Van Yüzüncü Yıl Üniversitesi Eğitim Fakültesi Dergisi, 2024; 21(3), s. 840-868.

Van Yüzüncü Yıl University Journal of Education, 2024; 21(3), p. 840-868. DOI: 10.33711/yyuefd.1539072

A Comparative Analysis of PISA 2015 Türkiye Studies: Introducing A Variable Selection Model to International Large-Scale Assessments

Sinem DEMİRCİ[1], Özlem İLK[2]

Abstract: International large-scale assessments have a key role in improving educational, economic, and political systems. By using the data of these assessments, countries can draw conclusions about the status of educational systems. Studies and reports generally tend to choose variables available in data set to model the relationships among the variables. In this study, we aimed to introduce a variable selection method to analyze large-scale assessments to be able to decide which variables might be included in modelling country data. We used the entire data set of Türkiye PISA 2015 through elastic net regression to decide which variables should be selected for further analysis. We also provided a summary of the available studies based on Türkiye PISA 2015 data and compared the results. Based on the series of analyses, this study revealed that test anxiety, environmental awareness, interest in broad topics in science, playing video games after school, mathematics literacy, reading literacy, and collaborative problem-solving skills were the explanatory variables that contributed most to the degree of scientific literacy of students. This study has a potential to provide an example of shrinkage methods applied in educational context and offer another standpoint for providing a rationale to select which variables can be included.

Keywords: PISA 2015, scientific literacy, elastic net regression, shrinkage methods

PISA 2015 Türkiye Çalışmalarının Karşılaştırmalı Bir İncelemesi: Uluslararası Büyük Ölçekli Değerlendirmeler için Değişken Seçim Yöntemi Önerisi

Öz: Uluslararası büyük ölçekli değerlendirmeler, eğitim, ekonomik ve politik sistemlerin iyileştirilmesinde önemli bir rol oynamaktadır. Ülkeler, bu değerlendirmelerin verilerini kullanarak, eğitim sistemlerinin mevcut durumu hakkında çıkarımlarda bulunmaktadır. Bu verileri kullanan bilimsel çalışmalar ve raporlar, genellikle veri setinde mevcut olan bazı değişkenleri seçerek bu değişkenler arasındaki ilişkileri modellemeyi amaçlar. Bu çalışmada, Türkiye PISA 2015 verisinin tamamını kullanarak ülke verilerini modellemede hangi değişkenlerin dahil edilebileceğine karar vermek amacıyla bir değişken seçim yöntemi denemeyi hedeflenmiştir. PISA 2015 verisinin tamamını kullanarak büzüşme regresyonlarından biri olan

Geliş tarihi/Received: 26.08.2024 Kabul Tarihi/Accepted: 02.12.2024 Makale Türü: Araştırma Makalesi

*Bu çalışma birinci yazarın yüksek lisans tezinden üretilmiştir.

[1] Lecturer, California Polytechnic State University, Statistics, sdemirci@calpoly.edu, 0000-0002-2095-0674

[2] Professor Dr, Middle East Technical University, Statistics, oilk@metu.edu.tr, 0000-0002-4179-0534

Atıf için/To cite: Demirci, S., & İlk, Ö. (2024). PISA 2015 Türkiye Çalışmalarının Karşılaştırmalı Bir İncelemesi: Uluslararası Büyük Ölçekli Değerlendirmeler için Değişken Seçim Yöntemi Önerisi. *Van Yüzüncü Yıl Üniversitesi Eğitim Fakültesi Dergisi, 21*(3), 840-868. https://doi.org/10.33711/yyuefd.1539072

elastik net regresyonu kullanılmış ve elde edilen sonuçlar, Türkiye PISA 2015 verilerine dayalı mevcut çalışmaların sonuçları ile karşılaştırılmıştır. Analizler sonucunda, test kaygısı, çevresel farkındalık, geniş kapsamlı bilim konularına ilgi, okul sonrası video oyunları oynama, matematik okuryazarlığı, okuma okuryazarlığı ve işbirlikçi problem çözme becerilerinin öğrencilerin fen okuryazarlığı düzeyine en çok katkı sağlayan açıklayıcı değişkenler olduğu ortaya konulmuştur. Bu çalışma, eğitim bağlamında büzüşme yöntemlerinin uygulanmasına bir örnek sunma potansiyeline sahip olup, hangi değişkenlerin dahil edilebileceğine yönelik bir gerekçe sunmak için alternatif bir bakış açısı önermektedir.

Anahtar kelimeler: PISA 2015, bilimsel okuryazarlık, elastik net regresyonu, büzüşme tahminleyicisi

Introduction

It has been acknowledged that scientific and technological progress is a fundamental basis for the economic development level of a country (Laugksch, 2000) which impels countries to invest effort and money to improve science education policies (Lewis, 1982; Coll & Taylor, 2012). Science literacy or scientific literacy is being one of the approaches in science education history considered as the ‘main drivers of science education policy’ (Roberts & Bybee, 2014). Recently, the importance of scientific literate societies is also highlighted to meet the demands of the 21st century (Choi et al., 2011; Valladares, 2021) and to promote sustainability of the planet and humanity (Carter, 2008).

The concept of scientific literacy originated in the late 1950s, but the body of literature on it expanded significantly over the past three decades (Roberts, 2007). Over time, the definition of scientific literacy has broadened to include various issues where science plays a role. Additionally, as Pedretti (2014) pointed out, two important themes emerged in science education: science, technology, society, and environment (STSE) education (Pedretti & Nazır, 2011; Solomon & Aikenhead, 1994) and socioscientific issues (SSI) (Zeidler et al., 2009). These themes have, in turn, reshaped both the goals of science education and the scope of scientific literacy. The Programme for International Student Assessment (PISA), an international education survey, has its own scientific literacy framework enabling us to compare the level of scientific literacy of 15-year-old students in different countries. Launched by OECD (2016) in 1997, PISA test is done once in every three years. In every cycle, PISA has a different focal point and in PISA 2015, this focal point was scientific literacy. Scientific literacy is described in PISA 2015 as following:

Scientific literacy is the ability to engage with science-related issues, and with the ideas of science, as a reflective citizen.

A scientifically literate person is willing to engage in reasoned discourse about science and technology, which requires the competencies to:

- Explain phenomena scientifically – recognise, offer and evaluate explanations for a range of natural and technological phenomena.
- Evaluate and design scientific enquiry – describe and appraise scientific investigations and propose ways of addressing questions scientifically.
- Interpret data and evidence scientifically – analyse and evaluate data, claims and arguments in a variety of representations and draw appropriate scientific conclusions (OECD, 2017, p.22).

Besides assessing the skills and knowledge of the students in this literacy frameworks, additional information is also collected from students, teachers, parents, and school principals about students' home possessions, some individual characteristics of the students, and their learning environment in their school. In this way, policymakers gain an opportunity (1) to examine the level of skills and knowledge of students; (2) to assess how these levels of skills and knowledge might be related to different variables in their home, classroom, and school as well as individual differences (OECD, 2009).

PISA 2015 data are being examined by policymakers and researchers to be able to draw conclusions about which demographic, economic, and educational variables interacted with scientific literacy and to compare the levels of countries. For the same reason, Türkiye PISA 2015 data are also widely analysed by researchers interested in the scientific literacy of the students in Türkiye (see Table 1). Different from the previous studies based on Türkiye PISA 2015 data, we aimed to explore the entire data set through elastic net regression, one of the shrinkage methods in linear model selection and regularization, to be able to select which subset of the variables explain the variance in scientific literacy of students most. In the next section, we provided a literature review on these previous Türkiye PISA 2015 studies to be able to portray similarities and differences between them.

Literature Review

In this part, the quantitative studies on scientific literacy based on Türkiye PISA 2015 data are summarized. While most of the studies preferred hierarchical linear modelling (e.g., Karslı et al., 2019; Dolu, 2020, Yıldız et al., 2020), some included exploratory data analysis such as correlation (e.g., Kaya, 2017) or other statistical analyses (e.g., Öztürk, 2018). In terms of variable selection, different subsets of variables were used in these studies. The rationales for variable selection in most of these studies were not explicitly stated (except Karslı et al., 2019) but implicitly attributed to 'know how' of the literature on scientific literacy. In other words, no preliminary variable selection process based on a statistical analysis has been reported. Table 1 summarized some of the Türkiye PISA 2015 studies, analysis techniques, and variables used.

Table 1

Türkiye PISA 2015 Studies Predicting Scientific Literacy, Analysis Techniques, and Variables

Study	Analysis	Variables*	
Akgenç & Yapıcı Pehlivan (2019)	Multilevel Structural Equation Modelling	*Independent Variables (Within-Group)* 1. Activity in Science Class** 2. Family Support** 3. Gender** 4. Interest In Science** 5. Science Activities** 6. Science Self-Concept** 7. Science Working Hours** 8. Teacher Support in Science Class** 9. Program Type 10. Teacher's Comment in Science Class	*Independent Variables (Between-Groups)* 1. Number Of Smart Boards** 2. Science Equipment** 3. Settlement** 4. Laboratory 5. Laboratory Material 6. School Type 7. Teaching Hours *Class Variable* 8. Number Of Schools

Dolu (2020)	Hierarchical Linear Modelling	*Student-Level Variables* 1. ESCS** 2. Gender** 3. Grade Level** 4. Grade Repetition** 5. Early Childhood Education and Care 6. Language Spoken at Home	*School-Level Factors* 1. Region** 2. School ESCS** 3. School Type** 4. Class Size 5. School Location 6. School Size 7. Student-Teacher Ratio
Karslı et al. (2019)	Hierarchical Linear Modelling	*Student-Level Variables* 1. Adaptive Instruction** 2. Disciplinary Climate** 3. Enquiry-based Instruction** 4. Perceived Feedback** 5. Teacher Support to Students** 6. Teacher-directed Instruction** *Control Variable* 1. Index of Economic, Social, and Cultural Status (ESCS)	*School-Level Factors* 1. Shortage of Educational Material** 2. Shortage of Education Staff** 3. Student Behaviour Hindering Learning** 4. Class Size 5. Teacher Behaviour Hindering Learning
Üstün et al. (2020)	Hierarchical Linear Modelling	*Individual Variables* 1. Gender** 2. Grade Level** 3. Grade Repetition** 4. ESCS** *Learning Time Variables* 5. After-school Study Time** 6. Total Number of Class Periods for Science** 7. Total # of Class Periods per Week** *Teaching-Learning Process Variables* 8. Attitudes Towards Cooperation** 9. Enquiry-based Instruction**	*School Resources Variables* 1. Index of computers connected to the internet** 2. Proportion of Science Teachers ** 3. Science Resources at School** *Learning Environment at School Variables* 4. Student Behaviour Hindering Learning**
Yetişir (2021)	Hierarchical Linear Modelling	*Student-Level Variables* 1. Disciplinary Climate** 2. Gender** 3. Arriving Late for School 4. ESCS 5. Skipping School 6. Teacher Support	*School-Level Variables* 1. Staff Shortage** 2. Student Behaviours** 3. Educational Material Shortage 4. Teacher Behaviour *Control Variable* 1. Aggregated ESCS**
Yıldız et al., (2020)	Hierarchical Linear Modelling	*Affective Characteristics* 1. Epistemological Beliefs 2. Enjoyment Of Science 3. Interest In Broad Science Topics 4. Instrumental Motivation 5. Student Attitudes, Preferences and Self-Related Beliefs, Achieving Motivation 6. Science Self-Efficacy *Learning Environment* 1. Adaption of Instruction 2. Disciplinary Climate in A Science Class	*Level-2 IVs* 1. School Type** 2. Student-Teacher Ratio**

3. Inquiry-Based Science Teaching and Learning Practices
4. Perceived Feedback
5. Students' Science Activities
6. Teacher Support in A Science Class
7. Teacher Directed Science Instruction

Demographics

8. Socio-Economic Status**
9. Grade Level**
10. Gender**

** To provide reproducibility, the names of the variables written in English directly copied from the original sources. Variables written in Turkish was translated as given in OECD (2016) documents although it might create some minor inconsistencies.*
*** statistically significant ($p<0.05$) except Yetişir (2021) and Üstün et al. (2020) whose level of significance were* $p<0.001$ at some variables

As summarized in Table 1, approximately 30 variables at the student level and 20 variables at the school level were used in various combinations. Among these, gender, teacher support in science class, ESCS, disciplinary climate in science class, grade level, and school type were the most frequently used variables that significantly explained the variability in students' scientific literacy scores. Interestingly, variables such as environmental awareness and test anxiety were not examined, even though they have been found to be correlated with scientific literacy in Turkish science education literature and in previous Turkish PISA studies (e.g., Erbas et al., 2012; Genc, 2017; Öztürk, 2018; Haşıloğlu & Göğebakan, 2021). Moreover, other literacy-related variables, such as mathematics literacy, reading literacy, and collaborative problem solving, were also not studied.

Aim of the Study

Different from the previous studies based on Türkiye PISA 2015 data, we aimed to explore the entire data set through elastic net regression, one of the shrinkage methods in linear model selection and regularization, to be able to select a subset of variables that explain the variance in scientific literacy of students most. Shrinkage method (also known as penalized regression) is one of the classes of methods in linear model selection and regularization including all p predictors by shrinking their coefficients towards zero compared with least squares estimates providing a decrease in variance (James et al., 2013). Accordingly, the research questions of this study were determined as follows:

1. Which variables are related to the level of scientific literacy of students most?
2. Is there any congruency between the variables that emerged in penalized regression and multiple linear regression?

Significance of Study

International large-scale assessments have a key role for countries in terms of educational, economical, and political aspects. By using the data of large-scale assessments, they are able to analyze and infer the relationships among the variables as well as draw conclusions about their national educational systems. At this point, this study may provide useful information to educational policymakers as well as educators, parents, and students themselves about how Turkish students' degree of scientific literacy might be related to various demographic, social, economic, and educational variables in Türkiye. Moreover, examining the degree of scientific literacy of Turkish students may also give some clues about how to enhance our students' scientific literacy. In the long run, it may provide an insight into establishing improvements in our national educational systems and for understanding the relative strengths and weaknesses of our own education systems. On the other hand, as OECD (2009) highlighted, the economic and social welfare of the nations are largely correlated with their citizens' level of knowledge and skills. Therefore, participating in international tests like PISA enable them to evaluate how their young population is ready for the future. Accordingly, the results of this study may give information and allow doing some projections about our national economy and social welfare in the long run.

It is important to acknowledge that our study is not the only study interested in explaining scientific literacy of Turkish students using PISA 2015 data. As given in the literature review section, there are available studies examining Turkish students' degree of scientific literacy through PISA 2015 data. While these studies provide some valuable insight pertaining to variance in scientific literacy of Turkish students, they often lack of providing a rationale for variable selection process. As a result, we see different variable combinations and different models which can be confusing to understand which explanatory variables contributed most to the degree of scientific literacy of Turkish students. To overcome this limitation, we benefited from elastic net regression for variable selection because this technique enabled us to (1) include a total of 246 variables from Türkiye PISA 2015 data and (2) determine which explanatory variables could possibly be related with the degree of scientific literacy of Turkish students among these variables. Thus, we believe that our study has a potential to contribute to the literature in a way that which variables would predict the degree of scientific literacy most when more than 200 variables are tested via penalized regression. Besides, shrinkage methods to analyse international large-scale assessment data has recently started to be used. For example, Santi et al. (2019) used several penalized likelihood approaches using Indonesian PISA 2015 data for math literacy to decide which of them had the best performance for variable selection and coefficient estimation processes. Thus, our study may be considered as an initial point to use elastic net regression to determine which variables may be related with the students' level of scientific literacy mostly among the other variables.

Materials and Method

Research Design of the Study

In this study, correlational research was employed to examine the relationships among variables without manipulating them. This methodology is generally used for describing phenomena by exploring the relationships of some variables that might affect the variability in these phenomena (Fraenkel et al., 2012). Within the context of this study, the researchers intended to explore the factors that might be related with Turkish students' scientific literacy and used PISA 2015 data.

Population and Sample

The actual population of 15-year-old students in Türkiye was reported as 1,324,089 students, whereas the accessible population was determined as 925,366 students (Ministry of National Education [MoNE], 2018). In this study, Turkish students who participated in PISA 2015 were chosen as sample. A total of 5,895 students from 187 schools representing 61 cities from 12 statistical regions was selected as a sample for PISA 2015. The detailed information is presented in Table 2.

Table 2

Descriptive Statistics of PISA 2015 Sample According to Regions

Name of the Region	Frequency (n)	Percentage (%)
Istanbul	1,070	18.15
West Marmara	245	4.16
Aegean	707	11.99
East Marmara	510	8.65
West Anatolia	553	9.38
Mediterranean	817	13.86
Middle Anatolia	334	5.67
West Blacksea	303	5.14
East Blacksea	194	3.29
Northeast Anatolia	199	3.38
Middle East Anatolia	276	4.68
Southeast Anatolia	687	11.65
TOTAL	5,895	100

Data Collection

In PISA 2015, students were evaluated in science, mathematics, reading, collaborative problem-solving and financial literacy. Türkiye did not take options of financial literacy, ICT literacy, educational career, parent, and teacher questionnaires. Besides, some additional information was collected from students, teachers, school principals and parents. The language of the PISA test was in Turkish. Further details related to questionnaires and data set can be found on PISA 2015 website.

PISA tests are generally composed of both multiple-choice questions and open-ended questions based on authentic scenarios related to real-life problems. Students also take an additional questionnaire related to themselves, facilities at home and at their school as well as learning experiences (OECD, 2016). For the first time, PISA 2015 included a bullying questionnaire and collaborative problem-solving questionnaire for the students. Türkiye did take these tests as well, besides the scientific literacy test. In PISA 2015, students answered the questions on the computer rather than taking a paper-pencil test.

Data Analysis Procedure

As PISA manual pointed out, in the case of answering all the items of PISA 2015, students should have spent about 810 minutes on test items. Since it is not possible and feasible as well, different students are administered different subsets of test items in PISA cycles. When students complete the subset of cognitive PISA items, Rasch model is used to estimate students' performance as if they took the whole test. PISA uses a version of generalized Rasch model to polytomous items proposed by Wright and Masters (1982). Among the Rasch ability estimators, PISA uses weighted likelihood estimate (WLE) and plausible values (PV). They produce 10 PVs for each student for their scientific literacy, math literacy, reading literacy, and collaborative problem-solving scores. In the PISA Data Analysis Manual (2009), it is highlighted that PVs should never be averaged at the student level because this has a potential to create bias. Instead, data analysis could be conducted based on a single PV during the exploration part but using all PVs is highly recommended for further analysis. In the data analysis of this study, single PV was used for the initial model selection in elastic net regression. For determining the final model, all PVs were included. Data analysis stages of this study are summarized in Table 3.

Data cleaning and data preparation processes were completed in IBM SPSS software program. Descriptive statistics, mean/mode substitution and penalized regression were done by using R Project for Statistical Computing (R Core Team, 2022). After selecting variables by using elastic net regression, IDB Analyzer Version 3.2 (IEA, 2018) was used to conduct multiple linear regression by using all plausible values and replicate weights. Backward elimination technique was used for deciding on the final model. Last, by using R Project for Statistical Computing program, Variance Inflation Factor (VIF) values were checked if there is a multicollinearity problem which IDB Analyser does not provide such analysis.

Data Preparation and Data Cleaning Processes

In the data preparation part, some variables were deleted from the data set for several reasons. For instance, variables that Turkish students did not answer were deleted from the data set. Moreover, if the percentage of missingness is higher than 50%, these variables were also removed from the data set. Besides, there are some variables that explain the same phenomenon in different types of data. These duplicate variables were also eliminated. Next, there are some variables such as entering one by one and their total scores or estimates as a factor as well. To overcome this, total scores/estimates were retained in the data file and the others were eliminated. Next, the nominal variables that have many categories were removed from the data set to simplify analysis. For example, job occupations of parents have more than 15 categories and generating dummy variables for each job for both mother and father may not be feasible while conducting analysis. Moreover, the categorical variables that have highly unbalanced levels were also eliminated. Last, the variables that were related to mathematic literacy, reading literacy, and other lectures, such as attitudes towards mathematics, were also removed from the data set since the response variable of this study was scientific literacy.

After this process was completed, the data set was converted into .csv file to administer imputation methods for missing data in R Project for Statistical Computing. For the missing data of nominal and ordinal data, mode substitution was used as an imputation method. For continuous variables, mean imputation was carried out. The reason to use this imputation method is that the proportion of missing data was quite low in the data set. After data cleaning and preparation processes were completed, 246 variables remained to be analysed.

Shrinkage Methods / Penalized Regression

Shrinkage method (also known as penalized regression) is one of the classes of methods in linear model selection and regularization including all *p* predictors by shrinking their coefficients towards zero compared with least squares estimates providing a decrease in variance (James et al., 2013). There are three kinds of shrinkage methods as *ridge regression*, *lasso regression* and *elastic net regression* (Zou & Hastie, 2005). In this study, elastic net regression was employed. Zou and Hastie (2005, p.302) described elastic net regression by making an analogy as '…*It is like a stretchable fishing net that retains "all" the large "fish"'*. This means that elastic net regression removes unimportant variables, and this leads to improve prediction accuracy. This shrinkage method was chosen for this study since the data of this study have more than 200 variables with 5,895 number of observations. Hence, we need to have a parsimonious model that can explain the degree of scientific literacy of the students with fewer parameters.

In order to select a suitable model in elastic net regression, some criteria are considered such as mean squared error (MSE), mean prediction error and deviance ratio. Choosing a model having lower MSE values would indicate predicted response values is closer to observational values (James et al., 2013). Thus, we chose relatively lower MSE value. As in MSE values, having relatively lower mean prediction error is also preferred. Additionally, deviance ratio can be regarded as R-square for elastic net regression (Friedman et al., 2018); hence, having the highest deviance ratio was one of the criteria in this study.

Table 3

Stages of the Data Analysis

Stages of Data Analysis	Aim of the Stage	Steps Followed	Data Analysis Software Program
Data Preparation	• Preparing data set for analysis merging different data files, choosing suitable variables, cleaning data, handling missing data	• Data Consolidation • Variable Selection • Data Cleaning • Missing Imputation	R Project for Statistical Computing
Cross-validation	• Examining test errors of elastic net regression • Choosing a model that meets criteria	• 20 different data sets were formed ○ 10 data sets for every PVs of science + *30 PVs* for math, reading and collaborative problem solving + other variables ○ 10 data sets for every PVs of science + *3 random PVs* for math, reading and collaborative problem solving + other variables • Each 20 data sets were divided randomly into two data sets as 'training data' (%80 of the data) and 'test data' (20% of the data) • A model was chosen providing the highest deviance ratio; the smallest MSE values and mean prediction error to be refitted for the full model	R Project for Statistical Computing
Refitting Elastic Net Regression	• Refitting elastic net regression for the full models.	• 3 data sets were formed ○ Data Set 1: 71 Variables ○ Data Set 2: 60 Variables	R Project for

		○ Data Set 3: 57 Variables (highly correlated variables were omitted from Data Set 2) • Elastic net regression was refitted for data sets	Statistical Computing
Multiple Linear Regression	• Testing the models emerged from elastic net regression by obtaining *R*-Square, standard error of the estimate, and *t*-values • Working with plausible values, replicate weights, and students' final weights by using IDB Analyser	• Backward Stepwise Elimination Technique was used manually for each 3 data sets ○ The nonsignificant variables were removed from the model and multiple linear regression was performed repeatedly until all explanatory variables were observed as statistically significant. ○ The final model was chosen	IDB Analyzer
Deciding on the Final Model	• Checking the assumptions of the model as IDB Analyser did not provide ANOVA table, VIF values and *p*-values.	• A multiple linear regression analysis was reconducted for the final model	R Project for Statistical Computing

Cross-validation

As plausible values were generated by PISA for each student, different data sets were formed (Table 3) for applying elastic net regression in R. Each 20 data sets were divided randomly into two data sets as 80% of the data was for 'training data' and 20% for 'test data' to examine the test error of elastic net regression. After splitting the data sets, elastic net regression model was fitted on the train data, and some criteria such as MSE, deviance ratio, and number of parameters were compared to choose the best data set for refitting the model on the full data set. Additionally, a grid search with the possible tuning parameters (λ) were done to obtain an interval between 50 and 100 parameters in the model. Among them, the tuning parameter that gives models with number of parameters close to 50 parameters was selected. The results are given in Table 4. The technical details of this procedure are available in the first authors' thesis (Demirci, 2018).

Table 4

Cross-validation of 10 Plausible Science Values

		MSE		Mean Prediction Error		Deviance Ratio		Lambda (λ)		Number of Parameters	
		30 PV	3 PV	30 PV	3 PV	30 PV	3 PV	30 PV	3 PV	30 PV	3 PV
PV1	*Train*	841.18	1040.12	-	-	0.863	0.823	2.66	2.01	50	60
	Test	-	-	916.94	1091.86	-	-	-	-	-	-
PV2	*Train*	788.64	1381.70	-	-	0.867	0.768	2.66	2.66	66	52
	Test	-	-	817.10	1446.35	-	-	-	-	-	-
PV3	*Train*	792.58	1406.92	-	-	0.864	0.766	2.66	2.01	58	66
	Test	-	-	831.19	1426.30	-	-	-	-	-	-
PV4	*Train*	753.53	1004.25	-	-	0.877	0.837	2.66	2.01	53	64
	Test	-	-	768.24	1092.29	-	-	-	-	-	-
PV5	*Train*	768.09	1370.12	-	-	0.870	0.764	2.66	2.66	55	53
	Test	-	-	811.51	1444.51	-	-	-	-	-	-
PV6	*Train*	797.97	1442.52	-	-	0.873	0.767	2.01	2.66	64	52
	Test	-	-	817.16	1451.37	-	-	-	-	-	-

PV7	*Train*	784.79	1197.16	-	-	0.863	0.802	2.66	2.01	58	67
	Test	-	-	799.26	1238.49	-	-	-	-	-	-
PV8	*Train*	814.90	1382.09	-	-	0.860	0.760	2.66	2.66	66	54
	Test	-	-	804.48	1514.40	-	-	-	-	-	-
PV9	*Train*	808.57	1424.09	-	-	0.868	0.761	2.66	2.66	55	54
	Test	-	-	793.63	1464.86	-	-	-	-	-	-
PV10	*Train*	771.00	1413.80	-	-	0.869	0.766	2.66	2.66	60	58
	Test	-	-	812.92	1478.65	-	-	-	-	-	-

Based on Table 4, it can be observed that the data set which has Plausible Value 4 in Science (PV4) has the lowest MSE, lowest Mean Prediction Error, and highest deviance ratio among all other data sets. Therefore, the data set that has PV4 was decided to be used as full data set to refit the elastic net regression.

Results

In line with the research questions of this study, the results were presented in this section.

Results of Refitting Elastic Regression

After deciding which plausible value of science will be used for refitting based on the cross-validation results, elastic net regression model on the full data set was repeated. Three full data sets were used; one had 30 PV within the explanatory variables (Data Set 1) while the other included 3 random PV (Data Set 2). Besides, one additional data set was generated by eliminating highly correlated explanatory variables from Data Set 2 before fitting the model (Data Set 3). Model for Data Set 1, with the lowest MSE and highest deviance ratio, indicated a better fit. However, it was evident that the models for the other data sets were more parsimonious. The results are reported in Table 5.

Table 5

Refitting Elastic Net Regression for the Full Model

	MSE			Deviance Ratio			Lambda			Number of Parameters		
	Data Set 1	Data Set 2	Data Set 3	Data Set 1	Data Set 2	Data Set 3	Data Set 1	Data Set 2	Data Set 3	Data Set 1	Data Set 2	Data Set 3
PV4	732.74	1055.02	1061.76	0.879	0.825	0.824	2.01	2.01	2.01	72	61	58

As R Project for Statistical Computing does not provide a multiple linear regression package including the analysis of plausible values, these subsets of variables in three data sets were modelled in IDB Analyser program by using a manual version of backward stepwise elimination technique. In the next section, detailed results were presented for all three models.

Results of Multiple Linear Regression in IDB Analyser

Results of Model 1

A total of 71 variables were refitted in IDB Analyser. Backward elimination was done manually to decide on the final version of Model 1. At first, all variables were included in the model by using Data Set 1. After eliminating non-significant variables from the model, multiple linear regression was performed twice more to acquire a final model. The alpha value (α), the risk of committing Type I Error, was decided as to be 0.05. The related output is given in Figure 1.

Figure 1

The Final Output of Model 1

Variable	Regression Coefficient	Regression Coefficient (s.e.)	Regression Coefficient (t-value)	Stndrdzd. Coefficient	Stndrdzd. Coefficient (s.e.)	Stndrdzd. Coefficient (t-value)
(CONSTANT)	10.92	6.03	1.81	.	.	.
ANXTEST	-2.58	.85	-3.02	-.03	.01	-2.97
ENVAWARE	1.77	.43	4.09	.03	.01	4.07
INTBRSCI	2.15	.78	2.76	.03	.01	2.75
PV_MATH	.42	.02	23.56	.43	.02	25.96
PV_READ	.36	.02	22.48	.37	.02	21.51
PV_CLPS	.21	.02	13.35	.20	.02	13.22

A multiple linear regression analysis was conducted to evaluate how well the variables obtained in the second version explain the degree of scientific literacy of Turkish students. According to the results given in Figure 1, a total of 6 variables were statistically significant ($|t_{calculated}| > t_{critical} = 1.96$). The sample multiple correlation coefficient was .86. Test anxiety, environmental awareness, interest in broad science topics, mathematics literacy, reading literacy, and collaborative problem-solving skills of the students were found to make statistically significant contributions to explain the degree of scientific literacy.

After setting the final model, the regression equation can be written as following:

$$\hat{Y}_{SCILIT} = 10.92 - 2.58X_{ANXTEST} + 1.77X_{ENVAWARE} + 2.15X_{INTBRSCIE} + .42X_{MATH} + .36X_{READ} + .21X_{CLPS}$$

Results of Model 2

A total of 60 variables (Appendix 1) were refitted in IDB Analyser and backward elimination was used manually to decide on the final version of the Model 2. At first, all 60 variables from Data Set 2 were included. Multiple linear regression was performed repeatedly until all explanatory variables were observed as statistically significant. The alpha value (α), the risk of committing Type I Error, was decided as to be 0.05. The related output is given in Figure 2.

Figure 2

The Final Output of Model 2

Variable	Regression Coefficient	Regression Coefficient (s.e.)	Regression Coefficient (t-value)	Stndrdzd. Coefficient	Stndrdzd. Coefficient (s.e.)	Stndrdzd. Coefficient (t-value)
(CONSTANT)	17.35	6.10	2.84	.	.	.
ANXTEST	-2.51	.86	-2.94	-.03	.01	-2.89
ENVAWARE	1.74	.44	3.94	.03	.01	3.93
INTBRSCI	2.07	.78	2.66	.03	.01	2.65
ST078Q06NA	-5.32	2.50	-2.13	-.03	.02	-2.15
PV_MATH	.41	.02	22.72	.42	.02	24.49
PV_READ	.36	.02	23.01	.38	.02	22.13
PV_CLPS	.21	.02	12.96	.21	.02	12.89

A multiple linear regression analysis was conducted to evaluate how well the variables obtained in the second version explain the degree of scientific literacy of Turkish students. According to the results given in Figure 2, a total of 7 variables were statistically significant ($|t_{calculated}| > t_{critical} = 1.96$). The sample multiple correlation coefficient was .86. Test anxiety, environmental awareness, interest in broad science topics, playing video games after school (1 = Yes, 2 = No), mathematics literacy, reading literacy, and collaborative problem-solving skills of the students were found to make statistically significant contributions to explain the degree of scientific literacy. Contrary to its univariate relation, playing video games after school changed its effect on scientific literacy scores when multiple relations with other variables are introduced.

After setting the final model, the regression equation can be written as following:

$$\hat{Y}_{SCILIT} = 17.35 - 2.51X_{ANXTEST} + 1.74X_{ENVAWARE} + 2.07X_{INTBRSCIE} - 5.32\, X_{PLAY} + .41X_{MATH} + .36X_{READ} + .21X_{CLPS}$$

Results of Model 3

A total of 57 variables emerged from Data Set 3 were refitted and backward elimination was used manually as a technique to decide on the final version of the Model 3. The insignificant variables were removed from the initial model and multiple linear regression was performed repeatedly until all explanatory variables were observed as statistically significant. The alpha value (α), the risk of committing Type I Error, was decided as to be 0.05. The related output is given in Figure 3.

Figure 3

The Final Output of Model 3

Variable	Regression Coefficient	Regression Coefficient (s.e.)	Regression Coefficient (t-value)	Stndrdzd. Coefficient	Stndrdzd. Coefficient (s.e.)	Stndrdzd. Coefficient (t-value)
(CONSTANT)	17.35	6.10	2.84	.	.	.
ANXTEST	-2.51	.86	-2.94	-.03	.01	-2.89
ENVAWARE	1.74	.44	3.94	.03	.01	3.93
INTBRSCI	2.07	.78	2.66	.03	.01	2.65
ST078Q06NA	-5.32	2.50	-2.13	-.03	.02	-2.15
PV_MATH	.41	.02	22.72	.42	.02	24.49
PV_READ	.36	.02	23.01	.38	.02	22.13
PV_CLPS	.21	.02	12.96	.21	.02	12.89

A multiple linear regression analysis was conducted to evaluate how well the variables obtained in the second version explain the degree of scientific literacy of Turkish students. According to the results, a total of 7 variables were statistically significant (($|t_{calculated}| > t_{critical}=1.96$). The sample multiple correlation coefficient was .86. Test anxiety, environmental awareness, interest in broad science topics, playing video games after school, mathematics literacy, reading literacy, and collaborative problem-solving skills of the students were found to make a statistically significant contribution to explain the degree of scientific literacy. Contrary to its univariate relation, playing video games after school changed its effect on scientific literacy scores when multiple relations with other variables are introduced.

After setting the final model, the regression equation can be written as following:

$$\hat{Y}_{SCILIT}=17.35-2.51X_{ANXTEST} + 1.74X_{ENVAWARE} + 2.07X_{INTBRSCIE} -5.32\ X_{PLAY} + .41X_{MATH} + .36X_{READ} + .21X_{CLPS}$$

Deciding on the Final Model

Model 2 was chosen as the final model. The rationale is as follows:

- Model 2 included all the significant variables that Model 1 had, but additionally involved the variable of playing video games after school. This additional variable has subject specific importance, and hence we preferred to keep it in the model.

- The results of multiple linear regression analysis yielded the same results for Model 2 and 3.

In the former analyses, all three models were tested in IDB Analyser because elastic net regression does not include plausible values, replicate weights, and students' final weights. Nevertheless, IDB Analyser does not contain ANOVA table of the model, VIF values, and *p*-values. Therefore, to provide additional details on the final model for this study, *lm* () function in R program were performed for Model 2 to check those values that are not given by IDB Analyser. The output is given in Figure 4.

Figure 4

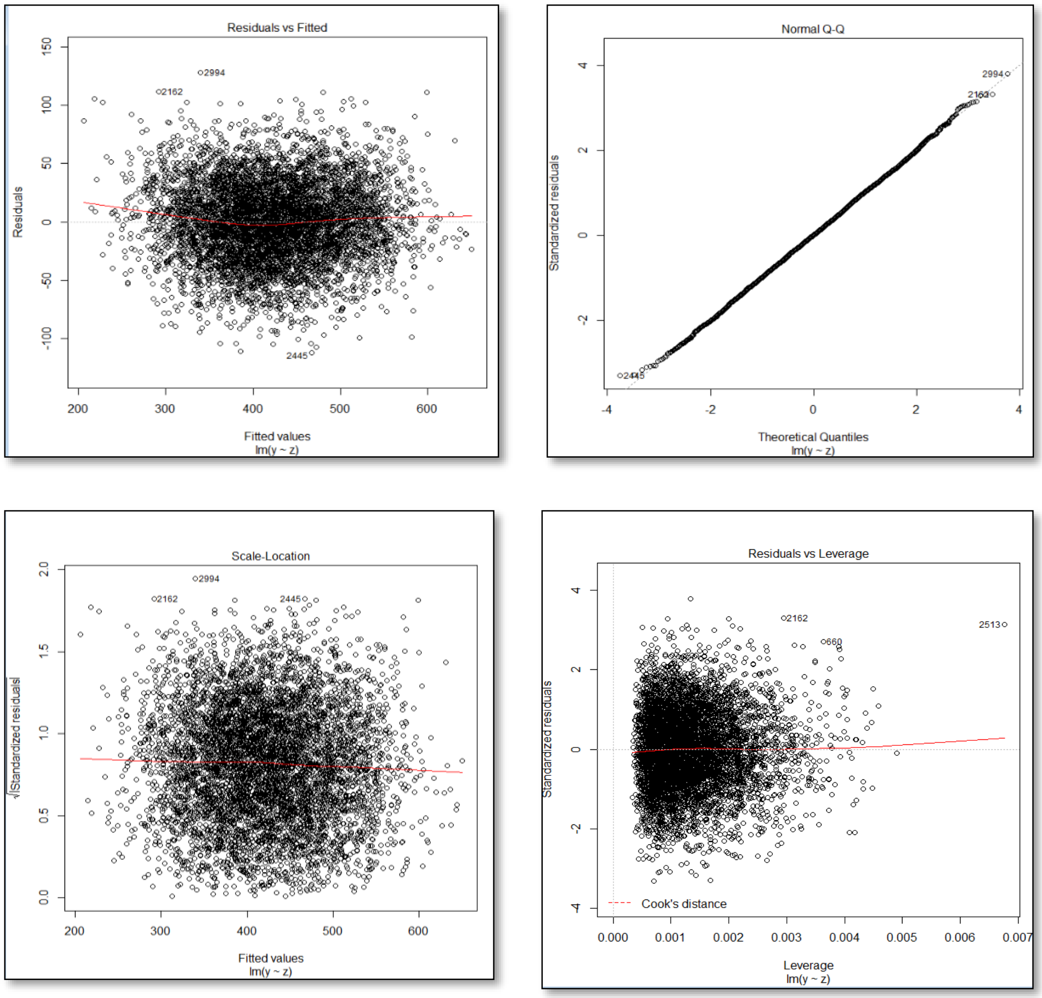
Output for Model 2

```
Call:
lm(formula = y ~ z)

Residuals:
     Min       1Q   Median       3Q      Max 
-112.022  -22.792   -0.096   22.741  128.058 

Coefficients:
             Estimate Std. Error t value Pr(>|t|)    
(Intercept) 30.672474   2.950817  10.395  < 2e-16 ***
zENVAWARE    1.797635   0.331430   5.424 6.06e-08 ***
zINTBRSCI    2.063025   0.467632   4.412 1.04e-05 ***
zANXTEST    -3.288669   0.420999  -7.812 6.64e-15 ***
zPV1MATH     0.284320   0.008157  34.856  < 2e-16 ***
zPV4READ     0.520932   0.008735  59.636  < 2e-16 ***
zPV7CLPS     0.145355   0.008099  17.948  < 2e-16 ***
zST078Q06NA -6.259642   0.903059  -6.932 4.61e-12 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 33.82 on 5887 degrees of freedom
Multiple R-squared:  0.8107,	Adjusted R-squared:  0.8105 
F-statistic:  3603 on 7 and 5887 DF,  p-value: < 2.2e-16

> vif(model22)
  zENVAWARE   zINTBRSCI    zANXTEST    zPV1MATH    zPV4READ    zPV7CLPS zST078Q06NA 
     1.1742      1.0706      1.0157      2.2797      2.5717      1.9654      1.0434 
```

A multiple linear regression analysis was conducted to evaluate how well the variables explain the degree of scientific literacy of Turkish students. The explanatory variables were test anxiety, environmental awareness, interest in broad science topics, playing video games after school, mathematics literacy, reading literacy, and collaborative problem-solving skills of the students. Preliminary analyses were conducted to ensure no violation of normality, linearity, homoscedasticity, independence of residuals assumptions and multicollinearity assumption (Figure 5). In addition, the data were inspected for outliers and no potential outliers were detected. According to the results given in Figure 4, the combination of the predictor variables was significantly related to the dependent variable (F (7, 5887) = 3603, *p-value* < 2.2 x10^{-16}). The sample multiple correlation coefficient was .82. All the coefficients were statistically significant (p<0.05).

Figure 5

Output for standardized residuals in Model 2



To sum up, Model 2 was decided as the final model for this study. The coefficients in IDB Analyser were used for the regression equation because this program includes plausible values, replicate weights, and students' final weights which reduce the bias. Accordingly, the regression equation can be written as following:

$$\hat{Y}_{SCILIT} = 17.35 - 2.51X_{ANXTEST} + 1.74X_{ENVAWARE} + 2.07X_{INTBRSCIE} - 5.32X_{PLAY} + .41X_{MATH} + .36X_{READ} + .21X_{CLPS}$$

Discussion and Conclusion

In this study, it was intended to determine a subset of variables that explain the variance in scientific literacy of Turkish students most. To be able to do this, one of the newly developed shrinkage methods was used to obtain a subset of 246 variables available in Türkiye PISA 2015 data. Before conducting the analysis, cross-validation was done for all 10 plausible values in science and 'Plausible Value 4 in Science' was chosen as a response variable for refitting elastic net regression model in the full data set because it has the lowest MSE and mean prediction error, highest deviance ratio, and tuning parameter which leads to approximately 50 parameters. Three different data sets were constructed and then, elastic net regression was refitted for each data set. After determining the number of variables, three models emerged from these three data sets were tested in IDB Analyzer. To conclude, 7 variables out of 246 variables were determined as variables that are statistically significant and correlated with the scientific literacy of the students living in Türkiye. We used the coefficients in IDB Analyzer in our final model. For further studies, we recommend that this final model can be fitted for all proficiency levels defined by OECD (2016) separately to see whether different proficiency levels have the same factors that we found in our final model.

Shrinkage methods are often used in statistical modeling to address issues such as multicollinearity, overfitting, and the inclusion of irrelevant predictors. These methods have both advantages and disadvantages. They can improve predictive performance by preventing overfitting and avoiding multicollinearity issues. Shrinkage can also reduce the variance of model coefficients, leading to more stable estimates, which allow for better generalization of the results. On the other hand, the choice of tuning parameters can be challenging and requires special attention. We addressed this challenge with a grid search algorithm in our study. Moreover, although these methods reduce the variance, they have the potential of introducing some bias into the estimates of coefficients. This is particularly true when the relationships between the response and covariates are complex, as shrinking the coefficients towards zero might lead to bias. Based on this model, one can argue that the degree of scientific literacy is directly proportional to environmental awareness, interest in broad science topics, playing videogames after school, the degree of mathematics literacy, the degree of reading literacy and collaborative problem-solving skills of the students. Test anxiety, on the other hand, led to a decrease in scientific literacy. However, the magnitude of regression coefficients should not be compared directly since the variables are not standardized. That is, it is not reasonable to conclude that test anxiety is more important than mathematics literacy in understanding scientific literacy. Standardized coefficients would be better in such comparisons.

These results are compatible with previous studies related to scientific literacy. In terms of the relationship between environmental awareness and scientific literacy, Öztürk (2018) reported similar interaction between environmental awareness and scientific literacy of Turkish students. This relationship is not surprising because especially in the last 25 years, science educators have been focusing on environmental issues which create a learning environment that fosters both raising environmental awareness (Wals, 2011) and level of knowledge of the students in science (Hadzigeorgiou & Skoumios, 2013). The positive relationship between interest in broad science topics and scientific literacy were evident in previous scientific literacy studies (e.g., Chang & Cheng, 2008; Grabau & Ma, 2017; Akgenç & Yapıcı Pehlivan., 2019). In our study, we provided evidence to support this positively correlated relationship in Türkiye context.

In our study, we used math literacy and reading literacy as explanatory variables different from previous PISA 2015 studies, and we found that they are positively correlated with scientific literacy. It has been evident that these three literacy frameworks are inherently correlated with each other (Arıkan et al. 2017, Bybee, 2010, Kullman, 1966). In PISA 2012 cycle, Arıkan et al. (2017) reported that reading literacy predicted mathematics literacy and scientific literacy of Turkish students in PISA 2012. Therefore, our results can be considered as compatible with previous PISA studies conducted in Türkiye context. In addition, we also included all plausible values of collaborative problem-solving as an explanatory variable one and detected a direct relation between scientific literacy and collaborative problem-solving. It was reported that problem-solving skill is one of the components of scientific literacy (Palincsar et al., 1993) and our study provided an indicator for this interaction.

Test anxiety was the only explanatory variable inversely proportional to scientific literacy. There are similar patterns in science education literature of Türkiye between these two variables (e.g., Genc, 2017; Haşıloğlu & Göğebakan, 2021). These studies generally reported that higher level of test anxiety lowers the academic performance of the students which were also apparent in our results.

Perhaps, playing video games after school and its positive relationship with scientific literacy was an unexpected result contrary to its univariate result. Whereas the univariate results indicated that playing video games after school decreases the scientific literacy scores, our final model implied quite the opposite. One of the most common reasons for such a change in sign is due to multicollinearity problem. However, our VIF results were smaller than 3 (see Figure 4), which provides an indication that there was no such problem with this model.

Another possible reason for this change in sign can be interpreted as multiple relations with other variables in the regression may contribute to reversing its relationship. In fact, we believe we are observing an example of Lord's paradox here (Tu et al., 2008). According to this paradox, the sign of the relationship between a continuous response (scientific literacy, in our case) and a categorical variable (playing video games after school, in our case) could be reversed with the introduction of a continuous covariate (e.g., reading literacy or collaborative problem solving). Following Tu et al. (2008) rationale, we claim that the relation between scientific literacy and playing video games after school would change because we have more than one continuous covariate in our model that positively correlated with both scientific literacy and playing video games after school.

The literature on playing games and academic achievement is growing and some trend studies (e.g., Young et al., 2012) reached a conclusion that playing video games influence language learning, history, and physical education but found little support promoting science and mathematics performance of K-12 students. Another study (Chaarani et al., 2022) aiming to explore the association between playing video games and cognition among 9- and 10-year-old children reported better cognitive performances for children playing video games.

In the context of our study, we believe that this relationship warrants further attention in future studies. It is important to note that merely playing video games after school may not lead to an improvement in scientific literacy, and instead may have a potential to imply some other latent relationships that require exploration. For example, this can be an indication of the effect of some other variables such as level of socio-economic status which did not appear as a variable in our study but in other PISA 2015 Türkiye studies used hierarchical linear modeling (e.g., Karslı et al., 2019, Üstün et al., 2020).

A Brief Discussion on Previous PISA 2015 Results of Türkiye and Possible Implications

Our study is not the only one that explores students' scientific literacy living in Türkiye through examining PISA 2015 data. In those studies, approximately 30 variables at the student level and 20 variables at the school level were used in different combinations (see Table 1) and discussed possible implications for science education policy in Türkiye. Among the variables, gender, teacher support in science class, ESCS, disciplinary climate in science class, grade level, and school type were most frequently used ones that significantly explained the variability of scientific literacy scores of the students some of which did not appear as significant predictors in our study. There might be possible explanations for this difference.

All previous studies related to PISA 2015 that we examined (except for Akgenç & Yapıcı Pehlivan, 2019 and Yıldız et al., 2020) did not test the variables that we found significant in our model to explain the degree of scientific literacy of Turkish students. For example, neither of these studies included math literacy and reading literacy in their analysis even though Karslı et al. (2019) acknowledged that it would be erroneous to expect a development in science and mathematics literacy without improving students' reading literacy. In our model, math literacy and reading literacy explained more than 60% of the variability. For this reason, the other variables that the previous studies used might be eliminated during the variable selection stage of our study due to the shrinkage nature of these methods.

In our study, we used a different analysis technique than the other available PISA 2015 studies conducted in Türkiye which might be considered as a reason for this divergence in our results. We gained the benefits of elastic net regression which does a feature selection on its own and used a total of 246 variables to be able to test further which explanatory variables significantly predicted scientific literacy. Previous PISA 2015 studies, on the other hand, employed Hierarchical Linear Modeling (HLM), chose different sets of variables that were different from each other and applied models on these assumedly related factors without providing a rationale for their variable selection method. What is more, some studies used an average of plausible values or selected one of them in their analysis which is not recommended by OECD (2009). At this point, we would like to note that we appreciate their choice of explanatory (hierarchical) variables, possibly depending on the know-how of the literature which have potentially provide useful insights. However, we would like to also emphasize that these studies were published between 2018 – 2021 and they did not provide a fruitful discussion to the science education community about the similarities and differences among their results and the other available PISA (2015) studies conducted in Türkiye. These inconsistencies have a potential to create challenges to inform policymakers about how to improve science education policies in Türkiye. Thus, we believe that introducing a variable selection method to the international large scale assessment literature might be useful to inform policymakers, educators, and societies on how to improve scientific literacy in Türkiye. Based on our model, we invite science education policymakers to formulate or amend policies on promoting scientific literacy of students by considering the interactions among scientific literacy, reading literacy, math literacy, and collaborative problem-solving skills.

Conclusion, Limitations, and Future Directions

There are some limitations in this study. First, we acknowledge that using hierarchical linear modeling tends to produce better results for education data sets due to its nature (Tat et al., 2019).

In this study, however, our main focus was to introduce a variable selection method for international large-scale assessments which have many variables to be tested. We are aware of the views on the limitations of using multiple linear regression for these types of data sets. For further studies, we recommend using hierarchical linear modeling after employing elastic net regression.

Secondly, we used IDB Analyzer because it includes students' weights and replicate weights as well as allow us to use every PV in multiple linear regression both as a single response and a single explanatory variable, but it also has a limited number of options. Moreover, it only accepts the variables already defined in PISA data set, does not offer all the diagnostics that is required to construct a valid multiple regression model, and do not allow users to introduce interaction terms. Therefore, we tried to compensate for this limitation by using R Project for Statistical Computing. Correspondingly, while we had 2 dummy variables (we divided the variable of stratum into two) in the analysis performed in R, it was not possible in IDB Analyzer. We accepted the coefficients in IDB Analyzer in our final model. For further studies, we recommend that this final model can be fitted for all proficiency levels defined by OECD (2016) separately to see whether different proficiency levels have the same factors that we found in our final model.

Some of the advantages of working with plausible values include capturing uncertainty and improving estimation for traits that are not directly measured. Latent variables, such as cognitive ability, are common in survey and assessment contexts. The use of plausible values and weighted likelihood estimates can improve estimation in the presence of underlying latent variables. However, combining shrinkage methods with plausible values complicates the analysis. These analyses require special software, and unfortunately, one software is not sufficient to handle all the analyses. The use of two software programs, R and IDB Analyzer, introduces some computational challenges.

In terms of data wrangling, possible limitations can be summarized as follows: for the missing values, mean/mode substitution were used as imputation methods which are one of the oldest imputation techniques that have many disadvantages. For further studies, we recommend using multiple imputation techniques where possible. On the other hand, deleting variables that have comparably higher percentage of missing data was another limitation for this study. Including them may improve the results of these kinds of studies in the future. Lastly, we excluded nominal variables that have too many categories and this resulted in loss of information.

Ethics Committee Permission Information: Not applicable because this study used the public PISA 2015 data. The datasets analyzed during the current study are available at PISA 2015 database https://www.oecd.org/pisa/data/2015database/

Acknowledgements: This paper was based on first author's master's thesis at Middle East Technical University.

Declaration of Interest: The authors declare that there is no conflict of interest.

Statement of Contribution Rate: Author 1 conducted literature review. Author 1 and Author 2 conceptualized the study and interpret the findings together and made equal contributions to methodology and writing the manuscript. Author 2 supervised the research process. All authors read and approved the final manuscript.

References

Akgenç, E. & Yapıcı Pehlivan, N. (2019). Analysis of PISA-2015 performance of Turkish students by multilevel structural equation modeling. *Mugla Journal of Science and Technology, 5*(1), 43–51. https://doi.org/10.22531/muglajsci.484469

Arıkan, S., Yıldırım, K., & Erbilgin, E. (2017). Exploring the relationship among new literacies, reading, mathematics and science performance of Turkish students in PISA 2012. *International Electronic Journal of Elementary Education*, *8*(4), 573–588. https://www.iejee.com/index.php/IEJEE/article/view/133

Bybee, R. W. (2010). What is STEM education? *Science, 329*(5995), 996. https://doi.org/10.1126/science.1194998

Carter, L. (2008). Sociocultural influences on science education: Innovation for contemporary times. *Science Education*, *92*(1), 165–181. https://doi.org/10.1002/sce.20228

Chang, C. Y., & Cheng, W. Y. (2008). Science achievement and students' self-confidence and interest in science: A Taiwanese representative sample study. *International Journal of Science Education*, *30*(9), 1183–1200. https://doi.org/10.1080/09500690701435384

Chaarani, B., Ortigara, J., Yuan, D., Loso, H., Potter, A., & Garavan, H. P. (2022). Association of video gaming with cognitive performance among children. *JAMA Network Open*, *5*(10), e2235721. https://doi.org/10.1001/jamanetworkopen.2022.35721

Choi, K., Lee, H., Shin, N., Kim, S. W., & Krajcik, J. (2011). Re-conceptualization of scientific literacy in South Korea for the 21st century. *Journal of Research in Science Teaching*, *48*(6), 670–697. https://doi.org/10.1002/tea.20424

Coll, R.K., Taylor, N. (2012). An international perspective on science curriculum development and implementation. In Fraser, B., Tobin, K., McRobbie, C. (Eds.) *Second international handbook of science education*. Springer. https://doi.org/10.1007/978-1-4020-9041-7_51

Demirci, S. (2018). *A Closer look to Turkish students' scientific literacy: what do pisa 2015 results tell us?* [Master's thesis, Middle East Technical University].

Dolu, A. (2020). Sosyoekonomik faktörlerin eğitim performansı üzerine etkisi: PISA 2015 Türkiye örneği. *Journal of Management and Economics Research, 18*(2), 41–58. https://doi.org/10.11611/yead.607838

Erbas, A. K., Tuncer Teksoz, G., & Tekkaya, C. (2012). An Evaluation of Environmental Responsibility and Its Associated Factors: Reflections from PISA 2006. *Eurasian Journal of Educational Research*, *46*, 41–62. https://eric.ed.gov/?id=EJ1057292

Fraenkel, J. R., Wallen, N. E., & Hyun, H. H. (2012). *How to design and evaluate research in education* (8th ed.). McGraw-Hill.

Friedman, J. Hastie, T., Tibshirani, R., Simon, N., Narasimhan, B. & Qian, J. (2018). Lasso and Elastic-Net Regularized Generalized Linear Models. https://cran.r-project.org/web/packages/glmnet/glmnet.pdf

Genc, A. (2017). Coping strategies as mediators in the relationship between test anxiety and academic achievement. *Psihologija*, *50*(1), 51–66. https://doi.org/10.2298/PSI160720005G

Grabau, L. J., & Ma, X. (2017). Science engagement and science achievement in the context of science instruction: a multilevel analysis of US students and schools. *International Journal of Science Education*, *39*(8), 1045–1068. https://doi.org/10.1080/09500693.2017.1313468

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning with applications in R*. Springer.

Hadzigeorgiou, Y., & Skoumios, M. (2013). The development of environmental awareness through school science: problems and possibilities. *International Journal of Environmental and Science Education*, *8*(3), 405–426. https://files.eric.ed.gov/fulltext/EJ1016851.pdf

Haşıloğlu, M. A. & Göğebakan, S. (2021). Ortaokul 8. sınıf öğrencilerinin fen bilimleri dersine yönelik kaygılarının bazı değişkenler açısından incelenmesi. *Fen Matematik Girişimcilik ve Teknoloji Eğitimi Dergisi,* *4*(2), 141–154. https://dergipark.org.tr/en/pub/fmgted/issue/62218/888624

IEA Data Processing and Research Center (2018). New Features and Installation Guide for the IEA's IDB Analyzer (Version 3.2). https://www.iea.nl/fileadmin/user_upload/IEA_Software/Help_Manual_for_the_IDB_Analyzer__Version_3_2__-_Features_and_Install_Guide_01.pdf

Karslı, N., Berberoğlu, G. & Çalışkan, M. (2019). Türkiye'de PISA fen okuryazarlık puanlarını yordayan değişkenler. *Uluslararası Bilim ve Eğitim Dergisi, 2* (2), 38–49. https://dergipark.org.tr/en/pub/ubed/issue/50464/567861

Kaya, V. H. (2017). Okuma becerilerinin fen bilimleri okuryazarlığına etkisi. *Milli Eğitim Dergisi*, *46*(215), 193–207. https://dergipark.org.tr/en/pub/milliegitim/issue/36134/405905

Kullman, D. E. (1966). Correlation of mathematics and science teaching. *School Science and Mathematics*, *66*(7), 645–649. https://doi.org/10.1111/j.1949-8594.1966.tb13617.x

Laugksch, R. C. (2000). Scientific literacy: A conceptual overview. *Science Education*, *84*(1), 71–94.https://doi.org/10.1002/(SICI)1098-237X(200001)84:1%3C71::AID-SCE6%3E3.0.CO;2-C

Lewis, J. D. (1982). Technology, enterprise, and American economic growth. *Science*, *215*(4537),1204–1211. https://doi.org/10.1126/science.215.4537.1204

Ministry of National Education [MONE] (2018). *PISA 2015 ulusal raporu*. http://pisa.meb.gov.tr/wp-content/uploads/2014/11/PISA2015_UlusalRapor.pdf

Organisation for Economic Co-Operation and Development [OECD] (2009). *PISA Data Analysis Manual SPSS, Second Edition*.

http://archivos.agenciaeducacion.cl/Manual_de_Analisis_de_datos_SPSS_version_ingles.pdf

Organisation for Economic Co-Operation and Development [OECD] (2016). *PISA 2015 Results (Volume I) Excellence and Equity in Education. OECD Publishing*. http://dx.doi.org/10.1787/9789264266490-en

Organisation for Economic Co-Operation and Development [OECD] (2017). *Education at a Glance 2017: OECD Indicators*. OECD Publishing. http://dx.doi.org/10.1787/eag-2017-en

Öztürk, Ö. (2018). *Using PISA 2015 data to analyze how the scientific literacy of students from different socioeconomic levels can be predicted by environmental awareness and by environmental optimism* (Publication No. 29045663). [Master's Thesis, Ihsan Dogramaci Bilkent University]. ProQuest Dissertations and Theses Global.

Palincsar, A. S., Anderson, C., & David, Y. M. (1993). Pursuing scientific literacy in the middle grades through collaborative problem solving. *The Elementary School Journal*, *93*(5), 643–658. https://doi.org/10.1086/461745

Pedretti E. (2014). Environmental education and science education: ideology, hegemony, traditional knowledge, and alignment. *Revista Brasileira de Pesquisa em Educação em Ciências, 14*(2), 305–314.

Pedretti, E., & Nazir, J. (2011). Currents in STSE education: Mapping a complex field, 40 years on. *Science Education*, *95*(4), 601–626.

R Core Team (2022). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/.

Roberts, D. (2007). Scientific literacy/science literacy. In S. K. Abell & N. G. Lederman (Eds.), *Handbook of research on science education* (pp. 729–780). Mahwah, NJ: Lawrence Erlbaum Associates.

Roberts, D. A., & Bybee, R. W. (2014). Scientific literacy, science literacy, and science education. In Lederman N. G. & S. K. Abell (Eds.), *Handbook of research on science education, Volume II* (pp. 559–572). Routledge. https://doi.org/10.4324/9780203097267

Santi, V. M., Notodiputro, K. A., & Sartono, B. (2019, December). Variable selection methods applied to the mathematics scores of Indonesian students based on convex penalized likelihood. In *Journal of Physics: Conference Series* (Vol. 1402, No. 7, p. 077096). IOP Publishing. https://doi.org/10.1088/1742-6596/1402/7/077096

Solomon, J., & Aikenhead, G. (1994). *STS Education: International Perspectives on Reform. Ways of Knowing Science Series*. Teachers College Press, 1234 Amsterdam Ave., New York, NY 10027 (clothbound: ISBN-0-8077-3366-0; paperback: ISBN-0-8077-3365-2).

Tat, O., Koyuncu, İ. & Gelbal, S. (2019). The influence of using plausible values and survey weights on multiple regression and hierarchical linear model parameters. *Journal of Measurement and Evaluation in Education and Psychology*, *10*(3), 235–248. https://doi.org/10.21031/epod.486999

Tu, Y. K., Gunnell, D., & Gilthorpe, M. S. (2008). Simpson's Paradox, Lord's Paradox, and Suppression Effects are the same phenomenon–the reversal paradox. *Emerging Themes in Epidemiology*, *5*(1), 1–9. https://doi.org/10.1186/1742-7622-5-2

Üstün, U., Özdemir, E., Cansız, M., & Cansız, N. (2020). Türkiye'deki öğrencilerin fen okuryazarlığını etkileyen faktörler nelerdir? PISA 2015 verisine dayalı bir hiyerarşik doğrusal modelleme çalışması. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi, 35*(3), 720–732. https://doi.org/10.16986/HUJE.2019050786

Valladares, L. (2021). Scientific literacy and social transformation. *Science & Education*, *30*(3), 557–587. https://doi.org/10.1007/s11191-021-00205-2

Wals, A. (2011). Learning our way to sustainability. *Journal of Education for Sustainable Development, 5*, 177–86. https://doi.org/10.1177/097340821100500208

Wright, B. D., & Masters, G. N. (1982). *Rating scale analysis: Rasch measurement*. Mesa Press.

Yetişir, M. İ., & Kaan, B. A. T. I. (2021). The effect of school and student-related factors on PISA 2015 science performances in Türkiye. *International Journal of Psychology and Educational Studies*, *8*(2), 170–186. https://dergipark.org.tr/en/pub/pes/issue/62298/935968

Yıldız, M., Erdas Kartal, E., & Mesci, G. (2020). Investigation of Türkiye's PISA 2015 Science Performance and Associated Variables Using Hierarchical Linear Modeling. *Necatibey Faculty of Education Electronic Journal of Science & Mathematics Education*, *14*(1). https://doi.org/10.17522/balikesirnef.663737

Young, M. F., Slota, S., Cutter, A. B., Jalette, G., Mullin, G., Lai, B., ... & Yukhymenko, M. (2012). Our princess is in another castle: A review of trends in serious gaming for education. *Review of Educational Research*, *82*(1), 61–89. https://doi.org/10.3102/0034654312436980

Zeidler, D. L., Sadler, T. D., Applebaum, S., & Callahan, B. E. (2009). Advancing reflective judgment through socioscientific issues. *Journal of Research in Science Teaching*, *46*(1), 74–101.

Zou, H., & Hastie, T. (2005). Regression shrinkage and selection via the elastic net, with applications to microarrays. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, *67*(2), 301–320. https://doi.org/10.1111/j.1467-9868.2005.00527.x

Appendix 1: List of Variables of Model 2

1.	IBTEACH	Inquiry-based science teaching and learning practices (WLE)
2.	ENVAWARE	Environmental Awareness (WLE)
3.	ENVOPT	Environmental optimism (WLE)
4.	INTBRSCI	Interest in broad science topics (WLE)
5.	EPIST	Epistemological beliefs (WLE)
6.	SCIEACT	Index science activities (WLE)
7.	SMINS	Learning time (minutes per week) - <science>
8.	ANXTEST	Personality: Test Anxiety (WLE)
9.	unfairteacher	Teacher Fairness (Sum)
10.	SC016Q01TA	Percent. total funding for school year comes from? Government
11.	SC019Q02NA01	<School science> teachers <fully certified> by <the appropriate authority>: Full-time
12.	SC064Q03TA	<the last academic year>, what proport. of parents part. school-related activity? Partici. in local school government
13.	RATCMP1	Number of available computers per student at modal grade
14.	LEADCOM	Curricular development (WLE)
15.	SCHAUT	School autonomy (Mean)
16.	TOTST	Total number of science teachers at school
17.	STRATIO	Student-Teacher ratio
18.	PV1MATH	Plausible Value 1 in Mathematics
19.	PV4READ	Plausible Value 4 in Reading
20.	PV7CLPS	Plausible Value 7 in Collaborative Problem Solving
21.	ST004D01T	Student (Standardized) Gender
22.	ST125Q01NA	How old were you when you started <ISCED 0>? Years
23.	ST031Q01NA	On avg, how many days do you attend physical education classes each week?
24.	ST063Q06NA	Which <school science> course did you attend? <General, integrated, or comprehen science> course: This year
25.	ST063Q06NB	Which <school science> course did you attend? <General, integrated, or comprehen science> course: Last year
26.	ST064Q01NA	<school science> courses? I can choose the <school science> course(s) I study.
27.	ST076Q01NA	Before going to school did you: Eat breakfast
28.	ST076Q06NA	Before going to school did you: Play video-games
29.	ST076Q07NA	Before going to school did you: Meet friends or talk to friends on the phone

30.	ST076Q11NA	Before going to school did you: Exercise or practice a sport
31.	ST078Q06NA	After leaving school did you: Play video-games
32.	ST078Q07NA	After leaving school did you: Meet friends or talk to friends on the phone
33.	ST078Q08NA	After leaving school did you: Talk to your parents
34.	ST078Q09NA	After leaving school did you: Work in the household or take care of other family members
35.	ST016Q01NA	Overall, how satisfied are you with your life as a whole these days?
36.	ST038Q04NA	Other students made fun of me.
37.	ST038Q05NA	I was threatened by other students.
38.	ST038Q06NA	Other students took away or destroyed things that belonged to me.
39.	ST038Q07NA	I got hit or pushed around by other students.
40.	SC053D11TA	<This academic year>, follow. activities\school offers<national modal grade for 15-year-olds>? <country specific item>
41.	SC059Q01NA	Compared to other departments, our school's <school science department> is well equipped.
42.	SC059Q06NA	We have enough laboratory material that all courses can regularly use it.
43.	SC052Q01NA	Does your school provide study help? Room(s) where the students can do their homework
44.	SC009Q02TA	Frequency of <the last academic year>. I make sure that the professional development activities of teachers are in
45.	SC009Q05TA	Frequency of <the last academic year>. I praise teachers whose students are actively participating in learning.
46.	SC009Q11TA	Frequency of <the last academic year>. I ask teachers to participate in reviewing management practices.
47.	SC009Q12TA	Frequency of <the last academic year>. When a teacher brings up a classroom problem, we solve the problem together
48.	SC027Q04NA	Our school organises in-service workshops for specific groups of teachers (e.g. newly appointed teachers).
49.	SC034Q01NA	How often are students assessed? Mandatory <standardized tests>
50.	SC034Q04TA	How often are students assessed? Teachers’ judgmental ratings
51.	SC035Q07TA	Are <standardized tests> used in school? To make judgements about teachers' effectiveness
52.	SC035Q09NA	Are <standardized tests> used in school? To adapt teaching to the students' needs

53.	SC037Q03TA	Does improvement exist at school? Written specification of the schools curricular profile and educational goals
54.	SC037Q08TA	Does improvement exist at school? Teacher mentoring
55.	SC040Q15NA	Did your school implement any measures in: Student achievement
56.	region2	Region 2
57.	region6	Region 6
58.	region7	Region 7
59.	region10	Region 10
60.	typofsch2	Type of School 2

Geniş Özet

Problem Durumu

Uluslararası büyük ölçekli değerlendirmeler, eğitim, ekonomik ve politik sistemlerin iyileştirilmesinde önemli bir rol oynamaktadır. Ülkeler, bu değerlendirmelerin verilerini kullanarak, eğitim sistemlerinin mevcut durumu hakkında çıkarımlarda bulunmaktadır. Bu büyük ölçekli değerlendirmelerden biri olan PISA, ülkeler arası yapılan, 15 yaş öğrencilerinin belirli konularda bilgi ve beceri seviyelerini ölçen bir test olarak 2000 yılından beri uygulanmaktadır. Her üç yılda bir uygulanan bu test, her uygulandığı zaman aralığında bir konuyu ana tema olarak belirlemektedir. 2015 yılında PISA testinin ana teması olarak bilimsel okuryazarlık belirlenmiştir.

Uluslararası büyük ölçekli değerlendirmelerin verilerini kullanan bilimsel çalışmalar ve raporlar, genellikle veri setinde mevcut olan bazı değişkenleri seçerek bu değişkenler arasındaki ilişkileri modellemeyi amaçlar. Bu çalışmada, Türkiye PISA 2015 verisinin tamamını kullanarak ülke verilerini modellemede hangi değişkenlerin dahil edilebileceğine karar vermek amacıyla bir değişken seçim yöntemi denemeyi hedeflenmiştir. PISA 2015 verisinin tamamını kullanarak büzüşme regresyonlarından biri olan elastik net regresyonu kullanılmış ve elde edilen sonuçlar, Türkiye PISA 2015 verilerine dayalı mevcut çalışmaların sonuçları ile karşılaştırılmıştır.

Buna göre, bu çalışmanın araştırma soruları şu şekilde belirlenmiştir:

1. Öğrencilerin bilimsel okuryazarlık düzeyi ile en çok ilişkili olan değişkenler hangileridir?

2. Elastik net regresyon ve çoklu doğrusal regresyon sonucunda ortaya çıkan değişkenler arasında bir uyum var mıdır?

Uluslararası büyük ölçekli değerlendirmeler, ülkeler için eğitimsel, ekonomik ve politik açıdan önemli bir rol oynamaktadır çünkü OECD'nin (2009) vurguladığı gibi, ülkelerin ekonomik ve sosyal refahı büyük ölçüde vatandaşlarının bilgi ve beceri düzeyleriyle ilişkilidir. Büyük ölçekli değerlendirmelerin verilerini kullanarak, ülkeler değişkenler arasındaki ilişkileri analiz edebilir ve ulusal eğitim sistemleri hakkında sonuçlar çıkarabilirler. Bu noktada, bu çalışma, Türk öğrencilerin bilimsel okuryazarlık düzeylerinin Türkiye'deki çeşitli demografik, sosyal, ekonomik ve eğitimsel değişkenlerle nasıl ilişkili olabileceği konusunda eğitim politikası yapıcılarına, eğitimcilere, velilere ve öğrencilere yararlı bilgiler sağlayabilir. Ayrıca, Türk öğrencilerinin bilimsel okuryazarlık düzeylerini incelemek, bu değişkenleri kullanarak öğrencilerimizin bilimsel okuryazarlıklarını nasıl artırabileceğimize dair ipuçları da verebilir. Uzun vadede, ulusal eğitim sistemlerimizde iyileştirmeler yapmak ve kendi eğitim sistemlerimizin göreceli güçlü ve zayıf yönlerini anlamak için bir fikir sunabilir.

Giriş bölümünde de belirtildiği gibi, PISA 2015 verileriyle Türk öğrencilerin bilimsel okuryazarlık düzeyini inceleyen çeşitli bilimsel çalışmalar mevcuttur. Alana katkı sağlayan bu değerli çalışmalardan farklı olarak bu çalışmada değişken seçimi için esnek net regresyonunu test edilmiş; bu sayede Türkiye PISA 2015 verilerindeki toplam 246 değişkeni dâhil edilerek bu değişkenler arasında Türk öğrencilerin bilimsel okuryazarlık düzeyiyle ilişkili olabilecek açıklayıcı değişkenler belirlenmiştir. Bu nedenle, çalışmamızın alanyazınına, büzüşme regresyonlarından biri olan elastik net regresyonu yöntemiyle 200'den fazla değişkenin test edildiği bir ortamda hangi değişkenlerin bilimsel okuryazarlık düzeyini en çok tahmin edebileceği konusunda katkı sağlayabilecek bir potansiyele sahip olduğuna inanıyoruz.

Yöntem

Bu çalışmada nicel bir araştırma yöntemi olan korelasyonel araştırma kullanılmıştır. Büzüşme regresyonlarından biri olan elastik net regresyonu, çalışmanın amacı kapsamında analiz yöntemi olarak seçilmiştir. Elastik net regresyonu sonucunda elde edilen değişkenler, çoklu doğrusal regresyon analizi kullanılarak modellenmiştir.

Türkiye'deki 15 yaşındaki öğrencilerin ulaşılabilir popülasyonu 925.366 öğrenci olarak belirlenmiştir (MEB, 2018). Bu çalışmada, PISA 2015'e katılan Türk öğrenciler örneklem olarak seçilmiştir. PISA 2015 için 12 istatistiksel bölgeden 61 şehri temsil eden 187 okuldan toplam 5.895 öğrenci örneklem olarak seçilmiştir.

Bulgular

Mevcut aşamada veri analizi çoklu basamaklar halinde gerçekleştirilmiştir. Öncelikle çapraz geçerlik sonuçlarına göre analizde kullanılacak *makul değer* belirlenmiştir. Bu aşamadan sonra *makul değerlerin* farklı kombinasyonlarını içeren üç farklı veri setine elastik net regresyonu uygulanmış ve belirlenen değişkenler, çoklu doğrusal regresyon analizi kullanılarak modellenmiştir. Analizler sonucunda Model 2, diğer modeller arasında en iyi çalışan model olarak belirlenmiştir. Çalışma sonucunda sınav kaygısı, çevresel farkındalık, bilim konularına ilgi, okuldan sonra bilgisayar oyunu oynama, matematik okuryazarlığı, okuma becerisi ve ortaklaşa problem çözme becerisinin fen okuryazarlığının belirlenmesinde en önemli değişkenler olduğu gözlenmiştir.

Sonuç ve Tartışma

Bu çalışmada, Türk öğrencilerin bilimsel okuryazarlıklarındaki varyansı en çok açıklayan bir değişken alt kümesini belirlemek amaçlanmıştır. Sonuç olarak, 246 değişken arasından 7'si istatistiksel olarak anlamlı ve Türkiye'de yaşayan öğrencilerin bilimsel okuryazarlığı ile ilişkili olarak belirlenmiştir. Çalışmada, önceki PISA 2015 çalışmalarından farklı olarak açıklayıcı değişkenler olarak matematik okuryazarlığı ve okuma okuryazarlığını kullanılarak bu değişkenlerin bilimsel okuryazarlıkla ilişkili olduğu gösterilmiştir. Bu üç okuryazarlık çerçevesinin birbirleriyle doğal olarak ilişkili olduğu önceki PISA çalışmalarında ortaya konulmuştur (Kullman, 1966; Bybee, 2010; Arıkan ve ark., 2017). Dolayısıyla, sonuçlarımız Türkiye bağlamında yapılan önceki PISA çalışmalarıyla uyumlu olarak değerlendirilebilir.

Türkiye'nin PISA 2015 Sonuçları ve Olası Çıkarımlar

Bu çalışma, Türkiye'de yaşayan öğrencilerin bilimsel okuryazarlığını PISA 2015 verilerini inceleyerek araştıran tek çalışma değildir. Önceki çalışmalarda, öğrenci düzeyinde yaklaşık 30 ve okul düzeyinde ise 20 değişken farklı kombinasyonlarda kullanılmış (bkz. Tablo 1) ve Türkiye'deki fen eğitimi politikası için olası çıkarımlar tartışılmıştır. Bu değişkenlerden bazıları araştırmamızda yordayıcı olarak ortaya çıkmamıştır. Bu farklılıklara ilişkin olası açıklamaları ayrı bir başlık altında değerlendirerek Türkiye’nin fen eğitimi politikalarına ve PISA fen okuryazarlığı veri setiyle yapılacak gelecekteki araştırmalarına katkı sağlanması hedeflenmektedir.

İncelenen PISA 2015 ile ilgili önceki çalışmaların tümü (Akgenç & Yapıcı Pehlivan, 2019 ve Yıldız ve ark., 2020 hariç), Türk öğrencilerin bilimsel okuryazarlığını açıklamak için modelimizde istatistiksel olarak anlamlı bulduğumuz değişkenleri test etmemiştir. Örneğin, bu çalışmaların hiçbiri matematik okuryazarlığını ve okuma okuryazarlığını analizlerine dâhil etmemiştir; oysa Karslı ve ark. (2019), öğrencilerin okuma okuryazarlığını geliştirmeden bilim ve matematik okuryazarlığında bir gelişim beklemenin hatalı olacağını ifade etmiştir.

Mevcut çalışmada, Türkiye'de yapılan diğer PISA 2015 çalışmalarından farklı bir analiz tekniği kullanılması, sonuçlardaki farklılıklar için başka bir neden olarak görülebilir. Diğer yandan, önceki PISA 2015 çalışmaları, Hiyerarşik Doğrusal Modellemeden (HLM) faydalanmış, birbirinden farklı değişken setleri seçmiş ve bu varsayılan ilişkiler üzerine modeller uygulamışlardır. Ancak analizlerinde kullandıkları değişken setlerinin neden birbirinden farklı olduğuna ilişkin bir dayanağa rastlanmamıştır. Bu noktada, alanyazındaki bilgi birikimine dayanarak kullanılan farklı değişken setlerinin alanyazın için değerli olduğunu ve bu seçimlerin potansiyel olarak faydalı içgörüler sağladığını belirtmek isteriz. Bununla birlikte, bu çalışmaların 2018-2021 yılları arasında yayımlandığını ve Türkiye'de yapılan diğer PISA (2015) çalışmaları ile sonuçları arasındaki benzerlikler ve farklılıklar hakkında fen eğitimi araştırmacılarına ve politika yapıcılara verimli bir tartışma sunmadıklarını da vurgulamak isteriz. Bu tutarsızlıklar, Türkiye'de fen eğitimi politikalarını nasıl iyileştireceğine dair politika yapıcıları bilgilendirmede zorluklar yaratma potansiyeline sahiptir. Bu nedenle, uluslararası büyük ölçekli değerlendirme literatürüne bir değişken seçim yöntemi sunmanın, politika yapıcıları ve eğitimcileri Türkiye'deki bilimsel okuryazarlığın nasıl geliştirilebileceği konusunda bilgilendirmesi açısından faydalı olabileceğine inanıyoruz.