Determining the Relationships between Selected Variables and Latent Classes in Students’ PISA Achievement

Seher Yalcin
Ankara University

To cite this article:


This article may be used for research, teaching, and private study purposes.

Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden.

Authors alone are responsible for the contents of their articles. The journal owns the copyright of the articles.

The publisher shall not be liable for any loss, actions, claims, proceedings, demand, or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of the research material.
Determining the Relationships between Selected Variables and Latent Classes in Students’ PISA Achievement

Seher Yalcin

Abstract
The purpose of this study was to identify the multilevel latent classes for reading, mathematics, and science success of the students, who participated in the Programme for International Student Assessment (PISA) 2012 from Turkey and to determine the predictive ability of i) students’ perseverance, ii) their openness to problem solving, iii) their economic, social, and cultural status (ESCS), and iv) resources of school in relation to the determined student and school classes using a multilevel approach. The population of this research was school principals and all the 15-year-old students, who attended PISA 2012 in Turkey. Analyses were conducted with the data obtained from a total of 3,196 students and 169 school principals. In the first step, a multilevel latent class analysis was used to investigate the number of these classes in schools reading, mathematics, and science success of the students. Then, a three-step analysis was undertaken to determine the predictive ability of the chosen variables for the identified classes. The results indicated that all the chosen variables significantly predicted the school-level latent class membership. Moreover, analyses suggested that the students’ ESCS was the most important factor affecting their achievement.

Keywords
Multilevel latent class analysis
PISA
Reading
Mathematics and science achievements

Introduction
Students’ academic achievement is perceived as an indicator of the efficiency of the education system in many countries. One of the approaches to determining student achievement is large-scale applications. The results of a large-scale international evaluation affect the decisions taken by policy makers all around the world (Adamson, 2012; Kirsch, Lennon, von Davier, Gonzalez, and Yamamoto, 2013), particularly in developing countries. An example of such applications is the Programme for International Student Assessment (PISA), which is conducted every three years to evaluate students’ skills and knowledge in the fields of mathematics, science, and reading skills (Organisation for Economic Co-operation and Development [OECD], 2014a). In addition to these general fields, a different focal subject is chosen for each assessment.

As a developing country, Turkey first participated in PISA in 2003 to compare the existing education system with that of other countries and to evaluate student success in an international dimension. According to the results of PISA 2003, students scored 423, 441, and 434 points in reading, mathematics, and science fields, respectively (Ministry of National Education [MoNE], 2005). A similar distribution was observed in PISA 2006 with students scoring 424, 447, and 424 points in the respective fields (MoNE, 2007). There was a slight increase in students’ scores in PISA 2009 with the scores being 464, 445, and 454 points for reading, mathematics, and science, respectively (MoNE, 2010). A similar increase was observed in students’ achievement in PISA 2012 with the students scoring 448, 475, and 463 in the respective fields (Yıldırım, Yıldırım, Ceylan, and Yetişir, 2013). However, despite the increase in the average points of Turkey, there has not been a significant development in its ranking among both OECD countries and all countries, and student percentage at sub-proficiency levels has not improved. Therefore, it is important to determine the reasons why student success in Turkey has not significantly improved in order to take necessary actions. One of the possible causes for the lower ranking of Turkey is the economic situation. Turkey spends only 4% of its gross domestic products (GDP) on educational institutions at all educational levels, compared to an average of 6% for the OECD countries (OECD, 2014b). Another reason may be due to the PISA assessment being skill-based and Turkish students not having a high level of skills. This is clearly seen in the results of PISA 2015, which show that students had the lowest scores over the 12-year period; 428, 420, and 425 points in reading, mathematics, and science, respectively. In 2015, a computer-based evaluation was undertaken for the first time and students’
Factors that affect student achievement during the educational process demonstrate a multivariate structure. The most important components of this structure are student’s economic, social, and cultural status (ESCS) (Dinçer and Kołaşin, 2009; Ferrera, Cebada, Chaparro, and González, 2011; Finch and Marchant, 2013; Geske, Grinfelds, Dedze, and Zhang, 2006; Nonoyama, 2005; Wolfram, 2005; Xu, 2006), and problem-solving skills (Kutlu, Doğan, and Karakaya, 2008; OECD, 2014c; Scherer and Gustafsson, 2015). In Turkey, students’ success in the education area is mostly affected by their living conditions. The quality of education that students receive in Turkey is heavily influenced by the levels of education and income of their parents. This situation causes inequality of opportunity among students (Aslankurt, 2013). The study conducted by Yıldırım (2009) concerning the PISA 2006 results pointed out that the main elements determining the quality of education in Turkey were the occupation and education of the mother and father including the socio-economic and socio-cultural factors and resources they had at home. Sarier (2016) conducted a meta-analysis of 62 studies in order to specify the factors that affected student success in Turkey, and confirmed that socio-economic status was one of the most influential factors.

In order to underline the increasing importance of problem-solving skills in PISA 2012, the cognitive ability was measured using two new indexes; namely, openness towards problem-solving and perseverance (OECD, 2014c). Problem-solving is generally defined as a set of cognitive steps, which guide the person towards a solution. In addition to other cognitive behaviours such as memorising, understanding, and critical thinking, it also involves creativity when problems have multiple solutions (Haladyna, 1997). In PISA, this skill is defined as being open to problem-solving, being able to solve complex problems and a desire to explain the process (OECD, 2013a). In PISA 2012, students who were found to be more open to problem solving were from Thailand, Indonesia, and Malaysia that all had higher mathematics achievements (Thien, Darmawan and Ong, 2015). Similar findings were obtained in the studies by Demir (2005) and Yavuz, İlgün Dibek and Yalçın (2017) with students, who participated in PISA 2012 in Turkey. Another feature that plays an important role in student success and is related to problem-solving (Scherer and Gustafsson, 2015) is perseverance (OECD, 2013), which is the continuation of assiduous efforts for the achievement of an aim despite hardships and obstacles (Middleton, Tallman, Hatfield, and Davis 2015). The results of numerous studies in the field literature also suggest a positive relationship between students’ achievement and perseverance (Arikан, 2014; Chiu and Xihua, 2008; Demir, 2015; Duckworth, Peterson, Matthews, and Kelly, 2007). However, most of these studies focus on the problem-solving ability of students in relation to their mathematics achievement. To the best of my knowledge, the relationship of such skills with students’ achievements in reading and science as well as mathematics has not been examined. Therefore, it is of critical importance to evaluate students’ achievements in all three fields and address their relationship with students’ problem-solving skills.

Student achievement can be affected by various factors related to school as well as students’ personal characteristics. When the average achievement scores of different types of schools in Turkey are examined, it is seen that there is still a significant difference between schools (Berberoğlu and Kalender, 2005; Yalçın and Tavşancıl, 2014; Yildirim et al., 2013). This situation underlines the necessity of considering the effect of school features on student success. In the literature regarding school features, it is confirmed that the quality of school and educational resources (Acar and Öğretmen, 2012; Archibald, 2006; Nonoyama, 2005; Oral and Mcgivney, 2013; Özer-Ozkan, 2016) affect students’ achievements. Hanushek (1997) conducted a meta-analysis of 377 studies and reported that taking family characteristics into consideration, there was no strong and consistent relationship between student achievement and school resources. This situation indicates the possibility of another influential variable that cannot be explicitly observed due to being in the background of the relationship between the variables addressed. Different modelling studies are needed to further analyse such hidden but influential relationships.

Researchers generally focus on the relationship between two observable variables; however, some relationships cannot be measured directly under emerging relationships. When variable(s) that cannot be measured directly are identified and controlled, it is seen that the relationship between the two variables disappears. In order to detect these variable(s) with the help of patterns in the changes observed, the latent class analysis (LCA) method has been developed (Vermunt and Magidson, 2004).

Students’ achievements, attitudes, and motives are considered to be latent variables since they can be neither observed nor measured directly. The use of the concept of latent variable dates back to 1904, when Charles Spearman developed factor analysis models for continuous variables in intelligence tests (Kane and Brand
2003). These variables can only be measured via observable variables such as systematic and random errors; however, latent variables are assumed to have no error.

LCA is used to construct homogenous sub-classes from heterogeneous latent traits (Vermunt and Magidson, 2002) and individual observations that are free from other observations and the determined latent class memberships. In most applications in the educational fields, people (level-1) are sampled from clusters such as classrooms or societies (level-2). This situation results in correlations between observations from the same cluster (Asparouhov and Muthen, 2008). For these reasons, it is recommended that LCA is applied to multi-level models, in which it is accepted that membership possibilities and/or item response possibilities can change randomly between classes (Vermunt, 2003).

It has been known for many years that globally student achievement is lower than the desired level and therefore researchers have made suggestions to increase student achievement by focusing on its relations with various variables. In these studies, different models have been developed based on the latest technologies and building on previous work to statistically minimise errors. Multilevel approaches adopted in these models can be grouped into predictive regression (Anıl, 2009; Gülloğlu, Bilican Demir, and Demirtaşlı, 2014; Thorpe, 2006; Tomul and Çelik, 2009; Xu, 2006; Yıldırım, 2009), structural equation (Akyüz and Pala, 2010; Anıl, 2008, Özer and Anıl, 2011; Usta and Çıkrıkçı-Demirtaşlı, 2014), and hierarchical linear models (Akyüz, 2014; Atar, 2014; Atar and Atar, 2012; Geske et al., 2006; Tavşancıl and Yalçın, 2015). However, in the literature, there are only a limited number of studies which have evaluated student achievement as a latent variable and analysed it based on patterns identified in variables by creating homogenous sub-classes (Finch and Marchant, 2013; Lin and Tai, 2015). In this type of work, analysis is conducted by determining latent classes as well as including socio-economic level (SEL) or learning strategies variables in the model to make direct classifications.

In contrast to previous research, in this study, predictor variables were included in the model using a three-step analysis. In the literature, it is stated that there are disadvantages to including predictors in models when determining the number of latent classes and it is suggested that a three-step analysis is employed in order to compensate for these disadvantages (Gudicha and Vermunt, 2013; Vermunt, 2010). For this reason, this study is significant both for the way it departs from the work reported in the literature in terms of the process of including predictor variables in the model, and presenting the model used in the work. Therefore, this study is considered to remedy the deficiency in the existing literature.

Dividing student achievement into homogenous classes in itself makes it possible to determine the factors related to students with various levels of achievement and thus provides more detailed data regarding student achievement and allows making suggestions for the intended classes of students. For example, as commonly reported in the literature, if a student that has a high level of variable A is successful, when his/her achievements are classified and analysed using LCA, further data may suggest that not all students with a high level of A are successful; furthermore, those with a low level of A may also be successful albeit to a lower extent. Moreover, deductions can be made suggesting that hereditary factors play a greater role in the achievements of students with very high achievements rather than environmental factors that are analysed. In this sense, rather than approaching student achievement as a whole, examining relations by separating achievement into latent classes provides the opportunity to overcome stereotypes and helps determine more accurately what qualities students need or do not need.

**Education System in Turkey**

Upon examining the Turkish education system, it is seen that the learning approach adopted in the education programmes since 2005 requires teachers to implement a constructivist learning approach. “The central principles of this approach are that learners can only make sense of new situations in terms of their existing understanding. Learning involves an active process, in which learners construct meaning by linking new ideas to knowledge” (Naylor and Keogh 1999, p. 93). Constructivists acknowledge the central role of the learner and structure classroom experiences. These teachers seek and value their students’ points of view and assess student learning in the context of daily teaching (Brooks and Brooks, 1999). In addition to the constructivist approach, the Turkish education programmes aim to develop following skills in students: Critical and creative thinking, communication, problem-solving, effective use of information technologies, and using Turkish accurately and efficiently (MoNE, 2005).

The right to free education is given to every citizen in the constitution of the Republic of Turkey. Both girls and boys are required to attend twelve years of compulsory education divided into three parts (4 + 4 + 4 years) since
2012 (MoNE, 2012). MoNE is responsible for overseeing the management of education in Turkey, which includes setting curriculum, organising official, private and voluntary organisations’ responsibility, planning and building schools, and improving educational materials. The formal education provided in the Turkish National Education System comprises of pre-primary, primary, secondary and higher education.

Compulsory primary education starts at 5.5 years old; however, for children aged 3 to 5 years, there is the non-mandatory option of pre-primary education. Public schools provide eight years of education embracing the first two 4s (4 + 4). In addition, there are private and paid schools under the supervision of the state. After completing eight years, students who are successful can be accepted to the four-year secondary education programme (MoNE, 2012). If students are successful in the Transition from Basic to Secondary Education examination, they can be accepted by their chosen high schools. There are several types of high schools such as general, science, vocational, and technical, all providing compulsory four years of education. The students that participated in the PISA application are generally sophomore students attending these secondary schools. After the 12 years compulsory education, students who graduate from high school can be accepted by the universities on condition that they pass the university admission exams.

Objectives of the Study

The objectives of this study were to determine multilevel latent classes in reading, mathematics, and science achievements of students, who participated in PISA 2012 from Turkey and to determine the extent to which these latent classes can be predicted by i) students’ perseverance, ii) their openness to problem-solving, iii) their ESCS, and iv) school resources (physical infrastructure, teaching staff, and educational resources). In this context, the following research questions were formulated:

1. How many latent classes do students’ reading, mathematics, and science achievements in PISA 2012 have at student and school levels?

2. Do students’ perseverance, openness to problem-solving, ESCS, and school resources predict the identified latent classes?

Method

Population and Sample

This study was based on a descriptive survey model. The population of the study was 15-year-old students, who participated in PISA 2012 from Turkey and the principals of the schools attended by these students. The PISA student questionnaire consists of four forms; A, B, C and UH (OECD, 2013b). Since the student-level variables chosen in this study were related to the content of the A and B forms of the questionnaire; students who had completed these forms were selected. Appendix 1 presents the items chosen from the student questionnaire regarding students’ perseverance, openness towards problem-solving, social and cultural status with their reliability coefficients.

For the school-level variables, the school questionnaire completed by school principals was used. Appendix 1 presents the items chosen from the school questionnaire related to school resources; namely physical infrastructure, teaching staff, and educational resources with their reliability coefficients. Furthermore, Turkey’s Cronbach’s alpha regarding each index variable and the OECD median are given.

One of the schools was excluded from analysis since there was no questionnaire data was available at the school level. As a result, analyses were conducted with the data from 3,196 students (1,578 girls and 1,618 boys) and 169 school principals. Different weightings were used for the sample to represent the population in transnational and domestic comparisons in PISA. Since the student-level variables were required to represent the national results (OECD, 2014d), the final weight at the student level was utilised (W_FSTUWT). Schools were chosen randomly; thus, no weighting was used at the school level.
Data Collection Tool

The data regarding the students’ scores in reading, mathematics, and science literacy tests in PISA 2012 and two separate questionnaires for students and school principals were obtained from PISA’s international website (http://www.oecd.org/pisa/data/pisa2012database-downloadabledata.htm).

Data Analysis

At the first stage of the analysis, a multilevel LCA (MLCA) was used to identify the number of latent classes in students’ reading, mathematics, and science achievements at the student and school levels. MLCA is mostly preferred when people’s multi-item responses or repeated measures are nested (Bijmolt, Paas, and Vermunt, 2004; Vermunt, 2003, 2008).

In conventional LCA, model parameters are assumed to be the same for all individuals whereas in MLCA, some of these parameters can change between classes or clusters. This makes it possible to investigate how level 2 affects level 1 indicators, which describe the membership of latent classes. MLCA also lets the assessment of conceptual predictors (level 2) and expands the assessment of explanatory variables at an individual level (Henry and Muthen, 2010).

Let $h$ represent a specific higher-level group, $H$ the number of groups, and $n_h$ the number of individuals in group $h$. The index $h$ in used in $v_h$, $y_{hij}$, and $y_{hi}$ to demonstrate that it is about a quantity of an individual $i$ belonging to group $h$. Furthermore, $y_{hij}$ is employed to point to the item responses of all members of the higher-level unit $h$. There is a model which can be referred to as the basic variant of the multilevel latent class model, and the class membership probabilities vary across this model. That is, observations inside groups are considered to be dependent since there is a considerable possibility that they are part of the same latent class. Below is the formulated version of this model (Vermunt, 2016):

$$ P_h(y_{hi}) = \sum_{c=1}^{C} P_h(v_{ih} = c) P(y_{hi} | v_{hi} = c), $$

where the index $h$ in $P_h(\cdot)$ shows that probabilities are group-specific, and for the presented models, $P(y_{hi} | v_{hi} = c)$ has the following form:

$$ P(y_{hi} | v_{hi} = c) = \prod_{j=1}^{I} \pi_{cij}^{y_{hij}} (1 - \pi_{cij})^{1-y_{hij}}, $$

where the class-specific response probabilities $\pi_{cij}$ correspond to a probabilistic model, which was suggested by Guttman. The measurement model is accepted as being homogeneous across groups, and $\pi_{cij}$ values not having an index $h$ can be considered as evidence for it.

In one variant of the MLCA, Vermunt (2003) suggested modelling based on random-effects logistic regression, which is expressed as follows:

$$ \log \frac{P_h(y_{hi} = c)}{P_h(y_{hi} = c)} = \alpha_c + \tau_h u_h, \text{ for } c < C, $$

where $u_h$ is a normally distributed random effect with a mean value equal to 0 and a variance equal to 1.

Parameter prediction in is an iterative process, which involves the use of expectation maximisation and the Newton-Raphson algorithm. The simplest model with the minimum number of latent classes and the least predictive parameter is chosen (Vermunt, 2003; Vermunt and Magidson, 2004). In this study, to obtain the optimal number of clusters, log-likelihood (LL), Akaike information criterion (AIC), AIC3 and the Bayesian information criterion (BIC) were used. The model was chosen based on the results of the study of Lukočienė, Varriale, and Vermunt (2010), which includes a simulation demonstrating why BIC is the right fit index. Thus, the BIC value was used as a criterion for model selection.

At the second stage, a three-step analysis was undertaken to determine the ability of the chosen variables at student and school levels to predict the emerging multilevel latent classes. After assigning the average of the student-level variables to classes, the relation between these variable and the latent classes at the school level were determined (Bennink, Croon, and Vermunt, 2013). Latent Gold 5.1 package program was used for data analysis (Vermunt and Magidson, 2013).
Results

Latent Student and School Classes

First, MLCA was conducted taking the multilevel structure of the dataset into consideration. In order to determine the possible latent classes at the student and school levels, a total of 100 probabilistic models (0 to 10 clusters at each level) were tested. The model that best fits the data is preferred. Table 1 presents LL, BIC values, and number of parameters regarding latent classes that emerged at the student and school levels.

Table 1. Goodness-of-fit measures of the models

<table>
<thead>
<tr>
<th>Models</th>
<th>LL</th>
<th>BIC (LL)</th>
<th>Npar</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model75</td>
<td>-48438.803</td>
<td>97578.6553</td>
<td>87</td>
</tr>
<tr>
<td>Model76</td>
<td>-48375.684</td>
<td>97516.8816</td>
<td>95</td>
</tr>
<tr>
<td>Model77</td>
<td>-48371.366</td>
<td>97572.7109</td>
<td>103</td>
</tr>
<tr>
<td>Model85</td>
<td>-48389.982</td>
<td>97569.6508</td>
<td>98</td>
</tr>
<tr>
<td>Model86</td>
<td>-48325.656</td>
<td>97513.5225</td>
<td>107</td>
</tr>
<tr>
<td>Model87</td>
<td>-48312.192</td>
<td>97559.1155</td>
<td>116</td>
</tr>
<tr>
<td>Model89</td>
<td>-48373.626</td>
<td>97625.5780</td>
<td>109</td>
</tr>
<tr>
<td>Model91</td>
<td>-48333.729</td>
<td>97626.3638</td>
<td>119</td>
</tr>
<tr>
<td>Model97</td>
<td>-48273.894</td>
<td>97587.2757</td>
<td>129</td>
</tr>
</tbody>
</table>

According to the results of MLCA, as Lukočienė and others (2010) suggested, the best results were produced by the model with 9 clusters at the student level and the model with 6 classes at the school level (9Cluster-6GClass model). Table 2 gives the average scores of the students by class. The classes are ordered with respect to the increasing points, and in order to identify the emerging classes, the scores in classes corresponding to the PISA competence level are also given.

Table 2. Average values regarding classes that emerged at the student level and their competence levels

<table>
<thead>
<tr>
<th>Cluster</th>
<th>Size</th>
<th>Reading Mean</th>
<th>Reading Proficiency level</th>
<th>Mathematics Mean</th>
<th>Mathematics Proficiency level</th>
<th>Science Mean</th>
<th>Science Proficiency level</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.07</td>
<td>315.24</td>
<td>1a</td>
<td>308.04</td>
<td>Below of 1</td>
<td>323.67</td>
<td>Below of 1</td>
</tr>
<tr>
<td></td>
<td>0.14</td>
<td>368.32</td>
<td>2</td>
<td>346.18</td>
<td>Below of 1</td>
<td>365.00</td>
<td>Below of 1</td>
</tr>
<tr>
<td></td>
<td>0.15</td>
<td>417.570</td>
<td>1</td>
<td>373.049</td>
<td>Below of 1</td>
<td>401.594</td>
<td>Below of 1</td>
</tr>
<tr>
<td></td>
<td>0.20</td>
<td>449.648</td>
<td>2</td>
<td>412.131</td>
<td>Below of 1</td>
<td>435.319</td>
<td>Below of 1</td>
</tr>
<tr>
<td></td>
<td>0.18</td>
<td>487.05</td>
<td>3</td>
<td>453.12</td>
<td>1</td>
<td>472.92</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>0.14</td>
<td>525.829</td>
<td>4</td>
<td>496.244</td>
<td>2</td>
<td>512.883</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>0.07</td>
<td>561.98</td>
<td>4</td>
<td>543.83</td>
<td>3</td>
<td>547.39</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>0.07</td>
<td>599.28</td>
<td>4</td>
<td>597.79</td>
<td>4</td>
<td>586.68</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>0.02</td>
<td>643.49</td>
<td>5</td>
<td>647.99</td>
<td>5</td>
<td>629.21</td>
<td>5</td>
</tr>
</tbody>
</table>

As shown in Table 2, when cluster sizes are examined, there are three large clusters (Clusters 3, 4, and 5), two medium clusters (2 and 6) and four small clusters (1, 7, 8, and 9). Twenty per cent of the students were included in the fourth cluster, which was also the largest. The cluster with the least number of students was the ninth cluster, and the students in this cluster had the highest scores in all sub-tests. Students’ reading, mathematics, and science scores increased from Cluster 1 to 9. Based on the students’ competence levels in reading, mathematics, and science, the following 9 clusters emerged representing students with i) extremely low achievement (ELA), ii) very low achievement (VLA), iii) quite low achievement (QLA), iv) low achievement (LA), v) fairly low achievements (FLA), vi) medium achievement (MA), vii) fairly high achievement (FHA), viii) high achievement (HA) and ix) very high achievement (VHA).

It was determined that the six classes which emerged at the school level were ordered in line with students’ achievements similar to the clusters identified at the student level. The average values regarding the classes which emerged at the school level are given in Table 3.
Table 3. Average values regarding classes that emerged at the school level

<table>
<thead>
<tr>
<th>Variables</th>
<th>Cluster1</th>
<th>Cluster2</th>
<th>Cluster3</th>
<th>Cluster4</th>
<th>Cluster5</th>
<th>Cluster6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class Size</td>
<td>0.08</td>
<td>0.35</td>
<td>0.32</td>
<td>0.11</td>
<td>0.10</td>
<td>0.04</td>
</tr>
<tr>
<td>Reading Mean</td>
<td>356.016</td>
<td>427.484</td>
<td>463.721</td>
<td>527.666</td>
<td>575.260</td>
<td>617.356</td>
</tr>
<tr>
<td>Mathematical Mean</td>
<td>337.283</td>
<td>393.421</td>
<td>429.565</td>
<td>501.480</td>
<td>563.353</td>
<td>617.231</td>
</tr>
<tr>
<td>Science Mean</td>
<td>355.853</td>
<td>416.086</td>
<td>450.487</td>
<td>514.109</td>
<td>561.790</td>
<td>603.759</td>
</tr>
</tbody>
</table>

At the school level, similar to the classes that emerged at the student level, schools were grouped with regards to their achievement in reading, mathematics, and science. The following six clusters emerged; i) ELA, ii) LA, iii) medium-low achievement (MLA), iv) MA, v) medium-high achievement (MHA), and vi) HA. When the size of the classes was evaluated, it was seen that the widest was the second cluster and the smallest was the sixth cluster. Thirty-five per cent of the schools were in the second class, which was the largest. At 4%, the sixth class had the lowest number of schools. The average achievement scores of the students increased from Cluster 1 to Cluster 6.

Variables That Explained the Classes at the School Level

After determining the classes according to the model with 6 classes at the school level, a three-step analysis was conducted by adding the student level variables to the school level. As a result, the model’s LL value was calculated as -7107.2456 and BIC was calculated as 14429.9470. In this analysis, the first latent class was taken as a reference. This class consisted of the schools, in which students had the lowest achievement in mathematics. The results show that all the variables chosen at the student and school levels significantly predicted the latent class membership based on schools. Table 4 gives the probability of the students being included in a school class.

Table 4. Probability of finding variables which were investigated in latent classes at the school level

<table>
<thead>
<tr>
<th>Variables</th>
<th>ELA</th>
<th>LA</th>
<th>MLA</th>
<th>MA</th>
<th>MHA</th>
<th>HA</th>
</tr>
</thead>
<tbody>
<tr>
<td>All probability</td>
<td>0.08</td>
<td>0.35</td>
<td>0.32</td>
<td>0.11</td>
<td>0.10</td>
<td>0.04</td>
</tr>
<tr>
<td>Openness for problem solving</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-1.862  →  -0.0714</td>
<td>0.21</td>
<td>0.442</td>
<td>0.317</td>
<td>0.018</td>
<td>0.012</td>
<td>0.000</td>
</tr>
<tr>
<td>-0.0707  →  0.102</td>
<td>0.056</td>
<td>0.366</td>
<td>0.418</td>
<td>0.090</td>
<td>0.066</td>
<td>0.004</td>
</tr>
<tr>
<td>0.107  →  0.282</td>
<td>0.061</td>
<td>0.323</td>
<td>0.355</td>
<td>0.128</td>
<td>0.091</td>
<td>0.042</td>
</tr>
<tr>
<td>0.283  →  0.425</td>
<td>0.032</td>
<td>0.348</td>
<td>0.359</td>
<td>0.136</td>
<td>0.094</td>
<td>0.031</td>
</tr>
<tr>
<td>0.438  →  1.201</td>
<td>0.058</td>
<td>0.215</td>
<td>0.192</td>
<td>0.154</td>
<td>0.164</td>
<td>0.216</td>
</tr>
<tr>
<td>Perseverance</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-0.724  →  0.172</td>
<td>0.337</td>
<td>0.324</td>
<td>0.273</td>
<td>0.029</td>
<td>0.032</td>
<td>0.004</td>
</tr>
<tr>
<td>0.172  →  0.362</td>
<td>0.056</td>
<td>0.422</td>
<td>0.398</td>
<td>0.054</td>
<td>0.046</td>
<td>0.025</td>
</tr>
<tr>
<td>0.365  →  0.494</td>
<td>0.010</td>
<td>0.317</td>
<td>0.390</td>
<td>0.107</td>
<td>0.117</td>
<td>0.059</td>
</tr>
<tr>
<td>0.508  →  0.674</td>
<td>0.008</td>
<td>0.328</td>
<td>0.346</td>
<td>0.139</td>
<td>0.116</td>
<td>0.064</td>
</tr>
<tr>
<td>0.686  →  1.717</td>
<td>0.008</td>
<td>0.302</td>
<td>0.235</td>
<td>0.196</td>
<td>0.116</td>
<td>0.142</td>
</tr>
<tr>
<td>Economic social and cultural status</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-2.142  →  -1.740</td>
<td>0.053</td>
<td>0.532</td>
<td>0.391</td>
<td>0.019</td>
<td>0.004</td>
<td>0.000</td>
</tr>
<tr>
<td>-1.725  →  -1.509</td>
<td>0.010</td>
<td>0.389</td>
<td>0.466</td>
<td>0.101</td>
<td>0.028</td>
<td>0.007</td>
</tr>
<tr>
<td>-1.494  →  -1.140</td>
<td>0.002</td>
<td>0.213</td>
<td>0.488</td>
<td>0.181</td>
<td>0.109</td>
<td>0.007</td>
</tr>
<tr>
<td>-1.064  →  0.219</td>
<td>0.000</td>
<td>0.038</td>
<td>0.178</td>
<td>0.222</td>
<td>0.282</td>
<td>0.280</td>
</tr>
<tr>
<td>Quality of infrastructure</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-2.755  →  -1.464</td>
<td>0.192</td>
<td>0.411</td>
<td>0.330</td>
<td>0.055</td>
<td>0.0118</td>
<td>0.0000</td>
</tr>
<tr>
<td>-1.089  →  -0.462</td>
<td>0.112</td>
<td>0.381</td>
<td>0.356</td>
<td>0.087</td>
<td>0.0453</td>
<td>0.0173</td>
</tr>
</tbody>
</table>
When the probability of finding the variables in latent classes based on schools is examined (Table 4), it is seen that all the variables had more influence on the low and medium-low level achievements. This situation also resulted from the average of the students in these two classes constituting 67% of the whole group.

Of the students, 97.1% (0.212 + 0.442 + 0.317) who had the lowest value in openness to problem-solving (-1.862 to 0.0714) had ELA, LA and MLA. Those students who were found to be open to problem-solving (53%) had MA, MHA and HA. Ninety-three per cent of the students with the lowest perseverance value had ELA, LA and MLA while 46% with a high level of perseverance had MA, MHA and HA. Moreover, 30.2% of the highly perseverant students had LA. The students with a high ESCS value were likely to be in a class with MA, MHA and HA. However, 78.2% with the highest ESCS values were found to have MA, MHA and HA. This shows that having a high level of ESCS does not necessarily result in success for all students.

When the variables chosen at the school level were analysed, 93% of the students who attended schools with a poor physical infrastructure had ELA, LA and MLA, and half of the students whose schools had an adequate physical infrastructure had MA, MHA and HA. In addition, 91% of the students whose schools did not have sufficient educational resources had very ELA, LA and MLA. The availability of educational resources was found to directly affect the level of success. Approximately, 57% students who had MA, MHA and HA had access to sufficient educational resources. Sixty-four per cent of the students attending schools with a shortage of teachers had ELA, LA and MLA. Thus, a teacher shortage in schools can be interpreted as not bringing success. However, 85% of the students who stated that their schools had no shortage in teaching staff also had very ELA, LA and MLA. This indicates that it is the quality rather than the physical existence of teachers that has an influence in student achievement.

In brief, according to the results of this study, as the students’ openness to problem-solving, perseverance, ESCS, and school resources (physical infrastructure, teaching employees and educational resources) increased, the percentage of students with a high level of achievement also increased. This situation indicates that the variables addressed in this study influences student achievement directly and school achievement indirectly.

**Discussion, Conclusion, and Suggestions**

This study aimed to determine multi-level latent classes in relation to the reading, mathematics and science achievements of students, who participated in PISA 2012 from Turkey and to discover the status of the prediction of the determined latent classes by i) students’ perseverance, ii) openness of students to problem solving, iii) ESCS of students and, iv) school resources (physical infrastructure, teaching staff, and educational resources). The results indicated that the 9Cluster-6GClass model had the best fit; thus, this was the preferred model for this study.
At the student-level, the students’ reading, mathematics, and science scores were divided into nine clusters. The smallest cluster had students with the highest scores in all sub-tests and the largest cluster contained students with the lowest scores in all sub-tests. It was determined that the six classes which emerged at the school level were ordered in line with students’ achievements similar to the case for the student–level clusters. The classes were ordered with respects to the increasing scores, and in order to define the emerging classes, the scores of the classes’ equivalence in the PISA competence level were also given. In various studies in the literature, models containing four or five classes are generally considered to be more appropriate. Finch and Merchant (2013) explored the relations between the academic achievement of 20 wealthy countries participating in PISA with regards to their SEL and their socio-economic typologies with MLCA. As a result of the analyses, five latent classes (from the lowest achievement/SEL to the highest achievement/SEL) were defined at the student level. At the school level, the schools were clustered in four classes with a pattern similar to the student level. In their study, Lin and Tai (2015) used latent class analysis to determine which mathematics learning strategies are influential on the level of students’ mathematics literacy in Thailand. In the study, the PISA 2012 mathematics achievement test items and mathematics learning strategy items were used.

The analyses showed that the model with four classes was best accommodated to the data. In both studies, as in the current research, the grouping was organised according to achievement; yet, they obtained a smaller number of classes. The reason for this might be the inclusion of variables such as SEL and learning strategies into the classification process. Moreover, there are different criteria in the literature to determine the number of classes. Adopting a more conservative approach or using more flexible criteria to obtain the fit of the model might also cause the number of classes to vary. Additionally, the wide range of levels of achievement of the students in Turkey might have resulted in a larger number of classes in this study.

All the chosen variables significantly predicted the school-level latent class membership. Moreover, the analyses suggested that the students’ ESCS was perceived to be the most important effect on their reading, mathematics, and science achievement. This finding is also supported by various studies (Dincêr and Kolaşin, 2009; Ferrera et al., 2011; Finch and Marchant, 2013; Geske et al., 2006; Nonoyama, 2005; Oral and Mcgivney, 2014; Sarer, 2016; Xu, 2006; Wolfram, 2005). For example, in the study by Yildirim (2009) concerning the PISA 2006 results, the primary elements that determined the quality of education in Turkey were reported to be the educational level and profession of the mother and father as well as the resources students had at home including socio-economic and socio-cultural factors. These results are also in parallel with those reported in the international literature.

Nonoyama (2005) evaluated the PISA 2000 and PISA 2003 results and found that the socio-economic conditions of the students’ family had a strong influence on their achievement since families that had a high SEL in developed countries could provide more resources for their children’s education. Finch and Merchant (2013) explored the relations between academic achievements of 20 wealthy countries among those that participated in PISA with regards to their SEL and typologies using MLCA. The educational level of family at the student level was found to have a significant relationship with the five chosen latent classes. As the educational level of families increased, the students’ achievement and socio-economic conditions also improved. In Turkey, the quality of education students receive is strongly affected by the educational status and the income level of their parents. This situation causes inequality of opportunity among students (Aslankurt, 2013). In the literature, the findings regarding the influence of SEL over student achievement are commonly reported. However, contrary to the previous studies, the results of the present research showed that a high SEL does not always result in higher success. Some of the students with a high ESCS had low achievement. This situation may be due to the lower level of concerns these students have for their future since their parents have the necessary financial means to ensure for their children or offer them future work opportunities. This may have put less pressure on students or reduced their motivation to study and succeed, thus resulting in lower achievement.

Furthermore, since ESCS did not have a direct effect on the students belonging to the very low achievement group, there is a need to examine other variables e.g. students’ attitude, motivation, and affective features in terms of their effect on achievement. Taking the relation of the students’ achievement to their ESCS, educational levels and professions of the family, and the educational resources at home into consideration, it can be suggested that investing in ways to increase the educational level of parents should be taken as a prospective investment. In this way, the parents’ job status will rise as well as their salaries leading to an increase in the educational resources available to the students at home. In the short term, this situation can be resolved by ensuring that the opportunities provided at school are at a level that would reduce the differences in the students’ socio-economic situation.
The findings of the present study are consistent with those reported in the international literature in terms of determining that perseverance is related to student achievement, and students who are open to problem-solving are more successful (Arıkan, 2014; Chiou and Xihua, 2008; Demir, 2015; Duckworth et al., 2007; OECD, 2014c; Scherer and Gustafsson, 2015; Thien et al., 2015; Yavuz et al., 2017). It is seen that students who are open to problem-solving in PISA 2012 participant countries Thailand, Indonesia and Malaysia have higher mathematics achievement (Thien et al., 2015). Moreover, being among the 21st-century skills, openness to problem-solving and perseverance are also considered to be related to students’ creative problem-solving skills as well as their motivation and willingness to succeed (Scherer and Gustafsson, 2015). Studies in the literature generally focus on the relationship between mathematics achievement and problem-solving skill; however, one of the conclusions of this study was that students’ openness to problem-solving and level of perseverance are also influential in increasing their reading and science achievement. Different from the field literature, this study showed that while students with low or medium level achievement have a low level of perseverance and openness to problem-solving, those with a high level of perseverance generally have MA; however, they can be included in any of the achievement classes (from low to high) except for the ELA class. This situation is considered to result from the social desirability effect.

According to the literature, in the measurement of perseverance, students generally give biased responses for social desirability reasons. For this reason, the relationships between achievement and perseverance may be higher than found; therefore, it is suggested that measurements be performed using a multi-method and multi-resource basis (Duckworth et al., 2007). In this sense, it is suggested that deductions should not be made based solely on responses given by students to the questionnaire items and the consistency of the results should be examined through teacher observations and student interviews. Moreover, based on the finding that students who are not open to problem-solving are not successful, further studies should be conducted to explore how to make it enjoyable for students to address complex problems by undertaking activities that they can link to daily life. In this sense, policymakers have certain duties. For example, in order for students to develop such skills, text books should be enriched including fun problem-solving activities, in-service training programs should be organised for teachers, and prospective teachers should be offered guidance and support throughout their undergraduate education.

The majority of the students, who were not open to problem-solving, and had a low level of perseverance or ESCS, were found to have ELA, LA and MLA. Nearly half of the students, who were open to problem-solving, perseverant or had a high ESCS, had MA, MHA and HA. This situation shows that students who have the chosen variables at a low level are not successful; yet, higher levels of these variables does not necessarily result in higher achievement every time. In other words, for successful students, having these variables at high levels does not always have a positive effect on their achievement. These students’ achievements are independent from their problem-solving skills or their family’s ESCS. For this reason, the results should be evaluated in the way that students with low and high success levels have different needs and require different approaches to increase their success.

When the results concerning the variables chosen at the school level are examined, it is seen that they are consistent with those reported in the field literature in terms of demonstrating the effect of the quality of school and educational resources (Acar and Öğretmen, 2012; Archibald, 2006; Nonoyama, 2005; Oral and Mcgivney, 2013; Özer-Özkan, 2016) on students’ achievements. Hanushek (1997) examined 377 studies taking students’ family features into consideration but did not find a strong and consistent relationship between student achievement and school resources. This situation may arise from the high level of relationship between the socio-economic situation of students’ family and the resources of their school. Since families with a high SEL can afford not only to send their children to schools with sufficient resources but also to provide their children with adequate educational resources at home, these students’ achievement may be independent from school resources.

The majority of the students attending schools with a poor quality physical infrastructure and insufficient educational resources had ELA, LA and MLA. Nearly half of the students, whose schools possessed better physical infrastructure and educational resources, had MA, MHA and HA. This situation shows that the school lacking proper physical infrastructure and educational resources generally has a negative impact on students’ success; however, the opposite case may not always be true; i.e. students attending schools with a high level of infrastructure and available resources may not be successful. Furthermore, these two variables do not make any difference for some of the successful students. Considering the fact that students in Turkey enrol in schools according to their achievement levels, the schools attended by highly achieving students tend to be better equipped in terms of physical conditions and educational resources.
One interesting result of the present study was related to the teaching staff variable at the school level. It was determined that the students, who stated that their school did not have a shortage in teaching staff, were usually from schools attended by students with very ELA, LA and MLA. This situation also indicates that the qualifications of teaching staff rather than their physical existence influences student achievement. In the literature, significant relationships have been identified between qualifications of teachers and student achievement (Darling-Hammond, 2000; Hanushek, 2002; Rivers and Sanders, 2002). Developing teaching quality through teacher education and in-service training is the correlative responsibility of educators and policymakers (Rivers and Sanders, 2002) and remains to be an area that requires more attention and efforts.

In addition to family-related features, schools’ educational resources, physical infrastructure, and quality of education were also found to have a significant effect on student success. The effects of such external factors should be reduced by ensuring that all schools have similar resources and provide equal opportunities. In addition, training programs should be organised to increase the awareness of teachers related to the effect of students’ socio-economic situation on their success. Furthermore, since students’ openness to problem-solving plays a role both in their daily and school lives, teachers are suggested to allocate more time for activities that would develop such skills in students and provide them with guidance.

The results found out that nearly half of the successful students were not influenced by their family’s ESCS and profession, educational resources provided at home, school’s physical infrastructure and educational resources, and teacher qualifications. These students can be described as naturally smart and/or academically resilient. It is also stated in the field literature that academically resilient students are more successful independently of SEL (Çokluk, Gül, and Kayri, 2016; Dinçer and Oral, 2013; Yaşar, 2016). Moreover, some students coming from families with a low educational level and poor economic conditions had low achievement due to these conditions. Because their SEL was low, these students also did not have sufficient educational resources at home and their parents could not afford private tutors. In addition, they attended schools with insufficient physical and educational resources (Yaşar, 2016). For this reason, providing equality of opportunity for all students is essential. Similarly, one of the essential reasons behind Finnish students’ higher achievements has been reported as equality of opportunity offered to all students in Finland (Çobanoğlu and Kasapoğlu, 2010). In this context, researchers should address the areas that need improvement to provide equality of opportunity in education for all students in Turkey.

To date, in the field literature, no study has approached student achievement as a latent variable and analysed it by creating homogenous subclasses based on patterns emerging in selected variables. The present study is also important with regards to presenting the model that was used in the work. Since separating student achievement into homogenous classes in itself makes it possible to determine factors related to different levels of student achievements, more detailed data can be gathered to make suggestions for improving the success of students with different achievement levels. However, the present study has certain restrictions. Firstly, the population was limited to Turkish students. In further studies about this topic, students from other countries with a wide range of low to high achievement levels should be included in a comparative analysis. Another limitation concerns the selection of variables being based on previous reports regarding their effect on student achievement. Furthermore, openness towards problem-solving and perseverance variables were only used in PISA 2012. Since these indexes were not used in PISA 2015, this study was limited to the data obtained from PISA 2012.

Note

This study was presented at Congress of Measurement and Evaluation in Education and Psychology in 1-3 September 2016 in Turkey.

References


Gudicha, D. W., & Vermunt, J. K. (2013). Mixture model clustering with covariates using adjusted three-step approaches. B. Lausen, D. van den Poel, and A. Ultsch (eds), Algorithms from and for nature and life; studies in classification, data analysis, and knowledge organization. 87-93. Heidelberg: Springer-Verlag GmbH.


Kane, H., & Brand, C. (2003). The importance of spearman’s ‘g’ as a psychometric, social and educational construct. The Occidental Quarterly, 3(1), 7-30.


Lukočienė, O., Varriale, R., & Vermunt, J. K. (2010). The simultaneous decision(s) about the number of lower and higher-level classes in multilevel latent class analysis. Sociological Methodology, 40(1), 247-283.


---

**Author Information**

Seher Yalcin  
Ankara University  
Faculty of Educational Sciences  
06590, Ankara, Turkey  
E-mail: yalcins@ankara.edu.tr