Research Article https://doi.org/10.12973/ijem.7.3.487

International Journal of Educational Methodology
Volume 7, Issue 3, 487 - 500.
ISSN: 2469-9632
https://www.ijem.com/

Educational Data Mining: The Analysis of the Factors Affecting Science Instruction by Clustering Analysis*

Mehmet Taha Eser**
Adnan Menderes University, TURKEY

Derya Çobanoğlu Aktan
Hacettepe University, TURKEY

Received: February 19, 2021 ▪ Revised: May 2, 2021 ▪ Accepted: July 31, 2021

Abstract: Science literacy, which is included in Programme for International Student Assessment (PISA) as an assessment area, is an important research and discussion area of science education literature with all its dimensions. In this study, the clustering results of the students from 34 Organization for Economic Cooperation and Development (OECD) countries participating in the PISA 2015 test and sampled by systematic sampling method are obtained by K-Means Clustering and Two-Step Cluster Analysis using the factor scores and PISA science literacy average scores. It is thought that the study is of great importance in terms of dividing individuals into clusters according to science instruction methods and the mean of plausible values and having an idea about how each cluster is defined. As a result of the K-means cluster analysis, it was determined that the input variable with the highest level of importance in the formation of the first and third clusters in which the students with the highest scores were included was teacher-directed science instruction, and after this variable, the input variable with the highest level of importance was the perceived feedback from science teachers. Within the scope of the Two-Step Clustering Analysis, it was determined that teacher-directed science instruction has the most importance in terms of the decomposition of clusters, followed by adaptive instruction in science lessons in terms of importance level.

Keywords: *PISA, science literacy, clustering.*

To cite this article: Eser, M. T., & Çobanoğlu Aktan, D. (2021). Educational data mining: The analysis of the factors affecting science instruction by clustering analysis. *International Journal of Educational Methodology*, *7*(3), 487-500. https://doi.org/10.12973/ijem.7.3.487

Introduction

The importance of large-scale tests in the global world is increasing everyday day in terms of comparing the education systems of countries. Large-scale tests are the applications that are used to evaluate student achievement both nationally and internationally and consequently that help to make international comparisons by obtaining important information (Dossey et al., 2006; Feuer, 2012). The Organization for Economic Cooperation and Development (OECD) conducts many studies to compare the education systems of countries, which affect the education policies of countries. Programme for International Student Assessment (PISA), Progress in International Reading Literacy Study (PIRLS), Trends in International Mathematics and Science Study (TIMSS) and The Teaching and Learning International Survey (TALIS) are examples of these applications.

PISA, an OECD application, is the world's most comprehensive survey research that contributes to the international evaluation of the knowledge and skills of 15-year-old students every three years (OECD, 2019). PISA, which is repeated every three years since 2000, focuses on one of the three literacy areas (mathematics, science and reading) in each application, collecting in-depth data in that area. In other words, in each application, the attitude of each student towards one of these areas is determined as well as his/her overall literacy in all three areas, and their opinions about themselves, their family and their school are taken. In PISA 2015, financial literacy was also added to three basic literacies in the measurement process and the focus area of the application was science literacy (OECD, 2016a).

PISA evaluates the performance of 15-year-old students on "real life" tasks, which are thought to be suitable for active participation in society and lifelong learning (OECD, 2013, 2019). The objective of PISA is not creating a competitive environment between countries but ensuring that participating countries evaluate their own education systems, and

* This article was produced from the doctoral dissertation entitled "Examination of the Program for International Student Assessment 2015 Data by Clustering Methods in Data Mining" written by Mehmet Taha ESER.

** **Corresponding author:**
Mehmet Taha Eser, Adnan Menderes University, Educational Measurement and Evaluation, Aydın, Turkey. ✉ taha.eser@hacettepe.edu.tr

the knowledge and skills of the students in mathematics, science and reading are monitored by years (Ministry of Education, 2010). In the literature, the data obtained as a result of the PISA test are considered as big data and the results obtained from this test are of great importance for many different institutions and organizations. These projects contribute to practitioners' and policy makers' efforts of improving educational practices, by offering new perspectives (Plomp, 1998; Schleicher, 2011). The data obtained from this test, which has important results for policy makers and practitioners in education, is an indicator that data mining methods can be used on PISA data sets in education. Educational data mining is about developing, researching and applying analysis methods on the computer, to detect the patterns in the collections of large-scale educational data, which is difficult and impossible to analyze due to the presence of huge volumes of data (Romero & Ventura, 2013). At the same time, educational data mining is a set of methods that help to search for correlations and rules through programming, which reveal valuable information included in very large data and which allows to draw inferences about the future (Kayri, 2008). Educational data mining is a research area that has gained importance in recent years and aims to analyze original data types to solve research problems in the field of education (Baker & Yacef, 2009). In this respect, educational data mining can also be defined as the application of data mining (DM) techniques to specific data types obtained from educational environments to address important educational questions.

Science literacy, which is included in PISA as an assessment area, is an important research and discussion area of science education literature with all its dimensions. There are many studies in the literature that contain suggestions for improving the science literacy of students of all age groups. According to Jurecki and Wander (2012), for the last 40 years the main method of giving science literacy to students in an education process that students learn through observations and experiments as scientists do, which is summarized as learning by doing. In addition, an interdisciplinary approach associated with students' interests is suggested to be more effective in providing science literacy (Ross et al., 2013). Moreover, although it is essentially similar in nature, science literacy should be defined in different contexts for different cultures, as the culture significantly affects the individual's view of nature and learning style (Seraphin, 2014). Even though students are traditionally expected to have a good level of science literacy before graduating from high school, the development of science literacy is accepted as a lifelong process in the 21st century (Liu, 2009).

Most of the scientific principles and theories of 15-year-old students are taught at school. As in other fields, the way science is taught in schools can affect not only the achievement of the students in science, but also those who want to be involved in science in advanced education and career planning. Considering the expected growth in science-related employment worldwide and the decline of students' interest towards science due to school reasons, examining why some students are more interested in science-related careers has become even more important. This has created the need to analyze in detail the resources offered for science such as science learning opportunities at school, laboratory applications, science teachers and science activities, and the ways in which science is instructed at school. In addition to past and present science education policy, the theory (s) about the nature of science teaching and learning have been and continue to be developed through international studies. At this point, the question "How can science literacy be improved among the students who are responsible for the science lesson" comes to mind. Most studies using the data obtained from the PISA application focused on the relationship between science teaching method used in a single country and science achievement for students in that country. Some of the studies using the data obtained through the PISA application are based on comparing the science education of countries with different science levels. However, such studies are limited due to the inability to control confounding variables such as the country-based school system and education expenditures. For this reason, it is thought that more research is needed on how groups with similar science achievement levels across countries are perceived when science teaching methods are considered. For this reason, studies on the factors affecting science instruction are considered important. Accordingly, various studies on the factors affecting science instruction have been conducted in the last 30 years (Langdon et al., 2011) Vedder-Weiss, & Fortus, 2011). In the light of this information, in this study the students of OECD countries participating in the 2015 PISA test were modeled with different clustering methods according to the factors affecting science instruction and the mean of science plausible values, and the importance of input variables in cluster formation were determined.

It is thought that the study is of great importance in terms of dividing individuals into clusters according to science instruction methods and the mean of plausible values and having an idea about how each cluster is defined. In this respect, it is predicted that the study will contribute to the relevant literature and future studies, and will guide all stakeholders in education systems, especially policy makers. As a result of the literature review, there is no study focusing on different clustering models established by considering science instruction methods and students' science achievement, in which the results of the models are evaluated within themselves, and the factors that are most effective in the formation of the clusters are interpreted. At the same time, this study is important in terms of using data mining-based clustering methods together with classical clustering methods and analyzing the results obtained.

In this study, it was aimed to model the students in the OECD sample participating in the 2015 PISA test, according to the sub-dimensions of science instruction and the mean of science plausible values using different clustering methods, and to determine the importance of input variables in the formation of clusters regarding the sub-dimensions of science instruction and science achievement. The obtained cluster properties also provide information about the cluster profiles. The focus of the 2015 PISA survey was science literacy, which is thought to increase the importance of the research. The following research questions are addressed within the scope of the study:

How are the students who participated in the PISA 2015 student survey and who are in the OECD sample clustered according to the items related to science instruction and the mean of plausible values?

Methodology

Research Problem

Sub-problems of the research

Considering the purpose and importance of the study the following sub-problems are addressed within the scope of the study:

1. How are the students participating in the PISA 2015 student survey and included in the OECD sample clustered by k-means cluster analysis according to the items related to science instruction and the mean of plausible values?
2. How are the students participating in the PISA 2015 student survey and included in the OECD sample clustered by two-step cluster analysis according to the items related to science instruction and the mean of plausible values?

Research Group

In the study, 2015 PISA data obtained from students of OECD member countries except Slovenia were used. While performing cluster analysis with large data sets in the R program, calculation time may be delayed, and systemic problems may be encountered. In order to eliminate this delay, systematic sampling was applied to the data set and as a result, analyzes were carried out with the data of 9,870 student. Distribution of the number of students by country is shown in Figure 1.

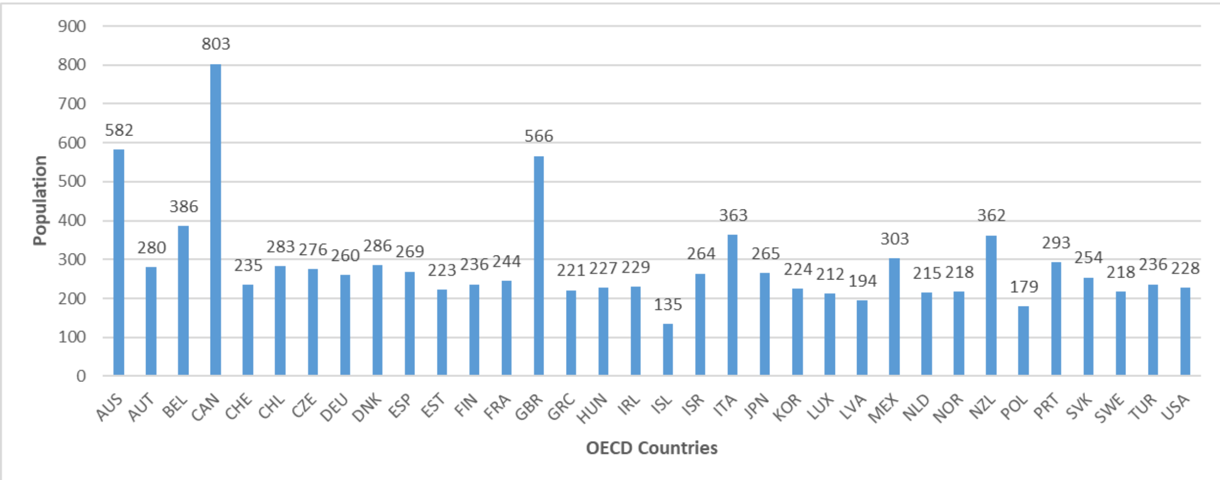

Figure 1. Distribution of the number of students by country

Data obtained from the students of 34 countries were included in the cluster analysis. Regarding Figure 1, which shows the distribution of the number of students by country, the most participants are from Canada and the least is from Iceland. The number of participants shows a noticeable decrease after Canada, followed by Australia and the UK. Poland and Lithuania follow Iceland as the countries with the least number of participants.

Variables

Inquiry-based science instruction was developed within the scope of PISA 2015 using exploratory factor analysis and confirmatory factor analysis (Lau & Lam, 2017). Within the scope of PISA 2015, there are a total of 4 science instruction structures: Inquiry-based science instruction, teacher-directed science instruction, perceived feedback from science teachers and adaptive instruction in science lessons. In the study, sub-dimensions of science instruction in the PISA 2015 student questionnaire and the mean of science plausible values were used as data input. The input variables used in the study are coded as: Factor 1 (Teacher-directed science instruction), Factor 2 (Perceived feedback from science teachers), Factor 3 (Adaptive instruction in science lessons), Factor 4 (Inquiry-based science instruction), Factor 5/PVSCIENCE (Mean of science plausible values) ”.

Analyzing of Data

In the research, the data has been subjected to the "Data Pre-Processing" process to get it ready for analysis before starting the analysis. In this process, missing data analysis was performed first. The missing data in the data set of PISA surveys have a random structure, thus multiple value assignment method, which is one of the missing data assignment methods, was applied to the data set used in the study (Adams et al., 2013; Kaplan & Su, 2016). The multiple value assignment was carried out based on logistic regression regarding the structure of the data set (The link related to data: https://figshare.com/articles/dataset/K-Means_Cluster_Analysis_Syntax_R/14979879). As a result of the missing data assignment, the data related to Slovenia, which has a large amount of missing data in the variables used in the research, was removed from the data set.

After missing data assignment, systematic sampling was applied to the data set. Considering that the data obtained from 253,140 students will delay computer's calculation time in the cluster analysis of the R program, a data set consisted of the answers of 10,000 individuals was thought to be sufficient and the proportionality constant k was set as k = 25 (253.140 / 10.128). A macro was created in the Excel program, and 1 out of every 25 students in the universe was included in the sample. Considering the systematic sampling and the data related to Slovenia were removed from the data set, the data set consisted of 9870 students. Systematic sampling is a non-random sampling method that includes people selected from the universe at regular intervals (Monette et al., 1990).

The total number of items in the sub-dimensions of science education, included in the study was 21; factor scores were used in the study, considering that interpretation will be difficult if the scores of all these 21 items are used as input. Factor scores are weighted combinations of associated variables, making this score type more reliable and better quality versus actual values, which is the main reason of preferring to use factor scores in the research (Fiedler & Mcdonald, 1993). There are different methods for estimating factor scores. Factor scores of input variables used in the study were obtained by regression analysis based on the least squares method. Regression at the point of estimating factor scores is a method that is very easy to apply and has higher statistical validity compared to factor scores estimated by other methods (Grice, 2001; Mulaik, 2009).

K-means and two-step cluster analysis were used in the study. Cluster analysis is an objective method that attempts to measure the structural characteristics of a set of observations. Regarding cluster analysis, although there is a normal distribution assumption for the data, the normality assumption usually remains in principle and is not considered much. In addition, there is no assumption about the covariance matrix in the cluster analysis, and the assumptions required in other multivariate analyzes, such as linearity and having covariance, are not requested. In short, cluster analysis is a type of analysis robust to the violations of assumptions (Garson, 2014; Hair et al., 2009; Tatlıdil, 1992). When it comes to violation of assumptions, although cluster analysis is a robust type of analysis, normality and linearity assumptions of the data set were checked in order to have more detailed information about the data used in the study and to prevent any data pre-processing errors.

The outliers included in the data sets are defined as noisy / dirty data within the scope of data mining. Within the scope of the research, it was aimed to reduce the input variables to the same scale level in order to prevent the outliers from adversely affecting the statistical analysis results. For this purpose, the scores related to the means of science plausible values were converted into standardized '' z '' values and included in clustering analysis.

As mentioned in the data analysis section, k-means and two-step cluster analysis are used in this study; and the most effective variables in the formation of these clusters are examined within the scope of the study. In addition, the clusters are used to generate outputs that can be used on determining the similarities and differences between countries. The findings obtained using the relevant cluster analysis are summarized below.

Findings about the First Sub-Problem

In the first sub-problem of the study, the outputs obtained by k-means method are discussed.

Since K-means algorithm does not certainly find the ideal number of clusters, the analysis was performed multiple times using different parameters and random samples. In the R program, the "nstart" parameter found under k-means algorithm allows to select random start for the trial, it also uses “iter.max” parameter to set the maximum number of iterations allowed by the algorithm for each random start. For this reason, the change of within-groups sum of squares should be checked for different number of clusters. Figure 2 shows the change of within-group sum of squares according to the number of clusters.

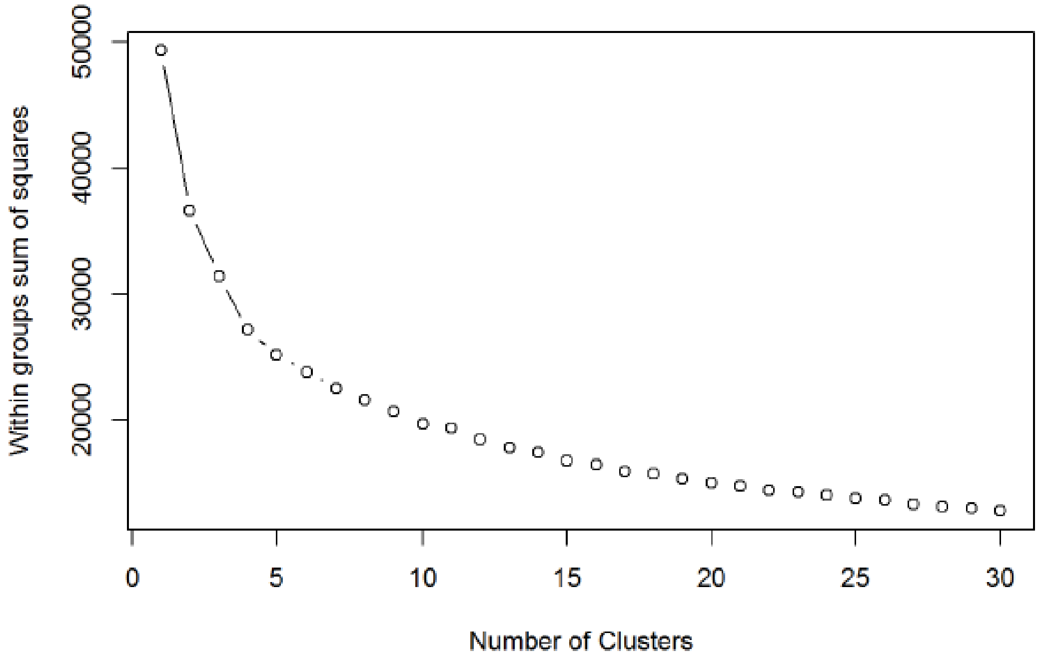

Figure 2. Change of within-group sum of squares according to the number of clusters

Regarding Figure 2, the sum of squares decreases as the number of clusters increases. Within-group sum of squares is also accepted as the error value and the distance between two points of the graph shows the number of clusters formed. Since the error values decrease as the number of clusters increases, the point where the error values do not change much can be taken as a reference in determining the ideal number of clusters. Accordingly, there is a rapid decrease in error values until the number of clusters reaches four, but then for five or more clusters, the change in error values is not too much. According to this result, the ideal number of clusters seems to be four. In order to verify this result, between clusters sum of squares should be checked as well. Figure 3 shows the change of between clusters error values according to the number of clusters.

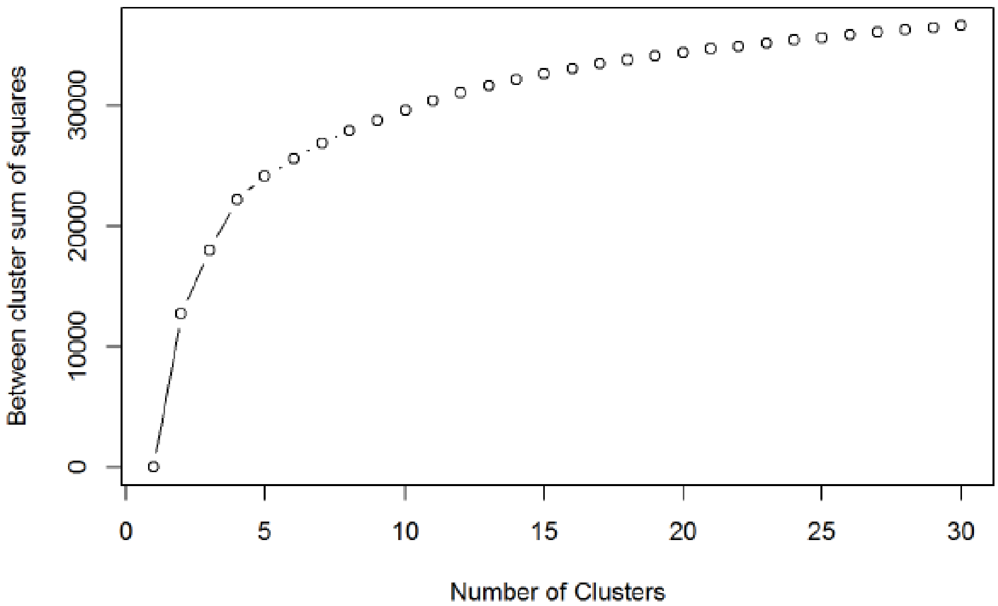

Figure 3. Change of between clusters error values according to the number of clusters

Regarding Figure 3, showing the change of between clusters error values obtained for different number of clusters, between clusters error values are observed to increase as the number of clusters increases. To determine the ideal number of clusters, the point where the difference between error values is not too much can be taken as a reference, as in the previous graphs. In this graph, which is accepted as a measure of how well and accurately a cluster differs from the others, the clusters stop to differentiate well when the number of clusters is five or more. According to this result, it is thought that the ideal number of clusters will be four. In addition to these, in order to determine the ideal number of clusters, it is necessary to examine how the AIC and BIC of the formed clusters vary for different number of clusters. The line graph obtained for both criteria is shown in Figure 4.

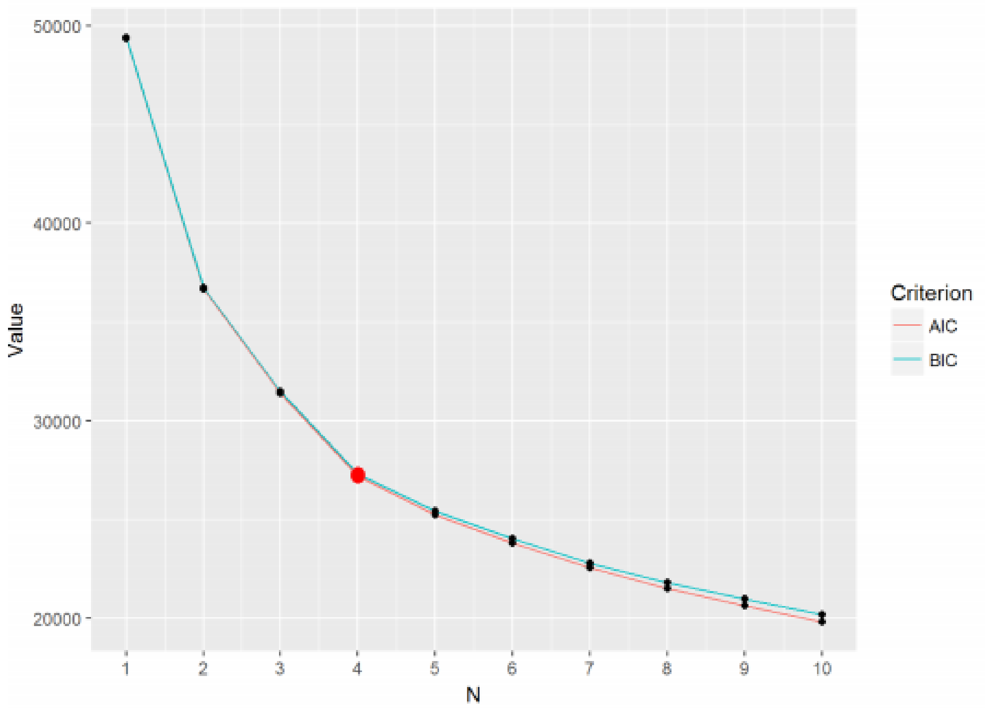

Figure 4. Change of criterion values according to the number of clusters

Regarding Figure 4, which shows the change of AIC and BIC criteria according to the number of clusters, both criteria are observed to decrease as the number of clusters increases. Information that will help in determining the ideal number of clusters is related to determining the point where the line graph reaches a flat plateau shape. Accordingly, it can be said that the amount of decrease in the graph is relatively less when the number of clusters is over four. Another method used to determine the ideal number of clusters is the gap statistics regarding the gap between clusters. The graphic regarding the gap between clusters is shown in Figure 5

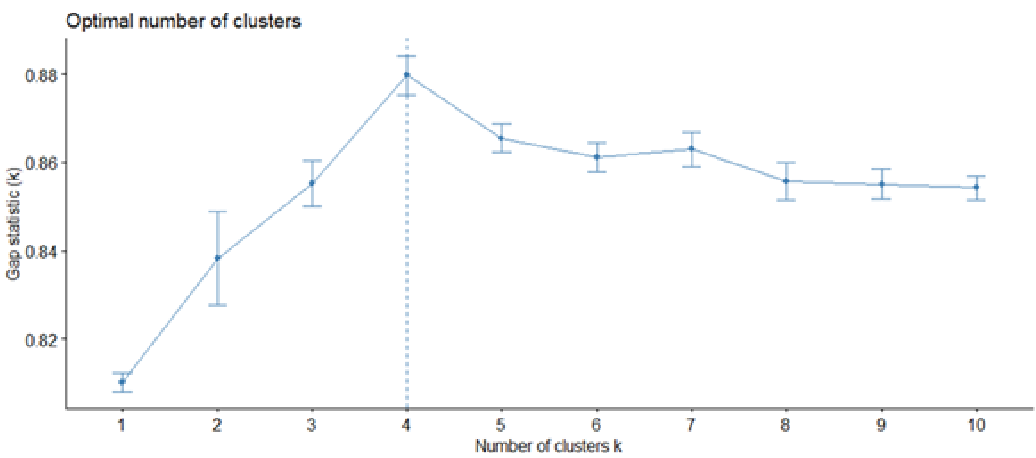

Figure 5. Graph of the gaps between clusters

Regarding Figure 5, showing the lower and upper values of the change of within-cluster coefficients for different number of clusters, intragroup variability is observed to gradually decreases after the number of clusters reaches k=4. It is known that the point where the Gap statistic is maximum will show the information about the ideal number of clusters. Accordingly, the ideal number of clusters is thought to be four. Although determining the ideal number is considered intuitive, the color scatter plot of the given clusters is seen as the most effective method in determining the ideal number of clusters. Figure 6 shows the scatter plot of the clusters obtained in the study.

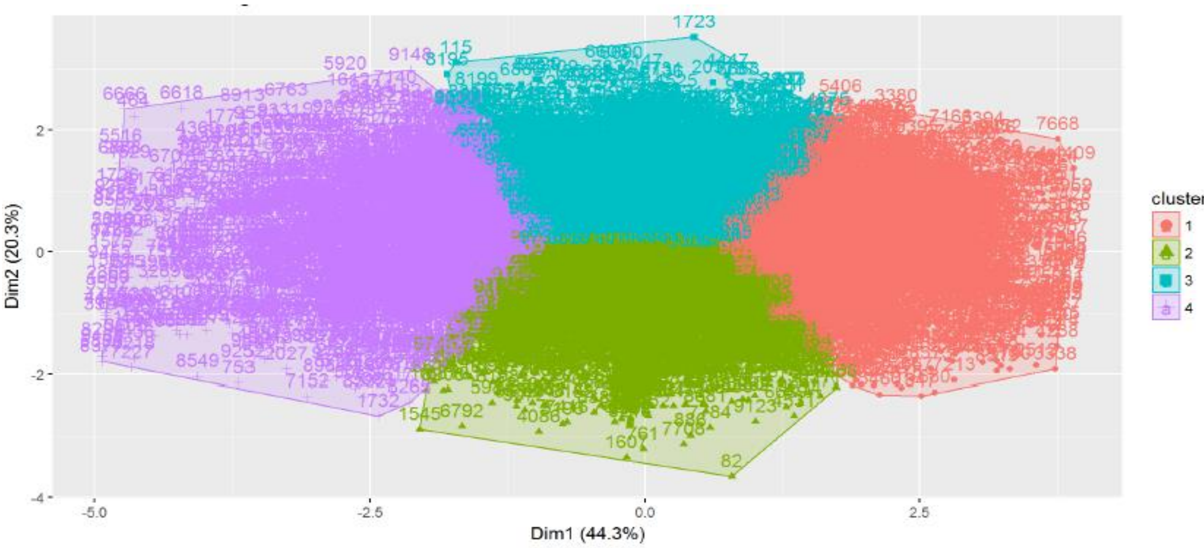

Figure 6. Scatter plot of the clusters

In Figure 6, the variables on the x and y axes show the clusters obtained for the first two variables Dimension1 (Dim1) and Dimension2 (Dim2), which are most effective in clustering analysis, in other words, which explain the maximum amount of variance. In the light of this information, it is seen that the first input variable explains 44.30% of the total variance, and the second input variable explains 20.30% of the total variance in clustering students into groups. The total amount of variance explained by the first two variables used in clustering is 64.60% and this is considered to be sufficient. Accordingly, a total of 4 clusters were formed, as shown in purple, green, blue and orange on the graph. When all the methods used to determine the ideal number of clusters in the scope of the study were evaluated in a holistic manner, it was concluded that the students were divided into 4 clusters. In addition to this, the column chart that gives information about how many students from each country in the sample among 4 clusters determined as the ideal number of clusters is shown in Figure 7.

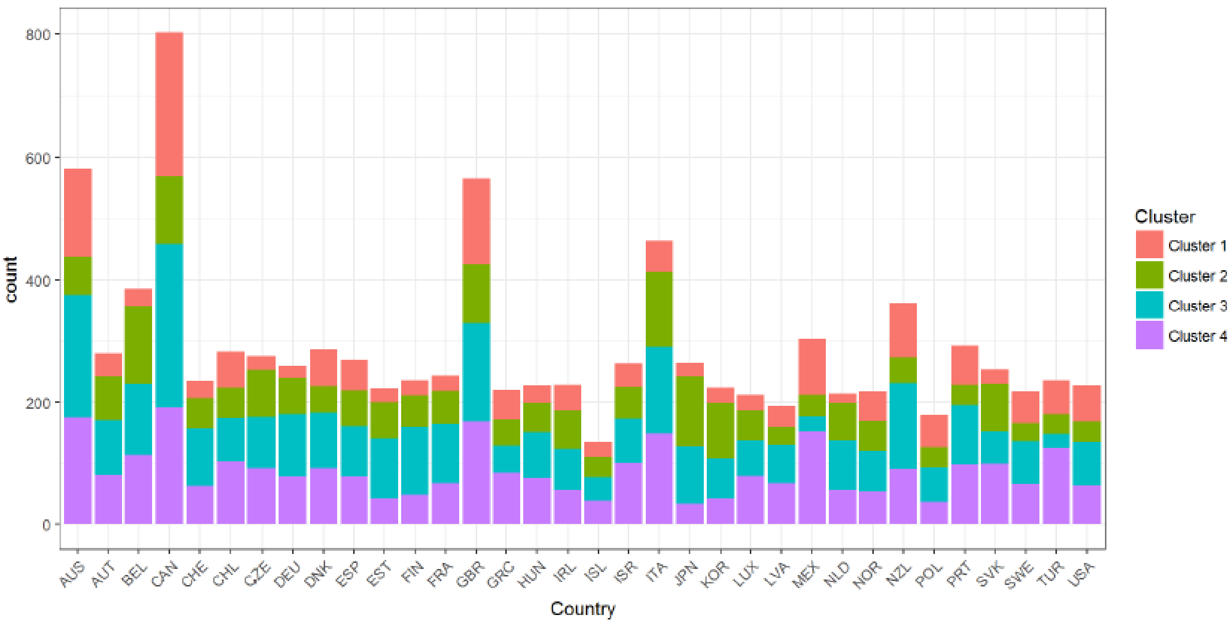

Figure 7. Student distribution of each country among clusters

Figure 7 gives information about how many students from each country are included in the four clusters determined by k-means cluster analysis. When Figure 7 is examined, it is seen that Canada, which is seen to have the highest number of students, gives the highest number of students to the third cluster and the least student to the second cluster. When Figure 7 is examined again, it is seen that, like Canada, Australia gave the highest number of students to the third cluster and the least students to the second cluster. It is observed that Israel, which has the least number of students within the scope of the research group, gives the most students to the third cluster and the least students to the first cluster. Cluster profiles were obtained in order to determine the effect of factor scores at the point of decomposition into clusters on the formation of each cluster. The effect of factor scores for four different clusters is shown in Figure 8.

Figure 8. Distribution of profile scores in clusters

Regarding Figure 8, teacher-directed science instruction (Factor 1) is the most effective factor in the first cluster. In addition, students with the highest science plausible values (factor 5) are in the first cluster. Teacher-directed science instruction (Factor 1) is the most effective factor in the second cluster as well. However, perceived feedback from science teachers (Factor 2), adaptive instruction in science lessons (Factor 3) and Inquiry-based science instruction (Factor 4) variables are observed to be effective in this cluster. Regarding the third cluster, teacher-directed science instruction (Factor 1) had the most effect. Compared to the second cluster, the students in the third cluster have higher science plausible values. In the fourth cluster, teacher-directed science instruction (Factor 1) had the most effect. However, the students with the lowest science plausible values are in this cluster.

The holistic evaluation of the obtained results shows that students in the first and third cluster generally adopt teacher-directed science instruction, and the first and third clusters consist of the students with higher science plausible values than the students in the second and fourth clusters. At the same time, it was determined that the perceived feedback from science teachers (factor2) for students in the first cluster and the third cluster, which consisted of individuals with the highest scores, respectively, had an important place in the formation of these two clusters after teacher-directed science instruction.

Findings about the Second Sub-Problem

In the second sub-problem of the study, the data obtained from 9,870 students are clustered by using two-step cluster analysis and the number of clusters is determined. First of all, the change of silhouette coefficients obtained for different number of clusters is examined.

Regarding the analysis results, in two-step cluster analysis silhouette coefficients are observed to be weak and different algorithms should be tested for 2, 3, 4, 5, 6 and 7 clusters. The highest silhouette value is obtained when the number of clusters is 2; they don't change considerably when the number of clusters is 3 and 4; but they decreases (<0.25) when the number of clusters is 5, 6 and 7. Accordingly, it can be said that the ideal number of clusters may be 2, 3 or 4. To support this result statistically, the results of the automatic clustering process obtained as the output of the analysis are shown in Table 1.

Table 1. Results of the Automatic Clustering Process

Number of Clusters	Bayesian Information Criteria (BIC)	Change of BIC	Percentage of Reduction in BIC	Ratio of Measured Distance
1	34296.286			
2	**28868.831**	**-5427.455**	**1**	**2.09**
3	26319.408	-2549.422	0.47	1.202
4	24213.022	-2106.387	0.388	1.595
5	22927.142	-1285.879	0.237	1.429
6	22054.826	-872.317	0.161	1.255
7	21378.379	-676.446	0.125	1.127
8	20788.732	-589.648	0.109	1.074
9	20245.913	-542.819	0.1	1.037
10	19725.474	-520.439	0.096	1.23
11	19319.553	-405.921	0.075	1.126

Table 1. Continued

Number of Clusters	Bayesian Information Criteria (BIC)	Change of BIC	Percentage of Reduction in BIC	Ratio of Measured Distance
12	18969.2	-350.353	0.065	1.119
13	18665.954	-303.246	0.056	1.031
14	18374.478	-291.476	0.054	1.044
15	18099.188	-275.289	0.051	1.046

BIC is one of the criteria in Table 1 that can be used to determine the ideal number of clusters, which can be set as the point where it has the smallest value. However, this value is sensitive to the number of clusters, therefore the percentage of reduction in BIC is a more stable criterion to be used. The point where this ratio is greater than the others shows the ideal number of clusters. Accordingly, the ideal number of clusters is thought to be two. Evaluating the obtained results as a whole and considering that silhouette value also reaches the highest value when the number of clusters is two, it is concluded that the units in the data set can be divided into two clusters. Two-step clustering analysis also provides an output about the importance of the variables in the formation of clusters. The results regarding the predictor importance of input variables in the formation of clusters in the Two-Step Clustering Analysis are shown in Figure 9.

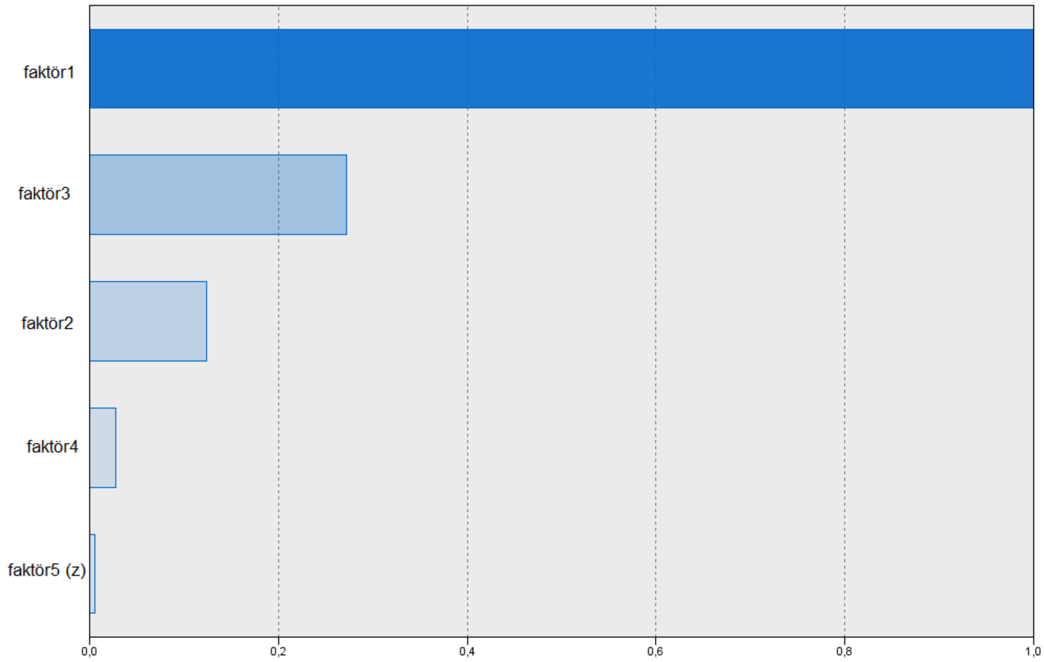

Figure 9. Predictor importance of the variables

Regarding Figure 9, teacher-directed science instruction (Factor 1) is determined to be the most effective variable in the formation of clusters. It is followed by adaptive instruction in science lessons (Factor 3), perceived feedback from science teachers (Factor 2), inquiry-based science instruction (Factor 4), and mean of science plausible values (Factor 5).

Discussion

In this study, the clustering results of the students from 34 OECD countries participating in the PISA 2015 test and sampled by systematic sampling method are obtained by K-Means Clustering and Two-Step Cluster Analysis using the factor scores and PISA science literacy average scores. Five input variables, namely teacher-directed science instruction sub-dimension, adaptive instruction in science lessons sub-dimension, perceived feedback from science teachers sub-dimension, inquiry-based science instruction sub-dimension, and mean of science plausible values are used and the clustering of the students is determined using different clustering methods.

The first sub-problem of the study addresses the number of clusters obtained by k-means clustering and cluster results. The ideal number of clusters is determined to be four, based on the computations of the changes in within-group sum of squares, between clusters error values, Gap statistics, AIC and BIC performed by K-Means Clustering in the R program. Students are determined to mostly fall in the third cluster and then in the fourth cluster. As a result of the K-means cluster analysis, it was determined that the input variable with the highest level of importance in the formation of the first and third clusters in which the students with the highest scores were included was teacher-directed science instruction, and after this variable, the input variable with the highest level of importance was the perceived feedback from science teachers. In k-means cluster analysis, the most effective variables in the formation of the first cluster are found to be teacher-directed science instruction and science achievement. This result is similar to the finding of the study conducted by Costa and Araújo (2018) that using teacher-directed science instruction method is associated with higher achievement in science.

The second sub-problem of the study addresses the number of clusters obtained by two-step cluster analysis and cluster results. In Two-Step Clustering Analysis, where the analysis results obtained for the first sub-problem are limited, the ideal number of clusters was determined regarding silhouette values, BIC and BIC change criteria. The results regarding silhouette value, BIC and BIC change criteria were examined and it was determined that the ideal number of clusters was similar and equal to two according to all three statistics. In the literature, there are studies in which the ideal number of clusters obtained by Two-Step Cluster Analysis and by K-Means Clustering are not equal (Ceylan, 2013; Shih et al., 2010). Within the scope of the Two-Step Clustering Analysis, it was determined that teacher-directed science instruction has the most importance in terms of the decomposition of clusters, followed by adaptive instruction in science lessons in terms of importance level.

The cluster profiles obtained in the study do not show a direct cause-effect relationship between individuals and variables. The cluster profiles obtained as a result of the cluster analysis performed within the scope of the study do not provide information about a direct cause-effect relationship between the individuals that make up the study group and the variables used in the study. Using the outputs obtained within the scope of the analysis, it gives comparable information (s) about the importance level of the variables within themselves at the point of dividing the students that make up the study group into groups. Therefore, there may be differences in the clusters in which the individuals are assigned according to the clustering method. This result is thought to be due to the high heterogeneity of the data set.

The review of the results of both clustering methods used in the research in a holistic manner shows that the students belonging to the cluster with the highest science plausible values adopted teacher-directed science instruction instead of inquiry-based science instruction. This may be due to the fact that inquiry-based science instruction is a more difficult teaching strategy, and its effective implementation requires certain resources and a specific school environment (Mostafa et al., 2018). Inquiry-based science instruction, in which students are guided at the lowest level due to its nature, requires students to have a high level of prior knowledge and self-discipline (Kirschner et al., 2006). Students who fails to meet these requirements may perform poorly when exposed to inquiry-based teaching (OECD, 2016a, 2016b). This suggests that using inquiry-based learning to 'engage' students in a classroom where noise, disorder, and wasted time are the norms is unlikely to increase the science performance of these students (Hattie & Yates, 2014; OECD, 2019). Another possible explanation is the excessive use of inquiry activities, in which minimal guidance is offered to students, by teachers. Lazonder and Harmsen (2016) conducted a meta-analysis study on inquiry-based methods and concluded that inquiry-based science teaching is the most effective method when appropriate guidance is provided by the teacher. A similar result was reached in another meta-analysis study conducted by Kirschner et al. (2006). In the meta-analysis study conducted by Furtak et al. (2012), the generalized effect size of student-directed course activities was determined to be .50; whereas the generalized effect size of teacher-directed course activities was .90.

The effectiveness of inquiry-based science instruction depends on the readiness of many factors. The most important of these factors can be listed as: the planning time of the teacher and intensive resources need for school-related materials, the readiness of the classroom environment for productive learning behavior, high level of prior knowledge, the correct adjustment of the amount of guidance that the teacher will provide, and the autonomy support that teacher provides to the student. The autonomy support provided by the teacher to the students is related to all other factors (Jang et al., 2010). There are studies suggesting that the autonomy support provided by the teacher to the students may not involve students in learning activities alone, and that this support should be turned into a structure (Hospel & Galand, 2016; Mouratidis et al., 2018; Sierens et al., 2009). Although many inquiry-based activities look excellent, they are not as effective as teacher-directed teaching activities alone. As a result of the study carried out by Flick (1995), when it is aimed to gain scientific knowledge and certain scientific process skills, teacher-directed science instruction is aimed to measure higher-level thinking skills, and inquiry-based science instruction is the basis for teachers with a high level of education provided they work in a supportive classroom environment that it should be taken.

Contrary to popular belief, the finding that successful students adopted teacher-directed science instruction more than inquiry-based science instruction is quite interesting in itself. Considering the undisputed opinions accepted in the context of education rhetoric today, this finding of the research is perhaps unexpected. This finding may lead to the conclusion that Piaget's notions of constructing one's knowledge is combined with the idea that inquiry-based learning methods are more desirable.

A teacher-directed formation which is structured in a more definite framework of what students should understand, in which the subjects and learning objectives are more clearly defined, that is containing much less uncertainty, can be considered to be more preferable for students (Areepattamannil et al., 2020).

When the relevant literature is examined, it is seen that there are experimental studies that suggest using teacher-directed science instruction instead of inquiry-based science instruction (Scheier et al., 2001) and also suggest using inquiry-based science instruction instead of teacher-directed science instruction (Klahr & Nigam, 2004). At the same time, it is striking that there are studies in the literature that indicate that there is no statistically significant difference regarding both methodologies (Cobern et al., 2010). At the point of science acquisition, there is a lack of evidence that teacher-directed science instruction is superior to inquiry-based science instruction (Areepattamannil et al., 2020). It is thought that the number of evidence can be increased with future studies that will yield similar results to this study.

When the results obtained within the scope of both clustering methods are examined in a holistic manner, perceived feedback from science teachers is the variable with the highest importance after teacher-directed science instruction for the k-means method, and adaptive instruction in science lessons is the variable with the highest importance after teacher-directed science instruction for two-step cluster analysis.

When the literature is examined, it is striking that only a few previous studies on PISA application dealt with the variables used within the scope of this research on science education. When the results of the related studies are examined, it is striking that teacher-directed science instruction and adaptive instruction in science lessons are generally associated with higher science achievement (Costa & Araújo, 2018; Forbes et al., 2020; Lau & Lam, 2017). These results are thought to be consistent with the results of the study. Contrary to inquiry-based science instruction, the concept of "teacher-directed science instruction" is not fully conceptualized in the literature. However, there is a consensus that teacher-directed science instruction should include teaching features such as regular feedback and adaptive teaching in achieving high levels of success in science (Forbes et al., 2020). It should be noted that 'experience-based' teaching and 'active student participation' are certainly advantageous for an effective science teaching process, but such 'hands-on' and 'catchy' aspects can arise when using inquiry-based or teacher-directed science instruction approaches (Costa & Araújo , 2018; Forbes et al., 2020).

A study conducted by McKinsey & Company reveals that a model that will be created by blending inquiry-based and teacher-directed science instruction by evaluating the PISA 2015 results for OECD countries can be used and needed at least within the education systems of European countries (Denoël et al., 2017). The approaches blending inquiry-based and teacher-directed science instruction approaches can be created and used. Findings of meta-analysis studies on the effectiveness of inquiry-based science instruction may help to find out which blended approach types are most promising in increasing students' learning, competencies and disposition. At this point, the most promising approach may be providing the student with a content knowledge and a sufficiently strong basis through teacher-directed science instruction and using inquiry-based science instruction methods on this basis after creating sufficient content knowledge.

Conclusion

In k-means cluster analysis, the most effective variables in the formation of the first cluster are found to be teacher-directed science instruction and science achievement. Within the scope of the Two-Step Clustering Analysis, it was determined that teacher-directed science instruction has the most importance in terms of the decomposition of clusters, followed by adaptive instruction in science lessons in terms of importance level. When the results obtained within the scope of both clustering methods are examined in a holistic manner, perceived feedback from science teachers is the variable with the highest importance after teacher-directed science instruction for the k-means method, and adaptive instruction in science lessons is the variable with the highest importance after teacher-directed science instruction for two-step cluster analysis.

Recommendations

Science literacy has always been and will be one of the main goals when considering international science studies (Waddington et al., 2007). At this point, the biggest problem in terms of achieving the basic goals is related to the selection and use of science-related teaching methods in achieving classroom success in a holistic sense. PISA is a large-scale and internationally representative application and therefore the data obtained as a result of this application contribute greatly to solve this problem. This study and similar studies contribute to theoretical perspectives on the nature of science teaching (Settlage, 2007). Such studies on science teaching methods can be helpful in terms of guidance for science teachers, experts developing science curriculum, and academic staff training teachers. As a result, it is thought that these and similar studies will help to solve problems related to science education and contribute to science literacy globally.

Limitations

In terms of variables, this research is limited to items related to science teaching strategies selected from the student questionnaire of the students in the PISA 2015 sample and items related to the possible average science achievement score. This research was also limited to 9870 students from OECD countries, which were obtained as a result of PISA results conducted in 2015, two different clustering methods and systematic sampling.

Authorship Contribution Statement

Eser: Concept and design, data acquisition, data analysis / interpretation, drafting manuscript, critical revision of manuscript, statistical analysis, securing funding. Çobanoğlu: Technical and material support, supervision, final approval

References

Adams, R. J., Lietz, P., & Berezner, A. (2013). On the use of rotated context questionnaires in conjunction with multilevel item response models. *Large-scale Assessments in Education, 1*(5), 1-27. https://doi.org/10.1186/2196-0739-1-5

Areepattamannil, S., Cairns, D., & Dickson, M. (2020). Teacher-directed versus ınquiry-based science ınstruction: ınvestigating links to adolescent students' science dispositions across 66 countries, *Journal of Science Teacher Education, 31*(6), 675-704. https://doi.org/10.1080/1046560X.2020.1753309

Baker, R. S., & Yacef, K. (2009). The state of educational data mining in 2009: A review and future visions. *Journal of Educational Data Mining*, *1*(1), 3-17. https://doi.org/10.5281/zenodo.3554657

Ceylan, H. H. (2013). Perakende sektöründe konjoint ve kümeleme analizi ile fayda temelli pazar bölümlendirme [Benefit-based market segmentation with conjoint and cluster analysis in the retail industry]. *Journal of Management and Economics/ Yönetim ve Ekonomi, 20*(1), 141-154.

Cobern, W. W., Schuster, D. G., Adams, B., Applegate, B., Skjold, B., Undreiu, A., Loving, C. C., & Gobert, J. D. (2010). Experimental comparisons of inquiry and direct instruction in science. *Research in Science & Technological Education, 28*(1), 81–96. https://doi.org/10.1080/02635140903513599

Costa, P., & Araújo, L. (2018). *Quality of teaching and learning in science (EUR 28865 EN)*. Publications office of the European Union. https://publications.jrc.ec.europa.eu/repository/handle/JRC109064

Denoël, E., Dorn, E. , Goodman, A. , Hiltunen, J. , Krawitz, M.& Mourshed, M. (2017). *Drivers of student performance: Insights from Europe*. McKinsey & Company Report. https://cutt.ly/cQkYPuU

Dossey, J. A., McCoren, S. S., & O'Sullivan, C. (2006). *Problem solving in the PISA and TIMSS 2003 assessments, U.S. Department of Education*. National Center for Education Statistics.

Feuer, M. J. (2012). *No country left behind: Rhetoric and reality of international large-scale assessment* (13rd ed.). ETS.

Fiedler, J. A., & Mcdonald, J. J. (1993). *Market figmentation: Clustering on factor scores versus individual variables*. AMA Advanced Research Techniques Forum. http://www.populus.com/files/Market%20Figmentation_f_1.pdf

Flick, L. B. (1995). *Complex instruction in complex classrooms: A synthesis of research on inquiry teaching methods and explicit teaching strategies* [Paper Presentation]. Annual Meeting of the National Association for Research in Science Teaching, San Francisco, CA, USA.

Forbes, C. T., Neumann, K., & Schiepe-Tiska, A. (2020). Patterns of inquiry-based science instruction and student science achievement in PISA 2015. *International Journal of Science Education, 42*(5), 783-806. https://doi.org/10.1080/09500693.2020.1730017

Furtak, E. M., Seidel, T., Iverson, H., & Briggs, D. C. (2012). Experimental and quasi-experimental studies of inquiry-based science teaching: A meta-analysis. *Review of Educational Research, 82*(3), 300–329. https://doi.org/10.3102/0034654312457206

Garson, D. G. (2014). *Cluster analysis (Statistical Associates Blue Book Series).* Statistical Associates Publishing.

Grice, J. W. (2001). Computing and evaluating factor scores. *Psychological Methods*, *6*(4), 430-450. https://doi.org/10.1037/1082-989X.6.4.430

Hair, J. F. Jr., Black, W. C., Babin, B. J., & Anderson, R. E. (2009). *Multivariate data analysis*. Prentice Hall.

Hattie, J. A. C., & Yates, G. C. R. (2014). Using feedback to promote learning. In V. A. Benassi, C. E. Overson, & C. M. Hakala (Eds.), *Applying science of learning in education: Infusing psychological science into the curriculum* (pp. 45–58). Society for the Teaching of Psychology.

Hospel, V., & Galand, B. (2016). Are both classroom autonomy support and structure equally important for students' engagement? A multilevel analysis. *Learning and Instruction, 41*(2), 1–10. https://doi.org/10.1016/j.learninstruc.2015.09.001

Jang, H., Reeve, J., & Deci, E. L. (2010). Engaging students in learning activities: It is not autonomy support or structure but autonomy support and structure. *Journal of Educational Psychology, 102*(3), 588–600. https://doi.org/10.1037/a0019682

Jurecki, K., & Wander, M. C. (2012). Science literacy, critical thinking, and scientific literature: Guidelines for evaluating scientific literature in the classroom. *Journal of Geoscience Education, 60*(2), 100-105. https://doi.org/10.5408/11-221.1

Kaplan, D., & Su, D. (2016). On matrix sampling and imputation of context questionnaires with implications for the generation of plausible values in large-scale assessments. *Journal of Educational and Behavioral Statistics. 41*(1), 57-80. https://doi.org/10.3102/1076998615622221

Kayri, Y. (2008). Elektronik portfolyo değerlendirmeleri için veri madenciliği yaklaşımı [Data mining approach for electronic portfolio reviews]. *Yüzüncü Yıl University Faculty of Education Journal/ Yüzüncü Yıl Üniversitesi Eğitim Fakültesi Dergisi, 5*(1), 98-110.

Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist 41*(2), 75-86. https://doi.org/10.1207/s15326985ep4102_1

Klahr, D., & Nigam, M. (2004). The equivalence of learning paths in early science instruction: Effects of direct instruction and discovery learning. *Psychological Science, 15*(10), 661-667. https://doi.org/10.1111/j.0956-7976.2004.00737.x

Langdon, D., McKittrick, G., Beede, D., Khan, B., & Dom, M. (2011). STEM: Good jobs now and for the future. *U.S. Department of Commerce Economics and Statistics Administration, 3*(11), 1-10.

Lau, K.-C., & Lam, T. Y.-P. (2017). Instructional practices and science performance of 10 top-performing regions in PISA 2015. *International Journal of Science Education*, *39*(15), 2128–2149. https://doi.org/10.1080/09500693.2017.1387947

Lazonder, A. W., & Harmsen, R. (2016). Meta-analysis of inquiry-based learning: Effects of guidance. *Review of Educational Research, 86(*3), 681-718. https://doi.org/10.3102/0034654315627366

Liu, X. (2009). Beyond science literacy: Science and the public. *International Journal of Environmental and Science Education, 4*(3), 301-311.

Ministry of Education. (2010). Milli Eğitim Bakanlığı uluslararası öğrenci değerlendirme programı ulusal nihai rapor [Ministry of Education programme for international student assessment national final report]. https://shorturl.at/drMO6

Monette, D. R., Sullivan, T. J., & Jong, C. R. (1990). *Applied social research*. Harcourt Broce Jovanovich, Inc.

Mostafa, T., Echazarra, A., & Guillou, H. (2018). *The science of teaching science: An exploration of science teaching practices in PISA 2015* (OECD Education Working Papers, No. 188). OECD Publishing. https://doi.org/10.1787/f5bd9e57-en

Mouratidis, A., Michou, A., Aelterman, N., Haerens, L., & Vansteenkiste, M. (2018). Begin-of-school-year perceived autonomy-support and structure as predictors of end-of-school-year study efforts and procrastination: The mediating role of autonomous and controlled motivation. *Educational Psychology, 38*(4), 435–450. https://doi.org/10.1080/01443410.2017.1402863

Mulaik, A. (2009). *Foundations of factor analysis*. Chapman & Hall/CRC Statistics in the Social and Behavioral Sciences Series.

Organisation for Economic Co-operation and Development. (2013). *PISA 2012 assessment and analytical framework: Mathematics, reading, science, problem solving and financial literacy*. OECD Publishing. https://cutt.ly/HQkYJyP

Organisation for Economic Co-operation and Development. (2016a). *Education at a glance 2016: OECD indicators.* OECD Publishing. https://cutt.ly/uQkYB9S

Organisation for Economic Co-operation and Development. (2016b). *Survey of Adult Skills Technical Report* (2nd ed.). OECD Publishing. https://cutt.ly/pQkY7uP

Organisation for Economic Co-operation and Development. (2019). *PISA 2018 Assessment and Analytical Framework*. OECD Publishing. https://cutt.ly/zQkUtBW

Plomp, T. (1998). The potential of international comparative studies to monitor the quality of education*. Prospects, 28,* 45–59*.* https://doi.org/10.1007/BF02737779

Romero, C., & Ventura, S. (2013). Data mining in education. Wiley Interdisciplinary Reviews: *Data Mining and Knowledge Discovery, 3*(1), 12-27. https://doi.org/10.1002/widm.1075

Ross, K., Hooten, M. A., & Cohen, G. (2013). Promoting science literacy through an interdisciplinary approach. *Bioscene*: *Journal of College Biology Teaching, 39*(1), 21-26.

Scheier, M. F., Carver, C. S., & Bridges, M.W. (2001). Optimism, pessimism, and psychological well-being. In E. C. Chang (Ed.), *Optimism and pessimism: Implications for theory, research, and practice* (pp. 189-216). American Psychological Association.

Schleicher, A. (2011). *Building a high-quality teaching profession: Lessons from around the world.* OECD Publishing.

Seraphin, K. D. (2014). Where are you from? Writing toward science literacy by connecting culture, person, and place. *Journal of Geoscience Education, 62*(1), 11-18. https://doi.org/10.5408/12-413.1

Settlage, J. (2007). Demythologizing science teacher education: Conquering the false ideal of open inquiry. *Journal of Science Teacher Education*, *18*(4), 461–467. https://doi.org/10.1007/s10972-007-9060-9

Shih, M. Y., Jheng, J. W., & Lai, L. F. (2010). A two-step method for clustering mixed categroical and numeric data. *Tamkang Journal of Science and Engineering, 13*(1), 11-19.

Sierens, E., Vansteenkiste, M., Goossens, L., Soenens, B., & Dochy, F. (2009). The synergistic relationship of perceived autonomy-support and structure in the prediction of self-regulated learning. *British Journal of Educational Psychology, 79*(1), 57-68. https://doi.org/10.1348/000709908X304398

Tatlıdil, H. (1992). *Çok değişkenli istatistiksel analiz* [Multivariate statistical analysis] Hacettepe Üniversitesi Yayınları.

Vedder-Weiss, D., & Fortus, D. (2011). Adolescents' declining motivation to learn science: Inevitable or not. *Journal of Research in Science Teaching, 48*(2), 199–216. https://doi.org/10.1002/tea.20398

Waddington, D., Nentwig, P., & Schanze, S. (2007). *Making it Comparable. Standards in science Education 1664.* Waxmann