majority of students with low levels of proficiency in reading, and whose skills, in accordance with the OECD definition, were limited to the simplest analysis of a text and observation of obvious relationships. The profound difference between the types of schools was also reflected in the fact that in 2000 Poland was characterized by a particularly strong variation between schools in comparison with other countries (Białecki and Haman 2000).

The question thus arises of whether changes in the educational structure, namely, extending comprehensive education by a year, and setting back the threshold for selection by introducing lower secondary schools, had an impact on improving the results in succeeding cycles of PISA. Some studies have confirmed the positive effect of this change (Jakubowski, Patrinos, and Porta 2010). Using the counterfactual method, which allows the achievements of students of like social and demographic characteristics to be compared at various points of time, researchers concluded that students who were in the 'old' vocational secondary school in the year 2000 would have had improved results had they begun their schooling in the 'new' system, with lower secondary school. The largest benefit would have been felt by the poorest performing students, whose results might have improved by one standard deviation in the course of six years, and to a lesser degree by students in the two remaining types of schools (Jakubowski, Patrinos, and Porta 2010).

Therefore, the hypothesis can be formulated that the source of the improvement was the extra year of general-knowledge education for all students. This extra year could have had a particularly beneficial impact on the results of less well-performing students—who instead of vocational education, were given the same curriculum as was reserved in the old system for the students of technical and general secondary schools (which does not mean they took full advantage of the opportunity). It also means that since PISA 2003, the level of abilities of that group of students reflected the effect of

teaching a general and not vocational curriculum (as was the case in PISA 2000). In other words, improvement in the results might be partly ascribed to the fact that PISA 2000 captured the performance of students who had already been sorted to one of the three paths in upper secondary schools. As a consequence, the results of the first PISA study reflected the effects of the mixed core curriculums, varied social backgrounds, and different ability levels of students from vocational, technical, and general upper secondary schools. In each of the next rounds of PISA (from 2003 to 2012), due to the delay in tracking, the entire target population was concentrated in lower secondary schools and was taught exactly the same core curriculum. Consequently students' performances had a more uniform distribution.

Another manner of grasping the 'pure' effect of introducing the *gymnasium* on improvement in the PISA survey is to compare the achievements at upper secondary school of two groups of students: those who graduated from lower secondary schools, after the reform, and the older cohort who went straight to the 'old' secondary school after finishing elementary education. The research on the subject indicates that students who studied in the new system of upper secondary schools—regardless of the type of school—have significantly higher scores than earlier generations at the same level of education (Dolata, Pokropek, Jakubowski 2013). This might be taken as proof that an additional year of comprehensive education and the delay of tracking have a positive impact on school effectiveness. However, the same study has shown that the secondary-school tracking system still strongly selects students according to their social origins and abilities (Dolata, Pokropek, Jakubowski 2013). Putting these two facts together it might be concluded that students in the new, three-step educational system are smarter than the students who studied before the introduction of the reform, but the gap between tracks has remained untouched (Dolata 2012). This suggests that the introduction of lower secondary schools would probably not in itself have had the beneficial effect observed in the PISA survey if it had not been accompanied by other changes introduced by the reform.

Another hypothetical explanation for the positive effects of the reform's structural changes on PISA scores involves the new procedure for enrolment in lower secondary schools. According to the principle introduced by the reform, lower secondary schools have been obligated to accept children living in their catchment area in first order (Białecki 2006). In Białecki's opinion, this principle has led to the 'mixing' of pupils in terms of competences and social backgrounds. His analysis of PISA data shows that in 2000, poorer students were concentrated chiefly in poorer schools (belonging to the bottom quarter of the distribution of scores), while in 2003, students with poor results were enrolled in both poorer and in average schools. In explaining that shift, he argues that the greater heterogeneity of students favoured the elimination of extremely low and high results to produce a less varied distribution of scores (Białecki 2006). This hypothesis argues for the positive impact of the 'peer effect', that is, the mechanisms of the influence of schoolmates—their aspirations, motivations, and abilities—on the results of the individual student. Other studies are not convincing, however, that the peer effect in lower secondary schools (understood as the effect of the school average on individual results) was significant (Pokropek 2013). This does not mean,

however, that the transmission of a system of values—invisible to statistical testing measurements—does not occur between able and less able students.

Changes in the Educational Process: External Examinations and New Core Curriculum

Not long after the reorganization of educational structures, the manner of evaluating students' knowledge was also changed. Since the school year 2001/2002, schooling at the lower secondary and elementary school levels finishes with an obligatory examination, which is a standardized nationwide test, conducted uniformly on the basis of provisions set forth in the applicable legislation. From that year on, the exit examination has been administered by teachers from outside the school the student attended, and the results are expressed on the same scale of points for all students in Poland. This has been a large change in comparison to the period before the reforms, when students after finishing their elementary education took entrance examinations as part of the admissions process to secondary school. Such exams consisted of answering essay questions and were organized by the school itself and marked by the teachers working there. In introducing the new principles, the reformers were guided by the necessity of providing more objective and comparable criteria of appraisal. This change, although significantly less often taken into account in the general evaluation of the reforms than the transformation of the educational structure, could have had an even larger influence in improving students' achievements in the PISA survey.

The theoretical bases for the idea that the reform of examinations might explain Poland's success in the PISA survey is provided by the signalling theory formulated by Michael Spence for economics (Spence 1973; cf. Bishop 1997). According to Spence, signalling theory explains the principles for employment contracts in the labour market, which are entered in conditions of an asymmetry of information. In concluding the contract, the employer does not have complete information about—to use the economic terminology—the employee's productiveness, even if a certain time has elapsed since his or her hiring. Ideas about the candidate's qualifications are based on signals, such as educational level, which are indicators of unobserved, 'hidden' qualities such as intelligence. On the 'education market' the role of such a signal could be played by the central exit examination result. It provides important information to various 'recipients' of educational services: it informs parents about their child's scores in comparison to other children, it helps school candidates orient themselves as to the level of the requirements, and informs supervisory organs about the effectiveness of schools. In other words, central, uniform, and external examinations break the asymmetry of information between actors in the educational market. Signalling theory has already been used to explain the relation between students' achievements and the existence of central examinations (Bishop 1997, Woessmann 2002, Białecki 2006).

Such conclusions emerged in two studies, PISA and TIMSS (Trends in International Mathematics and Science Study), comparing countries having a homogenous system of central examinations with other countries (Woessmann 2005). Although researchers in the study argue that the beneficial impact of central examinations motivates both students and teachers, who want to emerge better in comparison with

others (Woessmann 2005), another interpretation is obviously possible: the students in countries where examinations have a multiple-choice format manage better with the PISA or TIMSS examinations on account of having mastered the 'technique' of test taking and exam preparation. It appears, however, that students who take central examinations manage better after leaving school. German research on the relation between the form of exit exams and the occupational trajectories of graduates (Piopiunik, Schwerdt, Woessmann 2013) shows that students from states where there is a central system of examinations not only do better in school, but also obtain higher positions in the labour market (in terms of income size and ease of finding work) (Piopiunik, Schwerdt, Woessmann 2013). Another hypothesis suggests that in addition to setting in motion psychological mechanisms, the positive connection between central examinations and results is also to be explained by the greater amount of guidance in terms of the curriculum, which must correspond closely to the core curriculum (Bishop 1997).

Applying the results of the above-mentioned studies and referring to signalling theory could explain the systematic improvement of Polish students' scores on PISA tests. The new form of exit exam provides synthetic information, which can be compared and contrasted and which serves as a legible and credible signal of engagement in the educational process. Further, it might be assumed that if central examinations are indeed able to improve the effectiveness of schooling, then one of the signs would be greater differentiation of results between schools. This should be expected due to the fact that comparable and easily available information about schools' exam scores should increase the competition between schools: better schools, having greater incentives and resources, will improve their scores, and worse ones—on account of their limited possibilities—will maintain or have even lower scores as a result of the departure of the more talented students to more effective schools. Such processes have already been observed in Poland for several years—the percent of interschool variation (as a share in the total variation) has grown from 22 to 40 percent (Dolata et al 2012). The strongest differentiation has been observed in large cities, where the educational market is more competitive and thus the value of the signal is greater than in less urbanized areas.

Another sign of the signalling function of examination results should also be the adaptation of students and teachers to the new conditions. Research on this subject is only just getting under way, but certain tendencies are already observable. For example, it has been noted among history teachers, who have admitted that the form of the final examination has a strong impact on their choice of teaching materials and the educational methods they employ (Choińska-Mika 2013). The teachers admit they select content according to the probability of its appearance on the lower secondary school examination (Choińska-Mika 2013) (it is worth remembering here that questions from previous years and informative material with sample exams are publically available). The form and manner of the post-reform examination largely subordinates the didactic process, partly by the development of a more effective learning technique. It is worth drawing attention here to the fact that the multiple-choice method of examination could also explain why, in the course of recent years, it is the poorer students who have had the largest success in PISA studies—it is probable that

filling out a multiple-choice test causes them less trouble than the necessity of formulating long, spontaneous written answers or oral responses. It could be otherwise with students who are better equipped with cultural capital, who might prefer a freer, less classified manner of having their knowledge tested. Adapting to the 'test culture' does not change, however, the basic manner of organizing classes. From international studies on middle-school teachers conducted in 2009 it emerges that in Poland, as in other countries, a structured type of teaching predominates. The studies show that, regardless of the methods of appraisal, classes are commonly based on maintaining a certain routine, with minimal use of active-learning methods (Piwowarski, Krawczyk 2009). Research conducted in the United States, a country that has the largest experience in reforming its educational system, also shows that teaching practice over the course of recent decades has remained untouched, even if the circumstances of teaching have undergone fairly significant changes (Tyack, Cuban 1995).

Finally, in the context of the influence of exams on PISA scores, attention should be paid to their correlation with the content of PISA questions. Both the OECD studies and the new core curriculum introduced in 2009 adopt similar definitions of students' knowledge and competencies. A similar emphasis is also placed on understanding, flexibility, and the ability to make practical use of knowledge in various contexts. For example, the concept of mathematical literacy is defined in PISA documents of 2012 as, 'An individual's capacity to formulate, employ, and interpret mathematics in a variety of contexts. It includes reasoning mathematically and using mathematical concepts, procedures, facts and tools to describe, explain and predict phenomena' (PISA 2012: 17). In Poland's core curriculum, a similar understanding of mathematical thinking can be found: it is the 'ability to use mathematical tools in everyday life and to formulate judgments based on mathematical understanding' (Marciniak 2009:19). Accordingly, progress in the PISA results, and particularly in the results of 2012, reflects the convergence between the curriculum in Polish schools and the content of the questions the OECD uses for testing. In other words, it is likely that in 2012 students were being asked exactly what they had been taught in the course of the previous three years.

Reform of the System of Teachers' Remuneration and Promotions

Among the whole range of factors that influence the effectiveness of schools and have simultaneously undergone transformation in recent years is the role played by its key actors—the teachers. The year 2000, when new rules regarding promotion were introduced in order to make salary size closely dependent on competence, was an important point in the changes. From that year on, a teacher's career was divided into four levels: trainee teacher, contract teacher, appointed teacher, and finally, certified teacher. The adoption of this system has meant that since the year 2000 the professional structure for teachers in Poland has been dominated by teachers at the higher degrees of promotion, and in 2010 49% of all teachers were certified teachers (Herbst 2012, Kłos 2010). The new system of promotion went together with a rise in salary levels. In 2007, teachers' salaries exceeded the average salary of state

employees and since then the difference has increased (Kłos 2010). Other studies also show that in the years 2007–2010 the nominal value of a full-time teacher's salary increased by 33% on average, but the large differences between categories of teachers were maintained: during that time trainee teachers earned nominally more by 64%, and certified teachers about 26% more (Herbst 2012). Other research into teachers' salaries has shown that in comparison with other OECD countries, teachers in Poland began their careers by receiving the lowest salaries (in relation to national GDP), but after fifteen years of work, they earned relatively more than the average in those countries (Herbst, Levitas 2009). One of the causes of this fact could be the nature of the system of promotion in Poland, where there is a large difference in the salaries of teachers at the highest and lowest levels of advancement. For example in 2007, when the differences between salaries were at their height, appointed teachers earned about 250% more than trainee teachers. This difference was reduced to 190% after changes to the provisions of the Teachers' Charter in 2010[3] (Herbst 2012).

It is difficult to appraise the significance of the reform of teachers' salaries and promotion rules as a factor influencing the effectiveness of education in general and PISA results in particular. There is some evidence that financial incentives could have a positive impact on students' results, but the mechanism of this connection has not yet been precisely determined. It is usually assumed that in the short term, financial incentives may form a motivation for more productive work, and in the long term could increase the attractiveness of teaching as a career. Starting from the reasonable assumption that larger payment could be a motivation to larger engagement, some countries have introduced experimental programmes to make remuneration strongly dependent on students' examination results. One such study was conducted in Israel (Lavy 2009), where in 2000 nearly 49 schools were included in a programme to pay financial bonuses to teachers whose students had the best results in the school-leaving examinations (the amount of the bonus depended on the size of the difference between the actual result of the students and the predicted results based on a regression model, controlling for social factors affecting the schools and students). Analyses show that during the course of the programme the students' achievements were significantly higher, and the teachers were more engaged in their didactic work, staying with students after lessons and working to meet their needs (Lavy 2009). Lavy's research provided proof that a system of extra bonuses could bring a positive effect, at least in the short term. Conclusions about a long-term effect are less clear. Comparison (Woessmann 2011) of countries where teaching salaries reflect student achievement and those having a more egalitarian approach shows that students in the former had a quarter of a standard deviation higher achievement in mathematics and reading (controlling for relevant school and country characteristics) (Woessmann 2011). However, American researchers who have reanalyzed the effectiveness of these types of programmes in various countries and states point out that even though extra bonuses do have a positive effect on students' results, the improvement is usually modest. Moreover, when formulating educational policy, students' progress is not the

[3] The Teachers' Charter is an act setting forth the principles of employment, promotion, and remuneration of teachers.

only parameter taken into account. In the decided majority of cases, the costs and benefits of such solutions usually show it to be unprofitable in the long term (Haut, Elliott 2011).

The examples of bonus programmes given above do not entirely correspond to the situation of Polish teachers, who were encompassed as a group by the reform of the remuneration system. This is an important difference in terms of motivational effectiveness. The system for promoting teachers introduced by the reform probably did not have a larger impact on the effectiveness of schools either. The reformers assumed that a stronger correlation between promotion and salary size would motivate teachers to improve their skills through training courses and conferences. In practice, it emerged that the process of promotion depends primarily on intensive collecting of certificates, diplomas, and other documents confirming participation in various types of pedagogical enterprises, and rarely corresponds with real improvement in a teacher's work (NIK 2008). The process of climbing the career ladder is also recognized as fairly rapid—on average, teachers achieve the highest level in the teaching hierarchy at around 40 years of age, that is, around 13–16 years before retiring (Kłos 2010). Nor is there much evidence that the university faculties producing future teachers have rationalized their admissions systems. The needs of the educational system are still much smaller than the number of graduates with pedagogical degrees, who are additionally the most numerous group of graduates.

Decentralization of the Education System

Another factor that has a positive effect on the results of students and has simultaneously undergone reform in Poland is the decentralization of the education system (Fuchs, Woessmann 2004). The term 'decentralization' itself is fairly capacious and could refer to various aspects of the education system or indicate different degrees of dispersal. The idea behind decentralization was that the power and responsibilities of managing education at the local level should be transferred from the central to local authorities. The strongest argument behind the idea is that local authorities are better acquainted with the needs of the local community and can transfer funds directly where necessary. In Poland, this process of transferring school-management responsibilities accompanied the democratic transformations begun in the early 90s. The first decisions in the matter of increasing the influence of local authorities were taken in 1991, when the new act on education introduced a range of liberalizing changes to the education system. One important change was that non-state bodies were allowed, as they had not been before the democratic transition, to establish private schools. In terms of organizational process, decentralization in Poland can be divided into two basic stages (Herbst, Levitas 2012). Until the middle of the 1990s, schools and preschools were being transferred into the control of local governments. In the first years of the 1990s, the transfer occurred voluntarily, but after 1996 elementary schools were obligatorily under the curatorship of local governments, and after 1999 newly established local administrative units (*powiaty*) took over post-elementary and post-middle schools. The second stage, which began in 2000, occurred in

the context of far-reaching changes produced by the reform of 1999 and subsequent alterations. Local governments had been required to take on increasingly costly obligations in regards to the reorganization of the educational system (establishing lower secondary schools, for instance) and the increase in teacher salaries resulting from the reform of 2000. As the central-government subsidies had not increased as rapidly as expenditures, many local governments struggled with debt. Not infrequently, imbalances in local government budgets were caused by short-sighted decisions at the central level. One example is the requirement, imposed by the reform of 1999, that elementary schools and the newly established lower secondary schools be in separate school buildings. This decision did not take into account either the predictions that the student population would decrease in the coming years or the costs of maintaining the new schools and was withdrawn a few years later due to the insufficient financial means of many districts. Another example occurred in 2000, when the salaries of a large group of teachers increased beyond the subsidies provided for the purpose by the central government.

Succeeding changes to educational policies involving local governments included the introduction in 2000 of more market-oriented budget subsidies, dependent on the number of students (Herczyński 2012). This resulted in the necessity of rationalizing the school network, particularly by closing small schools in rural areas, whose operation was less profitable. The decisions to close schools were usually strongly denounced by the local communities, but in most cases were an inevitable by-product of increasing stress on local-government budgets.

Influence on personnel policy and the possibility of partly shaping salary sizes (from 2000 on) are the main local government instruments that could affect the quality of teaching in local schools. Theoretically, local governments' rights in this area permit them to have a real impact on educational policy, and indirectly also on the quality of teaching. Both of these prerogatives have been, however, closely restricted and limited by national legislation, especially the Teachers' Charter—an act containing general provisions concerning teachers' pay and defining other duties and privileges of teachers (Herbst 2010). For example, the minimum level of a teacher's basic pay is set annually by a decree of the minister of education. The influence of local governments on the size of teachers' salaries is limited in practice to paying supplements (for instance, for seniority bonuses or incentives) to basic salaries: 25% of teachers' entire pay in 2010. Another component remaining in the hands of local government is overtime pay (in 2010 it constituted 9% of entire pay) (Herbst 2010: 10). The amount of these salary components is dependent on the financial condition of the organ running the school, and thus certain territorial differences appear. In general, salaries are higher in more affluent large cities and in the west of Poland.

Another instrument potentially impacting the effectiveness of education is influence on hiring policy. The competences of local authorities include management of a network of schools and (since 1995) choice of school principal. The act of 1991 provided for school principals to be appointed on the basis of a competition and set limits on their term of office (Barański 2010). Its later amendments made the competition the responsibility of the school's managing organ, which was to form a competition

committee composed of representatives of the local government, the pedagogical supervisory body, and the local community. Changes introduced in succeeding years mainly concerned the relation between the managing organs and the pedagogical supervisory organs. These latter were given or deprived of the ability to veto a given candidacy and to increase the number of members of the competition committee. It was not until the regulation adopted in 2009 that the school managing bodies were given greater autonomy by eliminating the necessity of consulting the composition of the competition committee and by increasing the number of local government representatives (Barański 2010).

The question is whether the above changes in school management policy could have had a positive impact on the performances of individual students, or, to put it more generally, could have improved schools' effectiveness (including improvement in the PISA study). In the opinion of experts, the answer is unclear. It has been pointed out that changes in financial and organizational policy involving the transfer of responsibilities from central to local government could have had a positive impact on individual students' performances, but mainly as a matter of complex indirect relationships (Herbst 2012). Aside from the reforms, recent decades in Poland have seen important demographic changes, as well as general improvement in living standards. Table 2 shows the amount of progress made in the last decade in the sphere of infrastructure (the number of computers) and the spread of educational offerings (the universality of English-language instruction). The decreasing number of students, with a relatively stable number of teachers, must have had at least some influence on schools' effectiveness.

Table 2

Basic Information About Lower Secondary Schools

	2001	2003	2006	2009	2012
Ratio of students to teachers	13.7	13.3*	13.0	11.7	11.0
Ratio of students to schools	300	269	238	201	174
Town	398	—	—	—	231
Rural area	197	—	—	—	120
Ratio of students to departments	24.9	24.8	24.0	22.8	22.4
Number of students per computer	—	25.3	—	—	9.9
Percentage of students learning the obligatory English language	71.0	—	—	—	94.0

Source: Main Statistical Office, Oświata i wychowanie w roku 2011/2012 [*Education and Upbringing in the Year 2011/2*], *Małe roczniki statystyczne* [*Small Statistical Annual*] of the years 2001, 2003, 2006, 2009.

* Due to the lack of information from the year 2003, data for the year 2004 is given.

Other Examples of Success in the PISA Results

An alternative method of evaluating the impact of reform is to make an international comparison: in particular, a comparison with countries that also noted an improve-

ment in the period when their educational system was undergoing transformation. Besides Poland, in Europe, Latvia and Portugal can also pride themselves on improved results. The changes introduced in Latvia resemble to a large degree those implemented in Poland. They intensified toward the end of the 1990s, when it was decided to improve the quality of teaching, i.e., to increase the effectiveness of existing educational resources, increase access to education, and further the equalization of educational opportunities. As in Poland, the curriculum placed greater emphasis on problem-solving and on practical and interdisciplinary knowledge. The standards for appraising abilities and teaching goals were made uniform, and from 1997 examinations were central. These changes appeared without transformations of the school structure, that is, while preserving a nine-year elementary school (to age 16) (Kangro, James 2008).

The reforms that occurred in Poland and Latvia can be considered typical for post-transformation countries. Silova and Steiner-Khamsai (2008) indicate the considerable similarity of the reforms in all the countries of the former communist bloc. Typical changes include liberalization of curricula, increased length of education, modernization of standards, and greater freedom of teachers in directing the teaching process (Silova, Steiner-Khamasi 2008). This does not mean, however, that adopting such reforms is a recipe for success, which is rather a matter of the able application of methods to existing possibilities and problems. Among other countries, Portugal, which improved its results in all three areas of the PISA tests, is an example of this. 2005, when the reorganization of its secondary schools was initiated (in Portugal 15-year-olds attend secondary school), was an important point in the history of its educational policy. Many ineffectively operating schools, particularly those with too small a number of students, were closed at the time. The tempo of rationalizing the Portuguese school network was, however, significantly more rapid than in Poland: in the course of the decade between 1999 and 2009 over 5,700 schools were closed (Santiago et al. 2012: 35). Other activities were intended to increase decentralization. Special bodies, organizational units with their own administrative structure, were called into being to manage several schools—mainly elementary schools, and to a lesser degree, secondary schools. In succeeding years, the relationship between individual educational institutions was regulated, transferring larger responsibilities to the school directors. Care was also taken to introduce improved standards, which, as in other countries, consisted in precisely setting forth the results students should achieve at the end of the second stage of education (Santiago 2012). It is worth noting, however, that unlike in Poland, the scores of Portuguese teenagers had begun to improve before the introduction of the majority of these changes.

Conclusion

The aim of this analysis was to address the question of the causal relationship between Poland's educational reforms and the improvement in students' scores in the PISA tests. There are four main hypotheses explaining this relationship, i.e., it was caused

by: (1) change in the organizational structure of education and the delay of tracking; (2) introduction of external examinations and new core curriculum, which corresponds with the PISA test; (3) increase in teacher salaries and reform of promotions; and (4) decentralization of education and its processes. All these changes may have had some impact on the improvement of students' results in the PISA study, but some must be more important than others. The most convincing arguments are provided by the introduction of external examinations—whose positive impact on performances is also justified by theory. The psychological mechanism connected with the consciousness of making standardized test results public and subject to comparison probably changed the education system more than making a new stage of schooling or reforming the system for promoting teachers. Examinations have always been the link between the curriculum and students' and teachers' work, but in its present form it also shapes their ability to 'take tests', which becomes essential at various educational stages. It is also useful in taking the PISA test. However, an explanation of the links between the results of Polish students in the PISA surveys and educational reform in Poland must encompass both the cumulative changes already introduced by the reform and take the specific PISA round into account. Improvement in PISA scores in the years 2000–2003 was probably the result of delayed tracking and the 'mixing' of students in terms of social composition. The results of 2012 were more probably the result of the similarities between the core curriculum and the questions on the PISA test, as well as of the introduction of an obligatory school-leaving exam in mathematics.

These hypotheses may be combined with two methodological postulates concerning research on effects of the reform. Taking into account that reforms are usually broad plans for multidirectional changes, in studying their effectiveness it is necessary to 'decompose' their aims and look at each element separately. Application of this approach has shown that of all the changes introduced into the Polish educational system, only some could have had a direct influence on improving results. The remainder, such as the reform of teachers' promotions, either had none at all, or their effect was indirect, through processes not related with the reforms. Secondly, analysis has shown that explaining the results in the PISA tests between 2000 and 2012 by this same set of hypotheses might lead to less accurate conclusions. Individual changes during the given period could have been the effect of different, loosely related processes. This leads to the more general conclusion that the records in the 'black box' of the reform should be studied individually and in detail.

The example of the reforms in Poland, Latvia, and Portugal and simultaneous observation of the PISA results show that the effectiveness of an educational system is decided by many more factors than the architects of the reforms could wish. Often the most important stimuli may come from outside the educational institutions. This is well known by teenagers from the rich parts of China and their parents, who pay a high personal price to achieve the highest positions in PISA rankings. There is great pressure for achievement and a stronger subordination of life to the school regime. Doubtless the American government is also aware of this fact, as in spite of its relatively large expenditures on education, it is losing, for the moment, the 'Race to the Top'. Clearly, the Finns are also aware of it, as they have placed support for

poorer students and the development of talented ones higher as educational priorities than competitiveness and early selection.

References

Białecki, Ireneusz and Haman, Jacek. 2000. *Program Międzynarodowej Oceny Umiejętności Uczniów OECD/PISA. Wyniki polskie — raport z badań* [Programme for International Student Assessment OECD/PISA. Poland's results—report from the study] Access: http://www.ifispan.waw.pl/

Białecki, Ireneusz. 2006. *Uczniowie słabi w badaniu PISA 2003* [Poorer Students in the PISA 2003]. *Problemy Opiekuńczo-Wychowawcze 2*: 3–11.

Bishop, John H. 1997. The Effect of National Standards and Curriculum-Based Exams on Achievement, *American Economic Review* 87: 260–64.

Choińska-Mika, Jolanta et al. 2013. *Realizacja podstawy programowej z historii w gimnazjach* [Implementation of the Core Curriculum in History in Middle Schools]. Warszawa: Instytut Badań Edukacyjnych.

Dolata, Roman. 2011. *Analiza różnicowania się systemu oświaty w Polsce na poziomie szkół podstawowych i gimnazjów, edycja 2011* [Analysis of Variation in Poland's Education System in Elementary and Middle Schools, edition 2011]. Warszawa: Instytut Badań Edukacyjnych.

Dolata, Roman et al. 2012. *Analiza różnicowania się systemu oświaty w Polsce na poziomie szkół podstawowych i gimnazjów, edycja 2012* [Analysis of Variation in Poland's Education System in Elementary and Middle Schools, edition 2012]. Warszawa: Instytut Badań Edukacyjnych.

Dolata, Roman, Jakubowski, Maciej, Pokropek, Artur. 2013. *Polska oświata w międzynarodowych badaniach umiejętności uczniów PISA OECD* [Poland's Education System in PISA OECD]. Warszawa: Wydawnictwo Uniwersytetu Warszawskiego.

Herbst, Mikołaj. 2012. *Wynagrodzenia nauczycieli w Polsce — rozwiązania systemowe, dynamika i zróżnicowanie terytorialne* ['Teacher Remuneration in Poland—Systemic Solutions, Dynamics, and Territorial Variations'] in: M. Herbst (ed.), *Finansowanie oświaty* [Financing Education]. Warszawa: ICM Press of the University of Warsaw.

Herbst, Mikołaj, Herczyński, Jan and Levitas, Anthony. 2009. *Finansowanie oświaty w Polsce — diagnoza, dylematy, możliwości* [Financing Education in Poland—Diagnoses, Issues, and Possibilities]. Warszawa: Scholar.

Herbst, Mikołaj and Levitas, Anthony. 2012. Decentralizacja oświaty w Polsce 2000–2010: czas stabilizacji i nowe wyzwania [The Decentralization of Education in Poland, 2000–2010: The Period of Stabilization and New Challenges], in: M. Herbst (ed.), *Decentralizacja oświaty* [The Decentralization of Education]. Warszawa: Centrum Rozwoju Edukacji, pp. 118–153.

Herczyński, Jan and Herbst, Mikołaj. 2002. *Pierwsza odsłona. Społeczne i terytorialne zróżnicowanie wyników sprawdzianu szóstoklasistów i egzaminu gimnazjalnego przeprowadzonych wiosną 2002 roku* [First Scene: Social and Territorial Variation in Results on Tests for Sixth-Year Elementary School Students and Examinations for Middle School Students in Spring 2002]. Warszawa: Fundacja Klub Obywatelski.

Hout, Michael, Elliott, Stuart W. 2011. *Incentives and Test-Based Accountability in Education*. The National Academy of Sciences: Washington.

Jakubowski, Maciej et al. 2010. The Impact of the 1999 Education Reform in Poland, Education. Working Paper No. 49. OECD Publishing.

Kangro, Andris and James, David. 2008. Rapid Reform and Unfinished Business: the Development of Education in Independent Latvia 1991–2007, *European Journal of Education* (43) 4: 547–561.

Kłos, Bożena. 2010. *Nauczyciele w zreformowanym systemie edukacji* ['Teachers in the Reformed Education System']. Studia BAS (22) 2: 89–124.

Lavy, Victor. 2009. Performance Pay and Teachers' Effort, Productivity, and Grading Ethics, *American Economic Review* 99(5): 1979–2011.

Levin, Benjamin. 2001. *Reforming Education: from Origins to Outcomes*. London, New York: Routledge/Falmer.

Levitas, Anthony and Herczyński, Jan. 2012. *Decentralizacja oświaty w Polsce 1990–1999: tworzenie systemu* [The Decentralization in Poland 1990–1999: Creating the System], in: M. Herbst (ed.), *Decentralizacja oświaty* [The Decentralization of Education]. Warszawa: Centrum Rozwoju Edukacji, pp. 55–115.

Marciniak, Zbigniew. 2009. *O potrzebie reformy programowej kształcenia ogólnego* [On the Need to Reform the Curriculum of General Education], in: *Podstawa programowa wraz z komentarzami* [The Core Curriculum, with Commentaries]. Warszawa: Ministerstwo Edukacji.
OECD. PISA. 2009. *Assessment Framework: Key Competencies in Reading, Mathematics and Science*. OECD Publishing.
OECD. PISA. 2009. *Results: Learning Trends: Changes in Student Performance since 2000* (Volume V). OECD Publishing.
OECD. PISA. 2009. *Technical Report*. OECD Publishing.
OECD. 2010. *PISA 2009 Results: What Students Know and Can Do—Student Performance in Reading, Mathematics and Science*. Volume II. OECD Publishing.
OECD. PISA 2012. *Results: Excellence through Equity* (Volume II). OECD Publishing.
OECD. PISA 2012. *Assessment and Analytical Framework*. OECD Publishing.
OBOP [Public Opinion Research Centre]. 1999. *Reforma edukacji w opinii społecznej 1999* [Education Reform in Public Opinion].
Piopiunik, Marc, Schwerdt, Guido, Woessmann, Ludger. 2013. Central School Exit Exams and Labor-Market Outcomes, *European Journal of Political Economy* 31: 93–108.
Pokropek, Artur. 2013. Trafność egzaminów gimnazjalnych [The Fitness of Middle School Examinations], in: M. Karwowski (ed.), *Ścieżki rozwoju edukacyjnego młodzieży—szkoły pogimnazjalne* [Paths for the Educational Development of Youth-Middle Schools]. Warszawa: Wydawnictwo IFiS PAN, pp. 97–124.
Raport NIK [Supreme Audit Office Report]. 2008. *Informacja o wynikach kontroli funkcjonowania systemu awansu zawodowego nauczycieli* [Information on the Results of Monitoring the Operation of the Teachers' Promotion System].
Santiago, Paulo et al. 2012. *OECD Reviews of Evaluation and Assessment in Education*. Portugal OECD Publishing, http://dx.doi.org/10.1787/9789264117020-en
Śliwierski, Bogusław. 2009. *Problemy współczesnej edukacji. Dekonstrukcja polityki oświatowej III RP* [Problems in Contemporary Education: Deconstruction of Education Policy in the Third Republic]. Warszawa: WNiP.
Silova, Iveta and Steiner-Khamasi, Gita. 2008. *How NGOs React: Globalization and Education Reform in the Caucasus, Central Asia, and Mongolia*. USA: Kumarian Press.
Tyack, David and Cuban, Larry. 1995. *Tinkering toward Utopia: A Century of Public School Reform*. USA: Harvard University Press.
Piwowarski, Rafał and Krawczyk, Magdalena. 2009. *TALIS. Nauczanie — wyniki badań 2008. Polska na tle międzynarodowym* [Teaching—Results of the 2008 Assessment: Poland in the International Context]. Warszawa: Ministerstwo Edukacji, Instytut Badań Edukacyjnych.
Woessmann, Ludger. 2002. Central Exams Improve Educational Performance: International Evidence, *Kieler Diskussionsbeiträge* No. 397.
Wossmann, Ludger. 2005. The Effect Heterogeneity of Central Examinations: Evidence from TIMSS, TIMSS-Repeat and PISA, *Education Economics* 13: 143–69.
Woessmann, Ludger. 2011. Cross-Country Evidence on Teacher Performance Pay, *Economics of Education Review* 30: 404–418.

Biographical Note: Alicja Zawistowska (Ph.D.) is a sociologist working at the Institute of Sociology in the University of Bialystok. She specializes in issues of education and social inequalities.

E-mail: zawistowska.alicja@gmail.com