Science competencies across PISA OECD countries: comparing exceptionally high and low performers

Fabio Alivernini
National Institute for the Evaluation of the Education System - Frascati (RM)
Italy
fabio.alivernini@invalsi.it

Sara Manganelli
National Institute for the Evaluation of the Education System - Frascati (RM)
Italy
sara.manganelli@invalsi.it

Abstract

In PISA 2006 the largest proficiency gap is between students who do not show the competencies that are necessary to participate effectively in life situations related to science and technology and students who have competencies which could allow them to create innovative technology. The objective of this paper is to identify, for PISA 2006 OECD countries, distinct subgroups of students who share characteristics that are mostly associated with this proficiency gap.

Data were based on the answers of those PISA 2006 OECD students with scores classified below level 2 (N=50762) and above level 4 (N=21665), as well as the answers at the school questionnaire of their principals and the OECD indicators of financial and human resources invested in education at a national level for secondary school (year of reference: 2005).

The dependent variable of the analysis was a dichotomous variable the values of which represent the two different groups of students. The independent variables were the OECD indicators, items and indices derived from the student and school questionnaires. The analysis was based on classification and regression trees (CART; Williams, Lee, Fisher, & Dickerman, 1999), a method suited to detecting and interpreting complex interactions in large data sets. A multilevel logistic regression model was subsequently computed to replicate the CART findings. Results show that the teachers' salaries in lower secondary education for teachers with over 15 years of experience, is the country level factor that has the greatest importance to predict if these fifteen year-olds belong to the group of the "first" or the "last" in OECD countries in relation to their future ability to participate in life situations involving problems of a scientific nature. Additionally the analysis conducted by means of CART made it possible to point out that the factors associated with the school, the family and the student interact in a complex way.

Keywords: science proficiency levels, explorative study, PISA 2006

Aims of the Study

In PISA 2006 student science performance is classified according to six proficiency levels. The largest proficiency gap is between students who are below level 2 and students who are above level 4. Is a matter of fact level 2 has been indicated (OECD, 2007) as the baseline level, the point at which students begin to show the competencies that are necessary to effectively participate in life situations related to science and technology. On the other hand, the top science proficiency levels 5 and 6, refer to students who have competencies which could allow them to create innovative technology.

The objective of this paper is to identify, for PISA 2006 OECD countries, distinct subgroups of students who share characteristics that are mostly associated with the proficiency gap.

Theoretical framework

The theoretical framework of the study is consistent with the PISA 2006 assessment framework (OECD, 2006) which considers a large number of contextual factors influencing students' performance in science. These factors are related to the following contexts:

a) the context of the education system at a national level (e.g. financial and human resources invested in education; presence of immigrant students);
b) the school context including the socio-economic background of school peers and other factors that (on the basis of previous research) may be associated with student achievement, such as school autonomy in decision making and in the definition of the educational planning, the quality of human and material resources, the nature of public or private control and funding, decision-making processes within the school, class discipline, class size, and others;
c) the individual student context that includes factors such as:

- the student economic, social and cultural background (father and mother profession, father and mother level education, economic resources as well as educational resources possessed at home);
- student level variables, such as student self-efficacy, self-regulated learning, types of motivation and goals;
- various aspects of students' lives, such as their attitudes towards learning, their behavior and lifestyle at school.

One should note that the theoretical framework does not identify a precise pattern of interactions between these variables and an explorative approach is therefore appropriate. It is also important that the explorative objective of the present paper includes the identification of interaction effects, not completely predictable *a priori*, among a large number of variables of different types (Nominal, Ordinal, and Interval).

Data sources

Data is based on:

- the answers of those PISA 2006 OECD students with scores classified below level 2 (N=50762) and above level 4 (N=21665), as well as the answers at the school questionnaire of their principals;
- OECD indicators of financial and human resources invested in education at a national level for secondary school (year of reference: 2005; OECD, 2008).
- OECD indicators related to the learning environment and organization of schools (year of reference: 2005; OECD, 2008).

The dependent variable of the analysis was a dichotomous variable the values of which represent the two different groups of students. The independent variables were the

OECD indicators, items and indices derived from the contextual questionnaires (student and school questionnaires).

Table 1 shows the country level variables taken into consideration for the present study.

Table 1. The Country Level Variables Taken into Consideration for the Present Study.

Source	Variables
OECD educational indicators related to financial and human resources invested in education (year of reference 2005; OECD, 2008)*	▪ expenditure on educational institutions per student; ▪ proportion of national wealth spent on education; ▪ relative proportions of public and private investment in education; ▪ total public expenditure in education; ▪ tuition fees charged by institutions and public subsidies to students; ▪ services and resources in which education funding are spent; ▪ how efficiently the resources are used in education.
OECD educational indicators (year of reference 2005; OECD, 2008) related to the learning environment and organization of schools*	▪ time that students spend in the classroom; ▪ ratio of students to teacher staff and average class size; ▪ teacher salaries; ▪ time that teachers spend teaching; ▪ impact of evaluation and assessments within education systems; ▪ level of decision making in education systems.
Variables from **PISA 2006 Student Questionnaire** aggregated (mean) at the country level**	Index of Economic, Social and Cultural Status (ESCS) created on the basis of the following variables: ▪ home possession index: a summary index of family wealth possessions (e.g., cellular phones,), cultural possessions (e.g., classic literature, paintings), educational resources (e.g., educational software), and number of books at home (but recoded into three categories: 0-25 books, 26-100 books, and 101 or more books); ▪ highest occupational status of parents: occupational data for both the student's parents were obtained by asking open- ended questions and responses were coded to four-digit ISCO codes (ILO,1990) and then mapped to the international socio-economic index of occupational status (ISEI; Ganzeboom et al., 1992). The index corresponds to the higher ISEI score of either parent or to the only available parent's ISEI; ▪ highest educational level of parents expressed as years of schooling: educational levels of parents were obtained by recoding educational qualifications into the ISCED categories. The index corresponds to the higher ISCED level of either parent recoded into estimated years of schooling;

** For a detailed description and statistics about the variables please refer to OECD (2008).*
*** For a detailed description and statistics about the variables please refer to OECD (2009).*

Table 2 describes the school level variables and indices from PISA 2006 School Questionnaire and Student Questionnaire taken into consideration in this study.

Table 2. The School Level Variables Taken into Consideration for the Present Study.

Variables and indices from PISA 2006 School Questionnaire aggregated (mean) at the school level*

Variable or Index	**Description***
School size	The total enrolment at school based on the enrolment data provided by the school principal.
Class size	It is derived from one of nine possible categories, ranging from “15 students or fewer” to “More than 50 students”, it takes the midpoint of each response category, a value of 13 for the lowest category, and a value of 53 for the highest.
Availability of computers	The number of computers available at school
Student-teacher ratio	It is computed by dividing the school size by the total number of teachers
Index of school selectivity	How much consideration was given to the students’ academic record and the recommendation of feeder schools (computed by assigning schools to four different categories from “schools where none of these factors is considered for student admittance” to “schools where at least one of these factors is a pre-requisite for student admittance”)
Index of school responsibility for resource allocation	It is derived from six items measuring the school principals’ report on who has considerable responsibility for tasks regarding school management of resource allocation (e.g., “Selecting teachers for hire”; “Formulating the school budget”)
Proportion of fully certified teachers	It is calculated by dividing the number of fully certified teachers by the total number of teachers
Proportion of teachers with an ISCED 5A qualification	It is obtained by dividing the number of these kinds of teachers by the total number of teachers
Index of school responsibility for curriculum and assessment	It is obtained from four items measuring the school principal’s report concerning who had responsibility for curriculum and assessment (e.g., “Establishing student assessment policies”, “Choosing which textbooks are used”)
Index of teacher shortage	It is computed on the basis of four items measuring the school principal’s perceptions about how much the school’s capacity to provide instruction was hindered by the shortage of teachers.
Index of quality of educational resources	It is derived from seven items measuring the school principal’s perceptions of potential factors hindering instruction at school (e.g., “Shortage or inadequacy of science laboratory equipment”)
Index of school activities to promote students’ learning of science	It is computed on the basis of principal’s reports about school’s involvement in science activities (e.g., “Science clubs”, “Excursions and field trips”)
School activities for learning environmental topics	It is derived from principal’s reports on the occurrence at school of activities to promote students’ learning of environmental topics (e.g., Trips to museums, Extracurricular environmental projects)
Parental pressure on academic standards	It is derived from principals’ report about parental expectations towards the school in terms to set very high academic standards and to have the students achieve them

Variables and indices from PISA 2006 Student Questionnaire aggregated (mean) at the school level*

Index of Economic, Social and Cultural Status – ESCS	See the index description in Table 1

** For a detailed description and statistics about the variables please refer to OECD (2009).*

Table 3 describes the student level variables and indices from PISA 2006 Student Questionnaire taken into consideration in this study.

Table 3. The Student Level Variables Taken into Consideration for the Present Study.

Variables and indices from PISA 2006 Student Questionnaire*

Variable or Index	Description*
Index of Economic, Social and Cultural Status – ESCS	See the index description in Table 1
Index of interest in science learning	It is derived from eight items measuring student's interest about broad science topics (e.g., "topics in physics", "the biology of plants")
Index of enjoyment of science	It is computed on the basis of four items measuring student's enjoyment of science learning (e.g., "I enjoy acquiring new knowledge in broad science", "I like reading about broad science")
Index of instrumental motivation to learn science	It is derived from five item measuring student's motivation to learn science (e.g., I study school science because I know it is useful for me, "I will learn many things in my school science subject(s) that will help me get a job")
Index of future-oriented science motivation	It is computed on the basis of four items measuring expectations about science-related studies and careers (e.g., "I would like to work in a career involving broad science", "I would like to study broad science after secondary school")
Index of science self-efficacy	It is derived from eight items measuring student's confidence in performing science-related tasks (e.g., "Describe the role of antibiotics in the treatment of disease", "Interpret the scientific information provided on the labeling of food items")
Index of science self-concept	It is computed on the basis of six items about student's opinion about himself/herself (e.g., "I learn school science topics quickly", "I can easily understand new ideas in school science")
Index of general value of science	It is derived from five items measuring student's perceptions on the general value of science (e.g., "Broad science is valuable to society", "Broad science is important for helping us to understand the natural world")
Index of personal value of science	It is derived from five items measuring student's perceptions of the personal value of science (e.g., "Broad science is very relevant to me"; "I find that broad science helps me to understand the things around me")
Index of science-related activities	It is computed on the basis of six items measuring the frequency of student's participation activities related to science (e.g., "Watch TV programs about broad science", "Borrow or buy books on broad science topics")
Index of awareness of environmental issues	It is derived from five items about student's report about how much he/she is informed about several environmental issues (e.g., "nuclear waste", "acid rain")
Index of perception of environmental issues	It is derived from six items measuring the concern of the student about several environmental issues (e.g., "air pollution", "energy shortage")
Index of environmental optimism	It is derived from six items measuring student's perceptions about the improvement of problems related to environmental issues (e.g., "air pollution", "water shortage")
Index of responsibility for sustainable development	It is derived from seven item measuring student's support for sustainable development (e.g., "I am in favor of having laws that regulate factory emissions even if this would increase the price of products", "To reduce waste, the use of plastic packaging should be kept to a minimum")
Index of school preparation for science career	It is derived from four items measuring students' perceptions of the usefulness of schooling as preparation for science-related careers (e.g., "The subjects I study provide me with the basic skills and knowledge for a science-related career", "My teachers equip me with the basic skills and knowledge I need for a science-related career")
Index of student information on science careers	It is derived from four items measuring how much students' are informed about aspects of science-related careers (e.g., "Science-related careers that are available in the job market", "Where to find information about science-related careers")
Index of science teaching – interaction	It is computed on the basis of four items measuring students' reports on the frequency of interactive teaching in science (e.g., "Students are given opportunities to explain their ideas", "The students have discussions about the topics")
Index of science teaching - hands-on activities	It is computed on the basis of four items measuring students' reports on the frequency of hands-on activities (e.g., "Students spend time in the laboratory doing practical experiments", "Students are required to design how a school science question could be investigated in the laboratory")
Index of science teaching - student investigations	It is derived from three items measuring students' reports on the frequency of student investigations in science (e.g., "Students are allowed to design their own experiments", "Students are asked to do an investigation to test out their own ideas")
Index of science teaching - focus on models or applications	It is derived from five items measuring students' reports on the frequency of teaching in science lessons with a focus on applications (e.g., "The teacher uses science to help students understand the world outside school", "The teacher clearly explains the relevance of broad science concepts to our lives")

** For a detailed description and statistics about the variables please refer to OECD (2009).*

Methods

The analysis was based on classification and regression trees (CART) (Williams, et al., 1999), a method suited to detecting and interpreting complex interactions in large data sets that most traditional means of regression and classification analysis might ignore or find difficult to estimate and interpret (Allore et al., 2005). In fact tree methods can discover interactions during the growth of the tree whereas traditional regression techniques require *a priori* specification of interactions. It is important to note that CART is not affected by problems of multi-collinearity between predictors and that it is a truly non-parametric method, since it makes no assumptions regarding the underlying distribution from which the subjects are sampled. In addition, since classification trees derive their predictions from a few "if-then" conditions, a straightforward interpretation of results is possible in terms of the distinctive characteristics of the different groups. The CART algorithm proceeds by performing successive binary divisions of the subjects on the basis of a statistical criterion. Starting from the full sample (called root node or parent node) each independent variable is evaluated on the basis of the extent to which it is able to reduce the impurity of the parent node by dividing the subjects into two groups (called child nodes). The impurity consists in the degree to which the students at a node vary compared with the dependent variable: a minor impurity indicates a greater homogeneity of the subjects for the values of the dependent variable. In the case under examination a completely pure node would be one which includes students who only belong to the highest levels of skills or only belong to the lowest levels of skills. The end result of the process of partition is a tree organized in a hierarchical way, in which the root is the overall sample of individuals, the branches are the values of independent variables used in the analysis and the nodes are the subsets of individuals identified by some combination of values of the independent variables. If the nodes are located in a terminal position (i.e. they are not further split) they are called *leaves*.

Each individual is classified by following a pathway along the tree, leading from the root to a leaf.

The analysis was conducted using a hierarchical approach (Hox, 2002; Fabbris, 1997) in three stages. We developed a model:

1) with only country level variables;
2) with school level variables nested under the country model identified at stage 1;
3) with student variables nested under the country and school model identified at stage 2.

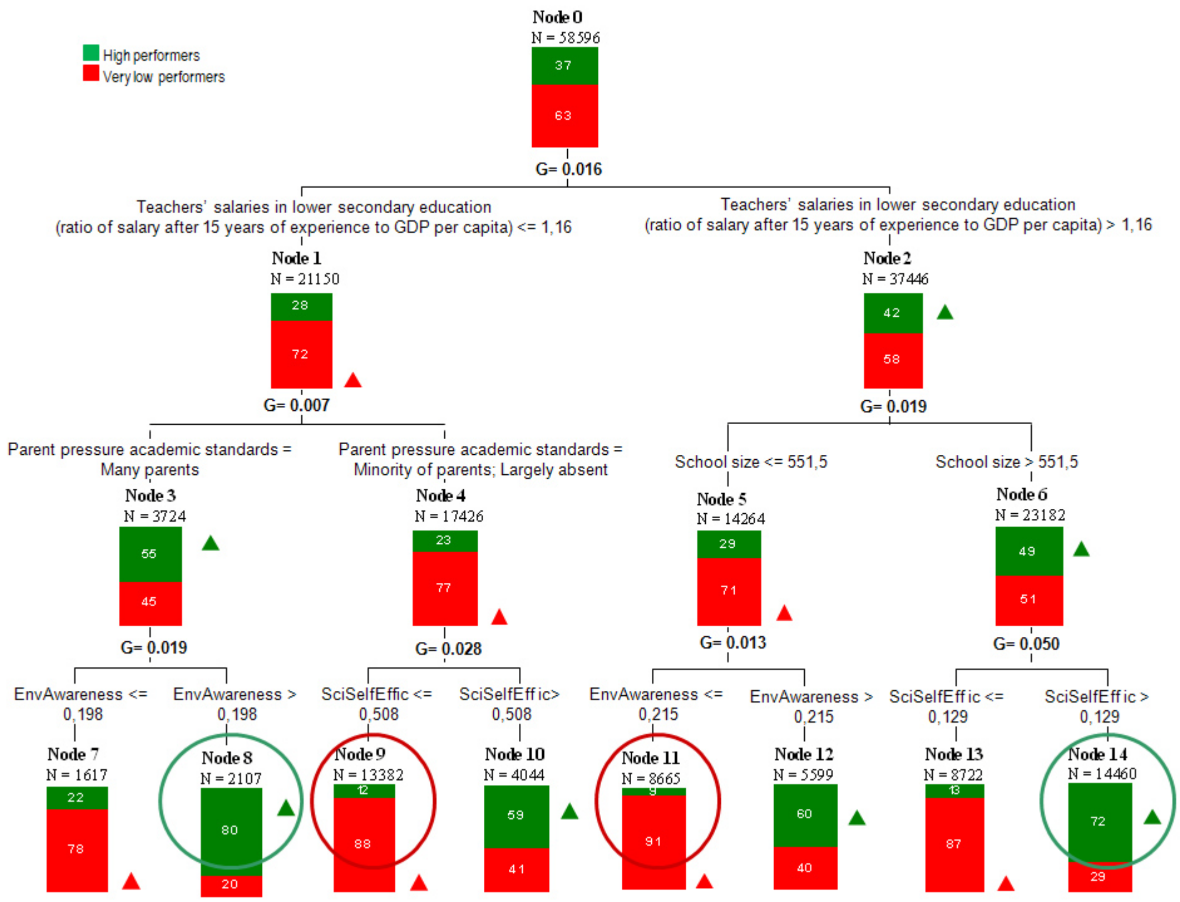
The classification model was developed on a random subset of the data (training sample) and then the results were validated on a separate random sample (test sample). The accuracy model was estimated using cross validation techniques (Breiman et al., 1984).

Additionally, in order to validate the model by means of more traditional techniques, once the CART model was created, a multilevel logistic regression model was computed to replicate the findings.

Results

Figure 1 shows the final tree produced by CART in order to identify the segments of students with the greatest disparities in their science proficiency levels.

Figure 1. The Final Tree Produced by CART (High Performer Students in Green, Very Low Performer Students in Red)



In order to replicate and extend the findings from the classification tree, a multilevel logistic regression model was tested. The following indicator variables (Table 4) represented in the logistic model the classification tree splits with the highest impurity reduction (i.e. divisions which led to a group with a high percentage of students who perform below the baseline level or at the top level).

Table 4. Indicator Variables in the Logistic Model

Indicator	Coding
Teacher Salaries	0 = Teacher Salaries<=1,16
School Size	0 = School Size<=551
Parent pressure	0 = Minority of parents or Largely absent
Awareness of environmental issues	0 = Awareness<=0,198
Science Self-Efficacy	0 = Science Self-Efficacy <=0,508

In the multilevel logistic model the Index of Economic, Social and Cultural Status (ESCS) was included at each level as a control variable. Table 5 shows the model specification.

Table 5. Multilevel Logistic Model Specification

Level-1 Model

Prob(Y=1|B) = P

log[P/(1-P)] = P0 + P1*(ESCS) + P2*(I Awareness of environmental issues) + P3*(I Science Self-Efficacy)

Level-2 Model

P0 = B00 + B01*(I School Size) + B02*(I Parent Pressure) + B03*(ESCS School Mean) + R0

Level-3 Model

B00 = G000 + G001(I Teacher Salaries) + G002(ESCS Country Mean) + U00

Results show that the findings from the classification tree have been replicated in the multilevel logistic model. In Table 6 coefficients, P-values and Odd Ratios from the results of the multilevel logistic model are reported.

Table 6. Results of the Multilevel Logistic Model

Indicator	Coefficient	Odds Ratio
Country level		
Teacher Salaries	0.90*	2.45
ESCS Country Mean	0.27	1.31
School Level		
School Size	0.27*	1.31
Parent pressure	0.33*	1.40
ESCS School Mean	1.6*	4.94
Student Level		
Science Self-Efficacy	1.44*	4.24
Awareness of environmental issues	1.61*	5.04
ESCS	0.5*	1.65

**p<.001*

Discussion

The analysis carried out in this study used an exploratory and hierarchical approach. In other words pre-formulated hypotheses were not tested, but a data driven model based on the theoretical framework of PISA 2006 (OECD, 2006), was developed. The multilevel structure of PISA 2006 data was taken into consideration by means of a hierarchical analysis of the country variables, school/teacher variables and parent/student variables. In order to assure the reproducibility of the findings, the classification model was developed on a training sample and tested on a control sample. Additionally the findings from the classification tree were replicated and extended by means of a multilevel logistic regression model.

Results showed that the country level variable that makes the greatest difference as regards students' performance groups is the teachers' salaries in lower secondary education for teachers with over 15 years of experience. Empirical evidence on the relationship between teachers' salaries and student performance is poor (Béteille and Loeb, 2009) and most of the research conducted has involved the comparison of states within the United States. The results of the present study, conducted on PISA OECD countries, suggest the hypothesis that teachers' salaries could be a relevant factor in order to explain large performances' gaps in science. If we look at the data from PISA 2006, in those countries where teacher salaries are higher (OECD, 2007) the students are 2.45 times likely to be top science performers.

As regards the variables relating to the characteristics of schools, a clear distinction can be made between students at the top level and students below the base level on the basis of the parental expectations towards the school in terms to set very high academic standards and to have the students achieve them. Results show that also the total enrolment at school is a variable relevant at the school level. Although the effect of school size is quite controversial in literature (Ahn and Brewer, 2009), this study suggests that there could be a cut-off point in investigating this phenomenon: those students who are in schools where the total enrolment exceeds 550 are more likely to be top science performers.

Student level variables that play an important role concerning the student performance groups are the student awareness of environmental issues and the science self-efficacy. Those results are consistent with previous findings (e.g., OECD, 2007) which showed a strong association between these two indices and science performance. Moreover, the present study shows that for a student is sufficient to be slightly above average (1/5 of a standard deviation) for the awareness of environmental threats and moderately (1/2 of a standard deviation) above average for science self-efficacy to be respectively 5 times and 4 times more likely to be a top performer.

In summary, this study provides a complete and detailed description of students with great disparities in science proficiency that is both informative and readily interpretable. The results are presented in a simple format, providing additional insights by pointing out the order in which different questions should be asked. In conclusion one should note that this study does not aim to explain the specific reasons why the phenomena that have been identified occur. Instead, rather than speculating on the causes of these phenomena, it provides a rich and reliable description of their characteristics. This description could be useful since it suggests a number of issues for further investigation and research.

References

Ahn, J., and Brewer, D. J. (2009). *What do we know about reducing class and school size?* In G. Sykes, B. Schneider & D. N. Plank (Eds.), AERA Handbook of Education Policy Research (pp. 426-437). New York: Routledge.

Allore, H., Tinetti, M. E., Araujo, K. L. B., Hardy, S., & Peduzzi, P. (2005). A case study found that a regression tree outperformed multiple linear regression in predicting the relationship between impairments and Social and Productive Activities scores. *Journal of Clinical Epidemiology*, *58*, 154–161.

Béteille, T., and Loeb, S. (2009). *Teacher quality and teacher labor markets*. In G. Sykes, B. Schneider & D. N. Plank (Eds.), AERA Handbook of Education Policy Research (pp. 426-437). New York: Routledge.

Breiman, L, Friedman, J., Olshen. R., & Stone, C. (1984). *Classification and Regression Trees*. Wadsworth International Group.

Fabbris, L. (1997). *Statistica multivariata: analisi esplorativa dei dati*. Milano: McGraw Hill.

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*, 1-56.

Hox, J. (2002). *Multilevel analysis: Techniques and applications.* Mahwah, NJ: Erbaum.

International Labour Organisation (ILO). (1990). *International Standard Classification of Occupations: ISCO-88*. Geneva: International Labour Office.

OECD (2006), *Assessing Scientific, Reading and Mathematical Literacy – A Framework for PISA 2006*, Paris: OECD.

OECD (2007). *PISA 2006 Science competencies for tomorrow's world*. Paris: OECD.

OECD (2008). *Education at glance 2008: OECD Indicators*. Paris:OECD.

OECD (2009). *PISA 2006 Technical Report*. Paris:OECD

Williams, C. J., Lee, S. S., Fisher, R. A. & Dickerman, L. H. (1999). A comparison of statistical methods for prenatal screening for down syndrome. *Applied Stochastic Models in Business and Industry, 15*, 89-101.