EURAMERICA Vol. 43, No. 4 (December 2013), 709-752
© Institute of European and American Studies, Academia Sinica
http://euramerica.org

Some Methodological Issues in Cross-National Educational Research —Quality and Equity in Student Achievement*

Robert M. Hauser
Division of Behavioral and Social Sciences and Education
National Research Council, U.S. National Academy of Sciences
500 Fifth St. N.W., Washington, DC 20001, USA
E-mail: rhauser@nas.edu

Department of Sociology, University of Wisconsin-Madison
1180 Observatory Drive, Madison, Wisconsin 53706, USA
E-mail: hauser@ssc.wisc.edu

Abstract

This article offers a critical review of measurement, analysis, and interpretation of international comparative data on socio-economic status and science achievement, as reported in the 2006 round of PISA, the Programme for International Student Assessment carried out by OECD, the

---

Received March 20, 2012; accepted October 23, 2013; last revised October 8, 2013
Proofreaders: Fang-Yi Chen, Hsih-Keng Yen, Chia-Chi Tseng

* Prepared for the NCES Conference on the Programme for International Student Assessment (PISA): What We Can Learn from PISA, Washington, DC, June 2, 2009. The research reported herein has been supported by the Research Triangle Institute, the Vilas Estate Trust at the University of Wisconsin-Madison, and a visiting fellowship at the National Research Council. I thank Fabian Pfeffer, Carl Frederick, and Min-Hsiung Huang for helpful comments and editorial suggestions. The opinions expressed herein are those of the author.

Organization for Economic Cooperation and Development. The OECD analysis overreaches by offering highly specific policy-related interpretations of correlations among socio-economic status and science achievement within and between schools among the large number of nations participating in PISA. The PISA measure of socio-economic status has no interpretable metric, and it is not strictly comparable across nations or years. The OECD analysis failed to identify independent effects of the components of socio-economic status or differences in those effects among nations or among areas of academic achievement. There were notable failures to identify and compensate for international differences in the reliability of the socio-economic data and to address the implications of international differences in the variability of socio-economic status. A primitive analysis of socio-economic effects on academic achievement, on between- and within-school differences in achievement, and differences in those effects among nations led to unwarranted inferences about national differences in academic achievement and factors affecting those differences. The article offers practical suggestions about ways to improve the analysis of the highly valuable PISA data.

Key Words: Academic achievement, contextual effects, Programme for International Student Assessment, school effects, socio-economic status

I. Introduction

The modern history of relationships among student background characteristics, school context, and academic performance begins with the Coleman-Campbell report of 1966, *Equality of Educational Opportunity* (Coleman et al., 1966).[1] As mandated by Section 402 of the Civil Rights Act of 1964 the Department of Health, Education, and Welfare commissioned a study of “the lack of availability of equal educational opportunities for individuals by reason of race, color, religion, or national origin in public educational institutions at all levels in the United States, its territories and possessions, and the District of Columbia.” What followed was a massive social and academic survey operation that covered almost every feature of American students, teachers, and schools.

The findings of the report were surprising:

(1) There was far greater equality in school resources and facilities than had been expected;
(2) Most of the variation in individual students’ academic performance occurred within, rather than between schools;
(3) Differentials in academic performance increased absolutely with grade level, while relative differences among social groups were maintained;
(4) The social and economic background of students contributed significantly to differentials in academic performance between schools and among students; and
(5) Neither school nor social background factors fully accounted for black-white differentials in academic performance.

---

[1] The Coleman-Campbell Report about the U.S. was soon followed by the parallel Plowden Report (1968) in Great Britain.

The Coleman-Campbell report was quickly suppressed—it was very soon out of print—for its findings did not suggest policy changes that would reduce inequalities in educational outcomes—especially those between blacks and whites.[2] The counter-intuitive findings of the report led both to a sustained and valuable critical literature, e.g., Mosteller and Moynihan (1972) and to a shift in the focus of educational research and policy from resources to outcomes. Perhaps the most valuable consequence of the report, however, was the understanding that educational differentials develop over time within schools, so an observational window limited to cross-sectional differences among individuals and schools could provide only the faintest hints about ways to reduce inequality in educational opportunities and outcomes.

On reading "Quality and Equity in the Performance of Students and Schools," Chapter 4 in *PISA 2006: Science Competencies for Tomorrow's World, Volume 1: Analysis* (Organisation for Economic Co-operation and Development [OECD], 2007c), I get the eerie feeling that nothing has changed in nearly 50 years. To be sure, the setting is quite different—narrower in content (science) and in age (15), yet far broader in geographic scope (30 OECD nations and 27 other "partner" nations). Moreover, there is real value in a comparative overview of social and economic differentials in academic performance between and within schools. The problem is that the chapter remains limited in its heavy reliance on a few key variables—the PISA index of economic, social, and cultural status (hereafter, the PISA SES Index) for individuals and schools and a composite measure of performance in science, along with the nation in which each student lived at the time of the study.[3]

---

[2] A copy of the report in PDF can be obtained as part of the documentation of the EEO data file at the Interuniversity Consortium for Political and Social Research (http://www.icpsr.umich.edu/cocoon/ICPSR/STUDY/06389.xml).

[3] To be sure, Chapter 4 also gives some consideration to immigrant status, non-native language use, and the valuation of science by students.

There is not much to be learned here beyond description, yet the text overreaches in its attempt to draw policy implications. Two examples of this stand out. First and most troublesome is the effort to impute specific meaning to effects of the socio-economic context of schools on students' performance in science. Such efforts have a long and undistinguished history, for average levels of socio-economic status in a student body may proxy for any number of causal processes or statistical artifacts (Hauser, 1969, 1970, 1972).

Second, the text attempts to adjudicate among future policies that might focus more directly on socially disadvantaged students or on low-performing students, based on the shapes of scatter plots of school levels of academic performance by individual and school values of the PISA SES Index (OECD, 2007c: 200-210, Figures 4.14a-e). Yet any such effort founders with the realization that "school" refers here only to the place of students at the time of their assessment and thus ignores the variety of lower-level schools in which their capacities were formed at younger ages. That is, in this respect the analysis ignores the prior influences of time and place in the process of schooling.

There are more problems than these, including incomplete explanations of procedures used in the study and statistics and statistical displays that are more likely to confuse or mislead the reader than to increase understanding of the size and sources of differentials in academic performance. On the other hand, the producers of the report deserve high praise for providing readily accessible spreadsheet sources of both the figures and data used throughout the report (OECD, 2007a, 2007b).

The elaboration and documentation of these observations is the substance of my discussion. In the following pages, I use Chapter 4 as an example of cross-national educational research. I briefly summarize the main features and findings of Chapter 4 and intersperse my discussion of each of them.

II. Measurement Issues: The PISA Indexes of Student Achievement and SES

A. The Index of Achievement in Science

What educational outcome or outcomes should be analyzed in a report of this kind? Chapter 4 immediately reports a choice and follows it consistently throughout:

> The overall impact of home background on student performance tends to be similar for science, mathematics and reading in PISA 2006. Therefore, to simplify the presentation and avoid repetition, this chapter limits the analysis to student performance in science, the focus area in 2006, and it considers the combined science scale (also referred to as, simply, the science scale) rather than examining the competency and knowledge area scales separately. (OECD, 2007c: 170)

This choice has strong implications. First, if it is truly the case that it does not matter whether one analyzes science, mathematics, or reading, and it does not matter whether one analyzes science competency and knowledge separately or jointly, then the analysis is truly not about science, but about some very general academic performance construct. In that case, one might ask whether it is really necessary to have developed all of the academic performance measures covered in PISA 2006—and thus burdened students with all of those assessments—and, also, why the reported analyses are couched in terms of "the science scale" rather than, simply, "academic achievement."

Second, even if it were the case that each of the separate academic achievement constructs responds similarly to school differences and to social, economic, and cultural background in the aggregate, one should ask whether the same holds across countries and regions. There are two aspects to this question, whether academic performance appears to be one-dimensional in relation to social background and schools and, even if it is one-dimensional, whether each of the academic achievement constructs responds similarly to social background and school factors in each country.

Third, even if performance on the composite science scale were fully representative of academic achievement in the same way in every country, an analysis using that variable alone has less statistical power than an analysis using more (or all) of the measures of academic achievement.[4] If one created an overall composite measure of academic achievement, it would undoubtedly be more reliable and thus more highly correlated with economic, social, and cultural status than is the composite science scale, and the measures of between- and within-school variance would also change.[5] But it would be both more powerful and more informative to estimate a multiple-indicator multiple-cause (MIMIC) model of academic achievement (Hauser, 1973; Hauser & Goldberger, 1971, 1975; Jöreskog & Goldberger, 1975). With such a model one could explicitly test whether the several academic achievement constructs respond similarly to variation in economic, social, and cultural background and variation among schools; one could also explicitly test whether those relationships vary among countries.

To be sure, one might legitimately ask why an analysis of achievement in science should be transformed into an analysis of overall academic achievement. But that question is begged in Chapter 4. By declaring that it is appropriate to ignore the differences among the several academic and scientific constructs, the chapter can only leave the reader wondering whether it is really about achievement in science or about a proxy for overall achievement.

---

[4] That is, standard errors of the estimated effects of explanatory variables would be smaller.

[5] If the measure of academic achievement were more reliable, within-school variance would decrease relative to between-school variance.

B. The PISA SES Index

Average values of the PISA SES Index account for half or more of the between-school variance in most nations, while the SES Index accounts for a small fraction of the variance in performance within schools in every country. This is taken at face value throughout the analysis, yet it raises serious questions of substance and method.

Aggregated to the school level, the PISA SES Index is doubtless highly reliable, but—as noted above—it is not at all clear what it means. School levels of socio-economic status are typically correlated positively with everything else that might be good about a school and negatively with everything else that might be bad about a school. Thus, the variance explained in the bivariate regression of average school performance on average school SES just doesn't tell us very much.

Individual student values of the SES Index are undoubtedly lower in reliability than school values of the same variable, and they are still less reliable within schools than in the total population (precisely because school-level reliability is higher). Thus, one should expect that within-school regressions of academic performance on the PISA SES Index underestimate the true effect of social and economic background. It is likely, also, that the amount of downward bias in the estimates varies across nations and among population groups within nations. In the U.S., for example, student reports of parents' educational attainments in the National Educational Longitudinal Study of 1988 are more reliable among whites than among blacks or Hispanics. Annex A1 of the Analysis report states, "The reliability of the index ranged from 0.52 to 0.80," but it provides no reliability estimates for specific countries (OECD, 2007b: 333, Annex A1). The reported range suggests the need for extreme caution in interpreting and comparing within school analyses of effects of the PISA SES Index across countries.

There are problems, also, with the construction of the PISA index of economic, social, and cultural status and with its comparability across years (2000, 2003, and 2006). In 2006, the index was based on separate IRT scaling of items that were common to each nation, plus 3 items that were potentially unique to each nation. The common items were the higher occupational status of parents on the International Socio-Economic Index of Occupational Status (Ganzeboom, De Graaf, & Treiman, 1992; Ganzeboom & Treiman, 1996, 2003), the higher educational level of parents, and an index of home possessions (including "a desk to study at, a room of their own, a quiet place to study, a computer they can use for school, an educational software, a link to the Internet, their own calculator, classic literature, books of poetry, works of art (e.g. paintings), books to help with their school work, a dictionary, a dishwasher, a DVD player or VCR, the number of cellular phones, televisions, computers, cars and books at home, and three other country-specific items"). Since the last collection of items was explicitly intended as a proxy for wealth, and the report says that "The rationale for the choice of these variables was that socio-economic status is usually seen as being determined by occupational status, education and wealth," it is not clear why the index is labeled as "cultural" rather than, simply, "socio-economic" (OECD, 2007b: 333, Annex A1). Finally, the index values were weighted in some relation to a principle component analysis—it is not clear what the weights were or whether they were common or unique to each nation—and the resulting index values were standardized to have a mean of zero and a variance of one in the combined OECD countries.

The PISA SES Index has several fatal flaws. First, it is not the same in each nation. The differences may not be large, but they are real. If the differences are small, then why not eliminate them entirely and trade minor differences in validity for strict comparability? And we do not know how large or consequential the differences are. If they are large, then cross-national comparisons of the within- and between-school regressions are not valid.

Second, similar observations hold for differences in the content of the PISA SES Index across years of the study, which are described in Annex A1. The report states that the correlation of the index between 2003 and 2006 is "very high (R of 0.96)." It does not state what the units were over which the index values were correlated. If the units were countries, a high correlation would provide no information about comparability of content, reliability, or metric from one year to the next. The lack of intertemporal comparability in the PISA SES Index is consequential because Chapter 4 offers comparisons of the effect of socio-economic background across years of the study.

Third, the report tells us that, "Since these various aspects of socio-economic background tend to be highly interrelated, most of the remainder of the report summarizes them in a index, the PISA index of the economic, social and cultural status of students, even though separate data for these are provided in the accompanying data tables . . . ." (OECD, 2007c: 174). This is an unpersuasive rationale for failing to analyze effects of the components simultaneously. Typical correlations among such variables are no larger than 0.6—having about a third of variance in common—and in large samples of students and schools, like those of PISA, comparisons of effects will be reliable.

Fourth, the index construction process traded the actual observable metrics of the variables—ISEI values, equivalent years of schooling, and (presumably) counts of various possessions—for the content-free metric of standard deviation units. Thus, the analysis does not tell us what difference any of the component variables actually makes in academic achievement, whether some of the components of the index either dominate or have negligible effects, or whether the effects of the components of the PISA SES Index vary among countries. Surely, in a study of so many thousands of youth the gain of a few degrees of freedom in a regression equation is not worth the loss of information about the effects of actual parental and family characteristics. Further, the weights of the index components were—incorrectly in my judgment—made using information about the relationships among the components, rather than their relationships with academic performance. To be sure, minor variations in the weights of the components are unlikely to have much effect on the overall **predictive** power of the index, but analyses using the index provide no information about the relative importance of the components. What aspects of socio-economic background really matter? Is it parents' educational attainments? Their occupational standing? What types of possessions actually make a difference in science achievement? For the same reason, the PISA analysis tells us nothing about cross-national differences in the effects of the components of the SES index. Again, estimation of a multiple-group MIMIC model would address these questions (Hauser & Goldberger, 1971; Jöreskog & Goldberger, 1975).

Chapter 4 wisely reports extensive analyses of the effects of immigrant status and non-native language on science achievement. Overall, first-generation (foreign-born) students are about a year and a half behind natives, but there are wide variations in the differentials among countries. There does not appear to be a correlation between the share of first-generation students in the population and the differential in performance; rather, my reading of the data is that the main source of cross-national differentials in the handicap of immigrant status is the cultural proximity of the immigrant and native populations—as often expressed in comparisons between the languages used at home and in the assessment. For example, in Canada, Australia, Macao-China, and Jordan, there are negligible differences in performance among foreign-born, second-generation, and other native-born students (OECD, 2007c: 177). On average, the PISA SES Index accounts for 36 percent of the differential between native students and those with "an immigrant background," and for 52 percent of the difference between native students and "Students with an immigrant background who speak a language at home that is different from the language of instruction" (OECD, 2007a: 121, Table 4.3c, reproduced in the Appendix).[6] Thus, economic, cultural, and social background does not account for the differential between immigrant and non-immigrant populations. It would be instructive to run a similar analysis, comparing native students with immigrant students who speak the language of assessment at home: Would the PISA SES Index then account for differences in science achievement? Or is immigration *per se* an educational handicap? And do the answers to these two questions differ among countries?

Chapter 4 also reports important findings about what does not explain immigrant-native differentials in science achievement. Although immigrant youths attend schools with lower values on the PISA SES Index than native youths in almost every country—often by half a standard deviation or more—there are only a few countries where immigrants attend schools with lower quality educational resources, higher student teacher ratios, or teacher shortages (OECD, 2007c: 179, Figure 4.3, reproduced in the Appendix). Moreover, immigrant students report levels of engagement with science—on several indices—that are higher or comparable to those of native students (OECD, 2007c: 180).

III. Analytical Issues: Academic Achievement and Socio-Economic Background

A. Regression Analysis of Science Achievement

Well over half of Chapter 4 is devoted to regression analyses of science achievement on the PISA SES Index, overall, between schools, and within schools. One 16-page section focuses primarily on the overall strength of the relationship between SES and achievement and secondarily on comparisons of between- and

---

[6] Both of these estimates, computed by the author, refer to subsets of OECD and partner countries for which relevant data are presented in Table 4.3c.

within-school effects of SES (pp. 181-196). The final 13 pages of the 41-page chapter are devoted to comparisons of total, between-school, and within-school regressions among the OECD and partner nations (pp. 198-210), and that section includes most of the policy recommendations in the chapter. Those two sections are punctuated by a brief passage about relationships between science achievement and parents' reports about their students and schools in 16 countries where such data were collected directly from parents (pp. 196-198).

The report on regression analyses starts well with an exposition of the zero-order regression of the combined science index on the PISA SES Index. There is, of course, a positive gradient in science achievement by socio-economic status; the gradient is close to linear across the observed range of the PISA SES Index; and there is a good deal of scatter of individual student achievement levels about the regression line (OECD, 2007c: 183, Figure 4.5, reproduced in the Appendix). This bivariate regression accounts for 20.2 percent of variance in science achievement among all OECD students and an average of 14.4 percent of variance across the 30 OECD countries (OECD, 2007c: 184, Figure 4.6, reproduced in the Appendix). The difference between these two statistics reflects the fact that there is variation in socio-economic levels among the countries.[7]

B. Science Achievement and the Effect of Socio-Economic Background

Figure 4.10 (reproduced in the Appendix) plots mean

---

[7] Here, and in other parts of the report, there is an arbitrary distinction between reported findings for the aggregate or average of OECD countries and findings for the partner countries, which are always reported separately, but never aggregated or averaged. There is perhaps a political or bureaucratic rational for this practice, but it surely detracts from the value of the analyses for science and policy.

national scores on the science composite by the percentage of variance explained by the PISA SES Index. The text accompanying that figure states:

> Figure 4.10 highlights that countries differ not just in their overall performance, but also in the extent to which they are able to moderate the association between socio-economic background and performance. PISA suggests that maximising overall performance and securing similar levels of performance among students from different socio-economic backgrounds can be achieved simultaneously. The results suggest therefore that quality and equity need not be considered as competing policy objectives. (OECD, 2007c: 190)

This discussion is problematic because neither axis of the diagram is well-chosen. In the context of the analysis, adjusted rather than observed mean levels of achievement should be used to indicate the quality of science education. In an ideal situation, one would base such an adjustment on a full model of achievement in science—including many more background, parental, and student characteristics beyond economic, social, and cultural status. At the least, the adjustment should take account of national differences in the PISA SES Index.[8] Then, as just explained, the second axis of the graph should be the variance about the regression line, indicating (inversely) how academic performance follows socio-economic status.[9] This relationship is shown in Figure 1. In that figure, unlike Figure 4.10, the horizontal and vertical lines mark the average values of performance in science and of error variance for all 55 countries, not just the OECD countries.

There is essentially no relationship between observed means

---

[8] To be sure, the text recognizes the import of socio-economic background for science achievement, and a consistent analysis would have taken that into account in the construction of Figure 4.10.

[9] The rationale for preferring absolute error variance to the percentage of variance explained (or unexplained) is demonstrated in the next section.

and percentages of variance explained in Figure 4.10 (r=-0.04). Thus, the discussion of this figure in the text points to examples of four types of nations, which appear in roughly equal numbers representing the four possible combinations of achievement in science and fit of the regression model. In contrast, there is a moderate relationship between the adjusted means and error variances in Figure 1 (r=0.34). That is, high performing countries tend to have greater equality of opportunity, in the sense that the scatter of individual observations about the regression line is greater, while low performing countries tend to have less equality of opportunity, less dispersion of individual observations about the regression line. Again, the position of nations on the vertical axis of Figure 1 (science achievement) is similar to that in Figure 4.10, with exceptions noted below, but as shown in Figure 2, there is very little relationship between the percentages of explained variance and the variances of observations about the regression of science achievement on the PISA SES Index.[10]

Figure 1 thus offers a very different picture from Figure 4.10 of the relationship between educational opportunity—lack of fit to the regression line—and national levels of academic performance in science. For example, in Figure 4.10, the United States appears near the center, slightly below the OECD average in science achievement and somewhat above average in percentage of variance explained. In Figure 1, the U.S. is slightly above average in science achievement (for all nations) and far above average in equality of educational opportunity, for there is a relatively high level of scatter of science achievement about the values predicted from the PISA SES Index. Why is a high percentage of variance explained in the U.S.? The variation in the PISA SES Index in the

---

[10] Since the X-axis of Figure 4.10 goes from high to low percentages of explained variance, while the X-axis of Figure 1 goes from low to high estimates of error variance, the spatial representation of effects in the two diagrams is the same. The strength of the relationship between the PISA SES Index and science achievement declines from left to right.

U.S. is the same as the OECD average, but the regression of science achievement on the PISA SES Index (49) is almost 25 percent steeper than the OECD average (40) (OECD, 2007a: 123-124, Table 4.4a, reproduced in the Appendix). Thus, the U.S. performs badly on one indicator of educational opportunity (the regression coefficient), but far better on another indicator, the scatter of individual student achievement about values predicted from socio-economic background. Israel appears as slightly below average both in science achievement and in the impact of socio-economic background in Figure 4.10, but Figure 1 shows Israel as far below average in the impact of socio-economic background. Bulgaria appears as below average in science achievement in both figures, but it is depicted as having very high dependence of science achievement on social background in Figure 4.10 and moderately low dependence of achievement on background in Figure 1. Plainly, it is possible to add to these examples of divergent findings.

One might imagine adding the regression coefficient of science achievement on the PISA SES Index as a third dimension of the display. In this way both aspects of the dependence of science achievement on SES would be represented, but the findings would not be confounded by statistically (though not substantively) irrelevant differences in the variability of socio-economic background. Unfortunately, it is not possible to distinguish between the effects of these two variables (the regression slope and the error variance) on mean country achievement levels. The correlation between the two is moderately high ($r=0.70$), while their correlations with mean science achievement are similar (0.35 and 0.34, respectively). That is, the error variances are **larger** in countries with steeper slopes of science achievement on the PISA SES Index. In a regression analysis of the adjusted means, the slope coefficient dominates, but there is actually no significant difference between the effects of the two explanatory variables. In other words, data are not available for a large enough number of countries to identify significant differences between the associations of the achievement-SES slope and the error variance with adjusted country means.

IV. Cross-National Differences in Achievement

Figure 4.6 (OECD, 2007c: 184, reproduced in the Appendix) offers a set of statistics about individual nations that apparently are intended to inform readers about the slope and strength of association between achievement and socio-economic status, about the extent to which socio-economic status differences may account for cross-national differences in science achievement, and about socio-economic differences among the several student populations. Unfortunately, this table and the accompanying discussion provide a blurred picture of cross-national differences in socio-economic effects on achievement and of the role of cross-national differences in socio-economic status in accounting for mean differences in achievement among countries.

The first two columns of Figure 4.6 report mean levels of achievement on the combined science scale as observed and adjusted for mean socio-economic differences among countries: “Mean score if the mean ESCS would be equal in all OECD countries.” However, neither Figure 4.6, nor the accompanying text, nor the source table describes how this regression standardization was actually carried out. In principal, it should be possible to reproduce the second column of Figure 4.6 from the source table (OECD, 2007a: 123-124, Table 4.4a, reproduced in the Appendix). That is, the table contains the observed and adjusted mean achievement scores and the mean values of the PISA SES Index (labeled ESCS in Figure 4.6 and Table 4.4a), along with the estimated regressions of achievement on ESCS.[11] However,

---

[11] Mean levels on the PISA SES Index would be more appropriate here than the percentages of students falling in the bottom 15 percent of the overall

Table 4.4a contains three different regression coefficients that might have been used to adjust mean national levels of achievement: the overall regression in OECD countries, the average regression within OECD countries, and the estimated regression within each country. None of these regression coefficients exactly reproduces the adjusted means reported in the second column of Figure 4.6 (and reported in somewhat different form in Figure 4.7, OECD, 2007c: 187, reproduced in the Appendix).

In my judgment, the closest approximation to the reported values uses the estimated regression coefficients for each country, and I assume this is the choice made in the analysis. This is the least desirable choice among the three alternatives. That is, it confounds the effect of mean differences among countries on the PISA SES Index with the effect of statistical interactions among countries in the association between socio-economic status and science achievement. To be sure, the report literally answers the question, "What is our best estimate of the mean level of achievement when the mean level of SES is the same as that for all OECD countries?" but it does so in a way that invalidates comparisons of the adjusted means across countries.[12] A better choice would have been the average within-country regression—and preferably that for all countries, not merely the OECD countries. All the same, in analyses reported here, I used the reported average regression for OECD countries to adjust the observed mean differences among countries.

In many cases, the country-specific regression was similar to the average within-country regression, so there was little difference between the two adjusted means. Indeed the overall correlation

---

distribution—reported in the last column of Figure 4.6—because the former enter directly into the adjustment of country means.

[12] This is more problematic in the partner countries than in OECD countries, for mean levels of the PISA SES Index are almost half a standard deviation lower in the partner countries than in the OECD countries.

between the two versions of the adjusted means is extremely high, 0.987. However, even with that large correlation, there were notable differences between the two sets of estimates. For example, in Mexico the observed mean was 410, and the adjusted mean was 435, but it should have been 449. In Turkey, Azerbaijan, Brazil, Colombia, and Thailand the adjusted means as reported were 11, 13, 11, 17, and 17 points lower than they would have been using the average within-country regression. And there are yet more extreme cases: Deviations of 29, 24, and 25 points in Indonesia, Macao-China, and Tunisia. What all of these countries have in common are relatively low slopes of achievement on the PISA SES Index and below average levels of the SES Index.

The problems with Figure 4.6 do not end here. The third column of the table reports "percentage of explained variance in student performance" (R^2), which is described as a measure of the strength of the association between the PISA SES Index and achievement on the combined science scale. The percentages of variance explained are compared among countries in the text:

> On average across OECD countries, 14.4% of the variation in student performance in science within each country is associated with the PISA index of economic, social and cultural status. This figure is significantly higher than the OECD average in Luxembourg, Hungary, France, Belgium, the Slovak Republic, Germany, the United States, New Zealand and the partner countries Bulgaria, Chile, Argentina and Uruguay. (OECD, 2007c: 185)

Unfortunately, this measure does not yield valid cross-national comparisons.[13] This follows from the definition of total variance in the regression model. The total variance in science achievement ($\hat{\sigma}_y^2$) has two components, explained and unexplained variance.

---

[13] There is also reason to be concerned about the robustness of the IRT model, especially with regard to the dispersion of test scores in less developed nations, but I have not pursued that issue here (Brown, Micklewright, Schnepf, & Waldmann, 2007: 643).

The former component is equal to the product of the variance in the regressor ($\hat{\sigma}_x^2$)—in this case the PISA SES Index—and the square of the regression coefficient of achievement on SES ($\hat{\beta}^2$). The second component is the error variance ($\hat{\sigma}_\varepsilon^2$), a measure of the scatter of observations about the regression line. Formally, that is:

$$R^2 = \hat{\beta}^2 \hat{\sigma}_x^2 / (\hat{\beta}^2 \hat{\sigma}_x^2 + \hat{\sigma}_\varepsilon^2)$$

For a fixed amount of scatter about the regression line, the percentage of variance explained will vary directly with both the absolute value of the regression coefficient and the amount of variance in the regressor. To be sure, the regression coefficient of science achievement on SES is an important indicator of the impact of social background, but the variance of SES is not. In fact, Chapter 4 reports measures of the variability in SES in Figures 4.8 and 4.9 (OECD, 2007c: 188, reproduced in the Appendix), but it does not use this information to refine its findings about the strength of association between socio-economic status and academic achievement.[14]

It would be more appropriate to compare the **accuracy** with which the PISA SES Index predicts achievement in science across countries by tabulating the actual variance about the regression line in each country ($\hat{\sigma}_{j\varepsilon}^2$), where the subscript *j* refers to a specific country. That is, the variance about the regression line is an inverse measure of goodness-of-fit; the larger the error variance, the less closely is social background related to science achievement.

Fortunately, the PISA 2006 Data volume provides enough information to calculate the error variances for 55 of the 57 countries (OECD, 2007a: 96, Table 4.1a; 123, Table 4.4a,

---

[14] Figure 4.8 is evidently mislabeled. Both Figure 4.8 and Figure 4.9 are labeled as reporting the interquartile range of the distribution of the PISA SES Index, but the former appears to report the location of the 5th and 95th percentiles, rather than the 25th and 75th percentiles.

reproduced in the Appendix).[15] There is scant relationship between the percentage of variance explained and the error variance about the country-specific regression lines. The correlation is just 0.13. As shown in Figure 2, the relationship between the two quantities is roughly linear, but very weak. By way of example, Israel ranks 19th from the bottom in the percentage of variance explained (10.9 percent), but the variance about the regression line is larger than in any other country. That is, there is great variation in science achievement in Israel that cannot be explained by socio-economic status. On the other hand, Indonesia is similar to Israel in the percentage of variance explained, but the variance about the regression line is only 40 percent as large as in Israel. Only one country, Azerbaijan, has less variance about the regression line than Indonesia. That is, in Indonesia and Azerbaijan, there is little variation in science achievement that cannot be explained by socio-economic status. In sum, the measure of strength of relationship used throughout Chapter 4 is utterly misleading with regard to comparisons among nations in the extent to which science achievement varies independently of socio-economic status. To be sure, the simple regression slope of science achievement on social background is also a very important indicator of educational opportunity, but I follow the text of Chapter 4 in focusing on the fit of the regression line in the following discussion.

V. Within-School and Between-School Regressions

A. Between- and Within-School Variance

Both among and within the 57 nations covered by PISA 2006, there are large differences in the organization of the schooling process:

---

[15] Data are missing for France and Qatar.

> Some countries have comprehensive school systems with no, or only limited, institutional differentiation. They seek to provide all students with similar opportunities for learning by requiring each school and teacher to provide for the full range of student abilities, interests and backgrounds. Other countries respond to diversity by grouping students through tracking or streaming, whether between schools or between classes within schools, with the aim of serving students according to their academic potential and/or interests in specific programmes. And in many countries, combinations of the two approaches occur. Even in comprehensive school systems, there may be significant variation in performance levels between schools, due to the socio-economic and cultural characteristics of the communities that are served or due to geographical differences (such as between regions, provinces or states in federal systems, or between rural and urban areas). Finally, there may be differences between individual schools that are more difficult to quantify or describe, part of which could result from differences in the quality or effectiveness of the instruction that those schools deliver. As a result, even in comprehensive systems, the performance levels attained by students may still vary across schools. (OECD, 2007c: 171-173)

Hence, one would expect to find cross-national differences in the variation in achievement, both within and between schools. This is masterfully documented in Figure 4.1 (OECD 2007c: 171, reproduced in the Appendix), which shows between- and within-school variance components of the composite science scale for each nation. These are expressed in relation to the average (total) variance in student performance in OECD countries, and within each of these two components, there is a further visual distinction between variance that is explained and unexplained by the PISA SES Index. Entries are ordered by the size of between-school variance components, and OECD countries are distinguished from partner countries.

The percentages of variance between schools vary dramatically across countries, from 69.6 percent in Bulgaria to 4.7 percent in Finland, each expressed relative to the average (total) variance in OECD countries (OECD, 2007a: 96, Table 4.1a, reproduced in the Appendix). This could be misleading because of the choice of reference value; the ordering of nations would vary somewhat if the reference value had been the percentage of between-school variance in each nation. For example, the U.S. appears in the middle of the pack with 29.1 percent of the OECD average variance between schools, but the actual percentage of between school variance in the U.S. is just 23.2 percent. Similarly, in the United Kingdom, the respective percentages are 23.5 and 18.9 percent. Inversely, the reported between-school variance in Hungary is 60.5 percent, while the actual figure is higher, 70.4 percent. The story is all the more confusing because partner countries do not contribute to the reference value of total variance. At the positive extreme, the percentage listed for Bulgaria is 69.6 percent when the actual value is 55.0 percent of the variance between schools. At the negative extreme, the percentage listed for Azerbaijan is 17.9 percent when the actual share of between school variance is 51.8 percent; this large disparity occurs because the total variance in science performance in that nation is unusually small. Thus, while the figure makes it possible to compare total variances to the OECD average, it distorts the shares of variance between and within schools in the several nations.

One of the side effects of the choice of reference values in Figure 2 is that it provides average shares of within- and between-school variance that are descriptively correct, but logically impossible: 68.1 and 33.0 percent, respectively. When data for each country are used, the average percentages across the 30 OECD countries are 66.9 percent within schools and 33.1 percent between schools. Across all 57 countries, they are 63.9 percent and 36.1 percent. The striking thing about these estimates is the extent to which individual differences among students dominate the decomposition, even when within- and between-country effects enter the picture.

There is a broader issue about the attention given here to between- and within-school variance. Suppose one were looking at a single school system—or even a state or national system—in which the assumption was that educational resources, opportunities, and outcomes were similar across the individual schools in the system. Then, a finding of substantial between- school variation in outcomes would carry a clear message, that the assumption of equality was wrong. But in a world-wide array of national systems, where there is clear acknowledgment that the organization and processes of schooling vary widely—and in a study that focuses on students of the same age, but multiple grade levels—there is much less information in the fact of large variations in outcomes among schools. Indeed, it is perhaps surprising that individual differences in academic performance remain far larger than variations among schools.

B. What Can We Learn from Within-School and Between-School Regressions?

Following the discussion of Figure 4.10, the text of Chapter 4 turns to two seriously flawed analyses, each based on comparisons of within- and between-school regressions of science achievement on social background. Figure 4.11 (OECD, 2007c: 192, reproduced in the Appendix) shows the total, within-school, and between-school estimates of the regression of science achievement on the PISA SES Index in each nation for which such data are available with few exceptions—Finland, Iceland, Poland, and, to a lesser extent, Norway and Spain—the estimated between-school regressions are much steeper than the within-school regressions. On average, in the OECD countries, the ratio of the two is roughly 3 to 1, and in many cases the ratios are much larger. Chapter 4 reports, "Socio-economic differences at student levels are much less predictive of performance than the schools' socio-economic context." The text goes on:

> Not all of the contextual effect is attributable to peer group effects, but socio-economic advantage of students and their families often also goes along with a better learning environment and access to better educational resources at school. Also, the manner in which students are allocated to schools within a district or region, or to classes and programmes within schools, can have implications for the contextual effect, in terms of the teaching and learning conditions in schools that are associated with educational outcomes. A number of studies (e.g. Baker et al., 2002) have found that schools with a higher average socio-economic status among their student intake are likely to have: fewer disciplinary problems, better teacher-student relations, higher teacher morale, and a general school climate that is oriented towards higher performance. Such schools also often have a faster-paced curriculum. Talented and motivated teachers are more likely to be attracted to schools with higher socio-economic status and less likely to transfer to another school or to leave the profession. Some of the contextual effect associated with high socio-economic status may also stem from peer interactions that occur as talented students work with each other . . . . For example, the parents of a student attending a more socio-economically advantaged school may, on average, be more engaged in the student's learning at home. This may be so even though their socio-economic background is comparable to that of the parents of a student attending a less-privileged school. (OECD, 2007c: 195)

To be sure, the text goes on to say, "the estimated contextual effects . . . are descriptive of the distribution of school performance, and should not necessarily be interpreted in a causal sense," but the text again turns a corner:

> In any attempt to develop education policy in the light of the above findings, there needs to be some understanding of the nature of the formal and informal

> selection mechanisms that contribute to between school socio-economic segregation and the effect of this segregation on students' performance. In some countries, socio-economic segregation may be firmly entrenched through residential segregation in major cities, or by a large urban/rural socio-economic divide. In other countries, structural features of the education system tend to stream or track students from different socio-economic contexts into programmes with different curricula and teaching practices. The policy options are either to reduce socio-economic segregation or to mitigate its effects. (OECD, 2007c: 196)

Chapter 4 immediately goes on to contrast the import of the two slopes in a fundamentally misleading way. Figure 4.12 (OECD 2007c: 194, reproduced in the Appendix) compares the "effects" on achievement of a one-half standard deviation change in a student's PISA SES Index:

> The lengths of the bars in Figure 4.12 indicate the differences in scores on the PISA science scale that are associated with a difference of one-half of an international standard deviation on the PISA index of economic, social and cultural status for the individual student . . . and for the average of the student's school . . . . One-half a student-level standard deviation was chosen as the benchmark for measuring performance gaps because this value describes realistic differences between schools in terms of their socio-economic composition: on average across OECD countries, the difference between the 75th and 25th quartiles of the distribution of the school mean PISA index of economic, social and cultural status is 0.63 of a student-level standard deviation. (OECD, 2007c: 193-194)

The problem with this comparison is that one-half a student standard deviation corresponds roughly to the difference between the 40th and the 60th percentiles of the distribution of student performance.[16] In other words, the text contrasts a large gap in the distribution of school-level SES with a much smaller gap in the distribution of individual-level SES and thus, unnecessarily, exaggerates the import of the obviously large difference between the two regressions.

Why are between-school regressions typically steeper than within-school regressions? Chapter 4 does not ask this question. Rather, it assumes that it has a clear sociological interpretation and strong policy implications, even while suggesting, rather ingenuously, that it is not proposing a causal interpretation. There are several reasons, some of which are addressed in Chapter 5, and others not. First, the regression model is woefully incomplete. By no means is socio-economic status the sole source of individual differences in academic achievement in science (or any other subject). Had the analysis included other social and psychological background characteristics, both the within- and between-school regressions of achievement on the PISA SES Index would have changed. Second, individual student values of the PISA SES Index are necessarily less reliable than (aggregate) school means. This effect is even larger when within-school differences are analyzed. This contributes a downward bias to the within-school SES-achievement regressions relative to the between-school regressions. Third—and this is the subject of Chapter 5—having observed differences in the between- and within-school regressions, one should immediately ask what school-level variables may explain the association between average school SES and achievement.[17] In short, Chapter 4 fails to grapple directly with either

---

[16] This assumes that the distribution of student achievement scores is approximately normal.

[17] I have not read Chapter 5 closely, but my initial impression is that it focuses heavily on the extent to which specific school organizational factors and resources affect achievement net of student intake, but it does not directly address the extent to which variation between schools in organization and resources explains the relationship between the socio-economic background composition of schools and their achievement in science.

methodological or substantive explanations for the observed differences in simple, bivariate between- and within-school regressions of science achievement on the PISA SES Index.

The final section of Chapter 4 turns to policy implications of its findings, largely focusing on the question whether it is better to direct reforms in science education to students from disadvantaged backgrounds or to low-performing students. Guidance in this matter is presumed to follow from comparisons of between- and within-school regressions and school-level residuals from those regressions, which are presented at length (OECD, 2007c: 198-210). In light of the preceding discussion, I am doubtful that these analyses are valid. The PISA SES Index obscures as much as it illuminates. Its content and reliability vary from country to country, and these affect the estimated slopes. If these matters were resolved, a simple model regressing science achievement on socio-economic background, could not possibly provide a sound or complete guide to the proximate sources of variation in students' achievements in science or in school differences in those achievements. And the effort is further compromised by the fact that achievement at age 15 represents the cumulative impact of schooling processes over about a decade of each student's life.

This is not to suggest that either PISA itself or the analysis of Chapter 4 lacks value for science or policy. The point of my observations is that Chapter 4 raises many more questions than it has answered. Some of these can be addressed by following the suggestions I have made throughout this discussion for additional or more refined analyses. Others could be answered by rearrangements or extensions of the analyses in Chapter 5. PISA is such a rich resource that, in my judgment, such additional analytic investments are worthwhile.

Figure 1 Performance in Science (SP) by Error Variance in Background Regression

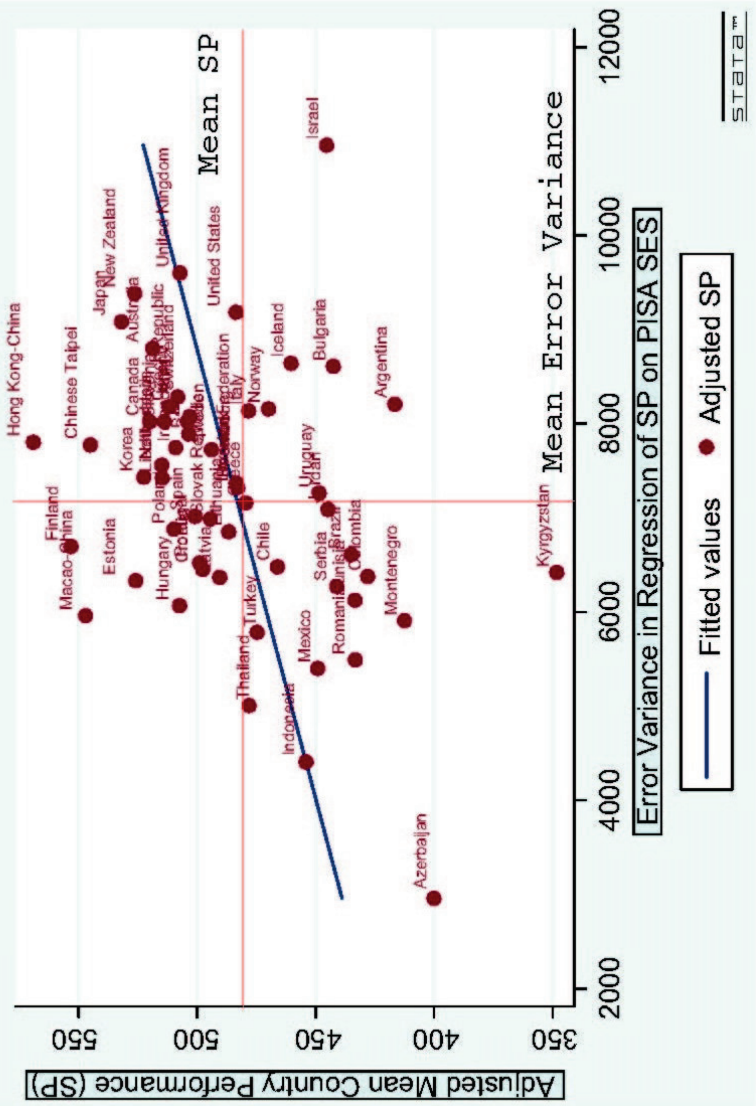
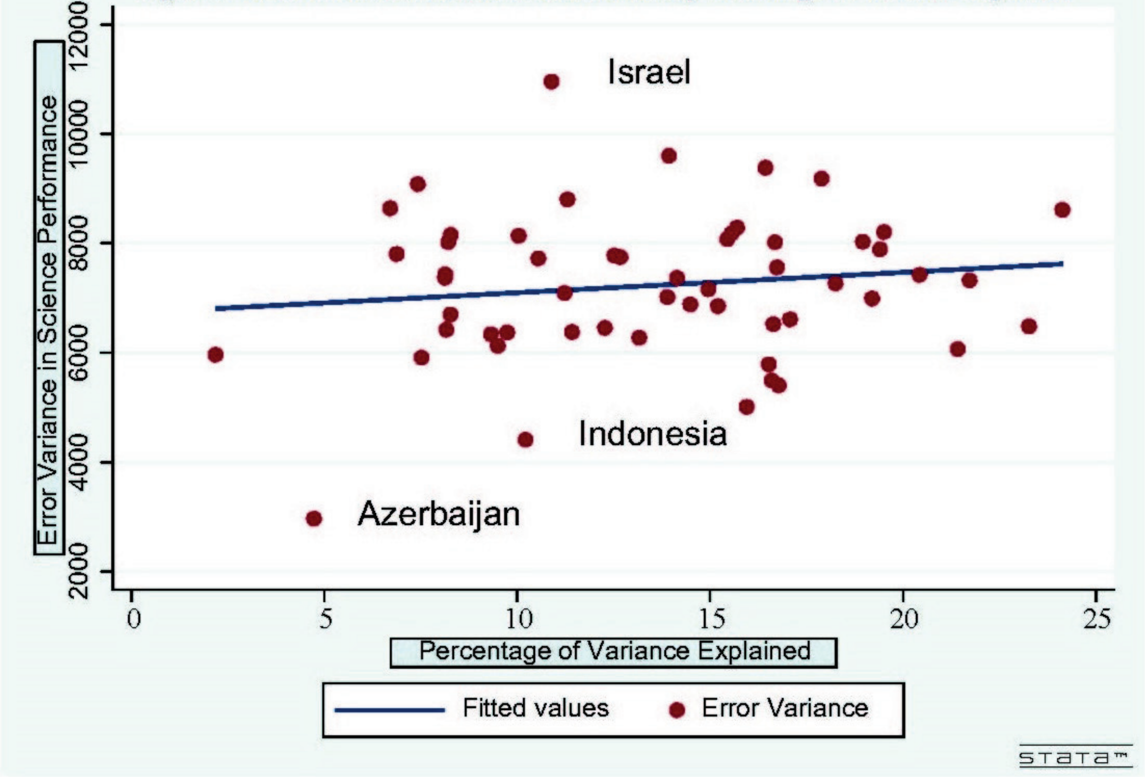


Figure 2 Error Variance in Science Achievement by Percentage of Variance Explained



APPENDIX

Table 4.1a Between-school and within-school variance in student performance on the science scale in PISA 2006

	Total variance in SP[2]	Variance expressed as a percentage of the average variance in student performance (SP) across OECD countries[1]											Total variance between schools expressed as a percentage of the total variance within the country[5]
		Total variance in SP expressed as a percentage of the average variance in student performance across OECD countries[3]	Total variance in SP between schools[4]	Total variance in SP within schools	Variance explained by the PISA index of economic, social and cultural status of students		Variance explained by the PISA index of economic, social and cultural status of students and schools		Variance explained by students' study programmes		Variance explained by students' study programmes and the PISA index of economic, social and cultural status of students and schools		
					Between-school variance explained	Within-school variance explained	Between-school variance explained	Within-school variance explained	Between-school variance explained	Within-school variance explained	Between-school variance explained	Within-school variance explained	
OECD													
Australia	9 926	110.6	19.8	91.1	7.8	4.3	12.5	4.4	1.9	3.9	13.0	7.9	17.9
Austria	9 551	106.5	60.7	50.7	7.9	0.6	40.1	0.6	45.2	0.3	49.5	0.8	57.0
Belgium	9 791	109.1	57.0	53.0	11.7	2.0	40.7	2.0	45.4	12.7	50.6	13.3	52.3
Canada	8 743	97.5	17.9	79.3	4.3	3.2	7.1	3.2	2.0	3.2	7.2	5.9	18.4
Czech Republic	9 687	108.0	62.4	55.9	12.7	1.7	43.5	1.8	50.2	0.4	52.2	2.0	57.8
Denmark	8 580	95.6	14.8	82.0	6.0	8.1	8.2	8.3	1.6	0.1	8.6	8.4	15.4
Finland	7 301	81.4	4.7	76.7	1.2	5.5	1.3	5.5	0.0	0.0	1.3	5.5	5.8
France	w	w	w	w	w	w	w	w	w	w	w	w	w
Germany	9 908	110.4	66.2	50.8	11.6	1.4	49.4	1.4	56.0	2.0	58.1	3.3	59.9
Greece	8 420	93.9	48.5	55.1	11.3	1.7	29.1	1.8	37.3	0.0	41.7	1.7	51.7
Hungary	7 720	86.1	60.5	38.5	9.4	0.2	47.5	0.2	46.2	0.0	51.6	0.3	70.4
Iceland	9 263	103.2	9.3	95.4	0.1	6.4	0.2	6.3	1.8	0.3	2.0	6.6	9.0
Ireland	8 871	98.9	16.9	82.6	7.4	4.9	11.4	5.0	1.1	3.6	11.4	8.3	17.0
Italy	9 045	100.8	52.6	51.8	4.8	0.4	27.6	0.5	26.4	0.1	31.9	0.5	52.1
Japan	9 812	109.4	53.0	59.4	2.9	0.1	29.0	0.1	9.7	0.0	30.2	0.1	48.5
Korea	8 093	90.2	31.8	59.3	3.8	0.4	16.9	0.4	15.2	0.4	20.9	0.8	35.3
Luxembourg	9 356	104.3	30.5	72.7	12.4	6.0	27.3	6.0	26.4	22.0	28.1	23.9	29.2
Mexico	6 490	72.3	25.5	38.2	4.2	0.3	13.3	0.4	9.1	0.0	16.8	0.4	35.3
Netherlands	9 081	101.2	59.6	40.0	6.8	0.7	41.1	0.8	55.7	8.8	56.3	9.1	58.9
New Zealand	11 230	125.2	20.0	106.0	10.6	10.1	14.9	10.2	0.2	1.9	14.9	11.7	15.9
Norway	8 894	99.1	9.9	88.8	2.8	5.2	3.7	5.2	0.8	0.1	4.0	5.2	9.9
Poland	8 047	89.7	12.2	78.9	5.5	8.6	5.8	8.7	1.0	0.5	6.0	8.9	13.6
Portugal	7 824	87.2	27.8	58.5	8.8	3.6	14.7	3.6	20.7	11.9	23.6	13.6	31.9
Slovak Republic	8 648	96.4	40.9	55.6	11.7	2.6	23.3	2.5	23.2	1.3	29.4	3.6	42.4
Spain	8 150	90.8	12.7	74.2	5.0	5.3	6.2	5.4	0.0	0.1	6.2	5.5	13.9
Sweden	8 635	96.3	11.5	85.8	4.4	6.2	6.1	6.1	4.2	0.0	6.7	5.9	12.0
Switzerland	9 830	109.6	37.5	66.7	8.0	4.8	17.0	4.8	5.9	1.0	18.0	5.6	34.2
Turkey	6 928	77.2	40.8	35.8	5.9	0.7	24.3	0.7	23.9	0.2	29.6	0.9	52.8
United Kingdom	11 156	124.4	23.5	97.8	8.6	6.1	14.8	6.4	0.6	1.2	14.9	7.4	18.9
United States	11 186	124.7	29.1	94.0	12.7	7.7	18.9	7.7	5.8	4.3	20.8	10.7	23.3
OECD average	8 971	100.0	33.0	68.1	7.2	3.8	20.5	3.8	17.8	2.8	24.3	6.1	
Partners													
Argentina	10 197	113.7	53.2	58.4	12.2	1.6	31.4	1.6	26.2	5.2	40.4	6.6	46.8
Azerbaijan	3 106	34.6	17.9	18.1	1.4	0.4	2.5	0.4	0.8	0.2	3.1	0.5	51.8
Brazil	7 970	88.8	41.4	46.6	8.2	0.6	24.1	0.7	14.5	3.8	28.7	4.5	46.6
Bulgaria	11 352	126.5	69.6	59.4	16.4	1.0	47.5	0.9	23.6	0.2	48.2	1.2	55.0
Chile	8 446	94.1	53.0	52.2	14.2	0.8	38.8	0.7	14.6	0.7	42.6	1.5	56.3
Colombia	7 200	80.3	25.2	57.0	7.5	1.3	14.1	1.4	6.5	6.6	15.0	7.3	31.3
Croatia	7 356	82.0	33.8	50.0	6.0	1.3	20.4	1.3	25.7	8.2	26.4	8.5	41.3
Estonia	6 986	77.9	16.0	61.5	3.8	2.9	6.5	2.9	0.1	0.5	6.4	3.3	20.5
Hong Kong-China	8 381	93.4	34.1	58.3	3.6	0.6	13.6	0.6	8.3	4.9	16.4	5.0	36.5
Indonesia	4 909	54.7	19.4	25.4	0.7	0.0	8.0	0.0	4.7	0.0	9.0	0.0	35.5
Israel	12 299	137.1	44.4	96.1	9.9	4.1	20.0	4.1	5.4	0.8	21.7	4.8	32.4
Jordan	7 989	89.1	19.7	67.5	5.1	3.3	7.8	3.3	0.0	0.0	7.8	3.3	22.1
Kyrgyzstan	6 991	77.9	30.7	48.3	3.0	0.2	17.4	0.2	0.0	1.0	17.0	1.1	39.4
Latvia	7 056	78.7	14.5	64.2	4.3	3.1	6.7	3.1	0.6	1.6	6.8	4.5	18.4
Liechtenstein	9 330	104.0	c	c	c	c	c	c	c	c	c	c	c
Lithuania	8 082	90.1	25.5	65.4	9.0	3.8	15.0	3.9	12.2	0.5	17.5	4.3	28.3
Macao-China	6 095	67.9	19.2	55.0	1.0	0.3	2.2	0.3	7.7	8.5	7.8	8.7	28.3
Montenegro	6 390	71.2	20.2	50.8	3.5	0.8	12.0	0.9	15.4	5.0	16.4	5.2	28.3
Qatar	7 012	78.2	47.3	41.9	c	c	c	c	17.6	0.6	c	c	60.5
Romania	6 585	73.4	35.5	37.7	6.8	1.0	19.8	1.0	19.5	0.0	25.2	1.0	48.3
Russian Federation	8 023	89.4	24.1	66.9	4.6	2.2	8.2	2.2	5.0	4.1	9.4	5.5	27.0
Serbia	7 224	80.5	34.3	48.7	6.6	1.0	22.9	1.0	22.2	3.2	25.5	3.7	42.6
Slovenia	9 628	107.3	64.8	42.8	6.2	0.3	46.2	0.3	52.0	0.1	54.3	0.4	60.4
Chinese Taipei	8 889	99.1	45.8	51.7	6.0	1.0	26.4	1.0	23.2	1.3	30.7	2.2	46.2
Thailand	5 958	66.4	25.6	43.6	7.7	0.4	18.0	0.6	7.4	1.3	19.4	1.9	38.5
Tunisia	6 768	75.4	32.3	43.9	3.0	0.2	12.6	0.2	25.0	2.1	26.5	2.2	42.8
Uruguay	8 887	99.1	39.6	57.7	11.8	1.9	23.9	2.0	26.3	2.6	32.8	4.2	39.9

1. The variance components were estimated for all students in participating countries with data on socio-economic background and study programmes.
2. The total variance in student performance is calculated from the square of the standard deviation for the students used in the analysis. The statistical variance in student performance and not the standard deviation is used for this comparison to allow for the decomposition.
3. The sum of the between- and within-school variance components, as an estimate from a sample, does not necessarily add up to the total.
4. In some countries, sub-units within schools were sampled instead of schools and this may affect the estimation of the between-school variance components (see Annex A2).
5. This index is often referred to as the intra-class correlation (rho).

Source: OECD (2007a: 96).

Table 4.3c Differences in science performance between students with an immigrant background (first- and second-generation) and native students associated with students' immigrant background and home language Results based on students' self-reports

	Difference in the science performance							
	WITHOUT ACCOUNTING for the economic, social and cultural status of students				WITH ACCOUNTING for the economic, social and cultural status of students			
	Students with an immigrant background minus native students		Students with an immigrant background who speak a language at home that is different from the language of instruction minus native students		Students with an immigrant background minus native students		Students with an immigrant background who speak a language at home that is different from the language of instruction minus native students	
	Difference	S.E.	Difference	S.E.	Difference	S.E.	Difference	S.E.
OECD								
Australia	-2.0	(5.0)	**-15.2**	(7.4)	-0.4	(4.4)	3.0	(3.9)
Austria	**-90.1**	(11.1)	**-96.4**	(13.2)	**-60.9**	(8.4)	**-36.8**	(9.4)
Belgium	**-86.4**	(6.5)	**-102.4**	(7.9)	**-57.2**	(5.4)	**-51.8**	(6.0)
Canada	**-16.9**	(4.5)	**-20.7**	(5.8)	**-12.8**	(4.1)	**-10.1**	(4.9)
Czech Republic	c	c	c	c	c	c	c	c
Denmark	**-86.9**	(7.7)	**-95.7**	(8.8)	**-48.9**	(7.6)	**-33.3**	(9.3)
Finland	c	c	c	c	c	c	c	c
France	**-53.1**	(9.2)	**-58.8**	(10.9)	**-18.1**	(7.6)	**-18.2**	(8.4)
Germany	**-85.4**	(6.7)	**-96.9**	(8.0)	**-45.8**	(6.5)	**-24.3**	(8.3)
Greece	**-44.3**	(9.6)	**-78.9**	(11.0)	**-25.1**	(8.6)	-10.4	(11.5)
Hungary	c	c	c	c	c	c	c	c
Iceland	c	c	c	c	c	c	c	c
Ireland	-10.5	(11.7)	c	c	-12.8	(10.3)	c	c
Italy	**-58.4**	(7.8)	c	c	**-46.9**	(8.1)	c	c
Japan	c	c	c	c	c	c	c	c
Korea	c	c	c	c	c	c	c	c
Luxembourg	**-66.5**	(3.3)	**-82.3**	(3.6)	**-31.7**	(3.9)	0.0	(5.4)
Mexico	c	c	c	c	c	c	c	c
Netherlands	**-75.5**	(9.7)	**-85.6**	(11.6)	**-41.0**	(7.9)	**-36.9**	(9.4)
New Zealand	-15.9	(6.0)	**-38.6**	(8.7)	**-16.7**	(4.5)	-7.4	(4.7)
Norway	**-58.6**	(8.5)	**-59.8**	(10.9)	**-35.3**	(9.0)	**-24.0**	(11.8)
Poland	c	c	c	c	c	c	c	c
Portugal	**-54.9**	(10.8)	c	c	**-56.5**	(8.3)	c	c
Slovak Republic	c	c	c	c	c	c	c	c
Spain	**-59.7**	(6.9)	c	c	**-48.2**	(6.1)	c	c
Sweden	**-60.8**	(5.1)	**-67.6**	(6.1)	**-43.4**	(4.5)	**-32.0**	(7.6)
Switzerland	**-81.4**	(4.2)	**-95.5**	(4.4)	**-56.3**	(4.1)	**-37.2**	(5.5)
Turkey	c	c	c	c	c	c	c	
United Kingdom	**-32.5**	(9.0)	**-49.1**	(14.3)	**-14.2**	(6.0)	-8.3	(6.7)
United States	**-48.3**	(6.4)	**-62.2**	(6.9)	**-16.8**	(6.1)	-9.5	(7.0)
OECD total	**-40.1**	(2.9)	**-51.7**	(3.7)	**-24.5**	(2.6)	**-16.4**	(2.9)
OECD average	**-54.4**	(1.8)	**-69.1**	(2.3)	**-34.4**	(1.5)	**-21.1**	(2.0)
Partners								
Argentina	c	c	c	c	c	c	c	c
Azerbaijan	c	c	c	c	c	c	c	c
Brazil	c	c	c	c	c	c	c	c
Bulgaria	c	c	c	c	c	c	c	c
Chile	c	c	a	a	c	c	a	a
Colombia	c	c	a	a	c	c	a	a
Croatia	**-19.4**	(4.4)	c	c	-7.1	(4.4)	c	c
Estonia	**-32.8**	(4.9)	c	c	**-30.2**	(4.8)	c	c
Hong Kong-China	**-9.1**	(4.1)	**-38.7**	(11.2)	**8.4**	(3.8)	c	c
Indonesia	c	c	c	c	c	c	c	c
Israel	-5.7	(5.4)	3.6	(7.6)	**9.0**	(4.9)	1.4	(5.4)
Jordan	**25.8**	(4.5)	c	c	**15.1**	(4.1)	c	c
Kyrgyzstan	c	c	c	c	c	c	c	c
Latvia	-3.3	(6.2)	c	c	-6.0	(5.8)	c	c
Liechtenstein	**-47.2**	(11.7)	**-107.3**	(20.1)	**-34.0**	(10.9)	-23.8	(12.3)
Lithuania	c	c	c	c	c	c	c	c
Macao-China	**11.0**	(2.6)	**10.1**	(2.6)	**17.3**	(2.7)	**34.7**	(14.5)
Montenegro	**16.7**	(5.9)	c	c	**16.0**	(5.8)	c	c
Qatar	**58.0**	(2.4)	**62.5**	(4.4)	**58.9**	(2.4)	**58.2**	(2.7)
Romania	c	c	c	c	c	c	c	c
Russian Federation	**-13.6**	(6.0)	c	c	**-11.7**	(5.5)	-1.1	(5.1)
Serbia	8.6	(4.7)	c	c	**10.8**	(4.0)	c	c
Slovenia	**-56.0**	(5.5)	**-78.5**	(7.1)	**-29.0**	(5.5)	-10.4	(7.7)
Chinese Taipei	c	c	c	c	c	c	c	c
Thailand	c	c	c	c	c	c	c	c
Tunisia	c	c	c	c	c	c	c	c
Uruguay	c	c	c	c	c	c	c	c

Note: Values that are statistically significant are indicated in bold (see Annex A3).
Source: OECD (2007a: 121).

Table 4.4a Relationship between student performance in science and the PISA index of economic, social and cultural status (ESCS)

	Unadjusted mean score		Mean score if the mean ESCS would be equal in all OECD countries		Strength of the relationship between student performance and the ESCS		Slope of the socio-economic gradient[1]		Length of the projection of the gradient line						ESCS mean		Variability in the ESCS		Index of curvelinearity[2]		Index of skewness in the distribution of the ESCS		Percentage of students that fall within the lowest 15 per cent of the international distribution on the ESCS	
									5th percentile of the ESCS		95th percentile of the ESCS		Difference between 95th and 5th percentile of the ESCS											
	Mean score	S.E.	Mean score	S.E.	Percentage of explained variance in student performance	S.E.	Score point difference associated with one unit on the ESCS	S.E.	Index	S.E.	Index	S.E.	Difference	S.E.	Mean index	S.E.	Standard deviation	S.E.	Score point difference associated with one unit on the ESCS squared	S.E.	Index	S.E.	Approximated by the percentage of students with a value on the PISA index of economic, social and cultural status smaller than -1	S.E.
OECD																								
Australia	**527**	(2.3)	**519**	(1.7)	**11.3**	(0.78)	**43**	(1.5)	-1.08	(0.02)	1.39	(0.03)	2.47	(0.03)	0.21	(0.01)	0.78	(0.01)	-1.23	(1.38)	-0.22	(0.03)	6.1	(0.3)
Austria	**511**	(3.9)	502	(3.7)	15.4	(2.02)	**46**	(3.1)	-1.04	(0.07)	1.63	(0.05)	2.67	(0.09)	0.20	(0.02)	0.83	(0.02)	**-7.84**	(1.73)	0.09	(0.08)	6.0	(0.7)
Belgium	**510**	(2.5)	503	(2.2)	**19.4**	(1.29)	**48**	(1.9)	-1.29	(0.04)	1.58	(0.02)	2.87	(0.05)	0.17	(0.02)	0.91	(0.01)	**-2.01**	(0.97)	-0.26	(0.04)	8.6	(0.5)
Canada	**534**	(2.0)	**524**	(1.8)	**8.2**	(0.68)	**33**	(1.4)	-0.99	(0.02)	1.60	(0.02)	2.59	(0.03)	0.37	(0.02)	0.81	(0.01)	**-2.57**	(1.14)	-0.29	(0.03)	4.7	(0.3)
Czech Republic	**513**	(3.5)	**512**	(3.2)	15.6	(1.35)	**51**	(2.6)	-1.14	(0.02)	1.30	(0.02)	2.44	(0.03)	0.03	(0.02)	0.76	(0.01)	-3.37	(1.96)	0.03	(0.06)	7.8	(0.5)
Denmark	496	(3.1)	**485**	(2.5)	14.1	(1.43)	39	(2.0)	-1.14	(0.04)	1.72	(0.03)	2.86	(0.04)	0.31	(0.03)	0.89	(0.01)	-1.00	(1.27)	-0.16	(0.05)	6.5	(0.5)
Finland	**563**	(2.0)	**556**	(1.8)	**8.3**	(0.87)	**31**	(1.6)	-1.04	(0.03)	1.48	(0.02)	2.52	(0.03)	0.26	(0.02)	0.79	(0.01)	1.89	(1.56)	-0.17	(0.04)	5.6	(0.4)
France	495	(3.4)	502	(2.7)	**21.2**	(1.77)	**54**	(2.5)	-1.50	(0.06)	1.30	(0.03)	2.81	(0.07)	-0.09	(0.03)	0.86	(0.02)	1.14	(1.88)	-0.19	(0.04)	14.1	(0.8)
Germany	**516**	(3.8)	505	(3.1)	**19.0**	(1.45)	**46**	(2.1)	-1.16	(0.05)	1.82	(0.04)	2.99	(0.06)	0.29	(0.03)	0.93	(0.01)	**-3.60**	(1.17)	-0.09	(0.05)	6.8	(0.6)
Greece	**473**	(3.2)	**479**	(2.6)	15.0	(1.72)	37	(2.2)	-1.72	(0.04)	1.45	(0.06)	3.18	(0.07)	-0.15	(0.03)	0.97	(0.02)	**-4.04**	(1.39)	0.04	(0.03)	20.2	(1.1)
Hungary	504	(2.7)	**508**	(2.2)	**21.4**	(1.58)	**44**	(1.8)	-1.53	(0.03)	1.50	(0.03)	3.02	(0.05)	-0.09	(0.03)	0.92	(0.02)	**-3.28**	(1.25)	0.12	(0.04)	15.4	(1.0)
Iceland	**491**	(1.6)	**470**	(2.1)	**6.7**	(0.80)	**29**	(1.8)	-0.67	(0.04)	2.11	(0.02)	2.79	(0.04)	0.77	(0.01)	0.87	(0.01)	-2.61	(1.69)	-0.24	(0.04)	2.4	(0.3)
Ireland	**508**	(3.2)	**510**	(2.5)	12.7	(1.37)	39	(2.2)	-1.38	(0.04)	1.43	(0.04)	2.81	(0.05)	-0.02	(0.03)	0.86	(0.01)	-1.05	(1.34)	0.02	(0.04)	12.0	(0.7)
Italy	**475**	(2.0)	**478**	(1.9)	**10.0**	(0.94)	**31**	(1.6)	-1.59	(0.03)	1.67	(0.04)	3.25	(0.05)	-0.07	(0.02)	0.98	(0.01)	**-4.57**	(0.94)	0.21	(0.02)	18.7	(0.6)
Japan	**531**	(3.4)	**533**	(3.1)	**7.4**	(0.95)	39	(2.7)	-1.08	(0.02)	1.13	(0.01)	2.22	(0.02)	-0.01	(0.02)	0.70	(0.01)	**-11.25**	(2.49)	0.06	(0.03)	6.9	(0.5)
Korea	**522**	(3.4)	**522**	(3.0)	**8.1**	(1.49)	**32**	(3.1)	-1.32	(0.05)	1.30	(0.04)	2.62	(0.07)	-0.01	(0.02)	0.81	(0.01)	2.51	(1.77)	-0.14	(0.04)	10.7	(0.6)
Luxembourg	**486**	(1.1)	**483**	(1.1)	**21.7**	(1.12)	41	(1.2)	-1.96	(0.02)	1.72	(0.02)	3.68	(0.03)	0.09	(0.01)	1.10	(0.01)	-1.71	(0.93)	-0.36	(0.03)	17.6	(0.5)
Mexico	**410**	(2.7)	**435**	(2.4)	16.8	(1.72)	**25**	(1.3)	-2.95	(0.06)	1.21	(0.06)	4.16	(0.08)	-0.99	(0.04)	1.31	(0.02)	**1.61**	(0.62)	0.20	(0.04)	52.5	(1.4)
Netherlands	**525**	(2.7)	**515**	(2.4)	16.7	(1.65)	44	(2.2)	-1.23	(0.06)	1.60	(0.03)	2.83	(0.06)	0.25	(0.03)	0.89	(0.02)	2.11	(1.65)	-0.12	(0.04)	7.5	(0.7)
New Zealand	**530**	(2.7)	**528**	(2.3)	**16.4**	(1.11)	**52**	(1.8)	-1.27	(0.04)	1.40	(0.04)	2.67	(0.05)	0.10	(0.02)	0.83	(0.01)	2.68	(1.61)	-0.28	(0.08)	9.0	(0.4)
Norway	**487**	(3.1)	**474**	(2.8)	**8.3**	(1.10)	36	(2.5)	-0.73	(0.03)	1.62	(0.03)	2.35	(0.04)	0.42	(0.02)	0.76	(0.01)	**-4.10**	(1.65)	-0.32	(0.05)	2.3	(0.3)
Poland	498	(2.3)	**510**	(2.1)	14.5	(1.13)	39	(1.8)	-1.56	(0.03)	1.31	(0.07)	2.87	(0.07)	-0.30	(0.02)	0.87	(0.01)	0.60	(1.09)	-0.25	(0.04)	20.8	(0.9)
Portugal	**474**	(3.0)	**492**	(2.3)	16.6	(1.50)	**28**	(1.4)	-2.46	(0.03)	1.70	(0.03)	4.16	(0.04)	-0.62	(0.04)	1.28	(0.02)	0.80	(0.78)	0.42	(0.03)	43.5	(1.5)
Slovak Republic	**488**	(2.6)	**495**	(2.2)	**19.2**	(1.96)	**45**	(2.6)	-1.40	(0.07)	1.48	(0.02)	2.88	(0.07)	-0.15	(0.02)	0.91	(0.02)	-3.39	(2.75)	0.20	(0.12)	13.5	(0.9)
Spain	**488**	(2.6)	499	(1.9)	13.9	(1.21)	**31**	(1.3)	-1.93	(0.05)	1.56	(0.01)	3.48	(0.05)	-0.31	(0.03)	1.07	(0.01)	**-2.44**	(0.99)	0.23	(0.03)	29.1	(1.0)
Sweden	503	(2.4)	**496**	(2.2)	**10.6**	(0.97)	38	(2.1)	-1.04	(0.03)	1.47	(0.04)	2.50	(0.05)	0.24	(0.02)	0.79	(0.01)	-1.49	(1.84)	-0.33	(0.09)	5.6	(0.4)
Switzerland	**512**	(3.2)	**508**	(2.6)	15.7	(1.20)	**44**	(1.8)	-1.37	(0.03)	1.54	(0.03)	2.91	(0.04)	0.09	(0.02)	0.89	(0.01)	**-2.30**	(1.24)	-0.04	(0.03)	11.7	(0.5)
Turkey	**424**	(3.8)	**463**	(6.4)	16.5	(2.96)	**31**	(3.2)	-2.85	(0.04)	0.77	(0.08)	3.62	(0.08)	-1.28	(0.04)	1.10	(0.03)	**5.72**	(1.39)	0.15	(0.05)	62.7	(1.6)
United Kingdom	**515**	(2.3)	**508**	(1.9)	13.9	(1.12)	**48**	(1.9)	-1.12	(0.03)	1.50	(0.01)	2.62	(0.03)	0.19	(0.01)	0.81	(0.01)	-0.33	(1.62)	-0.13	(0.05)	6.6	(0.5)
United States	**489**	(4.2)	**483**	(3.0)	**17.9**	(1.63)	**49**	(2.5)	-1.39	(0.06)	1.59	(0.04)	2.98	(0.07)	0.14	(0.04)	0.91	(0.02)	**3.30**	(1.38)	-0.21	(0.04)	11.0	(0.9)
OECD total	491	(1.2)	496	(0.9)	20.2	(0.57)	45	(0.6)	-2.00	(0.03)	1.47	(0.01)	3.47	(0.03)	-0.10	(0.01)	1.04	(0.01)	**-0.86**	(0.40)	-0.10	(0.02)	17.9	(0.3)
OECD average	500	(0.5)	500	(0.5)	14.4	(0.26)	40	(0.4)	-1.43	(0.01)	1.50	(0.01)	2.93	(0.01)	0.00	(0.00)	0.91	(0.00)	**-1.39**	(0.28)	-0.07	(0.01)	14.9	(0.1)
Partners																								
Argentina	**391**	(6.1)	**416**	(4.7)	**19.5**	(2.33)	38	(2.4)	-2.54	(0.06)	1.27	(0.07)	3.81	(0.08)	-0.64	(0.07)	1.16	(0.02)	3.11	(1.65)	-0.06	(0.06)	37.9	(2.2)
Azerbaijan	**382**	(2.8)	**388**	(2.7)	**4.7**	(1.71)	**11**	(2.0)	-2.06	(0.04)	1.31	(0.04)	3.37	(0.05)	-0.45	(0.03)	1.06	(0.02)	**3.86**	(1.16)	0.13	(0.05)	33.7	(1.2)
Brazil	**390**	(2.8)	**424**	(3.6)	17.1	(1.92)	**30**	(1.9)	-3.04	(0.02)	0.89	(0.04)	3.93	(0.05)	-1.12	(0.03)	1.25	(0.01)	**6.10**	(1.33)	0.05	(0.03)	52.9	(1.1)
Bulgaria	**434**	(6.1)	**446**	(4.4)	**24.1**	(2.76)	**52**	(3.6)	-1.77	(0.09)	1.44	(0.06)	3.20	(0.11)	-0.21	(0.05)	1.01	(0.02)	-1.55	(1.99)	-0.05	(0.08)	21.1	(1.4)
Chile	**438**	(4.3)	**465**	(3.3)	**23.3**	(1.92)	38	(1.8)	-2.55	(0.08)	1.30	(0.07)	3.85	(0.10)	-0.70	(0.06)	1.18	(0.03)	**4.30**	(1.12)	0.15	(0.05)	42.3	(2.2)
Colombia	**388**	(3.4)	**411**	(3.0)	11.4	(1.57)	**23**	(1.6)	-2.95	(0.07)	1.06	(0.08)	4.01	(0.10)	-1.00	(0.05)	1.23	(0.03)	**4.03**	(1.12)	0.04	(0.05)	49.9	(2.0)
Croatia	**493**	(2.4)	497	(2.3)	12.3	(1.21)	**34**	(1.9)	-1.46	(0.04)	1.46	(0.04)	2.92	(0.05)	-0.11	(0.02)	0.87	(0.01)	0.01	(1.16)	0.23	(0.03)	13.5	(0.6)
Estonia	**531**	(2.5)	**527**	(2.4)	**9.3**	(1.12)	**31**	(2.0)	-1.11	(0.03)	1.44	(0.02)	2.56	(0.03)	0.14	(0.02)	0.81	(0.01)	**5.04**	(2.20)	0.02	(0.04)	7.3	(0.7)
Hong Kong-China	**542**	(2.5)	**560**	(2.9)	**6.9**	(1.26)	**26**	(2.3)	-2.17	(0.04)	0.98	(0.08)	3.14	(0.09)	-0.67	(0.03)	0.93	(0.02)	-1.24	(1.50)	0.18	(0.03)	37.6	(1.2)
Indonesia	**393**	(5.7)	**425**	(7.5)	10.2	(2.31)	**21**	(2.6)	-3.11	(0.04)	0.35	(0.07)	3.46	(0.08)	-1.52	(0.05)	1.08	(0.02)	**4.01**	(1.25)	0.29	(0.07)	68.6	(2.1)
Israel	**454**	(3.7)	**448**	(3.5)	**10.9**	(1.10)	43	(2.7)	-1.29	(0.04)	1.50	(0.06)	2.79	(0.07)	0.22	(0.02)	0.86	(0.01)	**5.25**	(1.84)	-0.60	(0.05)	8.3	(0.6)
Jordan	**422**	(2.8)	**438**	(2.8)	**11.2**	(1.35)	27	(1.8)	-2.57	(0.09)	1.03	(0.05)	3.59	(0.09)	-0.57	(0.03)	1.11	(0.02)	**2.93**	(0.95)	-0.46	(0.05)	34.0	(1.2)
Kyrgyzstan	**322**	(2.9)	**340**	(2.8)	**8.2**	(1.42)	27	(2.6)	-2.02	(0.02)	0.83	(0.04)	2.85	(0.04)	-0.66	(0.02)	0.88	(0.01)	**4.65**	(1.54)	0.00	(0.05)	35.0	(1.1)
Latvia	**490**	(3.0)	**491**	(2.6)	**9.7**	(1.41)	**29**	(2.3)	-1.40	(0.03)	1.42	(0.04)	2.82	(0.05)	-0.02	(0.02)	0.90	(0.01)	-1.99	(1.84)	0.03	(0.04)	14.7	(0.8)
Liechtenstein	**522**	(4.1)	**513**	(4.3)	20.4	(4.42)	49	(5.5)	-1.34	(0.08)	1.70	(0.11)	3.04	(0.12)	0.19	(0.05)	0.89	(0.03)	**-9.11**	(4.43)	-0.09	(0.12)	9.2	(1.3)
Lithuania	**488**	(2.8)	**487**	(2.3)	15.2	(1.33)	38	(2.0)	-1.37	(0.02)	1.49	(0.03)	2.86	(0.04)	0.04	(0.03)	0.92	(0.01)	-2.32	(1.72)	0.04	(0.03)	14.6	(0.6)
Macao-China	**511**	(1.1)	**523**	(1.8)	**2.2**	(0.49)	**13**	(1.5)	-2.28	(0.02)	0.55	(0.03)	2.83	(0.03)	-0.91	(0.01)	0.87	(0.01)	**-2.81**	(1.10)	0.23	(0.03)	48.6	(0.8)
Montenegro	**412**	(1.1)	**412**	(1.1)	**7.5**	(0.90)	**24**	(1.4)	-1.44	(0.03)	1.42	(0.02)	2.87	(0.03)	-0.02	(0.01)	0.90	(0.01)	0.30	(1.36)	-0.82	(0.02)	14.4	(0.5)
Qatar	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m
Romania	**418**	(4.2)	**431**	(3.9)	16.6	(3.15)	35	(3.4)	-1.89	(0.06)	1.27	(0.06)	3.16	(0.09)	-0.37	(0.04)	0.95	(0.03)	1.05	(1.69)	0.08	(0.05)	24.1	(1.3)
Russian Federation	**479**	(3.7)	**483**	(3.2)	**8.1**	(1.23)	**32**	(2.6)	-1.31	(0.03)	1.18	(0.01)	2.48	(0.03)	-0.10	(0.03)	0.79	(0.01)	0.28	(2.32)	0.20	(0.04)	12.6	(0.9)
Serbia	**436**	(3.0)	**440**	(2.5)	13.2	(1.27)	**33**	(1.8)	-1.56	(0.03)	1.52	(0.04)	3.08	(0.05)	-0.14	(0.03)	0.94	(0.01)	-1.21	(1.61)	-0.03	(0.06)	16.9	(0.9)
Slovenia	**519**	(1.1)	**513**	(1.2)	16.7	(1.11)	**46**	(1.6)	-1.25	(0.04)	1.57	(0.02)	2.82	(0.04)	0.13	(0.01)	0.87	(0.01)	-1.09	(1.71)	0.09	(0.03)	8.7	(0.4)
Chinese Taipei	**532**	(3.6)	**546**	(2.8)	12.5	(1.19)	42	(2.1)	-1.60	(0.04)	1.04	(0.03)	2.63	(0.05)	-0.31	(0.02)	0.80	(0.01)	1.38	(1.32)	0.41	(0.04)	20.3	(1.1)
Thailand	**421**	(2.1)	**461**	(3.3)	15.9	(2.00)	**28**	(1.6)	-2.84	(0.04)	0.77	(0.07)	3.62	(0.07)	-1.43	(0.03)	1.11	(0.02)	**4.96**	(1.15)	0.01	(0.03)	69.4	(1.1)
Tunisia	**386**	(3.0)	**408**	(4.4)	**9.5**	(2.11)	**19**	(2.2)	-3.26	(0.03)	1.08	(0.08)	4.34	(0.08)	-1.20	(0.07)	1.36	(0.03)	**4.59**	(0.96)	0.68	(0.04)	56.9	(2.3)
Uruguay	**428**	(2.7)	**446**	(2.5)	**18.3**	(1.23)	**34**	(1.4)	-2.47	(0.04)	1.43	(0.04)	3.90	(0.05)	-0.51	(0.03)	1.19	(0.01)	**3.70**	(0.98)	-0.05	(0.04)	34.7	(1.1)

Note: Values that are statistically significant are indicated in bold (see Annex A3).

1. Single-level bivariate regression of science performance on the ESCS, the slope is the regression coefficient for the ESCS.

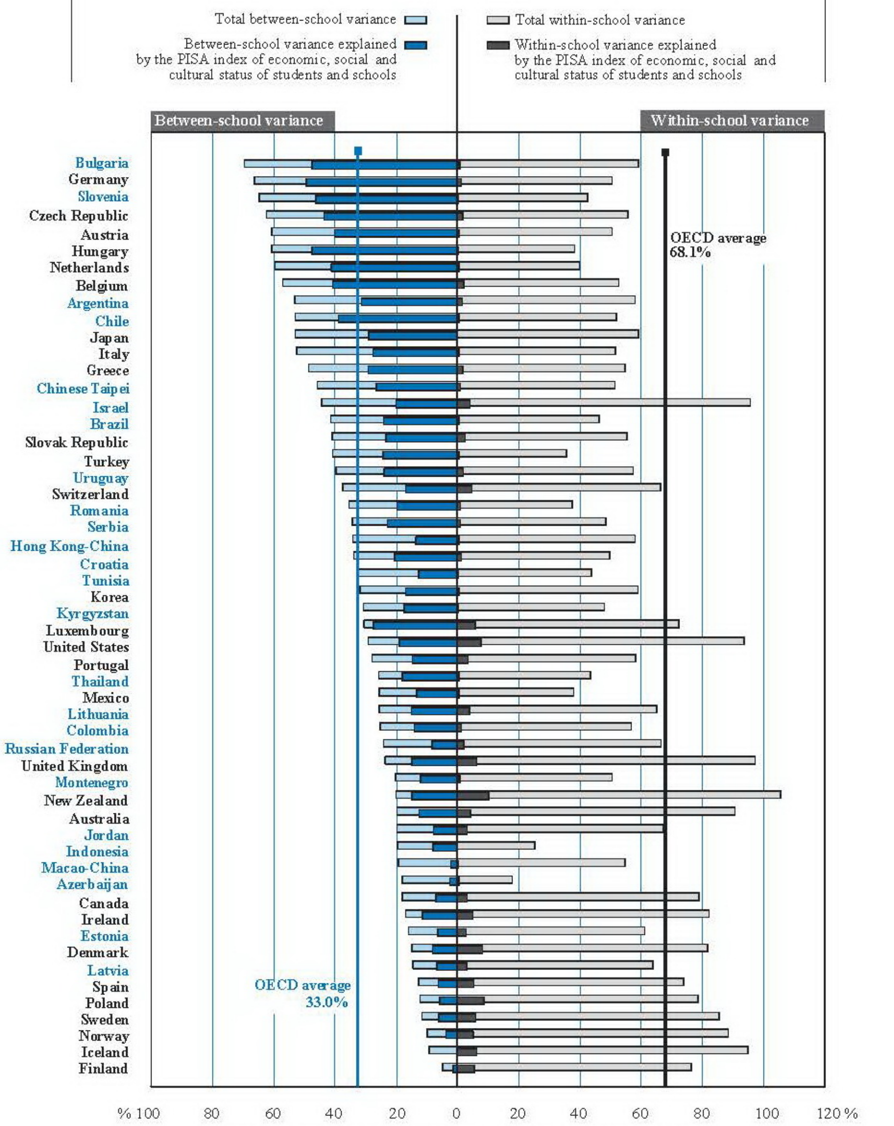
2. Student-level regression of science performance on the ESCS and the squared term of the ESCS, the index of curvelinearity is the regression coefficient for the squared term.

Source: OECD (2007a: 123-124).

Figure 4.1

Variance in student performance between schools and within schools on the science scale

Expressed as a percentage of the average variance in student performance in OECD countries



Source: OECD PISA 2006 database, Table 4.1a (as cited in OECD, 2007c: 171).

Figure 4.3

Characteristics of schools attended by native students and students with an immigrant background

School characteristics are LESS favourable for students with an immigrant background by:

School characteristics are MORE favourable for students with an immigrant background by:

<<<	at least 0.50 index points	>>>
<<	between 0.20 and 0.49 index points	>>
<	up to 0.19 index points	>

		Percentage of immigrant students[1]	Economic, social and cultural status[1]	Quality of educational resources[1]	Student/teacher ratio[1]	Teacher shortage[1]
OECD	Australia	22				
	Austria	13	< < <		>	
	Belgium	13	< < <		> >	< <
	Canada	21				
	Denmark	8	< < <			
	France	13	< < <	w	w	w
	Germany	14	< < <			< <
	Greece	8	< <	< <		
	Ireland	6				
	Italy	4	< <		>	
	Luxembourg	36	< < <	>		>
	Netherlands	11	< < <			
	New Zealand	21			< <	
	Norway	6	< < <			
	Portugal	6				
	Spain	7	< <		> >	
	Sweden	11	< <		> >	
	Switzerland	22	< < <			
	United Kingdom	9	< <		> >	
	United States	15	< < <		< <	
Partners	Croatia	12	< <			
	Estonia	12				> >
	Hong Kong-China	44	< < <			
	Israel	23	< <			
	Jordan	17	> >	> >	< <	
	Latvia	7		> >		
	Macao-China	74	< < <	>		>
	Montenegro	7	> >		<	<
	Qatar	40	>	>	< <	>
	Russian Federation	9				
	Serbia	9				
	Slovenia	10	< < <	<		

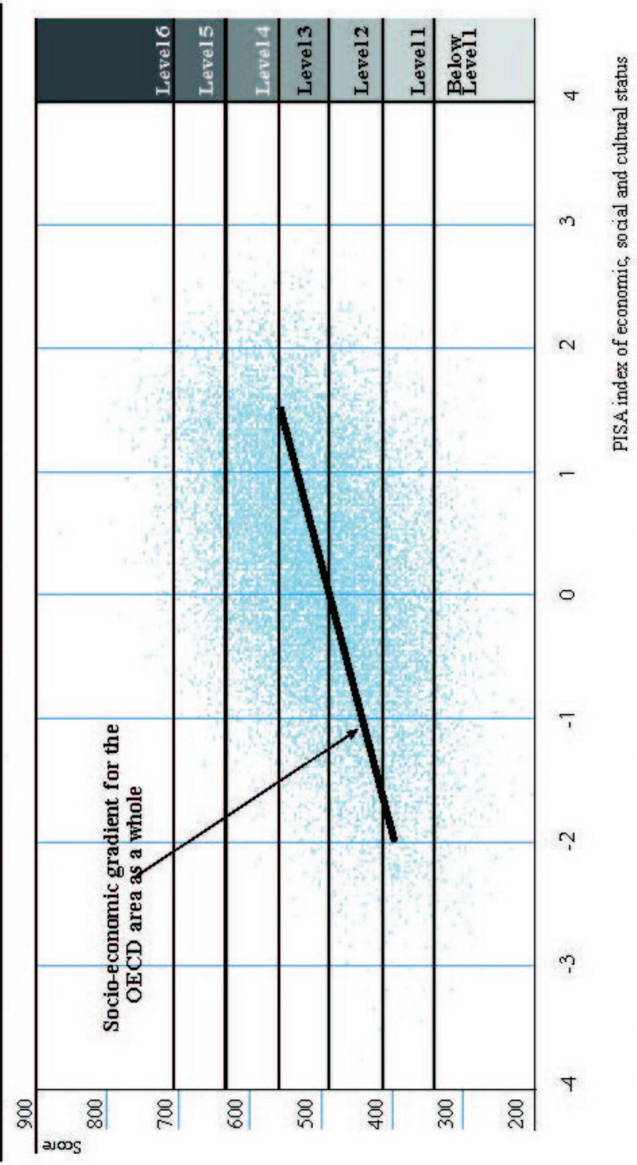
Schools have similar characteristics	9	24	20	24
Schools that immigrant students attend have more favourable characteristics	3	5	6	4
Schools that immigrant students attend have less favourable characteristics	20	2	5	3

1. Scores were standardised within each country sample to make an index which has 0 as the country mean and 1 as the standard deviation within the country.

Source: OECD PISA 2006 database, Table 4.3d (as cited in OECD, 2007c: 179).

Figure 4.5

Relationship between student performance in science and socio-economic background for the OECD area as a whole



Note: Each dot represents 497 students drawn randomly from the OECD area.
Source: OECD PISA 2006 database (as cited in OECD, 2007c: 183).

Figure 4.6

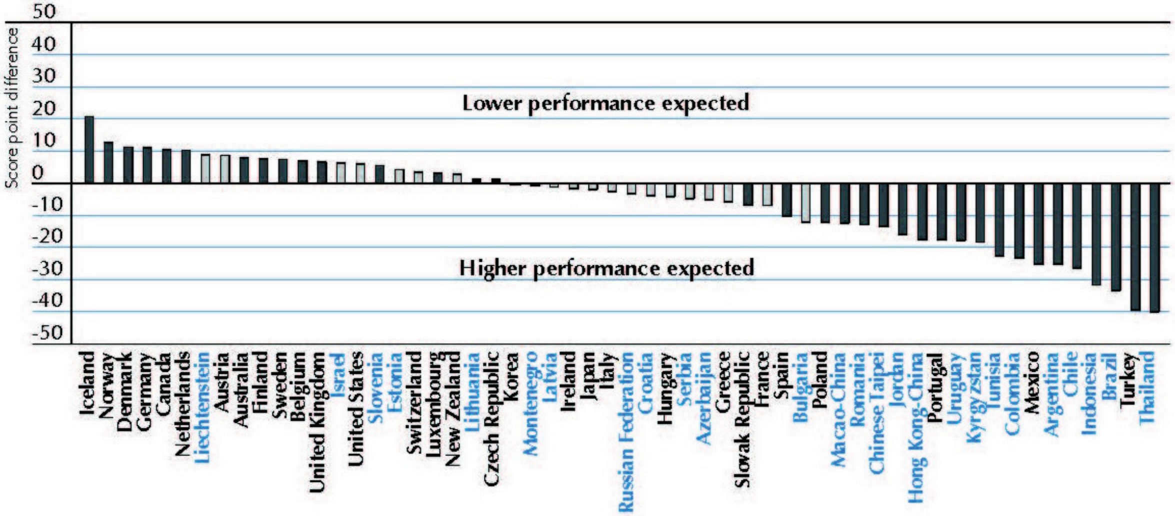
How socio-economic background relates to student performance in science

		Mean score	Mean score if the mean ESCS^1 would be equal in all OECD countries	Percentage of explained variance in student performance	Score point difference associated with one unit on the ESCS$^{1, 2}$ (gradient)	Percentage of students that fall within the lowest 15% of the international distribution on the ESCS^1
OECD	**Australia**	527	**519**	**11.3**	**43**	6.1
	Austria	**511**	502	15.4	**46**	6.0
	Belgium	**510**	503	**19.4**	**48**	8.6
	Canada	**534**	**524**	**8.2**	**33**	4.7
	Czech Republic	**513**	**512**	15.6	**51**	7.8
	Denmark	496	**485**	14.1	39	6.5
	Finland	**563**	**556**	**8.3**	**31**	5.6
	France	495	502	**21.2**	**54**	14.1
	Germany	**516**	505	**19.0**	**46**	6.8
	Greece	**473**	**479**	15.	37	20.2
	Hungary	504	**508**	**21.**	**44**	15.4
	Iceland	**491**	**470**	**6.7**	**29**	2.4
	Ireland	**508**	**510**	12.	39	12.0
	Italy	**475**	**478**	**10.**	**31**	18.7
	Japan	**531**	**533**	7.4	39	6.9
	Korea	**522**	**522**	**8.1**	**32**	10.7
	Luxembourg	**486**	**483**	**21.**	41	17.6
	Mexico	**410**	**435**	16.	**25**	52.5
	Netherlands	**525**	**515**	16.	44	7.5
	New Zealand	**530**	**528**	**16.**	**52**	9.0
	Norway	**487**	**474**	**8.3**	36	2.3
	Poland	498	**510**	14.	39	20.8
	Portugal	**474**	**492**	16.	**28**	43.5
	Slovak Republic	**488**	**495**	**19.2**	**45**	13.5
	Spain	**488**	499	13.9	**31**	29.1
	Sweden	503	**496**	**10.6**	38	5.6
	Switzerland	**512**	**508**	15.7	**44**	11.7
	Turkey United	**424**	**463**	16.5	**31**	62.7
	Kingdom United	**515**	**508**	13.9	**48**	6.6
	States	**489**	**483**	**17.9**	**49**	11.0
	OECD total	491	496	20.2	45	17.9
	OECD average	500	500	14.4	40	14.9
Partners	**Argentina**	**391**	**416**	**19.5**	38	37.9
	Azerbaijan	**382**	**388**	**4.7**	**11**	33.7
	Brazil	**390**	**424**	17.1	**30**	52.9
	Bulgaria	**434**	**446**	**24.1**	**52**	21.1
	Chile	**438**	**465**	**23.3**	38	42.3
	Colombia	**388**	**411**	11.4	**23**	49.9
	Croatia	**493**	497	12.3	**34**	13.5
	Estonia	**531**	527	**9.3**	**31**	7.3
	Hong Kong-China	**542**	**560**	**6.9**	**26**	37.6
	Indonesia	**393**	**425**	10.2	**21**	68.6
	Israel	**454**	**448**	**10.9**	43	8.3
	Jordan	**422**	**438**	**11.2**	27	34.0
	Kyrgyzstan	**322**	**340**	**8.2**	**27**	35.0
	Latvia	**490**	**491**	**9.7**	**29**	14.7
	Lithuania	**488**	**487**	15.	38	14.6
	Macao-China	**511**	**523**	**2.2**	**13**	48.6
	Montenegro	**412**	**412**	7.5	**24**	14.4
	Romania	**418**	**431**	16.	35	24.1
	Russian Federation	**479**	**483**	**8.1**	**32**	12.6
	Serbia	**436**	**440**	13.	**33**	16.9
	Slovenia	**519**	**513**	16.	**46**	8.7
	Chinese Taipei	**532**	**546**	12.	**42**	20.3
	Thailand	**421**	**461**	15.	**28**	69.4
	Tunisia	**386**	**408**	**9.5**	**19**	56.9
	Uruguay	**428**	**446**	**18.**	**34**	34.7

Note: Values that are statistically significant are indicated in bold (see Annex A3).
1. ESCS: the PISA index of economic, social and cultural status.
2. Single-level bivariate regression of science performance on the ESCS, the slope is the regression coefficient for the ESCS.
Source: OECD PISA 2006 database, Table 4.4a (as cited in OECD, 2007c: 184).

Figure 4.7

Difference between the unadjusted mean score and the mean score on the science scale if the mean PISA index of economic, social and cultural status were equal in all OECD countries



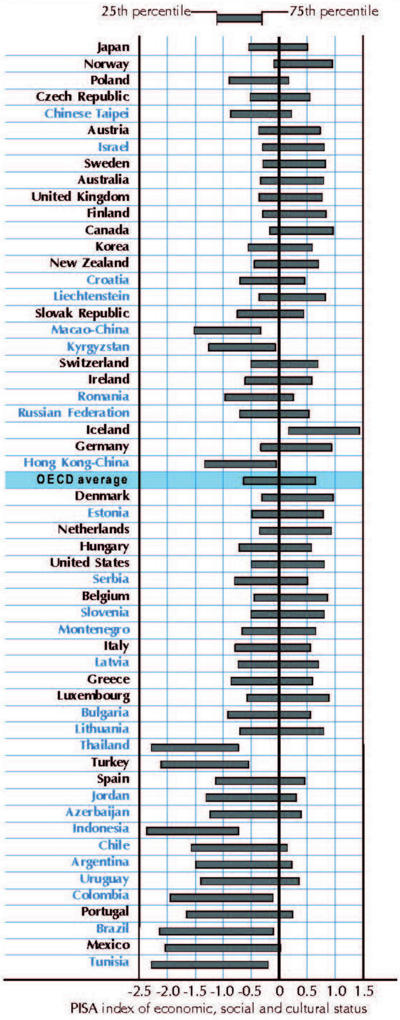
Countries are ranked in descending order of the difference between the unadjusted mean score and the mean score if the mean PISA index of economic, social and cultural status would be equal in all OECD countries.

Note: Statistically significant differences are marked in a darker tone.

Source: OECD PISA 2006 database, Table 4.4a (as cited in OECD, 2007c: 187).

Figure 4.8

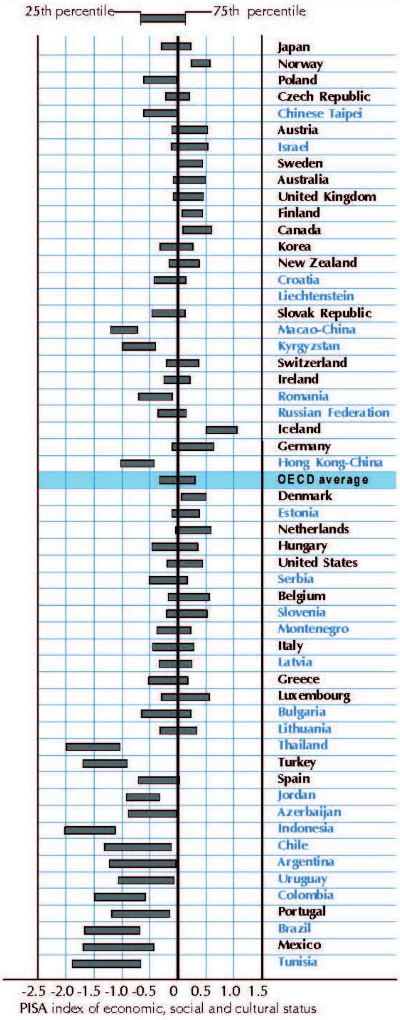
Student variability in the distribution of the PISA index of economic, social and cultural status (ESCS)



Countries are ranked in ascending order of the interquartile range of the distribution of the student-level ESCS.
Source: OECD PISA 2006 database, Table 4.4b (as cited in OECD, 2007c: 188).

Figure 4.9

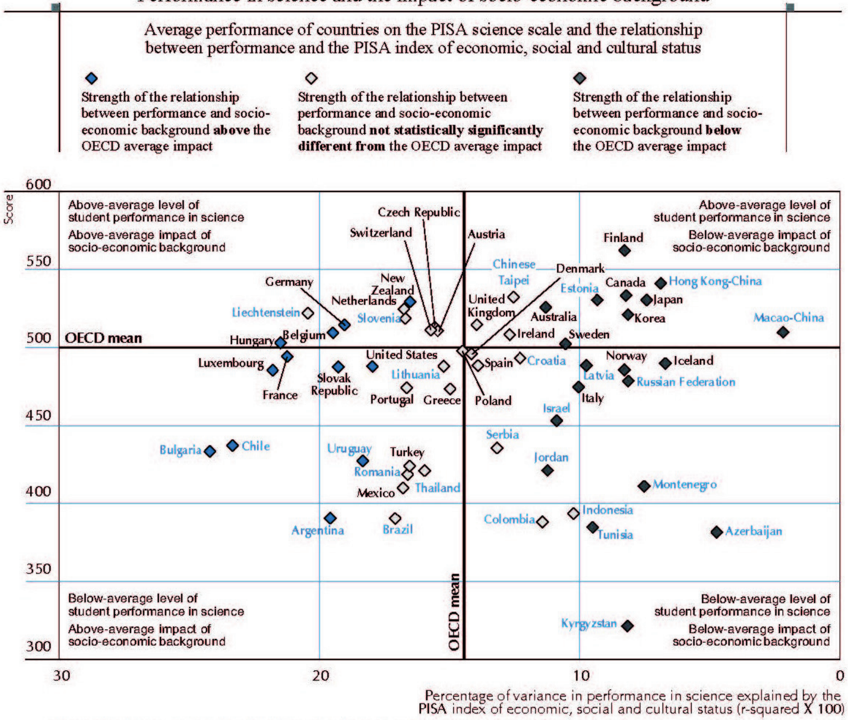
School variability in the distribution of the PISA index of economic, social and cultural status (ESCS)



Countries are ranked in ascending order of the interquartile range of the distribution of the student-level ESCS.
Source: OECD PISA 2006 database, Table 4.4b (as cited in OECD, 2007c: 188).

Figure 4.10

Performance in science and the impact of socio-economic background



Note: OECD mean used in this figure is the arithmetic average of all OECD countries.
Source: OECD PISA 2006 database, Table 4.4a (as cited in OECD, 2007c: 189).

Figure 4.11

Within-school and between-school socio-economic effect[1]

		Effect of the PISA index of economic, social and cultural status (ESCS)			Index of inclusion[5]
		Overall effect of ESCS[2]	**Within-school effect of ESCS[3]**	**Between-school effect of ESCS[4]**	
		Student-level score point difference associated with one unit of the ESCS	**Student-level score point difference associated with one unit of the student-level ESCS**	**School-level score point difference associated with one unit of the school mean ESCS**	**Proportion of ESCS variance within schools**
OECD	**Australia**	43	29	56	0.77
	Austria	46	10	110	0.71
	Belgium	48	17	102	0.73
	Canada	33	23	44	0.81
	Czech Republic	51	19	120	0.73
	Denmark	39	32	41	0.87
	Finland	31	30	10	0.91
	France	w	w	w	w
	Germany	46	14	114	0.75
	Greece	37	16	66	0.66
	Hungary	44	7	85	0.54
	Iceland	29	29	-5	0.85
	Ireland	39	28	48	0.79
	Italy	31	7	87	0.76
	Japan	39	5	133	0.76
	Korea	32	9	80	0.74
	Luxembourg	41	24	69	0.77
	Mexico	25	6	37	0.60
	Netherlands	44	11	123	0.78
	New Zealand	52	41	55	0.82
	Norway	36	31	29	0.88
	Poland	39	35	21	0.76
	Portugal	28	17	32	0.69
	Slovak Republic	45	21	56	0.63
	Spain	31	24	21	0.76
	Sweden	38	32	34	0.87
	Switzerland	44	26	70	0.82
	Turkey	31	9	65	0.69
	United Kingdom	48	32	71	0.83
	United States	49	34	51	0.74
	OECD total	45			
	OECD average	40	21	64	0.76
Partners	**Argentina**	38	13	57	0.61
	Azerbaijan	11	7	15	0.63
	Brazil	30	8	48	0.61
	Bulgaria	52	13	68	0.49
	Chile	38	11	54	0.47
	Colombia	23	11	31	0.60
	Croatia	34	14	83	0.78
	Estonia	31	22	42	0.81
	Hong Kong-China	26	9	64	0.76
	Indonesia	21	1	42	0.67
	Israel	43	26	69	0.76
	Jordan	27	18	28	0.75
	Kyrgyzstan	27	6	75	0.74
	Latvia	29	21	35	0.80
	Liechtenstein	49	c	c	c
	Lithuania	38	24	47	0.73
	Macao-China	13	7	15	0.67
	Montenegro	24	11	65	0.80
	Qatar	m	m	m	m
	Romania	35	12	60	0.66
	Russian Federation	32	20	39	0.76
	Serbia	33	12	75	0.74
	Slovenia	46	7	121	0.74
	Chinese Taipei	42	14	107	0.77
	Thailand	28	8	42	0.50
	Tunisia	19	4	36	0.64
	Uruguay	34	14	45	0.62

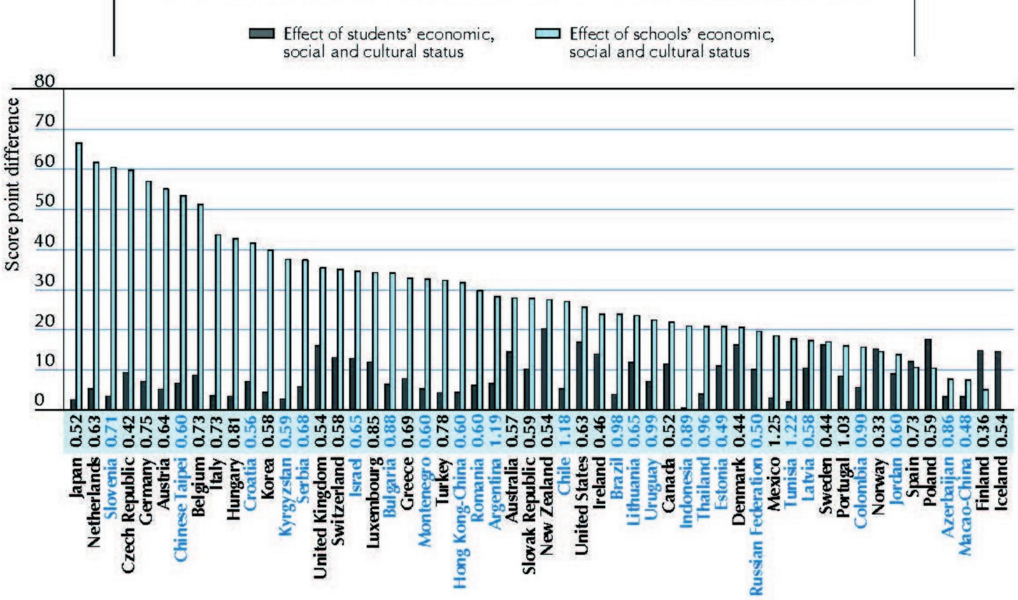
1. In some countries, sub-units within schools were sampled instead of schools as administrative units and this may affect the estimation of school-level effects.
2. Single-level bivariate regression of science performance on the ESCS, the slope is the regression coefficient for the ESCS.
3. Two-level regression of science performance on student ESCS and school mean ESCS: within-school slope for ESCS and variance explained by the model at the student level.
4. Two-level regression of science performance on student ESCS and school mean ESCS: between-school slope for ESCS and variance explained by the model at the school level.
5. The index of inclusion is derived from the intra-class correlation for ESCS as 1- the intra class correlation coefficient.

Source: OECD PISA 2006 database, Table 4.4b (as cited in OECD, 2007c: 192).

Figure 4.12

Effects of student's and schools' socio-economic background on student performance in science

Differences in performance on the science scale associated with one-half of a student-level standard deviation on the PISA index of economic, social and cultural status



Note: Data on blue background are values of the interquartile range of the school-level average PISA index of economic, social and cultural status.
Source: OECD PISA 2006 database, Table 4.4b (OECD, 2007c: 194).

References

Brown, G., Micklewright, J., Schnepf, S. V., & Waldmann, R. (2007). International surveys of educational achievement: How robust are the findings? *Journal of the Royal Statistical Society: Series A* (*Statistics in Society*), 170, 3: 623-646.

Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, J., Mood, A. M., Weinfeld, F. D., et al. (1966). *Equality of educational opportunity*. Washington, DC: Government Printing Office.

Ganzeboom, H. B. G., De Graaf, P. M., & Treiman, D. J. (1992). A standard international socio-economic index of occupational status. *Social Science Research*, 21, 1: 1-56.

Ganzeboom, H. B. G., & Treiman, D. J. (1996). Internationally comparable measures of occupational status for the 1988 international standard classification of occupations. *Social Science Research*, 25, 3: 201-239.

Ganzeboom, H. B. G., & Treiman, D. J. (2003). Three internationally standardised measures for comparative research on occupational status. In J. H. P. Hoffmeyer-Zlotnik & C. Wolf (Eds.), *Advances in cross-national comparison: A European working book for demographic and socio-economic variables* (pp. 159-193). New York: Kluwer Academic.

Hauser, R. M. (1969). Schools and the stratification process. *American Journal of Sociology*, 74, 6: 537-611.

Hauser, R. M. (1970). Context and consex: A cautionary tale. *American Journal of Sociology*, 75, 4: 645-664.

Hauser, R. M. (1972). *Socio-economic background and educational performance*. Washington, DC: American Sociological Association.

Hauser, R. M. (1973). Disaggregating a social-psychological model of educational attainment. In A. S. Goldberger & O. D. Duncan (Eds.). *Structural equation models in the social sciences* (pp. 255-284). New York: Seminar Press.

Hauser, R. M., & Goldberger, A. S. (1971). The treatment of unobservable variables in path analysis. In H. L. Costner (Ed.), *Sociological methodology 1971* (pp. 81-117). San Francisco: Jossey-Bass.

Hauser, R. M., & Goldberger, A. S. (1975). Correction of "the treatment of unobservable variables in path analysis." In D. R. Heise (Ed.), *Sociological methodology 1975* (pp. 212-213). San Francisco: Jossey-Bass.

Jöreskog, K. G., & Goldberger, A. S. (1975). Estimation of a model with multiple indicators and multiple causes of a single latent variable. *Journal of the American Statistical Association*, 70, 351: 631-639.

Mosteller, F., & Moynihan, D. P. (1972). *On equality of educational opportunity*. New York: Vintage Books.

Organisation for Economic Co-operation and Development. (2007a). *PISA 2006: Science competencies for tomorrow's world, Vol. 2, Data*. Paris: Author.

Organisation for Economic Co-operation and Development. (2007b). *PISA 2006: Science competencies for tomorrow's world, Vol. 1, Analysis*. Paris: Author.

Organisation for Economic Co-operation and Development. (2007c). Quality and equity in the performance of students and schools. In *PISA 2006: Science competencies for tomorrow's world*, *Vol. 1, Analysis* (pp. 169-212). Paris: Author.

Plowden, B. (1968). *Children and their primary schools: A report*. London: HMSO.

跨國教育研究中一些方法上的問題：學生學習表現的優與均

羅伯特‧豪瑟
Division of Behavioral and Social Sciences and Education
National Research Council, U.S. National Academy of Sciences
500 Fifth St. N.W., Washington, DC 20001, USA
E-mail: rhauser@nas.edu

Department of Sociology, University of Wisconsin-Madison
1180 Observatory Drive, Madison, Wisconsin 53706, USA
E-mail: hauser@ssc.wisc.edu
(黃敏雄譯)

摘要

本文探討在跨國資料中如何測量、分析及詮釋學生家庭社經背景與科學表現之間的關係，並針對2006年PISA所提出的研究報告做評論。PISA是由OECD所執行的一項跨國學生學習評量。OECD對於PISA評量結果所做的分析與詮釋言過其詞，對於學生科學成就與家庭背景之間的關係在學校之內或在學校之間的展現，和其跨國之間的差異，做過度的政策性解釋。PISA的家庭社經地位量表並沒有可詮釋的度量單位，也無法做嚴謹的跨國與跨年比較。OECD的分析未能辨別家庭社經地位量表裡各指標的獨立效果，也無法區辨這些效果在不同國家或在不同學科上的差異。OECD的分析沒有嘗試克服不同國家家庭社經背景資料的信度高低有所差異的問題，也未曾說明國與國之間在家庭社經地位差異程度上有顯著差別這其中的意涵。OCED過於粗略的分析導致我們無法正確評斷國與國之間在學生學習表現上的差異與造成這些差異的原因，為了使深富研究價值的PISA資料得到更完善的分析，本文提供一些實質建議。

關鍵詞： 學業成就、情境效應、國際學生能力評量計畫、學校效應、社經地位