



HAL
open science

Organisation and Context, Efficiency and Equity of Educational Systems: what PISA tells us

Marie Duru-Bellat, Bruno Suchaut

► **To cite this version:**

Marie Duru-Bellat, Bruno Suchaut. Organisation and Context, Efficiency and Equity of Educational Systems: what PISA tells us. *European Educational Research Journal*, 2005, 4 (3), pp.181-194. halshs-00006354

HAL Id: halshs-00006354

<https://shs.hal.science/halshs-00006354>

Submitted on 8 Jun 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Organisation and Context, Efficiency and Equity of Educational Systems: what PISA tells us

MARIE DURU-BELLAT et BRUNO SUCHAUT
IREDU-CNRS, Université de Bourgogne, Dijon, France

Article publié dans European Educational Research Journal, Vol.4; N.3, 2005, pp.181-194

ABSTRACT

After describing both average scores, dispersion, and social inequalities in achievement in the various countries included in the latest Programme for International Student Assessment (PISA) study, this article relates those 'products' to country economic and cultural characteristics. It then explores relations between student scores and a number of institutional characteristics of countries' educational systems. Results show that relations exist between average scores and certain institutional or pedagogical practices such as grade repeating or tracking. A high degree of social inequality in achievement proves to be associated with overall score dispersion and degree to which educational system differentiates among students.

Comparative education has been analysing the rise of assessment and benchmarking for the last twenty years, linking those trends with changes in national regulation procedures and globalisation (see, for instance, Broadfoot, 1996). This viewpoint may lead to international surveys such as those of the International Association for the Evaluation of Educational Achievement (IEA) being considered with some suspicion; but if these can actually be analysed as a part of some global trends demanding sociological scrutiny, they can also be considered as a heuristic tool for international comparisons, with the cautiousness required in any empirical research. The available data are sparse to investigate what is an overarching question in comparative education, that is, the influence of educational structures and context upon pupils' achievement. To find out why nations differ, it is a classic procedure in comparative education to consider differences in results obtained for different countries (with regard to various 'products' of schooling such as students' acquired knowledge or their attitudes) and relate them to structural characteristics of the various educational systems. This allows for testing of the possible influence of one or another characteristic, which would have been impossible if the study were confined to a single framework (most often national) where those characteristics do not vary.

Recent studies, making use of either IEA surveys (cf. Lassibille & Navarro Gomez, 2000, among others) or other international sources, such as UNESCO (see Willms & Somers, 2001, on Latin America), attest to the usefulness of this way of proceeding. Such comparisons must obviously take into account the social and economic contexts in which schooling occurs, which, as is easily imagined, affect student's academic achievements, though it is often impossible to take the exact measure of this influence. It is important, however, from both scientific and political perspectives, to distinguish between what pertains to the extra-school-system context and what is properly a matter of the system and its institutional characteristics (structures, pedagogical organisation, etc.) Asking the question this way implies accepting the validity of term-to-term comparison of traits that have been isolated from their contexts, along the lines of 'cross-national' approaches, though it is probably necessary, dialectically in the course of the analysis, to adopt a more 'cross-cultural' perspective, one that attributes special status to the national entity within which compared elements are organised and acquire meaning.

In this text we take data from the most recent PISA (Programme for International Student Assessment) survey of reading literacy skills and relate them to characteristics that concern either the country as a whole or its specific educational system, with the aim of understanding what influences student performance (both its mean level and disparity), and above all what explains the

fact that the range of social inequality varies appreciably by country. This second focus is not as readily found in the literature.

Relevance and Limitations of PISA Data

In the results presented here, the analytic unit was the country. The basic data were results obtained in PISA and PISA+ sample surveys (Organisation for Economic Cooperation and Development [OECD], 2001, 2003) by 15-year-olds schooled in the various countries on a variety of tests aimed at apprehending life skills, namely, reading literacy. In addition to PISA data, we assembled and used supplementary information from international data banks on economic contexts and educational systems in the various countries.

Comparisons always run up against a certain number of difficulties. One problem is whether or not the elements used to compare educational system 'products' are themselves comparable. PISA data are precious here in that they provide measurements of skills displayed by 15-year-olds in a high number of countries on the basis of the same exercises. One particularity of the data is that they bear on the total number of young persons in school at that age, regardless of grade or course of study; this was not always the case in previous comparative studies. The 1994 Third International Mathematics and Science Study (TIMSS) survey, for example, observed mathematics and science skills of students in the third year of lower secondary level, rather than those of all students of a given age, thereby giving countries where grade repeating is practised an advantage over countries where it is not. Conversely, according to PISA data, countries where grade repeating is practised have a percentage of 15-year-old students at lower levels than almost all students of countries that do not use repeating. This particularity of the PISA survey is responsible for another limitation: data for students (still) in school at age 15 concern a generation percentage that varies by country: while the figure is close to 100% in the wealthiest OECD countries, it is much lower in countries such as Thailand, Brazil and Mexico. This generates problems that will be considered below.

Another drawback is the standard difficulties of measuring students' acquired knowledge: psychometric literature shows that we do not know how to measure acquired knowledge without margins of error. We know, for example, that exercise content and presentation can work to certain students' advantage: the differential in achievement between girls and boys may be systematically affected by this variable. There is nothing new in the observation that any manifest performance is contingent to some degree on the textual material and type of questioning used to evaluate it; the whole concern is of course to evaluate the degree of this contingency. Clearly, skill is not directly observable, and this requires anyone investigating it to use an operative definition of the sort classically attributed to the psychologist Binet in the early twentieth century: 'Intelligence is what my test measures'. What unbiased measuring of skill would be (i.e. measurement not systematically influenced by the measuring instrument) is an entirely open question.

Though in international comparisons we can accept the idea that, regardless of country, school prepares students for fairly comparable occupational and, more broadly, social situations that require and mobilise the same skills, this approach can hardly claim to be bias free. In certain countries, such as France, critical reference is made to 'cultural bias' in the tests: students are understood to be less familiar with a certain type of exercise or the multiple-choice question format, itself understood as culturally Anglo-Saxon. But actually, countries with the highest performing students belong to very different cultural areas (Korea, Finland, Canada), and students in culturally similar countries may perform relatively unevenly (the USA and Canada, for example). If the question of the reality of cultural biases remains a valid one, the essential question in the present study is whether student scores are still reasonably comparable from one country to the next. We can reasonably answer in the affirmative at the focal distance from which comparative education generally views things. Our comparison here aims primarily at establishing relations between a

given score in literacy [1] (and mostly score dispersion as a function of variables such as social origin) and given educational system characteristics.

The exercise of relating these measurements to systems' structural characteristics (with the implicit aim of shedding light on certain education policy issues) remains rife with stumbling blocks. Differences among countries, whether they involve average scores or distributions, require extremely careful interpretation, because they are not in any immediate, straightforward way a reflection of differences in system efficiency (or other system parameters). For 15-year-olds, those differences result jointly and cumulatively from the curricula in use, but also from the educational structures and type of teacher–student relations that have prevailed throughout schooling, as well as from all learning experiences from birth to the student's present age. In other words, they not only reflect what students have acquired through schooling but what they have learned from their family, peers, the media, daily life, etc. Major cross-sectional IEA surveys such as PISA, then, do not allow for evaluating the real efficiency of educational systems.

Our research here into correlations between skills displayed at age 15 and system organisation modes is guided by certain research results. We can, for example, expect higher levels of social inequality when structures differentiate early (due to the inequality characteristic of the tracking process) and when instruction content is close to classic culture (a state of affairs which benefits students whose families master this type of culture). But in reality, these diverse system characteristics are likely to be linked together and form a system. A classic stumbling block in international comparison is to attribute too much importance to an element isolated from its context. In general it is risky to work from correlations that bring together term for term a given performance and a given educational system characteristic, since what is probably at work is the whole set of traits that constitutes each country's societal coherence. Ideally, comparison should bear on these overall structures; but in reality, in comparative studies, imputing cause on the basis of observed correlations is quite common. Such imputations are as widespread as they are ideologically tinted. For instance, this is the case for a 2002 OECD study that suggests, despite rhetorical precautions, a cause–effect relation between good performances in certain countries and the fact that schools there are granted a certain autonomy, whereas only one correlation in this direction was observed and a number of other characteristics of strongly performing countries where schools may or may not be autonomous were not taken into account. It is crucial not to conclude too hastily that a causal relation exists between this factor and that phenomenon, and we will be extremely cautious here in this respect.

A number of other difficulties, also classic in comparative studies, should be noted briefly. One methodological problem involves the quality of data available on system characteristics. For unspecified reasons, data are occasionally missing from UNESCO or OECD files. It is therefore necessary to fill in the gaps by enriching existing bases, but this leads to 'mixing' data of different origins that are perhaps not perfectly comparable. A more delicate question is that of the potential disparity between a formal characteristic and the way a given system actually works: existing administrative data may make it appear that tracking is not practised in a given lower secondary school, whereas in fact the school uses some form of ability grouping.

Last but not least, in comparative education as in other social science fields, results are contingent on sample used, namely, degree of sample variety. This means that we cannot estimate the influence of a characteristic universally spread throughout a country sample. Likewise, if in one analysis we combine factors that vary little from one country to the next with others that vary sharply, the weight of the second will be more manifest than that of the first. In other words, not only would we be evaluating only what varies, but we would evaluate it that much 'better' because it varies. If we include both rich and very poor countries in the sample of countries studied, country's wealth can show up as influential with regard to high scores whereas it would not be for a sample of more homogeneous countries. Likewise, and unsurprisingly, context factors associated significantly with student performance can vary for poor and rich countries (see Mingat & Suchaut, 2000, for

examples), since many factors only come into play below a minimum threshold. Indeed, some studies show that schooling organisation factors weigh more on student performance in poor countries than rich ones (Heyneman & Loxley, 1983).

In this respect, the sample extended, in ‘PISA Plus’ to 40 countries (OECD, 2003) still does not reflect the variety that exists at the scale of the world as a whole, as it includes no African or Middle Eastern countries and many Asian countries are missing (India; China, with the exception of Hong Kong, Pakistan, Malaysia, Bhutan, etc.) The sample is representative of industrialised and emerging countries. It should be stressed, however, that the results of analyses based on PISA countries cannot be generalised to apply to the world’s poorest countries.

Overview of Differences in Average Level and Inequalities among Countries

A preliminary step in understanding the structural factors associated with observed differences is to describe differences in average performance and dispersion, and social inequality: are social inequalities in achievement less important when the average level of pupils is high, or when the scores are on the whole more homogeneous? But though this exercise has real significance, it also has its limitations: approximately one-tenth of total variance in PISA student performance is explained by membership in a given country (inter-country variance). Only this tenth, then, can be related to differences in averages among countries (these in turn explicable by national context characteristics and schooling organisation factors). The remaining nine-tenths of variance concerns individual countries themselves (intra-country variance). This means that variable structural factors from one country to the next have little import with regard to skills mastered by students at age 15 compared to what hinges on inequalities among schools or students, these in turn due in large part to family inequalities. It is also possible that PISA countries share all characteristics that affect efficiency. Moreover, the one-tenth figure cannot be considered valid at the world scale in any case, and it is likely that if countries differing in even greater degree in terms of educational system, wealth, education level of population at large, etc. were taken into account, structural factors would increase in importance. As we shall see, certain relations may vary appreciably depending on whether the entire set of countries is considered or only richest countries.

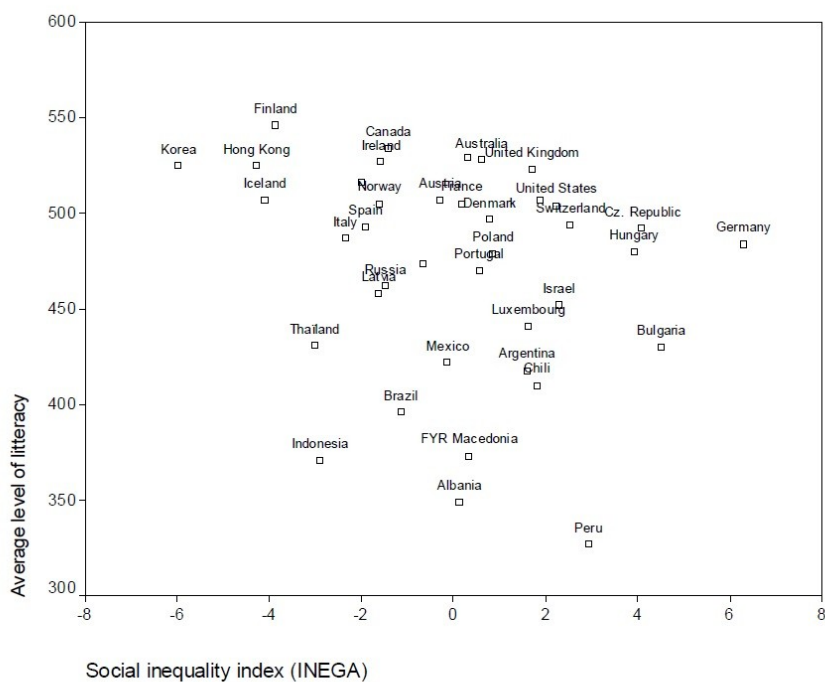


Figure 1. Social inequality index (INEGA) and average literacy level.

The first observation is that reading literacy performance levels differ greatly among countries. Whereas the average for all countries taken together is 500, with a standard deviation of 100, average scores range from 540 for Finland to 327 for Peru (see Figure 1 for different countries' average scores). In some countries, performance is both high on average and relatively homogeneous (Korea, and less strongly Finland and Japan). In others, on the contrary, performance is both low and dispersed (Luxemburg, Argentina, Bulgaria, and, at an extremely low level, Albania). But all sorts of situations exist. There are many countries where students perform both well and unevenly (New Zealand, the United Kingdom) and countries where the opposite is the rule, i.e. low, homogeneous performances (Indonesia). It can also be stressed that at the same average level, inequality among students is more or less marked from one country to another. Contrary, then, to the commonly accepted idea that attaining a certain performance level is 'paid' for by marked result dispersion, there is no automatic relation between students' average level and disparities among them. In other words, a high average level is not necessarily associated with greater inequality. Moreover, if we analyse only the richest countries in the sample [2], we see a slight negative relation between average level and dispersion, namely in science, where the highest performing countries also have a general tendency to show the least student score dispersion (correlation coefficient for reading literacy is -0.29 ; mathematics, -0.54 ; science, -0.57).

Another approach to evaluating reading literacy performance is to adopt a criteria-based perspective that takes into account the content of skills at each level. PISA has identified six threshold scores, corresponding to a hierarchy of degrees of language command. Here we focus on indicators representing extreme groups, where student percentages fall either below level 1 proficiency [3] or above level 5 [4], i.e. students in 'trouble' when it comes to using reading as a tool and the scholastic 'elite'. These two indicators are negatively linked overall (correlation of -0.66), meaning that number of students in 'trouble' increases as number of 'elite' students falls. But there are certain specificities within the sample. Korea shows clear compression of scores around the average; it has both very few excellent students (5.7%) and almost no students in major difficulty (0.9%), whereas Belgium has high numbers of both excellent and weak students (12% and 7.7% respectively). Some countries manage not to have any students at an extremely low level in reading literacy (Korea, Finland, Canada, etc.), an attestation if one were needed that there is no reason to believe scholastic difficulties are inevitable. This also means a priori that all countries can attain the minimum skill level, even if, as will be demonstrated further on, certain socio-economic or pedagogical conditions are more or less favourable. The fact remains that the greatest differences among countries concern the weakest students. As we know, context variables weigh heaviest on the weakest.

There are also marked variations among countries in social inequality in performance. The PISA survey provides indicators of the family's material or cultural wealth that enable us to evaluate impact of social background thus defined on student performance within the different countries. Here what we are trying to explain is variation by country of the strength of relation between a student's sociocultural background and test performance. This relation can be apprehended in various ways. PISA data provide several non-independent indicators [5], and in fact the rankings for a good number of countries are fairly similar regardless of indicator used. To give more consistency to the social inequality in achievement measure, we have constructed a synthetic, reduced indicator (average 0, standard deviation 1), calculated from the average of three indicators standardised to the same scale (for the construction, see note 5). This indicator, called INEGA, is used in our analyses further on.

First, however, we should point out that this study, focused on social inequality in achievement, does not take into account all types of social inequality relevant to schooling in the countries in question. The very fact of still being in school at age 15 refers to a major feature of social inequality. We would also stress the fact that our object of study here, social inequality affecting acquired skills, does not cover all social inequality operative in scholastic careers. We are well

aware that at comparable achievement levels, there is also social inequality with regard to tracking and course of study every time a study or training specialisation choice is to be made. 'Total' social inequality level is thus likely to be much higher than level of inequality affecting acquired skills, which is rather the 'base' level of social inequality.

We can now examine on the basis of the general indicator whether there is a relation between social inequality in reading literacy performance and average level in the different countries. We could hypothesise competition between policies aimed at improving average efficiency and those aimed at reducing social inequality – what economists classically call arbitrating between 'efficiency and equity.' Figure 1 shows that variety in range of social inequality (INEGA) is weakly associated with average performance level, but not in the direction expected.

The relation is negative ($r = -0.26$) but weak, only just significant (threshold of error: 11%). In other words, the most socially inegalitarian countries are also slightly more likely to have the weakest students. Conversely, countries where social equality is the strongest have, on average, the highest performing students. However, as shown in Figure 1, a number of countries with comparable inequality levels have highly varied performance levels: Finland versus Thailand, Denmark versus Albania, Hungary versus Peru, etc. Conversely, at the same performance level, there are highly variable social inequality levels: Germany versus Italy, United Kingdom versus Korea, Bulgaria versus Thailand, etc. Significantly, a number of poor countries (Indonesia, Thailand) have moderate levels of social inequality in performance, but this should be considered cautiously as education in these countries is not universal. It is therefore likely that students there are socially selected [6], which is likely to minimise the effect of social inequality on achievement. However, observed variety in range of social inequality at a given educational level makes clear the value of autonomous analysis of possible explanatory factors.

A further strong observation is the existence of a strong positive relation between level of social inequality in performance and dispersion of reading literacy results. Correlation between INEGA and literacy result dispersion is $+0.67$, meaning that countries where student performances are heterogeneous also have high levels of social inequality in performance. Conversely, countries whose students perform homogeneously are also countries where social inequality is least able to manifest itself. Moreover, there is a moderate relation between proportion of weakest students and strength of social inequality ($r = +0.27$, significant at the 10% level): countries where social inequality is most marked also tend to be those with the highest percentage of weak students. Conversely, when countries manage to reduce the percentage of weak students, achievement tends to be less affected by social inequality.

Overall Characteristics of National Contexts and Student Performance

The preceding results bring to light strong differences in performance from one country to another, suggesting that overall context is probably related to student performance. Before linking performance with country educational system characteristics, it is therefore important to take into account overall economic context and a population's – therefore parents' – education level. The question here is to try to disentangle the part played by the macroeconomic context and the characteristics of the educational system.

First, there is a statistical relation between students' average performance and country's gross domestic product (GDP). GDP statistically explains 46% of variance in average level. The relation is positive: the richer the country, the higher the scores. But this relation is 'disturbed' by one highly atypical country, Luxemburg, which is very rich but has a low performance level. One way [7] of circumventing this difficulty is to use the logarithm of GDP; economic wealth now explains 67% of variance in average literacy level. It should be underlined that if relatively rich OECD countries only are taken into account, the figure decreases to 13%. That said, at a given GDP level, countries'

average performances remain unequal. Eastern European countries, whose per capita GDP is fairly close to that for South American countries, have much higher performing students than the latter. Moreover, GDP is also linked to dispersion of results among students: student performances prove more heterogeneous the richer a country is, though the relation is moderate (percentage of explained variance or $R^2 = 9\%$).

On the contrary, for range of social inequality there is no significant relation with country wealth (correlation null). Likewise, and contrary to expectations, there is no relation between level of social inequality in scholastic achievement and adult population income inequalities (measured by the Gini index of income inequality). However, some countries are egalitarian in terms of both the Gini and student performance (Japan, Korea, Finland). But no country is both income-egalitarian and characterised by low social inequality in scholastic performance. In countries with the highest income inequality (Latin America), impact of social origin on performance is fairly average. What is most striking is the variety in influence of social origin in countries with similar Gini index ratings. This result indicates that some systems manage to compensate better than others for the inequality that exists in the societies in which they are embedded.

Another context indicator disposed to influence young people's performances is education level of the population at large, apprehended in PISA by mother's education level. There is in effect a negative relation between student performance and percentage of mothers with a low educational level (not above lower secondary school), though it is relatively moderate ($R^2 = 0.13$). The two GDP variables and mother's education integrated into a single explanatory model of average performance explain 70% of differences in literacy performance among countries ($R^2 = 0.70$). Each of the two variables exercises a specific influence, but the impact of GDP is sharply stronger. Adding mother's education increases model power by approximately 6%.

Relations between this sociocultural context indicator and social inequalities in performance are less clear. A negative relation is observed between range of social inequality in performance and students' mothers' education: the fewer well-educated mothers a country has, the weaker the impact of social origin ($r = -0.37$). This relation is primarily due to Asian countries, which are egalitarian but where mothers' level of education is low (Hong Kong, Indonesia, Thailand). The result is consistent with the preceding ones, however, in its affirmation that a country's economic development or overall education level are not at all automatically associated with reduction of social inequality in student performances.

While the major role of GDP is thus confirmed, the question remains as to what other paths its influence takes, in addition to the association between GDP and education level of population at large. One obvious point to investigate is how much funding a higher or lower level of wealth enables a country to allocate to schooling. Across countries, a fairly strong positive relation is observed between public expenditure on education and average performance level, but the relation weakens considerably and is no longer significant for countries situated above a certain wealth level (from Greece and the Czech Republic to the United States). Public expenditure on education as percentage of GDP follows this trend exactly, and thus is not a relevant variable for rich countries. [8] However, GDP in rich countries does have a positive if moderate impact on performances. Certain paths of GDP influence in relatively rich countries therefore remain somewhat mysterious, as they involve neither a higher percentage of the budget being allotted to education or increased expenditure amount. Wealth level would thus seem to work diffusely and in many ways through the multiple resources and stimulations a rich country can apply. This general area remains to be explored.

Lastly, we should note that expenditure or proportion of budget allocated to educational system does not seem any more closely correlated with range of social inequality than is GDP.

Selectiveness and Differentiation within Education Systems and Student Performance

In this section, we first consider the relations obtaining between certain system characteristics and students' displayed skills and subsequent inequalities, relying upon research in educational sciences, then use a constructed synthetic variable to evaluate their overall impact.

Number of Instruction Hours

Amount of time spent in school is among school organisation characteristics known to affect students' level of acquired knowledge : some opportunity to learn is obviously necessary for pupils to progress (see, for example, Hanusheck, 1997); but the relation between allocated learning time and attainment is currently established at the pupils' level, in a given country, and a question to investigate is this relation at the country level: is some part of the differences in achievement between countries explained by the number of hours devoted to instruction ? In fact, there proves to be no correlation between number of instruction hours for students aged 12 to 14 (information only available for certain countries) and average performance (or range of social inequality). It must be specified that the overall average data for number of instruction hours gives a highly imperfect account of actual study time inside and outside school. In some countries, students take classes at times other than the standard periods. Another possible explanation for the absence of correlation is that with decentralising reforms and increased school autonomy, hours in school by discipline are increasingly decided non-centrally, which in turn makes this information more difficult to capture statistically.

However, we know there are diminishing returns for hours spent in school, and the fact of grade repetition in secondary school is negatively, significantly related to average reading literacy results ($r = -0.46$). That relation does not change when a country's economic wealth is taken into account. Moreover, these results are also valid for mathematics and sciences. In other words, the practice of grade repetition not only does not improve student results, but is even associated with poorer performances for 15-year-olds. This holds in particular for the weakest students, whose proportion increases with grade repetition. Furthermore, percentage of grade repeaters in secondary school is positively, significantly related to social inequality in scholastic achievement ($r = +0.31$). This is probably explained by the fact that in countries where grade repetition exists, it is due not only to scholastic difficulties but also, at a comparable level of achievement, to social biases in teachers' decisions and/or families' reactive resistance to those decisions.

System Selectiveness and Performance

In all industrialised countries, the expansion of education systems is supposed both to increase the mean level of achievement and to loosen the relationship between social background and level of achievement. So, we would expect the enrolment rate for 15-year-olds to be linked to those two parameters. In fact, the average level is higher in proportion to enrolment rate for 15-year-olds. It can also be observed that the percentage of young people with scores below level 1 is negatively correlated with the enrolment rate for 15-year-olds (highly significantly: $r = -0.57$).

Selective admission, regardless of whether it corresponds to an explicit policy or is linked to a country's poverty level, not only does not guarantee a higher student level but actually threatens student performance. To use a sports metaphor, level rises with number of competitors in the pool. It is certainly so because in the countries where there are high numbers of drop-outs at an early age, the selection is not always made on the basis of merit (some social inequalities are generally interfering), thus moving away from school pupils who would have been able to achieve well. So, as the enrolment rates increase, the pool of competitors becomes bigger, the level of achievement rises and the social inequalities decline.

But, on the other hand, a certain amount of the selectiveness inherent in the fact of being in school at 15 in countries where not all children of this age are in school may reflect some inequalities of

attainment. In that case, in those countries, only the best students from working-class backgrounds are enrolled, minimising the range of social inequality in performance. This would explain the fact that there is no relation between range of social inequality in performance among enrolled students and educational participation for 15-year-olds: correlation between these two indicators is null. In fact, with the available data, it was not possible to assess the respective validity of these two hypotheses.

Structure of Secondary Education

There is much literature on the impact of a system's structure and schedule for channelling students on the range of inequality among students, namely, social inequality in scholastic careers (for a synthesis, see, for instance, Duru-Bellat, 2002, 2004). Scholastic careers are strongly affected by how open school systems are, and therefore by a country's 'survival pattern' (that is, the structure of pass rates and thus range of selection at the various levels). Inequality is strongest in countries where selection is strong and occurs early; conversely, a sure way of democratising scholastic careers is to attenuate or defer selection. We also know that the existence of early streaming points, followed by tracking, or an important weight of family in orientation process, are institutional parameters that increase social inequality. But all these relations concern above all social inequality in scholastic careers, and are perhaps less relevant when the focus is on inequality in 15-year-olds' acquired knowledge. Recently, researchers have used IEA international studies (cf. the 1995 TIMSS survey on mathematics and science knowledge of students aged 13-14) to analyse acquired knowledge (Vandenberghe, 2001; Zachary, 2001). Among other results, these studies indicate that several countries show good average performance without a strong correlation existing in them between performance and social origin. In other words, range of social inequality by country can vary by as much as 100% at a given average performance level. Among factors likely to explain such variation, the studies stress that countries where all students follow a common core curriculum for a relatively long time seem likely to come out ahead of countries that have opted for early, marked track differentiation. Other institutional characteristics, such as allowing for grade repetition or the existence of segregation among schools, are also factors for social inequality in achievement. Our results converge with this literature. Looking first at average student performance level in relation to age of first streaming, we see that compared to countries where streaming first takes place at age 15 or over, students in countries with earliest streaming (no later than 14) perform slightly less well on average, though the difference is not statistically significant. Likewise, no statistical relation can be observed between streaming at 14 and dispersion in student performances, whether we are looking at the proportion of students below skill level 1 or the proportion of excellent students. However, there is a slight negative relation for this variable: the proportion of excellent students appears lower in countries with early streaming. A mode of educational organisation that streams students not only does not improve their performance but does not produce a more substantial elite. Controlling for GDP in the model does not contradict these conclusions.

An alternative indicator of system structure is the length of time students have a common core curriculum. This can vary from 4 years (in German-speaking countries) to 13 years (Hong Kong). There is a moderate (positive) relation between length of time and average scores (which is nonetheless maintained, though becoming slightly weaker, if a country's wealth is taken into account). By contrast, and surprisingly, length of time with a common core curriculum is not related to score dispersion, namely, for the proportion of students below performance level 1.

Relations between these indicators of system differentiation and range of social inequality are sharper. Social inequality is greater in countries with streaming, and the relation is statistically significant: streaming at 14 alone accounts for 12% of variation in the social inequality index. It is especially important to underline this relation given that, as we have seen, the existence of early streaming is not significantly linked to average performance level. Convergetly, if we take into

account length of common core curriculum study, we find a significant correlation with the range of social inequality for scholastic achievement: the longer all students follow the same curriculum, the smaller the degree of inequality; conversely, inequality tends to be more marked when common core curriculum time is brief ($r = -0.40$).

Segregation among Schools

Depending on the education system, students are more or less likely to be grouped together in schools by scholastic or social characteristics. This is an important point, since research shows the impact of what is called 'school mix' (both social and academic) upon pupils' achievement and socialisation (see, for instance, Opdenakker & Van Damme, 2001). In particular, within schools where pupils come from a disadvantaged background, progress and attitudes towards schooling are generally dragged down.

Segregation among schools can be apprehended by proportion of student score variance explained by school attended. Student results become weaker as inter-school variance increases ($r = -0.35$); they are also proportionally more dispersed ($r = +0.39$), but these two relations are of moderate intensity. It turns out too that social inequality in achievement among students tends to be more marked in countries with strong segregation ($r = +0.66$). In places where, in fact or by law, schools enrol clearly distinct types of students, such scholastic segregation takes on a social dimension and students' acquired knowledge is more marked by their social origin than in countries where there is greater heterogeneity within schools.

Degree to Which Systems Differentiate among Students

Systems with streaming, distinct schools, or grade repeating (all of which work to homogenise classes) represent ways of grouping students that allow for establishing an opposition between systems in terms of 'integration' or 'differentiation' 'culture,' as suggested by Crahay & Delhaxhe (2003). Does this affect student performance and social inequalities? To investigate this question we constructed an aggregate variable that takes into account degree of grade repeating in primary school, length of common core curriculum time, and range of scholastic segregation among schools. The higher the value for this variable, the more homogeneous the pedagogical environment students are placed in. Conversely, the lower the value, the more system organisation is characterised by heterogeneity. A negative correlation, significant though of moderate intensity, is observed between the aggregate variable for system differentiation and average reading literacy score ($r = -0.33$). The same correlation is observed for mathematics.

Countries whose systems cultivate differentiation thus have somewhat lower-performing students. But contrary to what might be expected, no significant correlation exists between degree to which a system differentiates and performance dispersion. It is clear that strongly differentiating systems do not generate a scholastic 'elite' (apprehended by percentage of students at level 5); in fact the opposite is true: the least differentiating systems are ranked highest on this point ($r = -0.35$).

These two relations – between the degree to which a system differentiates and average score on the one hand, and weight of the 'elite' on the other – are worth testing by taking into account the primary variables we saw to be important in the preceding analysis; namely, economic wealth and universality of enrolment. The models show that the degree to which a school system differentiates continues to be negatively associated with both average literacy score and percentage of students above level 5, even at given levels of GDP and secondary school enrolment.

Another, more intense relation exists between the degree to which a system differentiates and range of social inequality (see Figure 2). The higher the level of heterogeneity maintained by school systems (low level of grade repeating, long same-programme period, highly similar schools), the lower the degree of social inequality in achievement among students ($r = +0.63$). Reciprocally, the fact that the amount of system differentiation is above all correlated with range of social inequality (and not average level or result dispersion) underlines how much is at stake in access to schooling.

In their strategies, families seize on any and all differences (choice of track, choice of school) and make them into criteria of distinction.

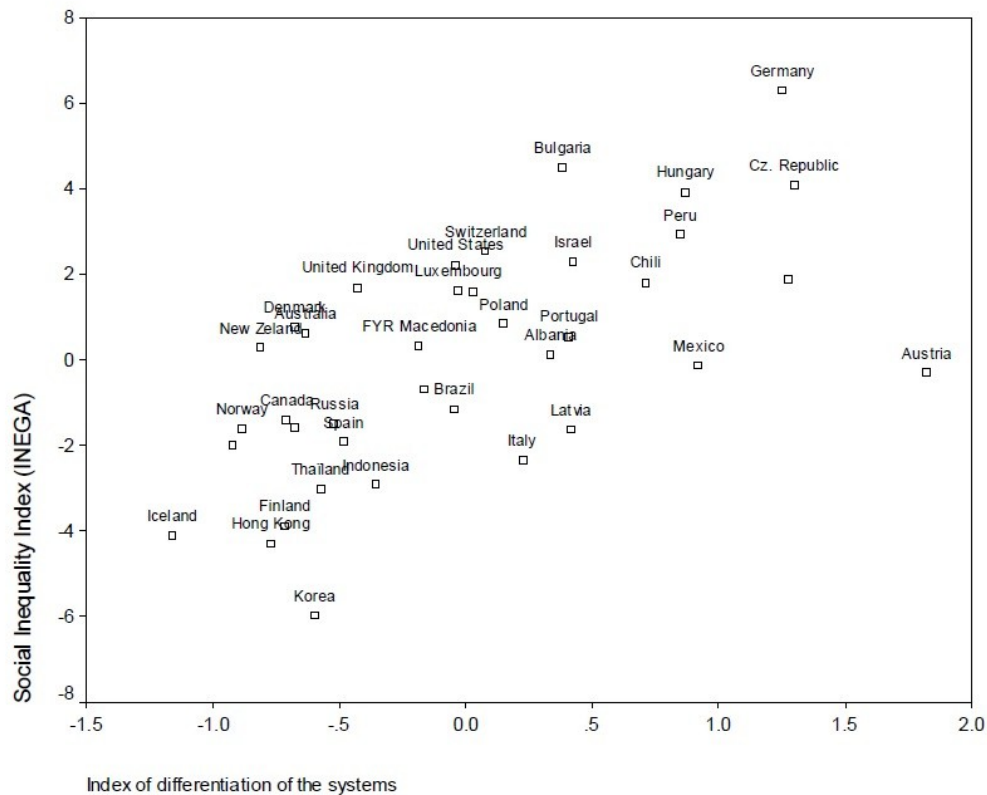


Figure 2. Social inequality index (INEGA) and degree to which education systems differentiate.

Lastly, we have constructed a synthetic model that combines score dispersion and differentiation indicator only. The model explains 67% of differences among countries for social inequality in performance. In other words, if countries display more or less social inequality in student performance, this is above all because of two factors of fairly similar importance which are fairly weakly correlated and therefore come to be added on : the existence of intra-country inter-individual inequality and the lesser or greater degree to which a country's educational system differentiates among students.

Variables	Coefficient	Significance
Differentiation indicator	+0.72	***
Dispersion of reading literacy results	+0.061	***
Constant	-5.85	***
R ²		0.67

*** Significant at 1 % level; ** significant at 5% level; * significant at 10% level.

Table I. Model explaining social inequality by degree to which systems differentiate and dispersion of student literacy results.

To provide an overall view of differences among countries, we did a principal components analysis combining the major institutional characteristics of an educational system with range of social inequality, a country's level of wealth, and reading literacy performance.[9] Figure 3 presents a horizontal axis that explains 23% of variance and contrasts countries with high enrolment and high performance (on the right in the figure) and that tend to be rich, with countries in the opposite situation (on the left), as expressed by the variables *gdp1*, *isei1*, *scol1*, *écrit1*, *niv03*. This offers confirmation that what structures differences among countries is first and foremost enrolment and performance level, together with wealth. The vertical axis explains 13% of variance and represents the degree of a system's structural homogeneity and range of inequality (*zinega1*, *red1*, *fill1*, *segreg1*). At the top of Figure 3 are variables representing the most egalitarian countries and least differentiating structures; countries with the opposite features are represented at the bottom.

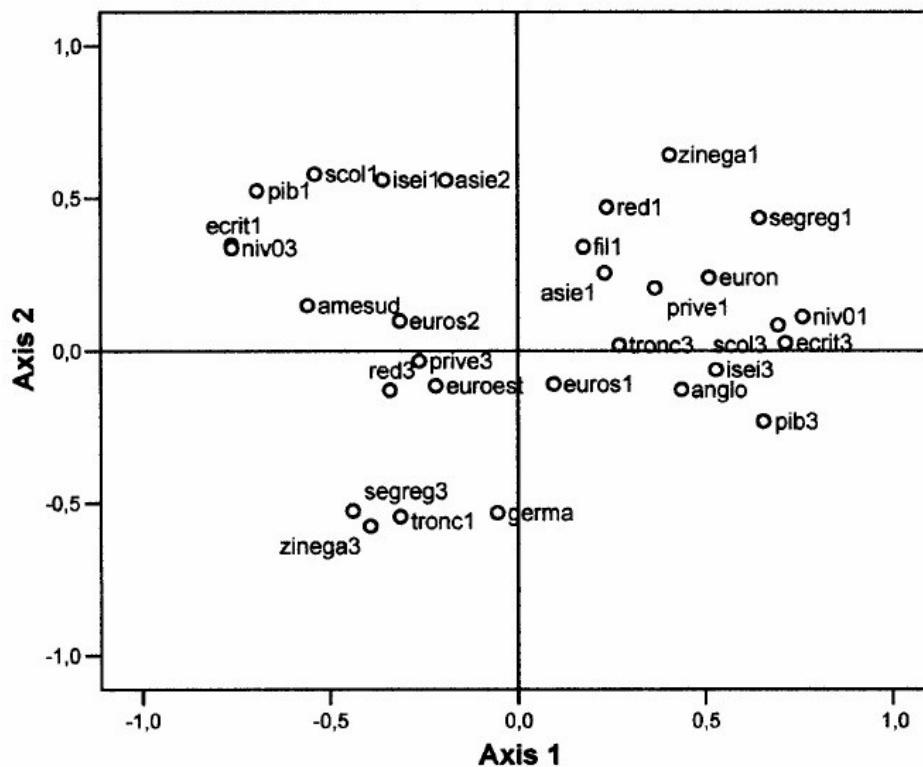


Figure 3. Principal components analysis of main characteristics of the systems.

In Figure 3 we have also projected groups of countries, identified as major regions of the world.[10] This allows for establishing a contrast, between fairly high-performing, egalitarian countries – northern European and rich Asian countries – which appear in the upper right-hand quadrant of the figure – and German-speaking countries, which display the opposite characteristics and appear in the lower left. In the upper left-hand quadrant we find lower performing countries that are also relatively egalitarian (namely, poor Asian countries); they may be contrasted with Anglo-Saxon countries, higher performing but less egalitarian, lower and to the right.

While factor analysis obviously does not produce results fundamentally different from those established previously, it does support the sense that a real coherence exists in relations between system characteristics and a country's level of wealth on the one hand, student performances and the social inequality affecting them on the other.

Conclusion

It would be premature to draw policy lessons from these observations, if only because they are only correlations, insufficient for establishing reliable cause–effect relations. Nonetheless, we should underline two points that are both substantial and not insignificant from the point of view of action, though they may seem rudimentary. First, the fact that students’ acquired knowledge and the types of inequality that affect it (inter-student inequality, social inequality) vary appreciably by country shows that these are issues for policy action, keeping in mind that the heavy weight of country wealth also points up the limited potential effect of policies restricted to the educational sphere.

Second, it clearly appears that the overall mode of a system’s organisation is correlated with its degree of equity. Any restriction on enrolment or early setting aside of certain students, any grouping by distinct level or streaming (in the framework of compulsory schooling), and all phenomena of segregation among schools (which create a ‘supply’ of uneven quality) tend to increase social inequality in performance without improving average level or even elite level. This observation is true when scholastic skills only are taken into account. It is surely even more suggestive for anyone also interested in a country’s young people’s social and civic integration.

This set of results, which meets a certain consensus among educational researchers, does not entail that to lower social inequalities towards school, the best policy would be to postpone all selection and to make the core curricula longer and longer. That is because unequal families strive to keep their children ahead, with those with the best educational level getting the best jobs, so that inequalities will probably only occur later, as it has been observed in most European countries in the last fifty years. It may be more important, but also more difficult to achieve, to try to reduce social inequalities occurring very early in the primary school, which accumulate years after year, resulting, at age 15, in inequalities of attainment that are significantly correlated with social ones. Another path, as PISA data suggest, is to struggle against the inequality of school supply, and, more broadly, to be very careful regarding the way pupils are grouped within classes and schools. Moreover, PISA data do not include observation of in-class practices. They therefore do not allow for exploring all possible avenues for action. However, they do show that ‘school matters’ and has an effect on inequality among students. This deserves to be explored in further research, and also replicated in the poorest countries. This needs also to be stressed, against some growing scepticism concerning the possible role of educational policies.

Notes

[1] In fact, our description of differences among countries concerns almost exclusively the average level of reading literacy, the primary focus of the present PISA survey. Moreover, correlation at the inter-country level is very high between this skill and the other two tested ($r = +0.95$ for mathematics culture, $r = +0.97$ for science culture); i.e. country rankings are very similar for all three of these dimensions.

[2] Australia, Austria, Belgium, Canada, the Czech Republic, Denmark, Finland, France, Germany, Greece, Hong Kong, Hungary, Iceland, Ireland, Israel, Japan, Korea, Latvia, Liechtenstein, Luxemburg, New Zealand, Norway, Poland, Portugal, the Russian Federation, Spain, Sweden, Switzerland, the United Kingdom, the United States.

[3] Students scoring below level 1 (335) cannot apply the most basic knowledge and skills. Those who reach level 1 are able only to perform the least complex PISA reading tasks, namely, locating a single fragment of information, identifying a text’s main theme, and establishing a relation between what they read and their everyday knowledge and experience.

[4] These are students scoring above 625, meaning they can successfully perform complex reading tasks (handle difficult-to-find information in unfamiliar texts, critically evaluate such information, etc.)

[5] First indicator of student’s social background is the ‘ISEI’ variable, an expression in continuous form of parents’ occupational qualifications. From ISEI, average level of students from families where the head of

household is in the lowest quartile can be compared with same for the head of household in the highest quartile, thereby producing the first indicator of social inequality in scholastic achievement for the country in question. There is also a composite indicator of a student's parents' social position, taking into account not only ISEI but also the education level of the more educated parent, the family's standard of living (measured on the basis of material goods possessed), and sum of cultural goods available in the household. This variable, called IESCS in PISA data, can be related to average performance; from that we can calculate a straight regression line whose slope, comparable from country to country, measures breadth of social inequality in performance. The last indicator of influence of a student's sociocultural background is difference in performance by whether the mother has primary or lower secondary education (ISCED 1 and 2) or higher education (ISCED 5 and 6). INEGA is the average of these three indicators.

[6] It is impossible to test this hypothesis precisely with the available data, but we did compare the correlation between average ISEI value and country's GDP ($r = +0.50$) with that between ISEI and secondary school enrolment rate ($r = +0.80$). In low-enrolment countries which also show strong social selectiveness, the average value of ISEI should be higher than that observed for countries at a comparable level of economic development. In fact, when poor countries are isolated, we do not observe a negative relation between ISEI and enrolment rate. If such a relation were found, this would have meant very strong social selectiveness, with the restriction that countries where not all 15-year-olds are in school also have a less skilled social structure. The question thus remains open for the present study.

[7] The relationship between educational inequality and inequality of income remains to be studied more thoroughly; for a perspective taking into account educational inequality among adults (on IALS data), see Green et al (2003).

[8] The fact that all comparative results are contingent on the population in question is clearly illustrated with regard to a central component of expenditure on education; namely, personnel as reflected in student-teacher ratios. If this indicator is combined with average reading literacy score for all countries, the correlation is negative and significant ($r = -0.45$). This means that the higher the student-teacher ratio, the lower the performance. On the other hand, if richest countries are isolated, the relation becomes positive ($r = +0.56$), meaning that performance tends to be better in countries where the number of students per teacher is high (rich Asian countries being largely responsible for this relation). For Latin American countries, however, Willms & Somers (2001) obtain the opposite result (the higher class size, the lower the performance).

[9] Most factor analysis variables were constructed by breaking down the initial variable into three thirds ranging from 1 (lowest values) to 3 (highest values). For example, 'niv01' corresponds to a low percentage of students below reading literacy skill level 1; likewise, 'zinegal' represents the most egalitarian countries. The variable expressing streaming is dichotomous: 'fill' means no streaming.

[10] The following groupings may be distinguished: northern Asian countries, marked Asia 1 on Figure 3 (Japan, Korea); poorer Asian countries, marked Asia 2 (Thailand, Indonesia); Anglo-Saxon countries (United Kingdom, Australia, United States, etc.); northern European countries (Finland, Sweden, etc.); eastern European countries (Hungary, Russian Federation, Poland, etc.); German-speaking countries (Germany, Luxemburg, Austria, etc.); richest southern European countries, marked 1 (France, Italy, etc.); least rich southern European countries (Bulgaria, Macedonia, etc.); and Latin American countries (Brazil, Argentina, Mexico, Peru).

References

- Broadfoot, P. (1996) *Education, Assessment and Society* (Buckingham, Open University Press).
- Crahay M. & Delhaxe A. (2003) *L'enseignement secondaire inférieur : entre culture de l'intégration et culture de la différenciation, Rapport pour le PIREF-Ministère de la recherche* (French Ministry of Research).
- Demeuse, M., Crahay, M. & Monseur, C. (2001), *Efficiency and Equity*, in W. Hutmacher, D. Cochrane & N. Bottani Eds) *In Pursuit of Equity in Education*. Dordrecht: Kluwer.
- Duru-Bellat, M. (2002) *Les Inégalités sociales à l'école: genèse et mythes*. Paris: PUF.
- Duru-Bellat, M. (2004) *Social Inequality at School and Educational Policies* (Fundamentals of Educational Planning, no. 78). Paris: UNESCO.

- Duru-Bellat, M., Mons, N. & Suchaut B. (2004) Caractéristiques des systèmes éducatifs et compétences des jeunes de 15 ans: l'éclairage des comparaisons entre pays. Cahier de l'IREDU, no. 66. Dijon: IREDU.
- Green, A., Preston, J. & Sabates, R. (2003) Education, Equality and Social Cohesion: a distributional approach, Working paper, Wider Benefits of Learning Centre, London.
- Hanusheck, E.A. (1997) Assessing the Effects of School Resources on Student Performance: an update, *Educational Evaluation and Policy Analysis*, 19(2), pp. 141-164.
- Heyneman, S. & Loxley, W. (1983) The Effects of Primary School Quality on Academic Achievement across Twenty-nine High and Low Income Countries, *American Journal of Sociology*, 88, pp. 1162-1194.
- Lassibille, G. & Navarro-Gomez, L. (2000) Organisation and Efficiency of Educational Systems: some empirical findings, *Comparative Education*, 36(1), pp. 7-19.
- Meuret, D. (2003) Pourquoi les jeunes de 15 ans ont-ils à 15 ans des compétences inférieures à celles de jeunes d'autres pays? *Revue Française de Pédagogie*, 142, pp. 89-104.
- Mingat, A. & Suchaut, B. (2000) Les systèmes éducatifs africains: une analyse économique comparative. Brussels: De Boeck.
- Mons, N. (2004) Politiques de décentralisation en éducation: diversité internationale, légitimation théorique et justifications empiriques, *Revue Française de Pédagogie*, 146, pp. 41-52.
- Organisation for Economic Cooperation and Development (OECD) (2001) Knowledge and Skills for Life. First Results from PISA 2000. Paris: OECD.
- Organisation for Economic Cooperation and Development (OECD) (2003) Literacy Skills for the World of Tomorrow. Further Results from PISA 2000. Paris: OECD.
- Opdenakker, M-C. & Van Damme, J. (2001) Relationship between School Composition and Characteristics of School Process and their Effects on Mathematics Achievement, *British Educational Research Journal*, 27(4), pp. 407-432.
- Vandenberghe, V. (2001) Efficacité-équité dans l'enseignement secondaire : quelle relation ? *Reflets et Perspectives de la vie économiques*, XL(4) (reproduced in *Problèmes économiques*, 2.753, 20 mars 2002, pp. 12-16).
- Willms, J.D. & Somers, M-A. (2001) Family, Classroom and School Effects on Children's Education outcomes in Latin America, *School Effectiveness and School Improvement*, 12(4), pp. 409-445.
- Zachary M-D. (2001) Structures scolaires et équité des systèmes d'enseignement en Europe, Wallonie, 66, pp. 93-102.