Education and Science

Vol 41 (2016) No 185 101-122

Classification of PISA 2012 Mathematical Literacy Scores Using Decision-Tree Method: Turkey Sampling

Gökhan Aksu [1], Cem Oktay Güzeller [2]

Abstract

The purpose of this study was to classify successful and unsuccessful students in terms of mathematical literacy according to interest towards the course, attitude, motivation, perception, self-efficacy, anxiety and studying discipline variables and to determine the effect of these variables on classification. The sampling of the study consisted of the students who participated in the Program for International Student Assessment (PISA) in Turkey. Data was collected from a total of 1391 15-year-old students. CHAID analysis, which is a decision-tree technique, and data mining were used for data analysis. SPSS and WEKA software were used to analyze data. Self-efficacy perception, attitude towards the course and studying discipline were found to be the most important affective characteristics in classification of successful and unsuccessful students. It was found that accurate classification percentage obtained by J.48 decision tree, which is a data mining method, was very close to the value obtained by CHAID analysis method. These results suggest that CHAID analysis can be considered as an alternative method to decision tree methods used in data mining. According to the findings obtained from the study, firstly, self-efficacy, attitude towards the course, anxiety and studying discipline should be concentrated on in mathematical literacy for the Turkey sampling. It is believed that success status of students can be changed and Turkey can rank higher in PISA exams through the arrangements to be made in these domains.

Keywords

CHAID

PISA

Mathematical literacy

Data mining

Decision tree

Article Info

Received: 16.06.2015

Accepted: 17.05.2016

Online Published: 09.06.2016

DOI: 10.15390/EB.2016.4766

Introduction

Large scale examinations are used in evaluating the student success both nationally and internationally (Dossey, McCoren, & O'Sullivan, 2006). Several countries apply large scale examinations to obtain significant information on their education systems and are therefore able to conduct international comparisons (Feuer, 2012). Programme for International Student Assessment (PISA) is one of the most comprehensive education studies in the world (Bulle, 2011). PISA 2012 test with a focus on mathematics was taken by 510 thousand students in the age group of 15 years old with minimum 7 years of education (MEB, 2013). In parallel to the recent changes and developments

[1] Adnan Menderes University, Aydın Vocational High School, Turkey, gokhanaksu1983@hotmail.com

[2] Hacettepe University, Faculty of Education, Deparment of Education Sciences, Turkey, cguzeller@gmail.com

in mathematics education, the concept of mathematics literacy emerged and studies were conducted in this field in the international large scale examinations (Yenilmez & Turgut, 2012). Literacy is defined as the ability of students to use the information and skills gained in basic courses when and where necessary, to analyize and judge the problems in various situations and to effectively use the results thus obtained (OECD, 2003). The literacy concept which involves various areas was defined as the capacity of an individual to understand and know the role of mathematics in the surrounding world by using the processes of mathematical thinking and decision making in the solution of present and future problems that may be faced as a citizen who thinks, produces and generates (OECD, 2006). By a different definition, mathematics literacy is described as the ability of students to solve problems, to make analyses, to judge, to create effective solutions in different areas and conditions (Özgen & Bindak, 2008). Various studies addressed to the relation between the variables in the PISA student test and the achievement in mathematics (Şengül Avşar, & Yalçın, 2015; Aydın, Erdağ, & Taş, 2011; Coşguner, 2013; Gürsakal, 2012).

Large amount of data is obtained with the advancement of technology; however a great deal of difficulty is experienced in extracting the meaningful and beneficial ones among these data (Alan, 2012). Different disciplines carry out statistical or mathematical analyses for different purposes on the data in databases (Alan, 2014). Data obtained from the studies in the fields of social sciences, economy and medicine aim to make inference on the unknown; to determine the factors affecting the analyzed phenomena and to determine their levels (Doğan & Özdamar, 2003). As it is not easy to analyze and interpret large amounts of data using traditional statistical methods, a need has arisen to eliminate these difficulties and data mining methods were introduced to address to this need (Özkan, 2008). Data mining is an analysis method used to discover the information in a database to reveal unknown patterns (Larose, 2005). In addition to being a rapidly progressing information technology in today's information age, data mining is described as a method that is used to determine important information that is hidden in large datasets (Fayyad, 1998).

Decision trees, which are one of the models formed by using J48, REPTree, RandomTree, DecisionStump, SimpleCART and NBTree algorithms among the most common data mining algorithms (Alan, 2014), are also termed as classification trees (Bramer, 2007, p. 6). Decision trees is an operation of classifying input data using a classification or clustering algorithm into tree-like sub-groups until all elements have the same class label (Orhan, 2012). Decision tree, which is an attractive classification method, consists of a collection of decision nodes connected by the branches that extend until leaf nodes where classes downward from the root node that is accepted as the start point of classification take place (Larose, 2005, p. 107). Different methods are used in decision trees for classification and splittingwhich are Chi-squared Automatic Interaction Detection (CHAID), Classification and Regression Trees (C&RT); Quick, Unbiased Efficient Statistical Tree (QUEST), C5.0 and ID3 methods (Elsayad & Elsalamony, 2013; Chang, 2011). CHAID analysis was first developed by Kass (1980) to perform analysis with categorical dependent variables (Magidson, 1982; Ratner, 2007). This analysis type is an approach that can be used when dependent variable classification or ranking scale and independent variables are continuous, categorical or ordinal (Kayri & Boysan, 2007). In CHAID analysis, modelling operation is performed by taking into account a group of independent variables and the interaction between them in such a way to provide optimum prediction of the dependent variable (Doğan & Özdamar, 2003). CHAID analysis divides the data cluster of categorical variables into detailed homogenous sub-groups in such a way to best explain the dependent variable (Erbaş & Güneş, 1998). The obtained sub-clusters consist of smaller satisfying groups and beginning variables are independently re-categorized until optimum estimation is obtained (Satıcı, Akkuş, & Alp, 2009). Similar category integration operation which is applied stepwise continues until it is statistically decided that no more can be performed between the variables (Doğan, 2003). A review of the literature revealed that classification based analyses are used together with the variables that are specific to affective domain to determine the variables affecting academic achievement in PISA and similar large scale examinations (Demir & Kılıç, 2010; Güzel İş & Berberoğlu, 2010; Güzel İş, 2014). This indicates that affective domain of learning has begun to gain more prominence (Lehman, 2006).

Considering that emotions affect our ideas and behaviors, affective characteristics are believed to have a major role in a learning environment (Picard, 1997; Broekens, Kosters, & Verbeek, 2007). Affective preparedness has a significant role in success or failure of students in mathematics course (Çoban, 1989). Considering that interest, attitude, self-efficacy, anxiety, perception and motivation are psychological constructs that cannot be directly observed, these are indicated to be highly important latent variables on mathematics success (Aşkar & Erdem, 1986). Pajares (1996) reported that self-efficacy has a mediatory effect on success and Zimmerman (2000) reported that there are statistically significant correlations between self-efficacy and academic success. In addition, a review of the literature carried out on large scale examinations like PISA 2003 revealed that mathematical self-efficacy had a mediatory effect between PISA mathematics score and independent variables such as gender, prior knowledge about the course, cognitive skill level and learning skills and a significant impact on achievement (Ferla, Valcke, & Cai, 2009). Similarly, Pajares and Miller (1994) found significant correlations between academic self-efficacy and academic performance. Güzeller and Akın (2014) attempted to determine the effect of knowledge and communication techniques variables which have sub-scales such as internet use and program use on mathematics success using data of 40 countries in PISA 2006 data. It is stated that the Internet use and ability to use software have a significant effect on the mathematics susccess while the variance values they explain are very low. Aypay (2010) conducted correlation and regression analyses to knowledge and communication techniques variables in PISA 2006 Turkey sampling. Yalçın, Aslan, and Usta (2012) analyzed whether there was a significant difference between mathematics, reading and sciences according to some demographic variables. In a similar study, Acar (2012) observed that variables such as educational resources at home and computer use had a significant impact on mathematics performance of students. Özdemir and Gelbal (2014) found correlations between variables group of reading skills and variable group of the opportunities of students and their families using canonical correlation analysis on PISA 2009 data. However, a review of the literature found no empirical study finding on the affective characteristics affecting mathematics success and their levels of effect on mathematics success.

The purpose of this study was to determine which independent variables, that were thought to have an effect on mathematical literacy, had a significant effect and to determine order of importance of these variables. Furthermore, the study analyzed how the students were classified in terms of their success status according to interest towards the mathematics course, attitudes towards mathematics, motivation, perception, self-efficacy, anxiety and studying discipline variables which are considered as independent variables. The reason for using CHAID analysis technique, which is a decision tree method, was that order of importance of predictive variables on dependent variable can be easily visualized and classification can be performed conveniently based on that visual structure in this method. This study is believed to differ from previous studies and contribute to the literature in terms of using decision trees and data mining methods which have not been much used in the field of education.

Depending on all of the above matters, the purpose of the study is to determine which independent variables, that are believed to have an effect on the mathematics literacy, have a significant effect and to reveal the order of significance of these variables. In addition, the study examined the success classification of students according to the independent variables of interest, attitudes, motivation, perception, self-efficacy, anxiety and study discipline with respect to the mathematics course. The problem sentence of the study within the framework of the specified purposes is "do the variables of interest, attitudes, motivation, perception, self-efficacy, anxiety and study discipline have a significant effect in classification with respect to the mathematics course in the sample of PISA 2012 Turkey?" The study attempted to find answers to the following research questions within the framework of the specified general purpose.

1. What is the effect of independent variables of interest, attitudes, motivation, perception, self-efficacy, anxiety and study discipline have a significant effect in classification with respect to the mathematics course?
2. How is the classification of mathematics literacy of students according to the independent variables of interest, attitudes, motivation, perception, self-efficacy, anxiety and study discipline with respect to the mathematics course?
3. How is the order of significance of the independent variables in classifying mathematics literacy of students?

Method

This is a general screening study aiming to determine to what extent the categorical dependent variable (mathematical literacy) is explained by the independent variables (interest, attitude, motivation, perception, self-efficacy, anxiety and studying discipline) in PISA 2012 Turkey sampling. The study used results of PISA 2012 national data. PISA exams were first conducted in 2000 and after that, in every three years on regular basis with the aim of determining level of basic knowledge and skills of 15 year-old students in OECD and other participating countries (OECD, 2013. In Turkey sampling, answers of a total of 1391 individuals for 7 affective areas obtained by exclusion of uncoded or blank items from the analysis were included in the analysis.

Target Population and Sample

The target population of the study consists of the 965.736 students from OECD member countries and other countries participating in the PISA 2012 student test within the scope of the study. There are 4848 students in the age of 15 years of old in the Turkey target population. Although all of the target population was included in the analysis, it was determined that there was a lost data problem in the responses given by the students for all of the 7 independent variables. Alpar (2003) emphasized that the lost data process should be random and therefore the obtained data should be a random sample of the complete and lost values set for the lost value could be considered to be negligible. When we study the responses given by the students, it was determined that students at different schools didn't reply to the scales of attitudes, interest and anxiety towards mathematics. There are 1391 students in the target population of Turkey as a result of deducting the empty replies for 7 subscales addressed within the scope of the study. Table 1 includes the distribution of the students in the target population per grades.

Table 1. Distribution of Participating Students Per Grades

Grades	Grade 7	Grade 8	Grade 9	Grade 10	Grade 11	Grade 12
Number of students	10	30	383	904	55	9
Percentage	0,72	2,16	27,53	64,99	3,95	0,65

When the Table 1 is studied, it was determined that the distribution of the students per grades within the scope of the study had great similarity with the target population of Turkey (MEB, 2013). Likewise, a positive and high relation was determined upon examining the order differences correlation coefficient between the percentages of the students within the study and the percentages of the students in the Turkey target population (r=.94, p<.01). According to this results, it is believed that the sample represent the target population. The STATISTICA program was used in order to determine the effect and power values that were calculated before the study. After entering the concerned data file within the scope of the analysis to the program, the power value was determined to be 0,52 for α=0,05 and β=0,20. As stated by Goodwin (2010), different criteria are used to interpret the Cohen's d value which calculated by different formula for each test. While the high effect size for the independent sample t test was accepted to be 0,8 as the bottom limit value, the high effect size bottom limit in the variance analysis was accepted to be 0,4, and 0,5 for the chi-square test (Cohen, 1992). The effect size value calculated for the CHAID analysis that was done by 1391 students based on the chi-square statistics was determined to be 0,52. According to the obtained value, it is believed that the findings from the sample could be generalized to the target population (Murphy & Myors, 2004).

Data Collection Tools

The PISA 2012 test studied particularly the student surveys used in the examination in order to determine the affective qualities which are believed to have an effect on the mathematics literacy. 7 different sub-scales were used in order to determine the affective qualities of the students after studying the questions in the PISA student test. Table 2 includes the affective qualities within the scope of the study in the PISA 2012 Turkey sample and items related to each scale.

Table 2. Items on Affective Qualities

Affective Area	**Number of Items**	**Items**
Interest	4	ST29Q01, ST29Q03, ST29Q04, ST29Q06
Motivation	4	ST29Q02, ST29Q05, ST29Q07, ST29Q08
Attitudes	6	IC22Q01, IC22Q02, IC22Q04, IC22Q06, IC22Q07, IC22Q08,
Self-efficacy	8	ST37Q01, ST37Q02, ST37Q03, ST37Q04, ST37Q05, ST37Q06, ST37Q07, ST37Q08
Anxiety	5	ST42Q01, ST42Q03, ST42Q05, ST42Q08, ST42Q10
Perception	5	ST42Q02, ST42Q04, ST42Q06, ST42Q07, ST42Q09
Studying discipline	9	ST46Q01, ST46Q02, ST46Q03, ST46Q04, ST46Q05, ST46Q06, ST46Q07, ST46Q08, ST46Q09

When the Table 2 is studied, it is seen that subscales of items in various numbers for each affective area are used as the data collection tool. All of the items in subscales are scored in 4 point Likert type like strongly agree=1, agree=2, disagree=3 and strongly disagree=4.

Data for 7 different affective characteristics were collected using the answers to the items related with interest, attitude, motivation, perception, self-efficacy, anxiety and studying discipline as data collection instrument. The questions in PISA student questionnaire about affective domains that were considered as predictive variable and the definitions of the variables are presented below.

Interest: Interest is defined as the will or tendency of a person to bond to an activity, person or an object for a long time even under restrictive conditions (Tan, 1972). 4 items in PISA student questionnaire that were scored in 4-item Likert type were included in the analysis to determine interest level. Reliability coefficient of the scale was found to be .91. Explanatory factor analysis was applied first to the items within the scope of the validity study of the interest subscale consisting of the 4 items used in the PISA 2012 test. Accordingly, KMO value of the scale was found to be 0,84 and Barlett test result was found to be statistically significant (χ2=3623,76, sd=6, p<.01). According to these

results, it was determined that the 4-item subscale had a single-factor structure and 77,82% of the total variance was explained by the scale items. Then, the explanatory factor analysis was applied to determine the validity of the single dimension factor structure of the scale and it was concluded that there were strong evidences regarding to the validity of the results obtained from the measurement tool according to the model fit indices (χ2/sd=8,28, RMSEA=0,07, NFI=1,00, GFI=0,99) obtained from the analysis (Kline, 2005).

Motivation: Motivation is defined as individuals' acting based on their desire and will to perform an objective (Ayaydın & Tok, 2015). 4 items in PISA student questionnaire were included in the analysis to determine motivation level. Reliability coefficient of the scale was found to be .87. Explanatory factor analysis was applied first to the items within the scope of the validity study of the single-dimension structure of the motivation subscale consisting of 4 items. Accordingly, KMO value of the scale was found to be 0,83 and Barlett test result was found to be statistically significant (χ2=2664,14, sd=6, p<.01). According to these results, it was determined that the 4-item subscale had a single-factor structure and 71,93% of the total variance was explained by the scale items. Then, the explanatory factor analysis was applied to determine the validity of the single dimension factor structure of the scale and it was concluded that there were strong evidences regarding to the validity of the results obtained from the measurement tool according to the model fit indices (χ2/sd=3,14, RMSEA=0,03, NFI=1,00, GFI=0,99) obtained from the analysis (Kline, 2005).

Attitude: Attitude is defined as affective preparedness or tendency of an individual which is observed in the form of acceptation or rejection of a certain person, organization or idea (Anderson, 1988). 6 items in PISA student questionnaire were included in the analysis to determine attitude level in the study. Reliability of the scale was found to be .69. Explanatory factor analysis was applied first to the items within the scope of the validity study of the single-dimension structure of the attitude subscale consisting of 4 items. Accordingly, KMO value of the scale was found to be 0,81 and Barlett test result was found to be statistically significant (χ2=326,08, sd=15, p<.01). According to these results, it was determined that the 4-item subscale had a single-factor structure and 51,36% of the total variance was explained by the scale items. Then, the explanatory factor analysis was applied to determine the validity of the single dimension factor structure of the scale and it was concluded that there were strong evidences regarding to the validity of the results obtained from the measurement tool according to the model fit indices (χ2/sd=14,40, RMSEA=0,29, NFI=0,63, GFI=0,80) obtained from the analysis (Kline, 2005).

Self-efficacy: Self-efficacy is defined as the belief in one's capabilities to organize and execute the courses of action required to manage prospective situations (Bandura, 1997). 8 items in PISA student questionnaire were included in the analysis to determine self-efficacy level. Reliability of the scale was found to be .83. Explanatory factor analysis was applied first to the items within the scope of the validity study of the single-dimension structure of the self-efficacy subscale consisting of the 8 items. Accordingly, KMO value of the scale was found to be 0,83 and Barlett test result was found to be statistically significant (χ2=3457,53, sd=28, p<.01). According to these results, it was determined that the 8-item subscale had a single-factor structure and 58,11% of the total variance was explained by the scale items. Then, the explanatory factor analysis was applied to determine the validity of the single dimension factor structure of the scale and it was concluded that there were strong evidences regarding to the validity of the results obtained from the measurement tool according to the model fit indices (χ2/sd=30,98, RMSEA=0,14, NFI=0,90, GFI=0,90) obtained from the analysis (Kline, 2005).

Anxiety: Anxiety is defined as the uneasiness or irrational fear in individuals as a reflection of fear of any danger (Manav, 2011). 5 items in PISA student questionnaire were included in the analysis to determine anxiety level. Reliability of the scale was found to be .80. Explanatory factor analysis was applied first to the items within the scope of the validity study of the single-dimension structure of the anxiety subscale consisting of 5 items. Accordingly, KMO value of the scale was found to be 0,80 and Barlett test result was found to be statistically significant (χ2=2184,72, sd=10, p<.01). According to these results, it was determined that the 5-item subscale had a single-factor structure and 56,63% of

the total variance was explained by the scale items. Then, the explanatory factor analysis was applied to determine the validity of the single dimension factor structure of the scale and it was concluded that there were strong evidences regarding to the validity of the results obtained from the measurement tool according to the model fit indices (χ2/sd=20,78, RMSEA=0,12, NFI=0,96, GFI=0,91) obtained from the analysis (Kline, 2005).

Perception: Perception is defined as the process of evolution of sensory stimuli from the environment into meaningful experiences in an individual's mind (Coren, Ward, & Enns, 1993). 5 items in PISA student questionnaire were included in the analysis to determine perception level. Reliability of the scale was found to be .50. Explanatory factor analysis was applied first to the items within the scope of the validity study of the single-dimension structure of the perception subscale consisting of 5 items. Accordingly, KMO value of the scale was found to be 0,86 and Barlett test result was found to be statistically significant (χ2=2962,19, sd=10, p<.01). According to these results, it was determined that the 4-item subscale had a single-factor structure and 63,77% of the total variance was explained by the scale items. Then, the explanatory factor analysis was applied to determine the validity of the single dimension factor structure of the scale and it was concluded that there were strong evidences regarding to the validity of the results obtained from the measurement tool according to the model fit indices (χ2/sd=8,79 RMSEA=0,07, NFI=0,99, GFI=0,99) obtained from the analysis (Kline, 2005).

Studying discipline: It can be defined as showing the best to effort by the students to be successful in classroom by completing necessary preparations for exams and the course (OECD, 2013. 9 items in PISA student questionnaire were included in the analysis to determine perception level. Reliability of the scale was found to be .91. Explanatory factor analysis was applied first to the items within the scope of the validity study of the single-dimension structure of the studying discipline subscale consisting of 9 items. Accordingly, KMO value of the scale was found to be 0,91 and Barlett test result was found to be statistically significant (χ2=7329,63, sd=36, p<.01). According to these results, it was determined that the 9-item subscale had a single-factor structure and 69,22% of the total variance was explained by the scale items. Then, the explanatory factor analysis was applied to determine the validity of the single dimension factor structure of the scale and it was concluded that there were strong evidences regarding to the validity of the results obtained from the measurement tool according to the model fit indices (χ2/sd=35,75, RMSEA=0,15, NFI=0,94, GFI=0,87) obtained from the analysis (Kline, 2005).

Collection of Data

Data about the PISA 2012 student test was obtained from https://pisa2012.acer.edu.au/downloads.php. Since the *Student questionnaire data file* on the page was in the text document (txt) format, the researchers converted the data in the SPSS programme by the command suggested on the PISA official page and made them ready for analysis. Conversion process took place by processing in the program the syntax that allowed saving of the files recorded as text document under the "SPSS™ control files" on the internet page of the data as SPSS extension (.sav). The data used in this study was obtained from the internet site including the PISA data as stated in the reports issued by the Ministry of National Education Department of Research and Development which is one of the official authorities of the Republic of Turkey. The data file handled within the study is used by the Ministry of National Education for similar objectives. In addition, the data file doesn't include any share on the personal information of the students and it is believed that the study has no ethical issue.

Analysis of Data

Total scores for each subscale was obtained by grading the responses to items on 7 affective areas by the researchers in the PISA 2012 test as strongly agree=1, agree=2, disagree=3 and strongly disagree=4. After determining the range values for each subscale, the range value was divided into the group number in determining the group ranges to create 3 groups with low, medium and high total scores. There are 4 items in the interest and motivation subscales and the minimum score to be

obtained from the scale was calculated as 4, and the highest score as 16 while the total scores between 4-8 are interpreted as low, between 8-12 as medium and between 12-16 as high. There are 5 items in the anxiety and perception subscales and the total scores between 6-12 are interpreted as low, between 12-18 as medium and between 18-24 as high. There are 8 items in the self-efficacy subscale and the minimum score to be obtained from the scale was calculated as 8, and the highest score as 32 while the total scores between 8-16 are interpreted as low, between 16-24 as medium and between 24-32 as high. There are 9 items in the studying discipline subscale and the minimum score to be obtained from the scale was calculated as 9, and the highest score as 36 while the total scores between 9-18 are interpreted as low, between 18-27 as medium and between 27-36 as high.

The study examined the presence of lost data which is believed to be problem to underline before mentioning about the findings obtained in the research (Demir & Parlak, 2012). As the systematic differences between those who response to the items and those who don't may cause bias, SPSS program was used for analysis in order to determine whether there is lost data in the data group (Allison, 2009). Alternative to the deletion of data loss or neglect approach is to assign the lost data Pigott, 2001). However, Little and Rubin (1987) and Allison (2003) indicate the traditional data assignment methods are dishonest methods. In the literature, there are two different opinions how to handle missing data and which of the two will be considered, it was decided after determining whether missing data has a random distribution.

Cronbach alpha coefficient was used within the scope of the reliability work of the measurement tools and explanatory and confirmatory factor analysis was used within the scope of the validity work. SPSS program factor analysis was used in factoring the items determined for sub scales while confirmatory factory analysis was done in the Lisrel program to confirm the obtained structure. Within the scope of the research questions, CHAID analysis was used to determine the relative effects and level of significance of the independent variables on the mathematics literacy that is determined is dependent variable. Data mining and J.48 classification tree among the decision tree methods were used in order to present evidence on the validity of the values obtained after the analysis.

PISA 2012 data was analyzed using CHAID analysis. Mean mathematics success score was obtained by taking the average of 5 different mathematics scores with PVMATH code from PISA data (Brown & Micklewright, 2004). Average of the obtained success scores was taken ($\bar{x}$=449.00) and this value was determined as the cutting value. The students over the average were categorized as 1, while those below the average were categorized as 0. Interest towards mathematics, attitude, motivation, perception, self-efficacy, anxiety and studying discipline scores were included in the analysis as independent variables.

Assumptions that are of great importance for many statistical methods such as normality, linearity and homogeneity of variances are not important in CHAID analysis (Alpar, 2011). It is recommended to analyze the quality of data and the validity of the obtained classification and decision-tree method to extract valid results from the data (Mertler & Vannatta, 2005). CHAID analysis gathers the lost values in the data set in a separate group, however, it can divide whole universe into stable sub-nodes by a strong shift logarithm. Therefore, a regression equation to be obtained by this analysis is kept dependent from conventional assumptions (normality, linearity, homogeneity etc) (Horner, Fireman, & Wang, 2010). The statistical test used in the CHAID analysis depends on the target variable with dependent variable being an F test when continuous and a chi-square (χ^2) test when categoric (Oğuzlar, 2003. The assumption necessary for the CHAID analysis is the determination of the scale types for the used variables. In addition, it is also required to state what the categories of the target variable for the categorical variables. The limitation of the analysis is that the dependent variable is a categoric variable. Another limitation is that the analysis provides good results in small sample groups (Ratner, 2015). Meanwhile, a strong aspect of the analysis is that it can visually present the relations the hierarchical relations of which are easily determined in the big data sets (Wilkinson, 1992). Although different classification algorithms are found in decision-tree practices in data mining, J48 algorithm is recommended to be used to obtain the highest classification

percentage (Kızılkaya Aydoğan, Gencer, & Akbulut, 2008). Therefore, this algorithm was used in the study. Java based WEKA 3.7 (Waikato Environment for Knowledge Analysis) package software was used to use this algorithm. WEKA program was developed by Waikato University in New Zealand to process agricultural data (Kuyucu, 2012). J48 decision-tree, which is believed to be the most suitable method to data structure of Naive Bayes, Logistics Regression, ID3, J48, JRIP, PART and neural networks algorithms, was used to determine the validity of the results obtained from CHAID analysis (Witten & Frank, 2000).

Confirmatory factor analysis was performed separately for each scale to determine the validity of the items of affective area in PISA 2012 student questionnaire that was used in the study. The results are presented in Table 3.

Table 3. CFA Results of Affective Characteristics

	χ^2	sd	χ^2/ sd	RMSEA	NFI	GFI	AGFI
Interest	16,57	2	8,28	0,072	1,00	0,99	0,97
Motivation	6,28	2	3,14	0,039	1,00	1,00	0,99
Attitude	1029,61	9	114,40	0,286	0,63	0,80	0,54
Self-efficacy	619,53	20	30,98	0,147	0,90	0,90	0,82
Anxiety	103,91	5	20,78	0,120	0,96	0,97	0,91
Perception	43,96	5	8,79	0,075	0,99	0,99	0,96
Studying Discipline	965,26	27	35,75	0,158	0,94	0,87	0,78

Evaluation of results of analysis of affective characteristics in Table 3 shows a general model data fit (Kline, 2005; Sümer, 2000; Şimşek, 2007). Internal validity means the extent of correct interpretation of the study results (Casady, 2005) and is defined as the ability to obtain causal results in studies (Seltman, 2014). The relevant literature emphasizes the fact that the individuals in the target population should be divided into two as test group and control group and analysis works should be done for both groups for consistent results in order to study internal validity in the relevant literature (Campbell & Stanley, 1963; Brossart, Clay, & Willson, 2002; Cook & Campbell, 1979). Therefore, 1391 students within the scope of the study were divided into two groups with odd and even line numbers considering their order in the data file (Casady, 2005). It was determined that there was no significant difference between the results of the analyses that were done separately for both groups. No change in the order of significance of the independent variables that are effective in classification and no change in the classes as a result of the CHAID analysis conducted for both groups demonstrate that the data have an internal consistency (Shadish, Cook, & Campbell, 2002). External validity refers to generality of the findings obtained within the scope of the study to bigger groups or universe (Brewer, 2000; Robson, 2002) and it is believed that there is external validity as there is sufficient number of sample to determine (Slack & Draugalis, 2001). According to a different perspective, it is recommended to carry out the study on more subjects (Büyüköztürk, 2014). On the other hand, it is believed that unbiased determination of groups remove biased choice which prevents external validity (Karasar, 2008).

Results

The study is primarily identified with the existence of a possible missing values sources in the statistical estimations. Due to interest to mathematics, attitude towards mathematics and motivation variables have missing data for more than 5%, correlation between the proposed variables were examined condition of randomness which recommended by Kalaycı (2014). It was determined that the lost values demonstrated a random distribution as there is low relation (r=.14, p>.05) between the variables (Allison, 2003; Little, 1988), after examining the randomness based on the correlation between the variables proposed by Kalaycı (2014). As a result of the obtained findings, it was determined that the lost data were random and a listwise method was applied which is one of the approaches to remove the lost data problem (Demir, 2013). The reason to do this process was that it was desired to include the students who responded to all of the 7 affective areas in the analysis. As a result of the analysis, the number of individuals who are included in the sample and who have full data matrix was determined to be 1391.

Following the testing of the assumptions, information on how the students are classified in terms of success using 7 affective characteristics and at what level the classification will end by frequency and percentage values of independent variables after determining order or importance of each independent variable on the dependent variable using CHAID analysis. In CHAID analysis, firstly the table including summary information of the model is presented. Accordingly, the dependent variable in the model is mathematics literacy status, while independent variables are interest to mathematics, attitude towards mathematics, motivation, perception, self-efficacy, anxiety and studying discipline total scores. However, it was found that among the independent variables only self-efficacy, attitude and studying discipline were included in the analysis as only they significantly predict success. Classification table of expected and observed values by CHAID analysis are presented in Table 4.

Table 4. Classification Table for Success Status

	Unsuccessful	Successful	Success Percentage
Unsuccessful	499	245	67,10
Successful	192	455	70,30
Total	49,70	50,30	68,60

As presented in Table 4, of 744 unsuccesful students, 499 (67.07%) were accurately classifid by the program; however, 245 (32.93%) were classified as succesful although they were in fact unscuccesful. Similarly, of a total of 647 succesful students, 455 (70.32%) were accurately classified by the program; however 192 (29.68%) students were classified as unsuccesful although they were succesful. It was observed that the program had a general success of 68.60% by taking the averages of succesful and unsuccesful classifications made by the program. Margin of error of the system is presented in the risk value produced by the program in addition to the classification table. Accordingly, risk value of the system was approximately 31.40% (1, 68, 60). Analysis results of affective characteristics that were determinative to classify succesful and unscuccesful students according to mathematics literacy scores and their order of importance are presented in Figure 1.

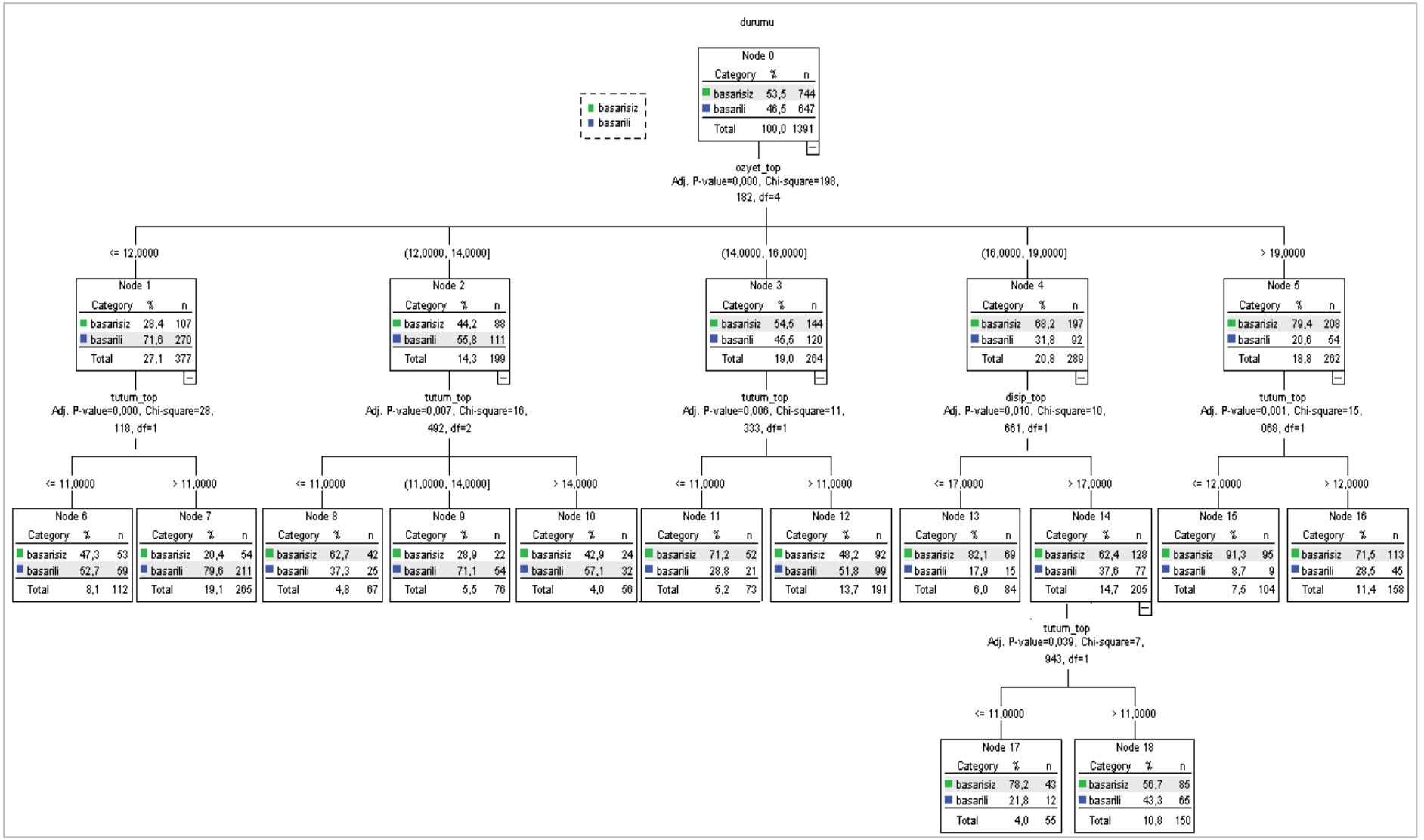

Figure 1. Decision-Tree Model for Success Status

It is understood from Figure 1 that 53.50% of students were classified as unsuccessful, while 46.50% were classified as successful. Self-efficacy, which consists of 5 sub-levels, was observed to be the independent variable that best explains success. Minimum and maximum possible scores for self-efficacy total scores was 8 and 32 respectively. Analysis of the sub-categories of the independent variable that best explains success showed that 377 students with self-efficacy score of below 12.00 (270 successful, 107 unsuccessful) formed node 1. Chi-square value (χ^2=28, p=.000) revealed that the first independent variable that has a statistically significant and highest levels of relation with the dependent variable is attitudes and attitudes best describe the cluster of self-efficacy scores equal to 12 points and small. Accordingly, cutting point for attitude was determined as 11. node was ended at this stage by forming two nodes, which are the ones with an attitude score below 11 and those with an attitude score above 11. The second variable with the highest relationship with the dependent variable was the individuals with a self-efficacy score of 12-14. In node 2, which consisted of a total of 199 individuals (111 successful, 88 unsuccessful) (χ^2=16, p=.007) attitude towards the course was determinative for a sub-node. The related branch was terminated by determining 3 different nodes for attitude scores, which are attitude scores of 11 and smaller, 14 and smaller and greater than 14. 264 individuals with a self-efficacy score of 14-16 were grouped for the third node (χ^2=11, p=.006). Attitude towards the course was determinative for a lower sub-node for this node. The branch was terminated by determining 2 different nodes for attitude scores including those with an attitude score of 11 and smaller and those with an attitude score of 11 and greater. Unlike other nodes, studying discipline was found to be determinative in a sub-node for node 4. 2 different nodes including those with studying discipline scores of 17 and lower and those with a studying discipline score of 17 and higher were determined for studying discipline scores. A sub-node occurred for the related branch. Independent variable that determined a sub-node for those with studying discipline scores of above 17 was found to be attitude towards the course (χ^2=7, p=.039). 2 different nodes, including those with an attitude score of 11 and smaller; and those with an attitude score of 11 and greater were determined and the related branch was terminated. A total of 262 individuals with a self-efficacy score of above 19 at node 5 were grouped for self-efficacy, which explained the dependent variable the best. Thus, the last node at level one for self-efficacy was completed. Attitude scores were determinative for a sub-node (χ^2=15, p=.001). Similarly, for fifth node, attitude towards the course was determinative for a sub-node. This time, 2 different nodes were determined including attitude scores of 12 and lower and those 12 and higher and the related node was terminated. Chi-square values showed that the independent variable that explained success the best was self-efficacy (χ^2=198, p<.05), followed by attitude towards the course (χ^2=10, p<.05) and studying discipline (χ^2=10, p<.05) variables. Gain values for the obtained nodes to determine what the best roots (nodes) were to classify successful students and which of those nodes provided more information are presented in Table 5.

Table 5. Gain Values for Success Status

Node	Node		Gain		Ratio of correct answer	Index
	n	%	n	%		
7. Node	265	19,10	211	32,60	79,60	171,20
12. Node	191	13,70	99	15,30	51,80	111,40
18. Node	150	10,80	65	10,00	43,30	93,20
6. Node	112	8,10	59	9,10	52,70	113,30
9. Node	76	5,50	54	8,30	71,10	152,80
16. Node	158	11,40	45	7,00	28,50	61,20
10. Node	56	4,00	32	4,90	57,10	122,90
11. Node	73	5,20	21	3,20	28,80	61,80
13. Node	84	6,00	15	2,30	17,90	38,40
17. Node	55	4,00	12	1,90	21,80	46,90
15. Node	104	7,50	9	1,40	8,70	18,60
8. Node	67	4,80	25	3,90	37,30	80,20

According to the values presented in Table 5, the best node to discriminate successful and unsuccessful students was node 7 (n=211, %32.60). This node includes the cluster including 265 individuals with attitude score of above 11 among those with a self-efficacy score below 12. Of these 265 individuals, 79.60% were accurately classified. Gain values were analyzed to determine the second best node and it was found that node 12 (n=99, %15,30) was quite successful in determining successful students according to mathematical literacy. This node was the cluster that consisted of 191 individuals with a self-efficacy score of 14-16 and attitude score of over 11. Of these 191 individuals, 51.80% were accurately classified. In addition, it was found that the node that gave the least information in discriminating students in terms of success was node 8 (n=25, %3,90). This node was the cluster that consisted of 67 individuals with a self-efficacy score of 12-14 and attitude score of below 11. Of these individuals, 37.30% were accurately classified.

An analysis was made to determine the validity of findings obtained in the study. Classification results of J.48 decision tree obtained to explain PISA mathematics success of students by their scores of 7 different affective domains is presented in Figure 2.

```
=== Classifier model (full training set) ===

J48 pruned tree
------------------

ozyeterlik = 8: 1 (96.0/25.0)
ozyeterlik = 9: 1 (52.0/14.0)
ozyeterlik = 10: 1 (74.0/16.0)
ozyeterlik = 11: 1 (72.0/24.0)
ozyeterlik = 12: 1 (83.0/28.0)
ozyeterlik = 13: 1 (99.0/42.0)
ozyeterlik = 14
|   tutum = 6: 0 (5.0/1.0)
|   tutum = 7: 1 (1.0)
|   tutum = 8: 0 (5.0/2.0)
|   tutum = 9: 0 (7.0/2.0)
|   tutum = 10: 0 (7.0/2.0)
```

Figure 2. J.48 Classification Results with J.48 Decision Tree

Figure 2 shows that the self-efficacy scores are effective on the first seven branches in classifying the students according to mathematics literacy. In addition, course attitudes are the second most effective independent variables in classifying the students. In order to better understand the three structures in Figure 2, the tree diagram obtained from the WEKA program was designed to show only nodes. Tree diagram of nodes are presented in Figure 3**.**

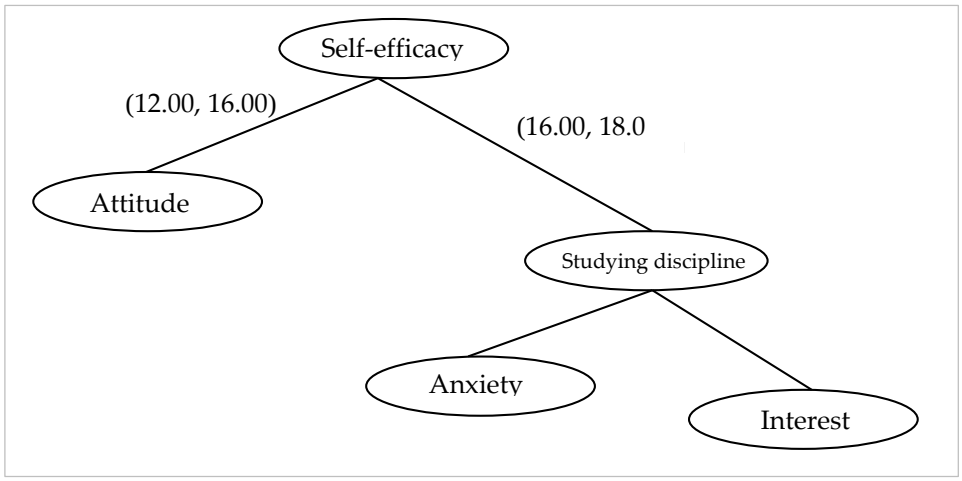

Figure 3. J.48 Decision Tree Nodes

As indicated in Figure 3, self-efficacy was the most important independent variable to classify successful and unsuccessful students in terms of mathematics scores. Since the nodes downwards in order of importance in WEKA program, attitude towards the course was found to be the second most important variable to determine success. It was observed that the determinative affective characteristic in the third node was anxiety. Interest was found to be effective in classification of successful and unsuccessful students. Ratio of correct classification by the program was 69.87%, while the ratio of inaccurate classification was 30.13%. These results are consistent with the results of CHAID analysis. Evidence for validity is believed to be adequate according to obtained results.

Discussion, Conclusion and Suggestions

Affective qualities of students like their attitudes, beliefs and self-confidence are found to be important in their mathematics performance (McLeod, 1992). There are various studies in the literature examining the relations of affective qualities between each other and with success (Kaiser & Willander, 2004). When we look at the studies on the relation of mathematics self-confidence and mathematics self-efficacy perceptions with the mathematics success, we find a positive relation between these variables (Shen, 2002). In addition "mathematics literacy" concept has been recently emphasized when evaluating the knowledge and skills of students in the PISA tests developed by the OECD countries (Uysal & Yenilmez, 2011).

This study was conducted to determine which independent variables have a significant effect on the mathematics literacy according to the total scores in mathematics literacy included in the PISA 2012 student test. In addition, evidence was collected in the study for the validity of results obtained by CHAID analysis using WEKA program.

The first sub problem of this study which aimed to determine the effects of the independent variables addressed within the scope of the study in PISA 2012 mathematics literacy showed that the independent variables of self-efficacy, mathematics attitude and studying discipline had a significant effect on mathematics literacy. It was concluded that these three independent variables were effective in classifying the students as successful and unsuccessful with respect to mathematics literacy. The mathematics literacy concept was also supported by NCTM (National Council of Teachers of Mathematics) In the end of 1990s and it was argued that the mathematics literacy concept was one of the basic goals of the mathematics education. Our country participated in the PISA test which is applied every three years with one focus area each time in 2003 for the first time and then in 2006, 2009 and 2012 (İskenderoğlu, Erkan, & Serbest, 2013). The mathematics literacy concept expressed by PISA deals with the students' capacities of analysing, reasoning and communication as well as solving and interpreting mathematical problems in different conditions including quantitative, spatial, probabilistic thinking and other mathematical concepts. To evaluate these skills, six levels included in the mathematical efficacy scale are used (OECD, 2007). There are numerous studies on the mathematics literacy especially in recent years (Harms, 2000; Kaiser, 2002; EARGED, 2008; Tekin & Tekin, 2004; Özgen & Bindak, 2008; Akay & Boz, 2011; Duran & Bekdemir, 2013). Various studies are also conducted to determine the variables which are believed to have impact on mathematics literacy (Dursun & Dede, 2004; Fisher, 1995; Savaş, Taş, & Duru, 2010; Özer & Anıl, 2011). In addition, Koğar (2015) studied the factors affecting the mathematics literacy by using the mediation model and determined that variables like sex, economic, social and cultural status indices have an important effect on the mathematics literacy. This study is believed to have a strong aspect as it highlights the affective qualities as well while previous studies on mathematics literacy focused on cognitive qualities. Thus, it is concluded that the variables of self-efficacy, attitude and studying discipline have great effect on mathematics literacy which is supported by the concerned literature.

The second sub problem of the study which determined the way of classification of students according to affective qualities with respect to PISA 2012 mathematics literacy found that the correct classification ratios of the seventh and twelfth nodes with the most data in classifying students as successful and unsuccessful to be 79% and 52% respectively. It was concluded that only self-efficacy and course attitudes were effective in this classification. Self-efficacy concept was studied so far in relation with numerous variables in different areas from development psychology to science education, from mathematics to computer (Özsoy Güneş, Çıngıl Barış, & Kırbaşlar, 2013). Self-efficacy refers to the perception and belief of an individual on his own skills and therefore it is believed to be an important variable to be studied in mathematics literacy (Yenilmez & Turgut, 2012).

Azar and Akıncı (2009) found a positive and statistically significant relationship between academic success and self-efficacy belief. Similarly, Randhawa, Beamer, and Lundeberg (1993) reported that self-efficacy perception has a mediatory effect on mathematics success. A study by Duran and Bekdemir (2013) found a positive and moderately significant relation between the self-efficacy perception of mathematics literacy and mathematics success. Pajares and Miller (1994) found that perceptions of students about self-efficacy had a positive impact on problem-solving skills. It was found in this study that self-efficacy was the most important predictive variable in mathematical literacy. This finding is consistent with the previous literature. These results indicate that self-efficacy perception is the variable that best predicts mathematics success in Turkey sampling.

The third sub problem of the study attempted to determine the order of significance of the independent variables which are effective in classifying students as successful and unsuccessful with respect to mathematics literacy as the focus area of PISA 2012 test was mathematics literacy (İnan & Bekler, 2014). As a result of the study, it was found that the independent variables of self-efficacy, mathematics attitude, studying discipline, anxiety and interest had a significant effect. In addition, it was concluded that course interest and student perception were not effective in classifying students with respect to mathematics literacy. Attitude towards the course was found to be the second best variable to predict mathematics literacy. Peker and Mirasyedioğlu (2003) reported a positive and moderate relationship between attitude towards mathematics course and success scores. Similarly, Saracaloğlu, Başer, Yavuz, and Narlı (2004) found a significant relationship between attitude towards mathematics and success. However, on the contrary to the relevant literature, a study by Keşan, Yetişir, and Kaya (2011) found that there was not a significant relationship between attitude towards mathematics and success. Yücel and Koç (2011) found a positive and moderately significant relationship between attitude towards the course and success. Previous literature and 16% predictive power of attitude on success support the result that the best two nodes in classification of individuals in terms of success were self-efficacy and attitude variables. In general, a review of the literature found that attitudes of students towards mathematics had a positive impact on their success (Minato & Yanese, 1984; Ethington & Wolfle, 1986; Cheung, 1988; Erktin, 1993). According to these results, it is expectable that attitude was the second best independent variable to predict success. Furthermore, Gülten, Poyraz, and Soytürk (2012) reported that there was a significant difference between the students who have the habit of studying on regular basis and the students who never study in favour of the students who regularly work in terms of mathematical literacy. Inadequate studying skills are known to be one of the most important reasons for school failure of students (Küçükahmet, 2000). These findings support the fact that studying discipline is the third most important node in classifying students in terms of success. Attention was drawn to the importance of developing positive attitude, emotion and belief in students within the framework of mathematics literacy evaluation results of the PISA 2012 test and it is intended to use these variables to explain the differences in the mathematics literacy success (MEB, 2011). Gülten et al. (2012) reported that mathematical literacy of prospective mathematics teachers who had a habit of studying on daily basis significantly differed from those of the prospective mathematics teachers who did not have that habit, in favour of those who studied regularly. This result supports the fact that studying discipline variable is the third best variable to predict success.

Bakioğlu and Yıldız (2013) emphasize that the students' fear of mathematics should be reduced and their interest in the subject should be increased. It is known that there are works to increase the interest in the subject in countries like Finland that are successful in PISA tests (LUMA, 2002). According to PISA 2012 data, Turkey is one of the countries with the greatest development like Brazil, Tunis and Mexico (OECD, 2014), these countries are far below the OECD average. Looking at the PISA test results, it is known that the science high schools are above the international average with respect to the mathematics literacy, while the state funded schools and vocational high schools have success levels which are well below the average (Berberoğlu & Kalender, 2005). These results reveal that affective qualities like self-efficacy, attitude, studying discipline, anxiety and interest should be developed to ensure higher positioning of our country in large scale tests like PISA.

The results of the study suggest that especially self-efficacy perception, attitude towards the course, anxiety and studying discipline should be concentrated on in mathematical literacy for Turkey sampling. We believe that improvements in these domains will help the students to be more successful and that Turkey will rank higher in PISA examinations. In addition, we recommend that motivation variable, which was found to have no significant effect on success in any of the methods be restudies by analyzing either scale items or responds to the scale.

These results reveal that affective qualities like self-efficacy, attitude, studying discipline, anxiety and interest should be developed to ensure higher positioning of our country in large scale tests like PISA. Particularly, teachers and authorities that draft the curriculum should carry out regulating works on these matters as soon as possible. In this concept, curriculum should be re-designed by also considering the affective qualities instead of focusing on merely cognitive qualities in course programs. Based on the findings, it is suggested that more emphasis should be placed on the self-efficacy perception, course attitude and anxiety for the Turkey sample for mathematics literacy. It is believed that arrangements in these areas will ensure that students will be more successful and our country will have higher ranks in PISA tests. In addition, it is suggested that either the scale items or the responses to the scale of the motivation variable which was determined to have no significant effect on success in any of the different methods should be re-examined and more psychometric works should be done about it.

This study contains some limitations as it only deals with the date in Turkey sample. Therefore, future studies can be conducted by date from different countries. Particularly, further studies are suggested with the PISA data of different years for the Turkey sample and the consistency of the measurement results should be studied. Furthermore, it is recommended to test the classification results with respect to mathematics literacy should be tested again with different analysis methods.

References

Acar, T. (2012). 2009 yılı uluslararası öğrenci başarılarını değerlendirme programında Türk öğrencilerin başarılarını etkileyen faktörler. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme, 3*(2), 309-314.

Akay, H., & Boz, N. (2011). Sınıf öğretmeni adaylarının matematiğe yönelik tutumları, matematiğe karşı öz-yeterlik algıları ve öğretmen öz-yeterlik inançları arasındaki ilişkilerin incelenmesi. *Türk Eğitim Bilimleri Dergisi*, *9*(2), 281-312.

Alan, M. A. (2012). Veri madenciliği ve lisansüstü öğrenci verileri üzerine bir uygulama. *Dumlupınar Üniversitesi Sosyal Bilimler Dergisi, 33*, 165-174.

Alan, M. A. (2014). Karar ağaçlarıyla öğrenci verilerinin sınıflandırılması. *Atatürk Üniversitesi İktisadi ve İdari Bilimler Dergisi, 28*(4), 101-112.

Allison, P. D. (2009). *Missing data: Quantitative methods in psychology*. London: SAGE Publication.

Allison P. D. (2003). Handling missing data by maximum likelihood. *SAS Global Forum 2012*.

Alpar, R. (2003). *Uygulamalı çok değişkenli istatistiksel yöntemlere giriş-1*. Ankara: Nobel Kitabevi.

Alpar, R. (2011). Uygulamalı çok değişkenli istatistiksel yöntemler. Ankara: Detay Yayıncılık.

Anderson, L. W. (1988). *Attitudes and their measurement: Methodology and measurement*. New York: Pergamon Pres.

Aşkar P., & Erdem M. (1986). *Öğretmen adaylarının öğretmenlik mesleğine yönelik tutumları*. Paper presented at the 1.Ulusal Eğitim Kongresi, Marmara Üniversitesi, İstanbul.

Ayaydın, M., & Tok, H. (2015). Motivasyonu etkileyen faktörlere ilişkin sınıf öğretmenlerinin görüşlerinin incelenmesi (Gaziantep örneği). *International Periodical for the Languages, Literature and History of Turkish or Turkic Volume, 10*(11), 187-200.

Aydın, A., Erdağ, C., & Taş, N. (2011). 2003 - 2006 PISA okuma becerileri değerlendirme sonuçlarının karşılaştırmalı olarak değerlendirilmesi (sınavda en başarılı beş OECD ülkesi-Türkiye örneği). *Kuram ve Uygulamada Eğitim Bilimleri Dergisi*, *11*(2), 651-673.

Aypay, A. (2010). Information and communication technology (ICT) usage and achivement of Turkish students in PISA 2006. *The Turkish Online Journal of Educational Technology, 9*(2), 116-124.

Azar, A. ve Akıncı, M. (2009), *Öğretmen Adaylarının Öğretmen Özyeterlik İnancı, Akademik Başarı ve KPSS Başarıları Arasındaki İlişkinin Çeşitli Değişkenler Açısından Yordanması*, 7. Ulusal Fen ve Matematik Eğitimi Kongresi, Gazi Üniversitesi, Ankara.

Bakioğlu, A., & Yıldız, A. (2013). Finlandiya'nın PISA başarısına etki eden faktörler bağlamında Türkiye'nin durumu. *Eğitim Bilimleri Dergisi*, *38*, 37-53.

Bandura, A. (1997). *Self-efficacy: The exercise of control.* New York: Freeman.

Berberoğlu, G., & Kalender, İ. (2005). Öğrenci başarısının yıllara, okul türlerine, bölgelere göre incelenmesi: ÖSS ve PISA analizi. *Eğitim Bilimleri ve Uygulama, 4*(7), 21-35.

Bramer, M. (2007). *Principles of data mining*. London: Springer.

Brewer, M. B. (2000). *Research design and issues of validity.* New York: Cambridge University Press.

Broekens J., Kosters W. A., & Verbeek F. J. (2007). Affect, anticipation and adaptation: Affect-controlled selection of anticipatory simulation in artifical adaptive agents. *Adaptive Behavior, 15*(4), 397-422.

Brossart, D. F., Clay, D. L., & Wilson, V. L. (2002). Methodological and statistical considerations for threats to ınternal validity in pediatric outcome data: Response shift in self-report outcomes. *Journal of Pediatric Psychology*, *27*(1), 97-107.

Brown, G., & Micklewright, J. (2004). *Using international surveys of achievement and literacy: A wiev from the outside.* Montreal, Rome: UNESCO Enstitute for Statistics.

Bulle, N. (2011). Comparing OECD educational models through the prism of PISA. *Comparative Education*, *47*(4), 503-521.

Büyüköztürk, Ş. (2014). *Deneysel desenler: Öntest sontest kontrol gruplu desen ve veri analizi.* Ankara: Pegem Akademi.

Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research*. Boston: Houghton Mifflin Co.

Casady, R. (2005). Internal validity. *AHA Research Guides*, 1-2.

Chang, T. S. (2011). A comparative study of artificial neural networks, and decision trees for digital game content stocks price prediction. *Expert Systems with Applications*, *38*, 14846-14851.

Cheungh, K. C. (1988). Outcomes of schooling: Mathematics achıevement and attitudes towards mathematics learning in Hong Kong. *Educational Studies in Mathematics* ,.19, 209-219.

Çoban A. (1989). *Ankara merkez ortaokullarındaki son sınıf öğrencilerinin matematik dersine ilişkin tutumları* (Master's thesis). Gazi University, Institute of Social Science, Ankara.

Cohen, J. (1992). Statistical power analysis. *Current Directions in Psychological Science*, *1*(3), 98-101.

Cook, T. D., & Campbell D. T. (1979). *Quasi-experimentation: Design and analysis issues for field settings.* Boston: Houghton Mifflin.

Coren, S., Ward, L.M., & Enns, J. T. (1993). *Sensation and perception.* Harcourt Brace College Publisers.

Coşguner, T. (2013). *Uluslararası öğrenci başarı değerlendirme programı (PISA) 2009 uygulaması okuma becerileri okuryazarlığını etkileyen faktörler* (Master's thesis). Akdeniz University, Institute of Educational Sciences, Antalya.

Demir, E. (2013). Kayıp verilerin varlığında çoktan seçmeli testlerde madde ve test parametrelerinin kestirilmesi: SBS örneği. *Eğitim Bilimleri Araştırmaları Dergisi*, *3*(2), 48-68.

Demir, E., & Parlak, B. (2012). Türkiye'de eğitim araştırmalarında kayıp veri sorunu. *Eğitim ve Psikolojide Ölçme ve Değerlendirme Dergisi, 3*(1), 230-241.

Demir, İ., & Kılıç, S. (2010). Using PISA 2003: Examining the factors affecting students' mathematics achievement. *H. U. Journal of Education*, *38*, 44-54.

Doğan, İ. (2003). Holştayn ırkı ineklerde süt verimine etki eden faktörlerin CHAID analizi ile incelenmesi. *Ankara Üniversitesi Veteriner Fakültesi Dergisi, 50*, 65-70.

Doğan, N., & Özdamar, K. (2003). CHAİD analizi ve aile planlaması ile ilgili bir örnek. *Türkiye Klinikleri Tıp Bilimleri Dergisi, 23*(5), 392-397.

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

Duran, M., & Bekdemir, M. (2013). Görsel matematik okuryazarlığı özyeterlik algısıyla görsel matematik başarısının değerlendirilmesi. *Pegem Eğitim ve Bilim Dergisi, 3*(3), 27-40.

Dursun, Ş., & Dede, Y. (2004). Öğrencilerin matematikte başarısını etkileyen faktörler: Matematik öğretmenlerinin görüşleri bakımından. *Gazi Eğitim Fakültesi Dergisi*, *24*(2), 217-230.

EARGED. (2008). *PISA'da okuma becerileri: PISA'da matematik okuryazarlığı*. Retrieved from http://pisa.meb.gov.tr/wp-content/uploads/2013/07/PISA-2009-Ulusal-On-Rapor.pdf

Erbaş, S., & Güneş, A. (1998). Chaid analizi. *Istatistik Konferansı Bildiri Kitabı* içinde (p. 381). Ankara.

Elsayad A. M., & Elsalamony, H. A. (2013). Diagnosis of breast cancer using decision tree models and SVM. *International Journal of Computer Applications, 83*(5), 19-29.

Erktin,E. (1993). The Relationship between math anxiety attitude toward mathematics and classroom environment. 14.*International Conference of Stress and Anxiety Research Society (STAR)*, Cairo, Egypt, April 5-7 1993.

Ethington, C.A. ve Wolfle, L.M. (1986). A structural model of mathematics achievement for men and women. *American Educational Research Journal*, 5-75.

Fayyad, U. (1998). Mining databases: Towards algorithms for knowledge discovery. *DE Bulletin, 21*(1), 41-48.

Ferla, J., Valcke, M., & Cai, Y. (2009). Academic self-efficacy and academic self-concept: Reconsidering structural relationships. *Learning and Individual Differences, 19*, 499-505.

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

Fisher, C. W. (1995). Academic learning time. L. W. Anderson (Ed.). *International encyclopedia of teaching and teacher education.* Oxford: Pergamon.

Goodwin, C. J. (2010). *Research in psychology methods and design*. New Jersey: John Wiley and Sons.

Gülten, D. Ç., Poyraz, C., & Soytürk, İ. (2012). Öğretmen adaylarının matematik okuryazarlığı öz-yeterliklerinin ders çalışma alışkanlıkları açısından incelenmesi. *Eğitim ve Öğretim Araştırmaları Dergisi, 1*(2), 143-149.

Gürsakal, S. (2012). PISA 2009 öğrenci başarı düzeylerini etkileyen faktörlerin değerlendirilmesi. *Süleyman Demirel Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, *17*(1), 441-452.

Güzel İş, Ç. (2014). The impact of student and school characteristics and their interaction on Turkish students' mathematical literacy skills in the programme for international student assessment (PISA) 2003. *Mediterranean Journal of Educational Research*, *15*, 11-30.

Güzel İş, Ç., & Berberoğlu, G. (2010). Students' affective characteristics and their relation to mathematical literacy measures in the programme for international student assessment (PISA) 2003. *Eurasion Journal of Educational Research, 40*, 93-113.

Güzeller, C. O., & Akın, A. (2014). Relationship between ICT variables and mathematics achievement based on PISA 2006 database: International evidence. *The Turkish Online Journal of Educational Technology, 13*(1), 184-192.

Harms, T. J. (2000). *Analysis of Minnesota students' mathematical literacy on TIMSS, NAEP and MNBST* (Yayımlanmış doktora tezi). North Dakota Üniversitesi, North Dakota, ABD.

Horner, S. B., Fireman, G. D., & Wang, E. W. (2010). The relation of student behavior, peer status, race and gender to decisions about school discipline using CHAID decision trees and regression modeling. *Journal of School Psychology*, *48*, 135-161.

İnan, C., & Bekler, E. (2014). PISA sınavlarında Türkiye'nin performansı ve öğretmen eğitiminde çözüm önerileri. *Turkish Studies*, *9*(5), 1097-1118.

İskenderoğlu, T. A., Erkan, İ., & Serbest, A. (2013). 2008-2013 yılları arasındaki SBS matematik sorularının PISA matematik yeterlik düzeylerine göre sınıflandırılması. *Turkish Journal of Computer and Mathematics Education*, *4*(2), 147-168.

Kaiser, G., & Willander, T. (2004). Development of mathematical literacy: Results of an empirical study. *Teaching Mathematics and Its Applications*, *24*, 2-13.

Kaiser, G. (2002). International comparison in mathematics education: An overview. *ICM, 1*, 631-646.

Kalaycı, Ş. (2014). *SPSS uygulamalı çok değişkenli istatistik teknikleri,* Ankara: Asil Yayın Dağıtım.

Karasar, M. (2008). *Bilimsel araştırma yöntemi*. Ankara: Nobel Yayın Dağıtım.

Kass, G. (1980). An exploratory technique for investigating large quantities of categorical data. *Applied Statistics*, *29*, 119-127.

Kayri, M., & Boysan, M. (2007). Araştırmalarda Chaid analizi kullanımı ve baş etme stratejileri ile ilgili bir uygulama. *Ankara Üniversitesi Eğitim Bilimleri Fakültesi Dergisi, 40*(2), 133-149.

Keşan, C.,Yetişir, Ş. ve Kaya, D. (2011). İlköğretim İkinci Kademe Öğrencilerinin Görsel, İşitsel ve Kinestetik Durumlarının Belirlenmesi ve Matematiğe Yönelik Tutumlarının Başarıya Etkisi. *e-Journal of New World Sciences Academy*, 6 (4), 2660-2674.

Kızılkaya Aydoğan, E., Gencer, C., & Akbulut, S. (2008). Veri madenciliği teknikleri ile bir kozmetik markanın ayrılan müşteri analizi ve müşteri bölümlenmesi. *Mühendislik ve Fen Bilimleri Dergisi, 26*(1), 42-56.

Kline, R. B. (2005). *Principles and practice of structural equation modeling* (2nd ed.). New York: Guilford Press.

Koğar H. (2015) Matematik okuryazarlığını etkileyen faktörlerin aracılık modeli ile incelenmesi, Educatıon And Science, 40 (179), 45-55.

Kuyucu, Y. E. (2012). *Lojistik regresyon analizi (LRA), yapay sinir ağları (YSA) ve sınıflandırma ve regresyon ağaçları (C&RT) yöntemlerinin karşılaştırılması ve tıp alanında bir uygulama* (Unpublished master's thesis). Gaziosmanpaşa University, Graduate School of Health Sciences, Tokat.

Küçükahmet, L. (2000). *Öğretimde planlama ve değerlendirme*. Ankara: Nobel Yayınları.

Larose, D. T. (2005). *Discovering knowledge in data.* New Jersey: Wiley Publication.

Little, R. J. A. (1988). A test of missing completely at random for multivariate data with missing values. *Journal of the American Statistical Association, 83*, 1198-1202.

Little, R. J. A., & Rubin, D. B. (1987). S*tatistical analysis with missing data*. New York: John Wiley and Sons.

LUMA. (2002). *Finnish knowledge in mathematics and sciences in 2002: Final report of LUMA programme.* Helsinki: Ministry of Education, Department for Education and Science Policy.

Magidson, J. (1982). Some common pitfalls in causal analysis of categorical data: Special issue on causal modeling. *Journal of Marketing Research, 19*(4), 461-471.

Manav, F. (2011). Kaygı kavramı. *Toplum Bilimleri Dergisi, 5*(9), 291-211.

McLeod, D. (1992). Research on affect in mathematics education: A reconceptualization. D. A. Grouws (Ed.). *Handbook of research on mathematics teaching and learning*. NewYork: MacMillan.

MEB. (2011). *PISA Türkiye.* Ankara: Milli Eğitim Bakanlığı Yenilik ve Eğitim Teknolojileri Genel Müdürlüğü.

MEB. (2013). *PISA 2012 ulusal ön raporu*. Ankara: MEB.

Mertler, C. A., & Vannatta, R. A. (2005). *Advanced and multivariate statistical methods: Practical application and interpretation*. Thousand Oaks, CA: Sage Publication.

Minato, S. ve Yanase, S. (1984). On the relarionship between students's attitudes toward school mathematics and their levels of intelligence. *Educational Studies in Mathematics*, 15, 313-320.

Murphy, K. R., & Myors, B. (2004). *Statistical power analysis, a simple and general model for traditional and hypothesis test.* London: Lawrence Erlbaum Associates.

OECD. (2014). *PISA 2012 results: What students know and can do student performance in mathematics, reading and science* (Vol. 1). Retrieved from http://www.oecd.org/pisa/keyfindings/pisa-2012-results-volume-I.pdf

OECD. (2003). *The PISA 2003 assessment framework – mathematics, reading, science and problem solving knowledge and skills*. Paris: OECD.

OECD. (2006). *Assessing scientific, reading and mathematical literacy: A framework for PISA.* Retrieved from http://www.oecd.org/

OECD. (2007). PISA 2006 database, [Online] Retrieved on 12-April-2010, at URL: http://pisa2006.acer.edu.au/downloads.php

OECD (2013). *PISA 2012 results: what students know and can do (volume I): student performance in mathematics, reading and science*. OECD publishing.

Oğuzlar, A. (2003). Veri önişleme. *Erciyes Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi, 21*, 67-76.

Orhan, U. (2012). *Makine öğrenmesi dersi: Entropi, karar ağaçları (ID3 ve C4.5 algoritmaları), sınıflandırma ve regresyon ağaçları, ders notları.* Retrieved from http://bmb.cu.edu.tr/uorhan/DersNotu/Ders03.pdf

Özdemir, B., & Gelbal, S. (2014). Investigating factors that affect Turkish students' academic success with canonical commonality analysis according to PISA 2009 results. *Education and Science, 39*(175), 41-57.

Özer, Y., & Anıl, D. (2011). Öğrencilerin fen ve matematik başarılarını etkileyen faktörlerin yapısal eşitlik modeli ile incelenmesi. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, *41*, 313-324.

Özgen, K., & Bindak, R. (2008). Matematik okuryazarlığı öz-yeterlik ölçeğinin geliştirilmesi. *Kastamonu Eğitim Dergisi, 16*(2), 517-528.

Özkan, Y. (2008). *Veri madenciliği yöntemleri*. İstanbul: Papatya Yayınları.

Özsoy Güneş, Z., Çıngıl Barış, Ç., & Kırbaşlar, F. G. (2013). Fen bilgisi öğretmen adaylarının matematik okuryazarlığı öz-yeterlik düzeyleri ile eleştirel düşünme eğilimleri arasındaki ilişkilerin incelenmesi. *Hasan Ali Yücel Eğitim Fakültesi Dergisi, 19*(1), 47-64.

Pajares, F. (1996). Self-efficacy beliefs and mathematical problem solving of gifted students. *Contemporary Educational Psychology*, 21, 325-344

Pajares, F. and Miller, M. D. (1994). The role of self-efficacy and self-concept beliefs in mathematical problem-solving: A path analysis. *Journal of Educational Psychology*, 86, 193-203.

Peker, M. ve Mirasyedioğlu, Ş. (2003). Lise 2.sınıf öğrencilerinin matematik dersine yönelik tutumları ve başarıları arasındaki ilişki. *Pamukkale Üniversitesi Eğitim Fakültesi Dergisi*, Sayı 14, 157-166

Picard, R. W. (1997). *Affective computing*. Massachusetts: MIT Press.

Pigott, T. D. (2001). A review of methods for missing data. *Educational Research and Evaluation, 7*(4), 353-383.

Randhawa, B. S., Beamer, J. E., & Lundberg, I. (1993) 'Role of mathematics self-efficacy in the structural model of mathematics achievement, *Journal of Educational Psychology*, 85 (1), 41-48.

Ratner, B. (2015) *Market segmentation: Defining target markets with CHAID, DM STAT-1*. Retrieved from http://www.dmstat1.com/res/MarketSegmentationWithCHAID.html

Ratner, B. (2007). *Statistical modeling and analysis for database marketing: Effective techniques for mining big data*. Boca Raton, FL: Chapman & Hall/CRC.

Robson, C. (2002). *Real world research: A resource for social scientists and practitioner-researchers*. Massachussetts: Blackwell Publishers.

Saracaloğlu, A. S., Başer, N., Yavuz, G. & Narlı, S., (2004). Öğretmen adaylarının matematiğe yönelik tutumları öğrenme ve ders çalışma stratejileri ile başarıları arasındaki ilişki. *Ege Eğitim Dergisi*, 5 (2), 53-64.

Satıcı, Ö., Akkuş, Z., & Alp, A. (2009). Tıp fakültesi öğretim elemanlarının teknolojiye ilişkin tutumlarının CHAID analizi incelenmesi. *Dicle Tıp Dergisi, 36*(4), 267-274.

Savaş, E., Taş, S., & Duru, A. (2010). Factors affecting students' achievement in mathematics. *İnönü University Journal of the Faculty of Education, 11*(1), 113-132.

Seltman, H. J. (2014). *Experimental design and analysis*. Retrieved from http://www.stat.cmu.edu/~hseltman/309/Book/Book.pdf

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston, MA: Houghton Mifflin.

Shen, C. (2002). Revisiting the relationship between students' achievement and their selfperceptions: A cross-national analysis based on TIMSS 1999 data. *Assessment in Education*, *9*(2), 161-184.

Slack, M. K., & Draugalis, J. R. (2001). Establishing the internal and external validity of experimental studies. *American Journal of Health System Pharmacy, 58*(22), 2173-2184.

Sümer, N. (2000). Yapısal eşitlik modelleri: Temel kavramlar ve örnek uygulamalar. *Türk Psikoloji Yazıları, 3*(6), 49-74.

Şengül Avşar, A., & Yalçın, S. (2015). Öğrencilerin okuma başarılarını açıklayan ailesel değişkenlerin CHAID analizi ile belirlenmesi. *Eğitim ve Bilim*, *40*(179), 1-9.

Şimşek, Ö.F. (2007). *Yapısal eşitlik modellemesine giriş*. Ankara: Ekinoks.

Tan, H. (1972). *Gazete haberleri testi: Geliştirilmesi ve standardizasyonu*. Ankara: TÜBİTAK Yayınları.

Tekin, B., & Tekin, S. (2004). Matematik öğretmen adaylarının matematiksel okuryazarlık düzeyleri üzerine bir araştırma. *Matematik etkinlikleri 2004: Matematik sempozyumu ve sergileri*. Ankara: MATDER.

Uysal, E., & Yenilmez, K. (2011). Sekizinci sınıf öğrencilerinin matematik okuryazarlığı düzeyi. *Eskişehir Osmangazi Üniversitesi Sosyal Bilimler Dergisi*, *12*(2), 1-15.

Wilkinson, L. (1992). *Tree structured data analysis: AID, CHAID and CART*. Chicago: SPSS Inc.

Witten I., Frank E. (2000). *Data mining: Practical machine learning tools and techniques with java implementations.* USA: Morgan Kaufmann Publishers.

Yalçın, M., Aslan, S., & Usta, E. (2012). Analysis of PISA 2009 exam according to some variables. *Mevlana International Journal of Education, 2*(1), 64-71.

Yenilmez, K., & Turgut, M. (2012). Matematik öğretmeni adaylarının matematik okuryazarlığı öz-yeterlik düzeyleri. *Eğitim ve Öğretim Araştırmaları Dergisi*, *1*(2), 253-258.

Yücel, Z., & Koç, M. (2011). İlköğretim öğrencilerinin matematik dersine karşı tutumlarının başarı düzeylerini yordama gücü ile cinsiyet arasındaki ilişki, *İlköğretim Online, 10*(1), 133-143.

Zimmerman, B. J. (2000). Self-efficacy: An essential motive to learn. *Contemporary Educational Psychology, 25*, 82-91.